

A HIGH-ORDER NUMERICAL METHOD FOR THE HELMHOLTZ EQUATION WITH NONSTANDARD BOUNDARY CONDITIONS*

D. S. BRITT[†], S. V. TSYNKOV[‡], AND E. TURKEL[§]

Abstract. We describe a high-order accurate methodology for the numerical simulation of time-harmonic waves governed by the Helmholtz equation. Our approach combines compact finite difference schemes that provide an inexpensive venue toward high-order accuracy with the method of difference potentials developed by Ryaben’kii. The latter can be interpreted as a generalized discrete version of the method of Calderon’s operators in the theory of partial differential equations. The method of difference potentials can accommodate nonconforming boundaries on regular structured grids with no loss of accuracy due to staircasing. It introduces a universal framework for treating boundary conditions of any type. A significant advantage of this method is that changing the boundary condition within a fairly broad variety does not require any major changes to the algorithm and is computationally inexpensive. In this paper, we address various types of boundary conditions using the method of difference potentials. We demonstrate the resulting numerical capabilities by solving a range of nonstandard boundary value problems for the Helmholtz equation. These include problems with variable coefficient Robin boundary conditions (including discontinuous coefficients) and problems with mixed (Dirichlet/Neumann) boundary conditions. In all our simulations, we use a Cartesian grid and a circular boundary curve. For those test cases where the overall solution was smooth, our methodology has consistently demonstrated the design fourth-order rate of grid convergence, whereas when the regularity of the solution was not sufficient, the convergence slowed down, as expected. We also show that every additional boundary condition requires only an incremental additional expense.

Key words. difference potentials, boundary projections, Calderon’s operators, regular grids, curvilinear boundaries, boundary value problems, variable coefficients, Robin boundary conditions, mixed/piecewise boundary conditions, compact differencing, high-order accuracy

AMS subject classifications. 65N06, 65N12, 65N22, 65N99, 65Z05

DOI. 10.1137/120902689

1. Introduction. High-order accuracy is desirable in the numerical solution of wave propagation problems. It is well known that to maintain the same level of error for higher wavenumbers k while solving the Helmholtz equation

$$(1.1) \quad \Delta u + k^2 u = f,$$

one needs to refine the grid faster than the wavelength $\lambda = 2\pi/k$ decreases. More precisely, the quantity $k^{p+1}h^p$ must remain constant, where h is the discretization size and p is the order of accuracy; this is known as the pollution effect [2, 1, 13]. To avoid this difficulty, the number of points per wavelength $\sim 1/hk$ should grow as $k^{1/p}$. The higher the order of accuracy p the slower this growth, which indicates that

*Submitted to the journal’s Methods and Algorithms for Scientific Computing section December 17, 2012; accepted for publication (in revised form) June 17, 2013; published electronically October 1, 2013. This work was supported by the US–Israel Binational Science Foundation (BSF) under grant 2008094 and by the US ARO under grant W911NF-11-1-0384.

<http://www.siam.org/journals/sisc/35-5/90268.html>

[†]Department of Mathematics, North Carolina State University, Raleigh, NC 27695 (dsbritt@ncsu.edu).

[‡]Corresponding author. Department of Mathematics, North Carolina State University, Raleigh, NC 27695, and Moscow Institute of Physics and Technology, Dolgoprudny, 141700, Russia (tsynkov@math.ncsu.edu, <http://www.math.ncsu.edu/~stsynkov/>).

[§]School of Mathematical Sciences, Tel Aviv University, Ramat Aviv, Tel Aviv 69978, Israel (turkel@post.tau.ac.il, <http://www.math.tau.ac.il/~turkel/>).

higher-order methods will be considerably more efficient than lower-order methods for solving (1.1) over domains which are large relative to the wavelength.

Finite difference schemes provide the easiest and least expensive way toward achieving high-order accuracy. Moreover, compact schemes [19, 51, 52, 4, 5, 58] offer the extra benefit of requiring no additional boundary conditions beyond those needed for the differential equation (1.1) itself. These schemes are built on a narrow stencil, and their high-order accuracy is achieved by means of equation-based differencing. However, the use of finite difference schemes may become highly nontrivial, and the approximation accuracy may suffer if the boundary of the domain of interest does not conform to the grid. In the literature, the adverse effect of a nonconforming boundary on the accuracy of a difference scheme is attributed to staircasing [9, 22].

An alternative to finite difference schemes is provided by finite volume, finite element, and boundary element methods (as well as various hybrid techniques). All of these methods have their own advantages and shortcomings, and we refer the reader to [35] for a brief comparative assessment.

An efficient approach that helps remove the geometric limitations pertaining to finite difference schemes is based on the method of difference potentials developed by Ryaben’kii. The theory of difference potentials is related to the theory of Calderon operators; see [8, 49]. A comprehensive account of the method of difference potentials can be found in the monograph [45] (see also [43]), while our recent paper [35] discusses its application to solving the Helmholtz equation. In particular, the algorithm presented in [35] maintains the design high-order accuracy of the chosen compact scheme even when treating nonconforming curvilinear boundaries on regular structured grids. A brief account of the method of difference potentials can also be found in [44, 46], whereas some of its other recent developments are presented in [31, 57, 59, 26, 47, 36].

The method of difference potentials provides a broad range of computational capabilities; see [35, 36]. It can handle differential equations with variable coefficients, which in the context of the Helmholtz equation may imply a variable propagation speed. It allows for both homogeneous and inhomogeneous equations, and either in the case of inhomogeneities or in the case of variable properties of the medium, there is no requirement of having a gap between the region of inhomogeneity/variation and the boundary of the domain. The method also works for curvilinear boundaries and interfaces of an arbitrary shape. In the current paper though we focus only on one particular feature of the method of difference potentials—its universal approach to the treatment of the boundary conditions.

The method is designed so that it does not require the approximation of the boundary conditions on the grid and so is applicable with nonconforming boundaries. Moreover, boundary conditions of any type are allowed—not only the simplest Neumann or Dirichlet, but also, for example, mixed or Robin, even with variable and discontinuous coefficients. In doing so, the core of the numerical algorithm always remains the same. Changes in the boundary conditions are accommodated by making only minor modifications to the computational procedure (unlike methods based on boundary integral equations).

In section 2 we introduce and discuss difference potentials and projections, which can be considered as generalized discrete counterparts of Calderon’s potentials and projections [8, 49]. We show how the finite difference equations on the domain can be reduced to equivalent discrete equations at the boundary. Our main objective is to demonstrate the versatility of the proposed algorithm in treating boundary conditions. Hence, we consider a computational setting which is otherwise very straightforward. Specifically, we use the fourth-order accurate compact scheme of [19, 51] to discretize

and solve the constant coefficient interior homogeneous Helmholtz equation (1.1), $f = 0$, on a uniform Cartesian grid in two space dimensions. The computational domain is chosen as a disk—the simplest nonconforming shape.

In section 3, we provide a detailed account of how various types of boundary conditions can be accommodated by the method of difference potentials. We demonstrate that this approach is completely general, which we illustrate by analyzing specific examples: Robin boundary conditions with variable coefficients and mixed Dirichlet/Neumann boundary conditions. We identify those parts of the overall numerical algorithm that need to be adjusted when changing the boundary conditions. We show that only minor changes are needed in the algorithm to handle changes in the boundary condition.

In section 4, we present the results of the numerical experiments. Our simulations corroborate the theoretical design properties of the algorithm. Specifically, when the solution is sufficiently smooth, the algorithm demonstrates fourth-order grid convergence, whereas when the solution is not sufficiently smooth, the grid convergence predictably slows down. We also show that changing the boundary condition and solving a new problem does not imply any substantial additional costs.

In section 5, we present conclusions and discuss future work.

2. Difference potentials and projections. Let Ω be a bounded domain on the Cartesian plane \mathbb{R}^2 , and let Γ be its boundary, $\Gamma = \partial\Omega$. Consider the following boundary value problem:

$$(2.1a) \quad \mathbf{L}u \stackrel{\text{def}}{=} \Delta u + k^2 u = 0, \quad \mathbf{x} \in \Omega,$$

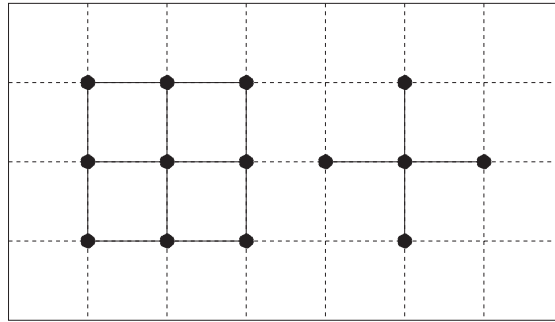
$$(2.1b) \quad \mathbf{l}_\Gamma u = \phi_\Gamma,$$

where $k = \text{const}$ in (2.1a). Problem (2.1) is required to have a unique solution u on Ω for a given ϕ_Γ and be well-posed. We discretize problem (2.1) on a Cartesian grid and solve it with high-order accuracy using the method of difference potentials for the case where Ω is a disk of radius $r = 1$ centered at the origin and Γ is a circle. Note, that (2.1b) is a generic boundary condition that will be specified later. The method allows for a broad variety of boundary conditions (2.1b).

2.1. The scheme. The method of difference potentials can be applied in conjunction with any finite difference scheme. A key advantage of high-order schemes is their improved efficiency in reducing the phase error; see section 1. Therefore, we have chosen to implement a fourth-order accurate approximation. A fourth-order compact scheme for the Helmholtz equation (1.1) was introduced in [19, 51]:

$$(2.2) \quad \begin{aligned} & \frac{1}{h^2} (u_{m+1,n} + u_{m,n+1} + u_{m-1,n} + u_{m,n-1} - 4u_{m,n}) \\ & + \frac{1}{6h^2} [u_{m+1,n+1} + u_{m+1,n-1} + u_{m-1,n+1} - u_{m-1,n-1} + 4u_{m,n} \\ & - 2(u_{m,n+1} + u_{m,n-1} + u_{m+1,n} + u_{m-1,n})] \\ & + \frac{k^2}{12} (u_{m+1,n} + u_{m,n+1} + 8u_{m,n} + u_{m-1,n} + u_{m,n-1}) \\ & = f_{m,n} + \frac{1}{12} (f_{m+1,n} + f_{m,n+1} - 4f_{m,n} + f_{m-1,n} + f_{m,n-1}). \end{aligned}$$

Scheme (2.2) is built on a square-cell Cartesian grid with size h . Its left-hand side involves a 9-node (3×3) compact stencil operating on u , and its right-hand

FIG. 2.1. *Stencils of the compact scheme (2.2).*

side uses a 5-node stencil operating on f ; see Figure 2.1. For the homogeneous Helmholtz equation (2.1a), the finite difference equation (2.2) simplifies and becomes homogeneous,

$$\begin{aligned}
 (2.3) \quad & \frac{1}{h^2} (u_{m+1,n} + u_{m,n+1} + u_{m-1,n} + u_{m,n-1} - 4u_{m,n}) \\
 & + \frac{1}{6h^2} [u_{m+1,n+1} + u_{m+1,n-1} + u_{m-1,n+1} - u_{m-1,n-1} + 4u_{m,n} \\
 & - 2(u_{m,n+1} + u_{m,n-1} + u_{m+1,n} + u_{m-1,n})] \\
 & + \frac{k^2}{12} (u_{m+1,n} + u_{m,n+1} + 8u_{m,n} + u_{m-1,n} + u_{m,n-1}) = 0,
 \end{aligned}$$

so that only a 9-node stencil on the left-hand side is used, while the 5-node stencil on the right-hand side is no longer needed.

Note that in [5] a similar fourth-order accurate scheme is derived for a more general form of the Helmholtz equation that has a variable coefficient Laplace-like term in place of the usual Laplacian and a variable wavenumber k . In [52], a sixth-order accurate scheme is constructed for the constant-coefficient Helmholtz equation using the same 9-node compact stencil on the left-hand side, and in [58] a sixth-order compact scheme is built for the Helmholtz equation with a variable wavenumber k . Hereafter, we restrict the discussion to the constant coefficient case, since our focus is on the treatment of the boundary conditions.

2.2. The auxiliary problem. In order to apply the method of difference potentials, we will embed the domain Ω (i.e., the unit disk) in a larger domain Ω_0 , which we choose to be a square of side length 2.2 centered at the origin. The larger domain Ω_0 will be used to formulate what is known as the auxiliary problem (AP). Following [5], we impose Dirichlet boundary conditions on the upper and lower edges of Ω_0 , and local Sommerfeld-type conditions on the left and right edges:

$$\begin{aligned}
 (2.4) \quad & \mathbf{L}u = f, & \mathbf{x} \in \Omega_0, \\
 & u = 0, & y = \pm 1.1, \\
 & \frac{\partial u}{\partial x} + iku = 0, & x = 1.1, \\
 & \frac{\partial u}{\partial x} - iku = 0, & x = -1.1.
 \end{aligned}$$

When formulating the AP, the only essential requirement is the existence and uniqueness of its solution on Ω_0 for any f and, of course, the well-posedness, i.e., the

continuous dependence of the solution on the data. Other than that, the AP is not going to affect the solution that we obtain inside Ω . We may therefore formulate the AP simply so that it is easy to solve. This is precisely the motivation behind our choice of a square domain, as well as of the boundary conditions in (2.4). Indeed, on one hand the AP (2.4) can be solved efficiently by means of the separation of variables. On the other hand, it is known that the Helmholtz equation is prone to resonances if only Dirichlet boundary conditions are used,¹ whereas Sommerfeld-type conditions of (2.4) make the spectrum complex and hence guarantee the uniqueness. The disadvantage of using Sommerfeld type conditions is that they introduce complex quantities into the calculation, which is not always necessary for interior problems (unlike for the exterior problems). Alternatively, one could use a Dirichlet or real Robin condition that was carefully chosen to avoid zero (or very small) eigenvalues and would keep the solution real; this approach was adopted for some of the computations in [35].

Applying the compact scheme of section 2.1 to the differential equation $\mathbf{L}u = f$ of (2.4), we have

$$(2.5a) \quad \begin{aligned} & \frac{1}{h^2} (u_{m+1,n} + u_{m,n+1} + u_{m-1,n} + u_{m,n-1} - 4u_{m,n}) \\ & + \frac{1}{6h^2} [u_{m+1,n+1} + u_{m+1,n-1} + u_{m-1,n+1} - u_{m-1,n-1} + 4u_{m,n} \\ & - 2(u_{m,n+1} + u_{m,n-1} + u_{m+1,n} + u_{m-1,n})] \\ & + \frac{k^2}{12} (u_{m+1,n} + u_{m,n+1} + 8u_{m,n} + u_{m-1,n} + u_{m,n-1}) = g_{m,n}, \end{aligned}$$

where we can formally think that [cf. formula (2.2)]

$$g_{m,n} = f_{m,n} + \frac{1}{12} (f_{m+1,n} + f_{m,n+1} - 4f_{m,n} + f_{m-1,n} + f_{m,n-1}).$$

We emphasize, however, that the actual governing equation (2.1a) is homogeneous, whereas the right-hand side f in (2.4) plays only an auxiliary role and does not represent any physical source term. As such, in our subsequent analysis the explicit form of $f_{m,n}$ will never be needed. What will rather be important for constructing the difference potentials and projections is the final discrete right-hand side $g_{m,n}$ of (2.5a). This right-hand-side g will be obtained directly, i.e., without having to relate it to any f by means of the 5-node stencil. The expression for g is given in (2.6) of section 2.3, and it is precisely this right-hand side that allows us to compute the discrete counterparts to Calderon's operators and equivalently reduce the governing equation from the domain to the boundary; see formula (2.8). We also note that even though we keep the right-hand side g as a key innate element of the method of difference potentials, the actual physical solutions that we obtain inside Ω in the form of difference potentials are those to the homogeneous finite difference equation (2.3). Their accuracy matches the design fourth-order accuracy of the scheme, as corroborated by the numerical simulations of section 4.

In order to maintain the overall accuracy of the solution, it is also important to have the boundary conditions of (2.4) approximated with fourth-order accuracy. This is trivial for the Dirichlet conditions:

$$(2.5b) \quad u_{m,0} = u_{m,N} = 0, \quad m = 0, \dots, M,$$

¹The Helmholtz equation is said to be at a resonance on Ω_0 if $-k^2$ is an eigenvalue of the Laplacian subject to zero Dirichlet boundary conditions at $\partial\Omega_0$. In this case, the solution to the Helmholtz equation is not unique.

where the grid nodes range between 0 and M in the x direction and between 0 and N in the y direction. (For square cells and square domain Ω_0 , obviously $M = N$.)

Obtaining a fourth-order accurate compact approximation of the Sommerfeld-type conditions requires some additional effort. In [5], we have derived such an approximation at the boundary (more precisely, at the midpoint of the outermost cell) for the Helmholtz equation with variable coefficients. Simplifying the result for constant coefficients, we arrive at the following discrete Sommerfeld-type conditions for the right and left edges of Ω_0 :

$$\begin{aligned}
 (2.5c) \quad & \left(\frac{u_{M,n} - u_{M-1,n}}{h} - \frac{1}{6h}(u_{M,n+1} - u_{M-1,n+1} + u_{M,n-1} \right. \\
 & \quad \left. - u_{M-1,n-1} - 2(u_{M,n} - u_{M-1,n})) - \frac{k^2 h}{24}(u_{M,n} - u_{M-1,n}) \right) \\
 & + ik \left(\frac{u_{M,n} + u_{M-1,n}}{2} + \frac{h^2 k^2}{8} u_{M-\frac{1}{2},n} \right. \\
 & \quad \left. + \frac{u_{M-\frac{1}{2},n+1} - 2u_{M-\frac{1}{2},n-1} + u_{M-\frac{1}{2},n-1}}{2} \right) = 0, \\
 (2.5d) \quad & \left(\frac{u_{1,n} - u_{0,n}}{h} - \frac{1}{6h}(u_{1,n+1} - u_{0,n+1} + u_{1,n-1} - u_{0,n-1} - 2(u_{1,n} - u_{0,n})) \right. \\
 & \quad \left. - \frac{k^2 h}{24}(u_{1,n} - u_{0,n}) \right) \\
 & - ik \left(\frac{u_{1,n} + u_{0,n}}{2} + \frac{h^2 k^2}{8} u_{\frac{1}{2},n} + \frac{u_{\frac{1}{2},n+1} - 2u_{\frac{1}{2},n-1} + u_{\frac{1}{2},n-1}}{2} \right) = 0.
 \end{aligned}$$

As in the case of the continuous AP (2.4), the overall discrete AP (2.5) is supposed to have a unique solution $u_{m,n}$, $m = 0, \dots, M$, $n = 0, \dots, N$, for any right-hand-side $g_{m,n}$ (defined on the interior subgrid $m = 1, \dots, M - 1$, $n = 1, \dots, N - 1$), and be well-posed. The discrete AP (2.5) can be solved by a sine FFT in the y direction combined with the tridiagonal elimination in the x direction. The complexity of this solution is log-linear with respect to the grid dimension N and linear with respect to M .

2.3. Grid sets and operators. Let \mathbb{N}_0 be a uniform Cartesian grid on the square Ω_0 with size h in both the x and y directions, and let $\mathbb{M}_0 \subset \mathbb{N}_0$ be the set of its interior nodes, i.e., all nodes of \mathbb{N}_0 except those on the edges of Ω_0 ; see Figure 2.2. The solution u to the discrete AP (2.5) will be defined on the grid \mathbb{N}_0 , while its right-hand-side g will be defined on the grid \mathbb{M}_0 .

Let $\mathbb{M}^+ \subset \mathbb{M}_0$ be the set of nodes of \mathbb{M}_0 that are inside the physical domain Ω , i.e., $\mathbb{M}^+ = \mathbb{M}_0 \cap \Omega$; see Figure 2.3(a). Since $\Gamma = \partial\Omega$ is not aligned with the grid, we define its discrete analogue, γ , which we refer to as the grid boundary. Let \mathbb{M}^- be the set of all nodes of \mathbb{M}_0 that lie outside Ω , i.e., $\mathbb{M}^- = \mathbb{M}_0 \setminus \mathbb{M}^+$; see Figure 2.3(b). Let \mathbb{N}^+ and \mathbb{N}^- be defined as the sets of nodes of \mathbb{N}_0 that are used when applying the 3×3 compact stencil (see Figure 2.1) to the nodes of \mathbb{M}^+ and \mathbb{M}^- , respectively. There will be an overlap in the sets \mathbb{N}^+ and \mathbb{N}^- when applying the stencil to nodes of \mathbb{M}^+ and \mathbb{M}^- which are adjacent to Γ . We refer to this overlap, i.e., the intersection of these sets, as the grid boundary, $\gamma = \mathbb{N}^+ \cap \mathbb{N}^-$; see Figure 2.3(c).

The solution of the discrete AP (2.5) plays a key role in the construction of the difference potentials and projections, which can be considered as discrete counterparts of Calderon’s potentials and boundary projections (pseudodifferential operators); see

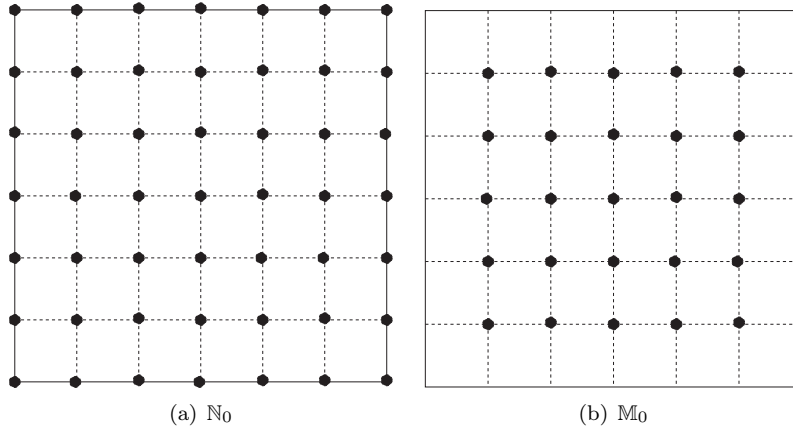


FIG. 2.2. Cartesian grids for the finite difference equation (2.5a).

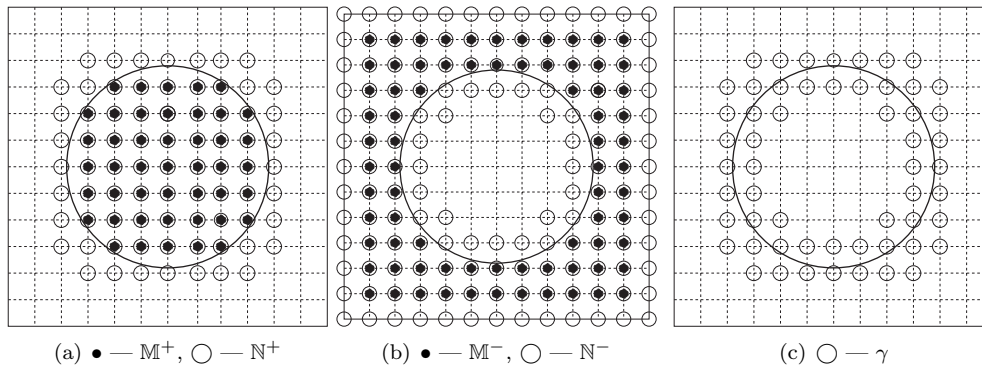


FIG. 2.3. Interior and exterior grid subsets and the grid boundary.

[8, 49, 45]. The difference potential will approximate the solution u of boundary value problem (2.1) on the grid N^+ . The density of the difference potential is a grid function defined on the grid boundary γ . It satisfies a special system of linear algebraic equations called the boundary equation with projection (BEP).

We denote the discrete operator on the left-hand side of (2.5a) by $L^{(h)}$. Then the discrete AP (2.5) consists of solving the finite difference equation $L^{(h)}u = g$ on the grid N_0 , subject to boundary conditions (2.5b), (2.5c), and (2.5d). Define the corresponding inverse operator $G^{(h)}$ as the solution of the discrete AP (2.5), so that $u = G^{(h)}g$. Also consider a grid function ξ_γ specified on the discrete boundary γ , Figure 2.3(c). The difference potential with density ξ_γ is given by

$$(2.6) \quad P_{N^+} \xi_\gamma \stackrel{\text{def}}{=} w - G^{(h)} \left(L^{(h)} w \Big|_{M^+} \right), \quad \text{where } w = \begin{cases} \xi_\gamma & \text{on } \gamma, \\ 0 & \text{on } N_0 \setminus \gamma. \end{cases}$$

The operation $L^{(h)}w \Big|_{M^+} := g$ in formula (2.6) denotes first the application of the operator $L^{(h)}$ to the auxiliary function w and then truncation of the grid function $L^{(h)}w$ to the grid M^+ ; see Figure 2.3(a). The difference potential $P_{N^+} \xi_\gamma$ is defined

on the grid \mathbb{N}^+ (that’s why we are using the subscript \mathbb{N}^+),² and at the nodes \mathbb{M}^+ it satisfies the homogeneous finite difference equation [cf. formula (2.3)]

$$\mathbf{L}^{(h)}(\mathbf{P}_{\mathbb{N}^+}\xi_\gamma) = 0.$$

Along with the grid function ξ_γ , consider a two-component vector function $\boldsymbol{\xi}_\Gamma = (\xi_0, \xi_1)$ defined on the continuous boundary Γ . This new function $\boldsymbol{\xi}_\Gamma$ can be used as the density of the genuine Calderon potential of the differential operator \mathbf{L} on the domain Ω [8, 49], in which ξ_0 is interpreted as the Dirichlet data and ξ_1 is interpreted as the Neumann data at the boundary Γ . Provided that the grid function ξ_γ is related to the continuous function $\boldsymbol{\xi}_\Gamma$ in some special sense (ξ_γ must be obtained from $\boldsymbol{\xi}_\Gamma$ as an equation-based extension based on the Taylor formula of order four; see section 3), the difference potential (2.6) approximates the continuous Calderon potential with density $\boldsymbol{\xi}_\Gamma$ with fourth-order accuracy on the grid \mathbb{N}^+ (the design accuracy of our compact scheme); see [35, 45, 39].

Truncating the difference potential (2.6) to γ , we obtain the difference projection,

$$(2.7) \quad \mathbf{P}_\gamma \xi_\gamma \stackrel{\text{def}}{=} (\mathbf{P}_{\mathbb{N}^+} \xi_\gamma)|_\gamma,$$

and then express the discrete BEP as

$$(2.8) \quad \mathbf{P}_\gamma \xi_\gamma = \xi_\gamma.$$

Its pivotal property (see [45]) is that those and only those grid functions ξ_γ satisfy the BEP (2.8) that can be obtained as the truncation to γ of a solution u (defined on \mathbb{N}^+ , Figure 2.3(a)) of the homogeneous difference equation (2.3): $\mathbf{L}^{(h)}u = 0$. Thus, the BEP (2.8) provides an equivalent reduction of the discrete equation $\mathbf{L}^{(h)}u = 0$ from the grid domain \mathbb{N}^+ to the grid boundary γ . If the grid function u satisfies $\mathbf{L}^{(h)}u = 0$, then its truncation $\xi_\gamma = u|_\gamma$ must satisfy the BEP (2.8). Conversely, if the grid function ξ_γ satisfies the BEP (2.8), then there exists a function u defined on \mathbb{N}^+ such that $\mathbf{L}^{(h)}u = 0$ and $u|_\gamma = \xi_\gamma$. In fact, this u is given by the difference potential (2.6): $u = \mathbf{P}_{\mathbb{N}^+}\xi_\gamma$.

Note that the difference potential (2.6) and, consequently, the projection (2.7) depend on the choice of the AP since changing the AP will change the inverse operator $\mathbf{G}^{(h)}$. However, the change of the AP does not affect the range of the projection \mathbf{P}_γ , i.e., it does not change the set of solutions to the BEP (2.8), as long as the AP remains uniquely solvable. In other words, when changing the AP one only changes the projection angle onto the same subspace; see [45].

3. Treatment of the boundary conditions. In this section, we show how to account for the given boundary condition (2.1b) in order to approximately reconstruct the data $(u, \frac{\partial u}{\partial n})|_\Gamma$ of the solution u to problem (2.1) at the boundary Γ and subsequently obtain the discrete solution of (2.1) on the grid \mathbb{N}^+ in the form of the difference potential (2.6). First, we need to build the equation-based extension of an arbitrary pair of functions, $\boldsymbol{\xi}_\Gamma = (\xi_0, \xi_1)|_\Gamma$, from the continuous boundary Γ to the grid boundary γ .

3.1. Equation-based extension to the grid boundary. The extension of a given $\boldsymbol{\xi}_\Gamma = (\xi_0, \xi_1)|_\Gamma$ from the smooth boundary Γ to the grid nodes γ (specifically the nodes adjacent to Γ ; see Figure 2.3(c)) is constructed using a truncated Taylor

²Even though both the auxiliary function w and the solution $\mathbf{G}^{(h)}g$ to the discrete AP (2.5) are defined on the entire grid \mathbb{N}_0 , the difference potential $\mathbf{P}_{\mathbb{N}^+}\xi_\gamma$ as introduced by formula (2.6) is of interest to us only on the grid \mathbb{N}^+ .

expansion with differentiation in the direction normal to Γ . Consider ξ_Γ as the data of some function $v = v(x, y)$,

$$(\xi_0, \xi_1)|_\Gamma = \left(v, \frac{\partial v}{\partial \mathbf{n}} \right) \Big|_\Gamma,$$

that we define near the curve Γ by means of the Taylor expansion:

$$(3.1) \quad v \stackrel{\text{def}}{=} v_\Gamma + \rho \frac{\partial v}{\partial \mathbf{n}} \Big|_\Gamma + \frac{\rho^2}{2} \frac{\partial^2 v}{\partial \mathbf{n}^2} \Big|_\Gamma + \frac{\rho^3}{6} \frac{\partial^3 v}{\partial \mathbf{n}^3} \Big|_\Gamma + \frac{\rho^4}{24} \frac{\partial^4 v}{\partial \mathbf{n}^4} \Big|_\Gamma.$$

In formula (3.1), ρ denotes the distance (with sign) from a given point near Γ to the curve Γ . We emphasize that while formula (3.1) takes the usual form of a Taylor approximation to the function v , it should not be interpreted this way. Instead, it should be thought of as the definition of v . The new function v can be evaluated at any point (x, y) which is sufficiently close to Γ . In particular, we call this new function ξ_γ when its domain is restricted to the nodes of the grid boundary γ (see Figure 2.3(c)):

$$(3.2) \quad \xi_\gamma \stackrel{\text{def}}{=} v|_\gamma.$$

However, the foregoing definition is not complete until we specify how to compute the normal derivatives of order 2 and higher required for formula (3.1). These will be obtained using equation-based differentiation applied to the homogeneous Helmholtz equation with constant coefficients (2.1a). (The extension to nonhomogeneous equations and variable coefficients is straightforward [35].) In doing so, we assume that the “input” functions v and $\frac{\partial v}{\partial \mathbf{n}}$ are known analytically on the circle Γ so that we can readily compute their tangential derivatives. Since we consider the simple case of a circular boundary Γ centered at the origin, the outward normal to Γ and the direction of the polar radius coincide. Hence, it is convenient to recast (2.1a) using polar coordinates (r, θ) :

$$(3.3) \quad \frac{1}{r} \frac{\partial v}{\partial r} + \frac{\partial^2 v}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2 v}{\partial \theta^2} + k^2 v = 0.$$

Equation (3.3) allows us to obtain the second derivative of v with respect to r :

$$(3.4) \quad \frac{\partial^2 v}{\partial r^2} = - \left(\frac{1}{r} \frac{\partial v}{\partial r} + \frac{1}{r^2} \frac{\partial^2 v}{\partial \theta^2} + k^2 v \right).$$

Recall that v and $\frac{\partial v}{\partial \mathbf{n}} = \frac{\partial v}{\partial r}$ are given on Γ and that $\frac{\partial^2 v}{\partial \theta^2}$ can be computed analytically as the second tangential derivative of the given function v . Hence, (3.4) allows us to compute the term $\frac{\partial^2 v}{\partial \mathbf{n}^2} = \frac{\partial^2 v}{\partial r^2}$ in the Taylor expansion (3.1).

We proceed to find the remaining normal derivatives, $\frac{\partial^3 v}{\partial \mathbf{n}^3} = \frac{\partial^3 v}{\partial r^3}$ and $\frac{\partial^4 v}{\partial \mathbf{n}^4} = \frac{\partial^4 v}{\partial r^4}$, via equation-based differentiation. In particular, we take the derivative of (3.4) with respect to r to obtain

$$(3.5) \quad \frac{\partial^3 v}{\partial r^3} = \frac{1}{r^2} \frac{\partial v}{\partial r} - \frac{1}{r} \frac{\partial^2 v}{\partial r^2} + \frac{2}{r^3} \frac{\partial^2 v}{\partial \theta^2} - \frac{1}{r^2} \frac{\partial^3 v}{\partial r \partial \theta^2} - k^2 \frac{\partial v}{\partial r}.$$

We are able to evaluate $\frac{\partial^3 v}{\partial r^3}$ using the given function $\frac{\partial v}{\partial r}$, the analytically computed second tangential derivative of $\frac{\partial v}{\partial r}$ for $\frac{\partial^3 v}{\partial r \partial \theta^2}$, and the representation (3.4) for $\frac{\partial^2 v}{\partial r^2}$. This way, we can compute the third normal derivative term of the Taylor expansion (3.1). To compute the next term, we differentiate equality (3.5) with respect to r :

$$(3.6) \quad \frac{\partial^4 v}{\partial r^4} = -\frac{2}{r^3} \frac{\partial v}{\partial r} + \frac{2}{r^2} \frac{\partial^2 v}{\partial r^2} - \frac{1}{r} \frac{\partial^3 v}{\partial r^3} - \frac{6}{r^4} \frac{\partial^2 v}{\partial \theta^2} + \frac{4}{r^3} \frac{\partial^3 v}{\partial r \partial \theta^2} - \frac{1}{r^2} \frac{\partial^4 v}{\partial r^2 \partial \theta^2} - k^2 \frac{\partial^2 v}{\partial r^2}.$$

Again, the term $\frac{\partial v}{\partial r}$ is given, the terms $\frac{\partial^2 v}{\partial r^2}$ and $\frac{\partial^3 v}{\partial r^3}$ are evaluated via (3.4) and (3.5), respectively, and the terms $\frac{\partial^2 v}{\partial \theta^2}$ and $\frac{\partial^3 v}{\partial r \partial \theta^2}$ are computed analytically as the tangential derivatives of the given pair of functions $(v, \frac{\partial v}{\partial r})$. The only remaining term of (3.6) that has not been accounted for yet is $\frac{\partial^4 v}{\partial r^2 \partial \theta^2}$. We evaluate it by differentiating equation (3.4) twice with respect to θ :

$$(3.7) \quad \frac{\partial^4 v}{\partial r^2 \partial \theta^2} = -\left(\frac{1}{r} \frac{\partial^3 v}{\partial r \partial \theta^2} + \frac{1}{r^2} \frac{\partial^4 v}{\partial \theta^4} + k^2 \frac{\partial^2 v}{\partial \theta^2} \right).$$

Note that all the terms required to compute $\frac{\partial^4 v}{\partial r^2 \partial \theta^2}$ by (3.7) are tangential (i.e., angular) derivatives of v and $\frac{\partial v}{\partial r}$. Therefore, substituting (3.7) for $\frac{\partial^4 v}{\partial r^2 \partial \theta^2}$ into (3.6) completes our ability to calculate the fourth normal derivative $\frac{\partial^4 v}{\partial r^4} = \frac{\partial^4 v}{\partial n^4}$ in the Taylor expansion (3.1).

Thus, given an arbitrary pair of functions $(v, \frac{\partial v}{\partial r})$ defined along the circle Γ , a fifth-order extension from the continuous circle to a nearby grid node along the normal direction is accomplished by substituting equations (3.4)–(3.7) into the Taylor expansion (3.1). In [35, Appendix A], we show how a similar extension can be built in the case of an arbitrary smooth curve Γ .

Hereafter, we use the notation \mathbf{Ex} for the equation-based extension operator defined by formulae (3.1), (3.2). It will act on an arbitrary pair of continuous functions defined on Γ : $\xi_\Gamma = (\xi_0, \xi_1)|_\Gamma$. This operator uses the truncated Taylor expansion (3.1) to construct a new function $v(x, y)$ near Γ , which is then sampled at the grid boundary γ according to (3.2). This yields the grid function that we refer to as ξ_γ :

$$\xi_\gamma = \mathbf{Ex} \xi_\Gamma = \mathbf{Ex} (\xi_0, \xi_1) |_\gamma.$$

We emphasize that while the normal derivatives of order 2 and higher in formula (3.1) are obtained by differentiation based on the Helmholtz (2.1a), the construction of the operator \mathbf{Ex} permits it to be applied to an arbitrary pair of functions. Thus, ξ_Γ does not need to represent the data of a solution to equation (2.1a) in order to apply the operator. However, if ξ_Γ happens to be the data of a solution u to (2.1a) on Ω , then formula (3.1) approximates this solution near $\Gamma = \partial\Omega$ with fifth-order accuracy with respect to n , and, in particular, it shall do so at the nodes of the discrete boundary γ .

Note that to ensure the design convergence rate of the overall method, the number of terms in the Taylor formula (3.1) that defines the operator \mathbf{Ex} must match in a particular way the accuracy of the finite difference scheme employed inside the computational domain. The corresponding relations follow from the approximation theorems of the continuous potentials of elliptic operators by difference potentials proved by Reznik [39, 40]. While the actual proofs are delicate, the end results can be used in their own right, and we reproduce some of those results in [35, section 4.4]. Moreover, in practice it turns out that taking fewer terms in (3.1) than prescribed by [39] may be sufficient. For the fourth-order accurate scheme (2.3) that we have implemented for the current study, the original Reznik theorem would suggest taking six terms, which is the sum of the accuracy of the scheme (fourth-order) and the order of the differential equation (second). However, it appears sufficient to truncate the Taylor expansion after the fourth derivative term to preserve the overall fourth-order accuracy. For other finite difference schemes, the number of terms in the expansion will need to be chosen accordingly; see [39], [40], and [35] for more detail.

3.2. Series representation of boundary functions. The next step is to select a basis for the space of smooth pairs of functions, ξ_Γ , on Γ :

$$(3.8) \quad \psi_n^{(0)} = (\psi_n^{(0)}, 0) \quad \text{and} \quad \psi_n^{(1)} = (0, \psi_n^{(1)}), \quad n = -\infty, \dots, \infty.$$

This basis will help us represent the data $(u, \frac{\partial u}{\partial n})|_\Gamma$ of the solution u to problem (2.1); see section 3.3. Provided that the expansion of a given ξ_Γ with respect to system (3.8) converges sufficiently fast, we may truncate it and replace the infinite series with a finite sum:

$$(3.9) \quad \xi_\Gamma = \underbrace{\sum_{n=-N}^N c_n^{(0)} \psi_n^{(0)}}_{(\xi_{0,0})} + \underbrace{\sum_{n=-N}^N c_n^{(1)} \psi_n^{(1)}}_{(0, \xi_1)},$$

where the number N that would guarantee the desired accuracy can be taken relatively small. This is the case, e.g., for the Fourier series when ξ_Γ are smooth periodic functions; see section 3.4 and also [35]. However, a different basis (3.8) may also be chosen, and the particular cases of both Fourier and Chebyshev bases are invoked and discussed in subsequent examples to suit the goal of solving specific nonstandard boundary value problems (2.1). In addition, we emphasize that the basis (3.8) may be selected independently of the discretization grid \mathbb{N}_0 , and this is accomplished by choosing the accuracy of representation (3.9) ahead of time so that it will match or exceed the accuracy that can be achieved on the grid (see [35] for details).

In the interest of solving problem (2.1), we will first apply the operator $\mathbf{E}x$ of section 3.1 and extend an arbitrary ξ_Γ represented by the truncated series (3.9) from the continuous boundary Γ to the grid boundary γ . Then we will substitute the resulting ξ_γ into the discrete BEP (2.8). This will yield a linear system in which the unknowns will be the coefficients $c_n^{(0)}$ and $c_n^{(1)}$, $n = -N, \dots, N$.

3.3. Applying the extension to form a linear system. Applying the extension operator $\mathbf{E}x$ to the series representation (3.9) of a pair of functions defined on the boundary, ξ_Γ , and noting that the operator is linear, we obtain

$$(3.10) \quad \mathbf{E}x \xi_\Gamma = \mathbf{E}x \left(\sum_{n=-N}^N c_n^{(0)} \psi_n^{(0)} + \sum_{n=-N}^N c_n^{(1)} \psi_n^{(1)} \right) = \sum_{n=-N}^N c_n^{(0)} \mathbf{E}x \psi_n^{(0)} + \sum_{n=-N}^N c_n^{(1)} \mathbf{E}x \psi_n^{(1)},$$

where $\mathbf{E}x \psi_n^{(0)} = \mathbf{E}x(\psi_n^{(0)}, 0)$ and $\mathbf{E}x \psi_n^{(1)} = \mathbf{E}x(0, \psi_n^{(1)})$ according to (3.8). Substituting expression (3.10) into the discrete BEP (2.8) yields a system of linear algebraic equations,

$$\sum_{n=-N}^N c_n^{(0)} P_\gamma \mathbf{E}x \psi_n^{(0)} + \sum_{n=-N}^N c_n^{(1)} P_\gamma \mathbf{E}x \psi_n^{(1)} = \sum_{n=-N}^N c_n^{(0)} \mathbf{E}x \psi_n^{(0)} + \sum_{n=-N}^N c_n^{(1)} \mathbf{E}x \psi_n^{(1)},$$

which we formalize by gathering the corresponding basis terms on the left-hand side:

$$(3.11) \quad \sum_{n=-N}^N c_n^{(0)} (P_\gamma - I_\gamma) \mathbf{E}x \psi_n^{(0)} + \sum_{n=-N}^N c_n^{(1)} (P_\gamma - I_\gamma) \mathbf{E}x \psi_n^{(1)} = 0.$$

In formula (3.11), I_γ is the identity operator in the space of grid functions ξ_γ defined on γ . In matrix form, the linear system (3.11) can be recast as

$$(3.12) \quad Qc = 0,$$

where the matrix \mathbf{Q} is given by

$$(3.13) \quad \mathbf{Q} = \underbrace{\left[(\mathbf{P}_\gamma - \mathbf{I}_\gamma) \mathbf{E} \mathbf{x} \psi_{-N}^{(0)}, \dots, (\mathbf{P}_\gamma - \mathbf{I}_\gamma) \mathbf{E} \mathbf{x} \psi_N^{(0)} \right]}_{\mathbf{Q}_0}, \\ \underbrace{\left[(\mathbf{P}_\gamma - \mathbf{I}_\gamma) \mathbf{E} \mathbf{x} \psi_{-N}^{(1)}, \dots, (\mathbf{P}_\gamma - \mathbf{I}_\gamma) \mathbf{E} \mathbf{x} \psi_N^{(1)} \right]}_{\mathbf{Q}_1},$$

and \mathbf{c} is a vector of unknown coefficients, $\mathbf{c} = [c_{-N}^{(0)}, \dots, c_N^{(0)}, c_{-N}^{(1)}, \dots, c_N^{(1)}]^T$. The dimension of the matrix \mathbf{Q} in (3.13) is $|\gamma| \times 2(2N + 1)$, where $|\gamma|$ is the total number of nodes in the grid boundary γ , and the dimension of the vector \mathbf{c} is $2(2N + 1)$. The first $2N + 1$ columns of the matrix \mathbf{Q} form the submatrix \mathbf{Q}_0 and correspond to the coefficients $c_n^{(0)}$, $n = -N, \dots, N$, while the last $2N + 1$ columns of \mathbf{Q} form the submatrix \mathbf{Q}_1 and correspond to the coefficients $c_n^{(1)}$, $n = -N, \dots, N$.

Any solution $\mathbf{c} = [\mathbf{c}^{(0)}, \mathbf{c}^{(1)}]$ to the linear system (3.12) furnishes $\boldsymbol{\xi}_\Gamma$ via formula (3.9), and the latter, in turn, yields $\xi_\gamma = \mathbf{E} \mathbf{x} \boldsymbol{\xi}_\Gamma$. As shown in [39], the corresponding difference potential (2.6) with the density ξ_γ provides a fourth-order accurate approximation to the continuous Calderon potential of the Helmholtz operator \mathbf{L} with the density $\boldsymbol{\xi}_\Gamma$; see also [45, Part III, Chapter 1] and [35, section 4.4]. Moreover, the continuous Calderon potential u solves the homogeneous Helmholtz equation (2.1a) on Ω , and the density $\boldsymbol{\xi}_\Gamma$ of (3.9) approximates its data $(u, \frac{\partial u}{\partial n})|_\Gamma$ on Γ .

However, the linear system (3.12) may have multiple solutions. Indeed, it does not take into account the boundary condition (2.1b) because it is derived from the discrete BEP (2.8) only. To take the boundary condition into account and make sure the overall solution is unique, system (3.12) needs to be modified and/or supplemented by additional equations. The simplest cases to analyze are those of the Dirichlet and Neumann boundary conditions; see [35].

When the boundary condition is Dirichlet, equality (2.1b) reduces to $u|_\Gamma = \phi_\Gamma$, and we can expand the given Dirichlet data function ϕ_Γ (i.e., the first component of the data) with respect to the chosen basis (3.8), yielding the coefficients $c_n^{(0)}$, $n = -N, \dots, N$. Recalling that $\mathbf{c} = [\mathbf{c}^{(0)}, \mathbf{c}^{(1)}] = [c_{-N}^{(0)}, \dots, c_N^{(0)}, c_{-N}^{(1)}, \dots, c_N^{(1)}]^T$ in (3.12), the vector $\mathbf{c}^{(0)}$ in the Dirichlet case can be considered given while the vector $\mathbf{c}^{(1)}$ is unknown. Accordingly, system (3.12) is recast as $\mathbf{Q}_1 \mathbf{c}^{(1)} = -\mathbf{Q}_0 \mathbf{c}^{(0)}$, where the right-hand side can now be thought of as a given vector of dimension $|\gamma|$. Then, the system is solved for $\mathbf{c}^{(1)}$ in the sense of the least squares.

In doing so, we choose the number of basis functions N independent of the size of the discretization grid \mathbb{N}_0 . This number is rather fixed ahead of time so that the accuracy of the truncated expansion (3.9) at the boundary would exceed any accuracy that one might expect to obtain on the grid. This is easy to achieve at a moderate cost, because when the boundary data are smooth and periodic, their Fourier expansion converges rapidly (see footnote 3) and the resulting N appears not very large. Once the dimension N of the boundary representation (3.9) has been fixed, the final accuracy of the solution on the domain is controlled only by the size of the grid \mathbb{N}_0 . For sufficiently fine grids, one should typically expect $|\gamma| \gg 2N + 1$. Hence, the system $\mathbf{Q}_1 \mathbf{c}^{(1)} = -\mathbf{Q}_0 \mathbf{c}^{(0)}$ is overdetermined and admits a robust solution by least squares. Moreover, as the original boundary value problem (2.1) has a unique solution, the discrete least squares solution is “almost classical” in the sense that as the grid is refined, the residual at the minimum decreases to zero with the rate determined by the accuracy of the finite difference approximation.

In the Neumann case, conversely, the vector $\mathbf{c}^{(1)}$ is given by expanding the boundary condition $\frac{\partial u}{\partial \mathbf{n}} = \phi|_{\Gamma}$ with respect to the basis (3.8), the vector $\mathbf{c}^{(0)}$ is unknown, and the system $\mathbf{Q}_0 \mathbf{c}^{(0)} = -\mathbf{Q}_1 \mathbf{c}^{(1)}$ is solved in the sense of the least squares for $\mathbf{c}^{(0)}$. In any case, once both $\mathbf{c}^{(0)}$ and $\mathbf{c}^{(1)}$ are known, the vector $\mathbf{c} = [\mathbf{c}^{(0)}, \mathbf{c}^{(1)}]$ is substituted into (3.9), and the resulting $\boldsymbol{\xi}_{\Gamma}$ is extended to γ via the extension operator: $\xi_{\gamma} = \mathbf{E}\mathbf{x}\boldsymbol{\xi}_{\Gamma}$. The difference potential (2.6) with density ξ_{γ} is then computed to approximate the solution u to the boundary value problem (2.1) with fourth-order accuracy on the grid \mathbb{N}^+ . A more detailed analysis of both the Dirichlet and Neumann boundary conditions, along with the corresponding numerical results that corroborate the grid convergence with the design rate of the scheme, can be found in [35], wherein both a circular and an elliptical boundary are considered with a Cartesian grid.

Our work [35] also provides an analysis of the Robin boundary condition, but only when the coefficients of the boundary condition are constant, and with no numerical computations. In sections 3.4 and 3.5, we present a more comprehensive analysis of the Robin boundary condition, which includes variable and/or discontinuous coefficients. In particular, this allows us to consider mixed boundary conditions, e.g., Dirichlet on one part of the boundary and Neumann on the other part of the boundary.

Finally, we emphasize that the reduction of problem (2.1) from its domain to the boundary based on Calderon's boundary equations with projections, regardless of the type of the boundary condition (2.1b), is always well-posed as long as the original problem (2.1) itself is well-posed; see [45, 35]. This is in contrast to methods based on boundary integral equations, for which care must be exercised, on a case-by-case basis, in choosing the equivalent boundary sources such that the resulting Fredholm integral equation is of the second kind (well-posed) rather than the first kind (ill-posed).

3.4. The Robin boundary condition with smooth variable coefficients.

Consider the case that formula (2.1b) represents a general Robin boundary condition with variable coefficients:

$$(3.14) \quad \alpha(\theta)u(\theta) + \beta(\theta)\frac{\partial u}{\partial \mathbf{n}}(\theta) = \phi(\theta),$$

where α , β , and ϕ are smooth periodic functions of the polar angle $\theta \in [0, 2\pi]$. We will expand each term of (3.14) with respect to the chosen basis (3.8) and obtain a set of linear algebraic equations that will supplement system (3.12). In doing so, it will be convenient, though not necessary, to consider the same basis functions for both u and $\frac{\partial u}{\partial \mathbf{n}}$, so that $\psi_n^{(0)} = \psi_n^{(1)}$, $n = -N, \dots, N$, in formula (3.8).

Since formula (3.14) is composed of smooth 2π -periodic functions, it is natural to choose an exponential Fourier basis, $\psi_n^{(0)}(\theta) = \psi_n^{(1)}(\theta) = e^{in\theta}$. To express the left-hand side of (3.14) in this basis, we will use a well-known convolution formula for the Fourier coefficients of a product of two functions. Let $f(\theta)$ and $g(\theta)$ be 2π -periodic, and denote by \hat{f}_n and \hat{g}_n their Fourier coefficients for the expansion with respect to the complex exponentials $e^{in\theta}$, $n = 0, \pm 1, \pm 2, \dots$. Then, it is easy to show that

$$(3.15) \quad (\widehat{fg})_n = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \hat{g}_m \hat{f}_{m-n}.$$

We now expand the boundary condition (3.14) using this result. Let $c_n^{(0)}$ and $c_n^{(1)}$ represent the Fourier coefficients of $u(\theta)$ and $\frac{\partial u}{\partial \mathbf{n}}(\theta)$, respectively, and let $\hat{\alpha}_n$, $\hat{\beta}_n$, and $\hat{\phi}_n$ be the coefficients of $\alpha(\theta)$, $\beta(\theta)$, and $\phi(\theta)$, respectively. Then, according to (3.15), formula (3.14) becomes

$$\frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \left(\sum_{m=-\infty}^{\infty} \hat{\alpha}_m c_{m-n}^{(0)} + \sum_{m=-\infty}^{\infty} \hat{\beta}_m c_{m-n}^{(1)} \right) e^{in\theta} = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \hat{\phi}_n e^{in\theta}.$$

By orthogonality of the basis functions, we obtain the following linear equation for each n :

$$(3.16) \quad \sum_{m=-\infty}^{\infty} \left(\hat{\alpha}_m c_{m-n}^{(0)} + \hat{\beta}_m c_{m-n}^{(1)} \right) = \hat{\phi}_n.$$

Moreover, since for sufficiently smooth functions their Fourier coefficients decay rapidly,³ it is sufficient to take only finitely many equations from (3.16) to supplement system (3.12). The summation range on the left-hand side of each of those equations can also be chosen finite. The specific number of equations needed will be determined based on the desired accuracy of the Fourier expansions.

Assume we are given a tolerance $\epsilon > 0$. Then, using the generic notation $f(\theta)$ and $g(\theta)$, we can find a positive integer number N [cf. formula (3.9)] such that

$$(3.17) \quad |\hat{f}_n| < \epsilon \text{ and } |\hat{g}_n| < \epsilon \text{ for } |n| \geq N.$$

Since the coefficients \hat{f}_n and \hat{g}_n decay rapidly, the number N in formula (3.17) is typically not large even if ϵ is taken on the order of the machine precision. (Specific choices that we have made for particular examples are discussed in section 4.) In general, assuming that ϵ in (3.17) is small, we can replace all the coefficients with indices $|n| \geq N$ in the Fourier expansions of f and g by zeros.

We therefore consider a pair of sufficiently smooth 2π -periodic functions $f(\theta)$ and $g(\theta)$ for which we set $\hat{f}_n = 0$ and $\hat{g}_n = 0$ for $|n| \geq N$. To find the coefficients $(\widehat{fg})_n$ of the truncated Fourier expansion for their product fg , we will identify and exclude from the last sum on the right-hand side of formula (3.15) all terms for which either $\hat{f}_{m-n} = 0$ or $\hat{g}_m = 0$ (i.e., those terms for which either $|m - n| > N$ or $|m| > N$, respectively).

If $n \geq 0$, then $m \geq m - n$; thus, the upper bound for the summation will be $m = N$, and the lower bound will be achieved when $m - n = -N$, which, solved for m , yields $m = n - N$. Similarly, when $n < 0$, we have $m < m - n$, which results in the lower bound being reached by $m = -N$, and the upper bound is reached by $m - n = N$, which implies $m = n + N$. Hence, we have

$$(3.18a) \quad (\widehat{fg})_n = \begin{cases} \sum_{m=n-N}^N \hat{g}_m \hat{f}_{m-n}, & 0 \leq n \leq 2N, \\ \sum_{m=-N}^{n+N} \hat{g}_m \hat{f}_{m-n}, & -2N \leq n < 0. \end{cases}$$

If $n > 2N$, then the summation range in the first sum on the right-hand side of (3.18a) becomes empty; if $n < -2N$, then the summation range becomes empty in the second sum. Empty summation ranges yield zero Fourier coefficients so that

$$(3.18b) \quad (\widehat{fg})_n = 0, \quad |n| > 2N.$$

³For an r -differentiable function with the derivative of order r in L_2 , the rate of decay of its Fourier coefficients is $o(n^{-r})$, and accordingly, the rate of convergence of its Fourier series is $o(n^{-(r-\frac{1}{2})})$; see, e.g., [48, section 3.1.3].

Applying formulae (3.18) to (3.16), we arrive at the following finite system:

$$(3.19) \quad \begin{cases} \sum_{m=n-N}^N (\hat{\alpha}_m c_{m-n}^{(0)} + \hat{\beta}_m c_{m-n}^{(1)}) = \hat{\phi}_n, & 0 \leq n \leq 2N, \\ \sum_{m=-N}^{n+N} (\hat{\alpha}_m c_{m-n}^{(0)} + \hat{\beta}_m c_{m-n}^{(1)}) = \hat{\phi}_n, & -2N \leq n < 0, \end{cases}$$

where $\hat{\phi}_n$ on the right-hand side become zero whenever $|n| \geq N$.

System (3.19) provides $4N + 1$ additional equations to supplement the $|\gamma|$ equations of system (3.12). The purpose of (3.19) is to take into account the boundary condition (3.14), which is a particular form of (2.1b), whereas system (3.12) is responsible for the differential equation (2.1a). Combining them yields a nontrivial solution which will satisfy both the differential equation and the boundary condition. The only inhomogeneity of the overall system comes from the coefficients $\hat{\phi}_n$ that are nonzero, i.e., from (3.19) with $n = -N, \dots, N$. The dimension of system (3.12), (3.19) is $[|\gamma| + (4N + 1)] \times 2(2N + 1)$, and its solution $\mathbf{c} = [c_{-N}^{(0)}, \dots, c_N^{(0)}, c_{-N}^{(1)}, \dots, c_N^{(1)}]^T$ is to be sought in the sense of least squares. This system simplifies in the previously analyzed cases of the Dirichlet and Neumann boundary conditions. The Dirichlet boundary condition is equivalent to $\alpha = 1$ and $\beta = 0$ in the general equation (3.14), yielding $u(\theta) = \phi(\theta)$. Then, $\hat{\alpha}_0 = 1$, $\hat{\alpha}_n = 0$ for $n = -N, \dots, -1, 1, \dots, N$, and $\hat{\beta}_n = 0$ for $n = -N, \dots, N$, so that system (3.19) reduces to

$$c_n^{(0)} = \hat{\phi}_n, \quad n = -2N, \dots, 2N.$$

Moreover, as $\hat{\phi}_n = 0$ for $|n| > N$, we can simply disregard the corresponding coefficients $c_n^{(0)}$ and keep only those $c_n^{(0)}$, for which $n = -N, \dots, N$. Substituting these $c_n^{(0)} = \hat{\phi}_n$ into (3.12), we get

$$\mathbf{Q}_1 \begin{bmatrix} c_{-N}^{(1)} \\ \vdots \\ c_N^{(1)} \end{bmatrix} = -\mathbf{Q}_0 \begin{bmatrix} \hat{\phi}_{-N} \\ \vdots \\ \hat{\phi}_N \end{bmatrix},$$

which is to be solved by least squares with respect to the unknown $c_n^{(1)}$, $n = -N, \dots, N$. This is precisely the approach that we employed in [35]; see also section 3.3.

Similarly, the Neumann boundary condition corresponds to $\alpha = 0$ and $\beta = 1$ in formula (3.14). Consequently, $\hat{\alpha}_n = 0$ for $n = -N, \dots, N$, $\hat{\beta}_1 = 1$, and $\hat{\beta}_n = 0$ for $n = -N, \dots, -1, 1, \dots, N$. This reduces (3.19) to

$$c_n^{(1)} = \hat{\phi}_n, \quad n = -2N, \dots, 2N.$$

Again, we disregard those $c_n^{(1)}$ for which $n = -2N, \dots, -N - 1, N + 1, \dots, 2N$, because the corresponding $\hat{\phi}_n = 0$ for $|n| > N$. Substituting the remaining $c_n^{(1)} = \hat{\phi}_n$, $n = -N, \dots, N$, into (3.12) yields

$$\mathbf{Q}_0 \begin{bmatrix} c_{-N}^{(0)} \\ \vdots \\ c_N^{(0)} \end{bmatrix} = -\mathbf{Q}_1 \begin{bmatrix} \hat{\phi}_{-N} \\ \vdots \\ \hat{\phi}_N \end{bmatrix},$$

which is equivalent to the approach implemented in [35].

In the cases other than those of the Dirichlet or Neumann boundary conditions, the full system (3.19) has $4N + 1$ equations, and it is not immediately obvious whether the $2N$ homogeneous equations can likewise be disregarded. Indeed, while it is possible that for $-2N \leq n \leq -N - 1$ or for $N + 1 \leq n \leq 2N$ the products of small terms on the left-hand side of the corresponding equations (3.19) will be $o(\epsilon)$ (i.e., asymptotically smaller than $\mathcal{O}(\epsilon)$), it is not automatically guaranteed. Experimentally, the setting with no homogeneous equations in (3.19) can be tested, and we found out that keeping or dropping those equations made very little difference for the cases computed in section 4.2. Theoretically, however, this issue requires more analysis, which we leave for future study.

3.5. The Robin boundary problem with discontinuous variable coefficients. Consider now the general Robin boundary condition (3.14), but with the relaxed assumptions that α , β , and ϕ are bounded and piecewise smooth rather than smooth. For simplicity, assume that $\Gamma = \Gamma_1 \cup \Gamma_2$, where Γ_1 includes all points on the circle with $\theta \in [0, a)$ and Γ_2 includes all points on the circle with $\theta \in [a, 2\pi)$ for some $0 < a < 2\pi$. Assume that α , β , and ϕ are smooth and bounded on either Γ_1 or Γ_2 but are not necessarily continuous on the entire circle Γ . For example, they may have a jump discontinuity at $\theta = 0$ and/or $\theta = a$.

Consider the standard Chebyshev polynomial basis, $\{T_n(x)\}_{n=0}^\infty$, $x \in [-1, 1]$, with the weight $\omega(x) = 2/\pi\sqrt{1-x^2}$. For a given function $f(x)$, denote its Chebyshev coefficients by \hat{f}_n , $n = 0, 1, 2, \dots$. To express the left-hand side of the boundary condition (3.14) in the Chebyshev basis, we need to find the form of the expansion for the products $\alpha(\theta)u(\theta)$ and $\beta(\theta)\frac{\partial u}{\partial n}(\theta)$. In Appendix A, the Chebyshev coefficients of a product are derived for a pair of arbitrary smooth functions f and g on $[-1, 1]$; see formulae (A.2). To implement this expansion in practice, one first needs to truncate it and replace the series by a finite sum. As Chebyshev coefficients of smooth functions decay rapidly (the relation between the rate of decay and smoothness is similar to that for the Fourier coefficients; see footnote 3), for a given $\epsilon > 0$ we can choose a relatively small number N such that

$$\begin{aligned} |\hat{f}_n| &< \epsilon, & n > N, \\ |\hat{g}_n| &< \epsilon, & n > N. \end{aligned}$$

We therefore set $\hat{f}_n = \hat{g}_n = 0$ for $n > N$. Then formula (A.2a) for $n = 0$ immediately yields

$$(3.20a) \quad (\widehat{fg})_0 = 2\hat{g}_0\hat{f}_0 + \sum_{m=1}^N \hat{g}_m\hat{f}_m.$$

Consider now the case $0 < n \leq N$. On the right-hand side of the last equality in (A.2b), we first replace the upper limit in the last sum by N because the factor \hat{g}_m in the product under the sum will be set to zero beyond $m = N$. Next, since $n > 0$, we notice that the index $m + n$ is the largest, and the corresponding terms become zero when $m > N - n$, which yields

$$(3.20b) \quad \begin{aligned} (\widehat{fg})_n &= \frac{1}{2} \left(\sum_{m=0}^{n-1} \hat{g}_m(\hat{f}_{n-m} + \hat{f}_{n+m}) + \hat{g}_n(\hat{f}_{2n} + 2f_0) \right. \\ &\quad \left. + \sum_{m=n+1}^N \hat{g}_m\hat{f}_{m-n} + \sum_{m=n+1}^{N-n} \hat{g}_m\hat{f}_{m+n} \right). \end{aligned}$$

We will use (3.20b) when $1 \leq n < \frac{N}{2}$. Clearly the term \hat{f}_{2n} in (3.20b) is zero whenever $n > \frac{N}{2}$ because its index will be $2n > N$. Moreover, the last sum will be zero for $n \geq \frac{N}{2}$ since the smallest index of \hat{f}_{m+n} will be $m+n = \frac{N}{2} + 1 + \frac{N}{2} > N$. To write the resulting formula in a convenient manner while eliminating the zero terms for $n > \frac{N}{2}$, we first rearrange (3.20b) so that the terms with the largest indices appear last:

$$\begin{aligned} (\widehat{fg})_n = \frac{1}{2} & \left(\sum_{m=0}^{n-1} \hat{g}_m \hat{f}_{n-m} + 2\hat{g}_n \hat{f}_0 + \sum_{m=n+1}^N \hat{g}_m \hat{f}_{m-n} \right. \\ & \left. + \sum_{m=0}^{n-1} \hat{g}_m \hat{f}_{m+n} + \hat{g}_n \hat{f}_{2n} + \sum_{m=n+1}^{N-n} \hat{g}_m \hat{f}_{m+n} \right). \end{aligned}$$

Then, we eliminate them as n becomes larger and they become zero. If $n \geq \frac{N}{2}$ (note that $n = \frac{N}{2}$ can occur only if N is even), then the summation range in the last sum becomes empty since

$$N - n \leq N - \frac{N}{2} = \frac{N}{2} \leq n < n + 1.$$

Thus, if N is even, we can write for $n = \frac{N}{2}$

(3.20c)

$$(\widehat{fg})_n = \frac{1}{2} \left(\sum_{m=0}^{n-1} \hat{g}_m \hat{f}_{n-m} + 2\hat{g}_n \hat{f}_0 + \sum_{m=n+1}^N \hat{g}_m \hat{f}_{m-n} + \sum_{m=0}^{n-1} \hat{g}_m \hat{f}_{m+n} + \hat{g}_n \hat{f}_{2n} \right).$$

Notice that the index $m+n$ in the last sum of (3.20c) exceeds N if $m > N - n$. At the same time, we see that if $\frac{N}{2} < n \leq N$, then $N - n < n - 1$. Thus, replacing the upper limit in the last sum of (3.20c) by the tighter bound $N - n$, we have, for $\frac{N}{2} < n \leq N$,

(3.20d)

$$(\widehat{fg})_n = \frac{1}{2} \left(\sum_{m=0}^{n-1} \hat{g}_m \hat{f}_{n-m} + 2\hat{g}_n \hat{f}_0 + \sum_{m=n+1}^N \hat{g}_m \hat{f}_{m-n} + \sum_{m=0}^{N-n} \hat{g}_m \hat{f}_{m+n} \right).$$

For $N < n \leq 2N$, the summation range in the second to last sum in (3.20d) is empty. In addition, $\hat{g}_n = 0$ and $\hat{f}_{m+n} = 0$ for any $m \geq 0$. Consequently, only the first sum will remain, with the upper bound replaced by N since now $n - 1 \geq N$:

(3.20e)

$$(\widehat{fg})_n = \frac{1}{2} \sum_{m=1}^N \hat{g}_m \hat{f}_{n-m}, \quad n = N + 1, \dots, 2N.$$

Finally, observe for $n > 2N$ that $n - m > 2N - m > N$, which leaves no nonzero terms so that

(3.20f)

$$(\widehat{fg})_n = 0, \quad n > 2N.$$

Altogether, the coefficients of the truncated Chebyshev expansion for the product fg are given by (3.20a) for $n = 0$, (3.20b) for $1 \leq n < \frac{N}{2}$, (3.20c) for $n = \frac{N}{2}$ (note that this occurs only if N is even), (3.20d) for $\frac{N}{2} < n \leq N$, (3.20e) for $N < n \leq 2N$, and (3.20f) for $n > 2N$.

We now derive the supplementary linear system by applying formulae (3.20) to the general boundary condition (3.14). Recall that we are considering a continuous

boundary Γ partitioned into two pieces, $\Gamma = \Gamma_1 \cup \Gamma_2$, which are two arcs of the circle of radius 1 on the intervals $\theta \in [0, a)$ and $\theta \in [a, 2\pi)$, respectively. We now recast the boundary condition (3.14) as

$$(3.21) \quad \begin{aligned} \alpha^{(1)}(\theta)u(\theta) + \beta^{(1)}(\theta)\frac{\partial u}{\partial \mathbf{n}}(\theta) &= \phi^{(1)}(\theta) \quad \text{on } \Gamma_1, \\ \alpha^{(2)}(\theta)u(\theta) + \beta^{(2)}(\theta)\frac{\partial u}{\partial \mathbf{n}}(\theta) &= \phi^{(2)}(\theta) \quad \text{on } \Gamma_2. \end{aligned}$$

To utilize the Chebyshev basis for (3.21), we perform a linear change of variables on Γ_1 and Γ_2 from θ to x so that $x \in [-1, 1)$ in each respective case. For Γ_1 , we have $\theta \in [0, a)$ and consequently,

$$(3.22a) \quad x = \frac{\theta}{a} + \frac{\theta - a}{a},$$

whereas for Γ_2 we have $\theta \in [a, 2\pi)$, and the transformation is given by

$$(3.22b) \quad x = \frac{\theta - a}{2\pi - a} + \frac{\theta - 2\pi}{2\pi - a}.$$

We consider two independent Chebyshev bases, one on each of the arcs, Γ_1 and Γ_2 , and denote their respective dimensions by N_1 and N_2 . As such, we have two double sets of coefficients,

$$\begin{aligned} [\mathbf{c}^{(0,1)}, \mathbf{c}^{(1,1)}]^T &= [c_0^{(0,1)}, \dots, c_{N_1}^{(0,1)}, c_1^{(1,1)}, \dots, c_{N_1}^{(1,1)}]^T, \\ [\mathbf{c}^{(0,2)}, \mathbf{c}^{(1,2)}]^T &= [c_0^{(0,2)}, \dots, c_{N_2}^{(0,2)}, c_1^{(1,2)}, \dots, c_{N_2}^{(1,2)}]^T, \end{aligned}$$

so that $\mathbf{c} = [\mathbf{c}^{(0,1)}, \mathbf{c}^{(1,1)}, \mathbf{c}^{(0,2)}, \mathbf{c}^{(1,2)}]^T$. Accordingly, instead of formula (3.9) we now have

$$\xi_\Gamma = \underbrace{\sum_{n=0}^{N_1} c_n^{(0,1)}\psi_n^{(0,1)} + \sum_{n=0}^{N_2} c_n^{(0,2)}\psi_n^{(0,2)}}_{(\xi_0,0)} + \underbrace{\sum_{n=0}^{N_1} c_n^{(1,1)}\psi_n^{(1,1)} + \sum_{n=0}^{N_2} c_n^{(1,2)}\psi_n^{(1,2)}}_{(0,\xi_1)},$$

where

$$(3.23) \quad \begin{aligned} \psi_n^{(0,1)} &= \begin{cases} (T_n, 0) & \text{on } \Gamma_1, \\ (0, 0) & \text{on } \Gamma_2, \end{cases} & \psi_n^{(0,2)} &= \begin{cases} (0, 0) & \text{on } \Gamma_1, \\ (T_n, 0) & \text{on } \Gamma_2, \end{cases} \\ \psi_n^{(1,1)} &= \begin{cases} (0, T_n) & \text{on } \Gamma_1, \\ (0, 0) & \text{on } \Gamma_2, \end{cases} & \psi_n^{(1,2)} &= \begin{cases} (0, 0) & \text{on } \Gamma_1, \\ (0, T_n) & \text{on } \Gamma_2. \end{cases} \end{aligned}$$

The extension of a given basis function from (3.23) to the discrete boundary γ is done according to the same formulae derived for the Taylor expansion in section 3.1. The matrix \mathbf{Q} is now partitioned into four blocks rather than two and will have the dimension $|\gamma| \times [2(N_1 + 1) + 2(N_2 + 1)]$:

$$(3.24) \quad \mathbf{Q} = \begin{bmatrix} \mathbf{Q}_0^{(1)} & \mathbf{Q}_1^{(1)} & \mathbf{Q}_0^{(2)} & \mathbf{Q}_1^{(2)} \end{bmatrix}.$$

As in the previous case (section 3.4), the corresponding homogeneous linear system (3.12) with the matrix \mathbf{Q} of (3.24) accounts for the differential equation (2.1a) but not for the boundary condition (2.1b). Thus, we supplement it with additional equations to account for the boundary conditions.

Transforming the boundary condition (3.21) according to (A.1) and taking into account formulae (3.20) for the coefficients of the truncated Chebyshev expansion of a product of two functions, we obtain a set of additional linear equations for each Γ_i , $i = 1, 2$. Specifically, we have for $n = 0$

$$(3.25a) \quad c_0^{(0,i)} \hat{\alpha}_0^{(i)} + \frac{1}{2} \sum_{m=1}^{N_i} c_m^{(0,i)} \hat{\alpha}_m^{(i)} + c_0^{(1,i)} \hat{\beta}_0^{(i)} + \frac{1}{2} \sum_{m=1}^{N_i} c_m^{(1,i)} \hat{\beta}_m^{(i)} = \hat{\phi}_0^{(i)};$$

for $1 \leq n < \frac{N_i}{2}$

$$(3.25b) \quad \begin{aligned} & \frac{1}{2} \left(\sum_{m=0}^{n-1} c_m^{(0,i)} \left(\hat{\alpha}_{n-m}^{(i)} + \hat{\alpha}_{n+m}^{(i)} \right) + c_n^{(0,i)} \left(\hat{\alpha}_{2n}^{(i)} + 2\hat{\alpha}_0^{(i)} \right) \right. \\ & + \sum_{m=n+1}^{N_i} c_m^{(0,i)} \hat{\alpha}_{m-n}^{(i)} + \sum_{m=n+1}^{N_i-n} c_m^{(0,i)} \hat{\alpha}_{m+n}^{(i)} \\ & + \sum_{m=0}^{n-1} c_m^{(1,i)} \left(\hat{\beta}_{n-m}^{(i)} + \hat{\beta}_{n+m}^{(i)} \right) + c_n^{(1,i)} \left(\hat{\beta}_{2n}^{(i)} + 2\hat{\beta}_0^{(i)} \right) \\ & \left. + \sum_{m=n+1}^{N_i} c_m^{(1,i)} \hat{\beta}_{m-n}^{(i)} + \sum_{m=n+1}^{N_i-n} c_m^{(1,i)} \hat{\beta}_{m+n}^{(i)} \right) = \hat{\phi}_n^{(i)}; \end{aligned}$$

for $n = \frac{N_i}{2}$ (if N_i is even)

$$(3.25c) \quad \begin{aligned} & \frac{1}{2} \left(\sum_{m=0}^{n-1} c_m^{(0,i)} \left(\hat{\alpha}_{n-m}^{(i)} + \hat{\alpha}_{n+m}^{(i)} \right) + c_n^{(0,i)} \left(\hat{\alpha}_{2n}^{(i)} + 2\hat{\alpha}_0^{(i)} \right) + \sum_{m=n+1}^{N_i} c_m^{(0,i)} \hat{\alpha}_{m-n}^{(i)} \right. \\ & \left. + \sum_{m=0}^{n-1} c_m^{(1,i)} \left(\hat{\beta}_{n-m}^{(i)} + \hat{\beta}_{n+m}^{(i)} \right) + c_n^{(1,i)} \left(\hat{\beta}_{2n}^{(i)} + 2\hat{\beta}_0^{(i)} \right) + \sum_{m=n+1}^{N_i} c_m^{(1,i)} \hat{\beta}_{m-n}^{(i)} \right) = \hat{\phi}_n^{(i)}; \end{aligned}$$

for $\frac{N_i}{2} < n \leq N_i$

$$(3.25d) \quad \begin{aligned} & \frac{1}{2} \left(\sum_{m=0}^{n-1} c_m^{(0,i)} \hat{\alpha}_{n-m}^{(i)} + 2c_n^{(0,i)} \hat{\alpha}_0^{(i)} + \sum_{m=n+1}^{N_i} c_m^{(0,1)} \hat{\alpha}_{m-n}^{(1)} + \sum_{m=n+1}^{N_i-n} c_m^{(0,i)} \hat{\alpha}_{n+m}^{(i)} \right. \\ & \left. + \sum_{m=0}^{n-1} c_m^{(1,i)} \hat{\beta}_{n-m}^{(i)} + 2c_n^{(1,i)} \hat{\beta}_0^{(i)} + \sum_{m=n+1}^{N_i} c_m^{(1,i)} \hat{\beta}_{m-n}^{(i)} + \sum_{m=n+1}^{N_i-n} c_m^{(1,i)} \hat{\beta}_{n+m}^{(i)} \right) = \hat{\phi}_n^{(i)}; \end{aligned}$$

and for $N_i < n \leq 2N_i$

$$(3.25e) \quad \frac{1}{2} \left(\sum_{m=0}^{N_i} c_m^{(0,i)} \alpha_{n-m}^{(i)} + \sum_{m=0}^{N_i} c_m^{(1,i)} \beta_{n-m}^{(i)} \right) = 0.$$

This gives us a total of $2N_i+1$ extra equations for the $2(N_i+1)$ coefficients $[\mathbf{c}^{(0,i)}, \mathbf{c}^{(1,i)}]$, where $i \in \{1, 2\}$, so that the augmented system comprising (3.12) with \mathbf{Q} given by

(3.24) and equations (3.25) will have the dimension $(|\gamma| + 2N_1 + 1 + 2N_2 + 1) \times [2(N_1 + 1) + 2(N_2 + 1)]$. It shall be solved in the sense of least squares. We also observe that the only inhomogeneity in the overall system arises from the supplemental equations (3.25a)–(3.25d), which account for the boundary condition, whereas the remaining supplemental equations (3.25e) are homogeneous.

Finally, we demonstrate that in the case of simple boundary conditions, such as Dirichlet or Neumann, the supplemental equations reduce to a diagonal system. For example, consider a Dirichlet boundary condition on Γ_1 and a Neumann boundary condition on Γ_2 . Then,

$$\begin{aligned}\alpha^{(1)}(\theta) &= 1, \beta^{(1)}(\theta) = 0, \\ \alpha^{(2)}(\theta) &= 0, \beta^{(2)}(\theta) = 1.\end{aligned}$$

Thus the Chebyshev coefficients of $\alpha^{(i)}$ and $\beta^{(i)}$, $i = 1, 2$, will be

$$\begin{aligned}\hat{\alpha}_0^{(1)} &= 1, \hat{\alpha}_n^{(1)} = 0 \text{ for } n = 1, \dots, N, \\ \hat{\beta}_n^{(1)} &= 0 \text{ for } n = 0, \dots, N, \\ \hat{\alpha}_n^{(2)} &= 0 \text{ for } n = 0, \dots, N, \\ \hat{\beta}_0^{(2)} &= 1, \hat{\beta}_n^{(2)} = 0 \text{ for } n = 1, \dots, N.\end{aligned}$$

Consequently, (3.25) for $i = 1$ reduce to

$$(3.26) \quad \begin{aligned}c_0^{(0,1)} &= \hat{\phi}_0^{(1)} \text{ for } n = 0 \text{ from (3.25a),} \\ c_n^{(0,1)} &= \hat{\phi}_n^{(1)} \text{ for } n = 1, \dots, N_1 \text{ from (3.25b)–(3.25d),} \\ 0 &= 0 \text{ for } n = N_1 + 1, \dots, 2N_1 \text{ from (3.25e),}\end{aligned}$$

while for $i = 2$ they reduce to

$$(3.27) \quad \begin{aligned}c_0^{(1,2)} &= \hat{\phi}_0^{(2)} \text{ for } n = 0 \text{ from (3.25a),} \\ c_n^{(1,2)} &= \hat{\phi}_n^{(2)} \text{ for } n = 1, \dots, N_1 \text{ from (3.25b)–(3.25d),} \\ 0 &= 0 \text{ for } n = N_1 + 1, \dots, 2N_1 \text{ from (3.25e).}\end{aligned}$$

Subsequently, we substitute (3.26) and (3.27) into system (3.12) as we did in the Fourier case (section 3.4) to obtain a reduced system for the coefficients that remain unknown, $\mathbf{c}^{(1,1)}$ and $\mathbf{c}^{(0,2)}$, that is also solved in the sense of least squares. Clearly, any combination of the Dirichlet and Neumann boundary conditions on Γ_1 and Γ_2 will yield a similar reduced system.

To conclude this section we mention that the same considerations as outlined in section 3.4 for the Fourier case apply to the Chebyshev case as well. Namely, the dimensions N_1 and N_2 should be chosen so as to have the accuracy of the truncated Chebyshev expansion on either Γ_1 or Γ_2 exceed the accuracy attainable on the grid, and the specific choices that we have made are discussed in section 4. Moreover, other than for the simple Dirichlet and Neumann cases it is not clear whether the homogeneous supplemental equations (3.25e) can be dropped from the overall system—this question requires further theoretical inquiry.

Another avenue for further consideration is the potential for the undesirable growth in the total number of basis functions needed when splitting the boundary

Γ into segments and using an independent basis for each piece. At first glance, it appears that we would accumulate basis functions linearly with respect to the number of partitions, eventually sacrificing the efficiency of the method. In fact, this should not be the case. A priori, any function on the continuous boundary Γ will experience less variation on a subinterval $\Gamma_i \subset \Gamma$ than on the whole of Γ , even if the boundary data are oscillatory. Thus, achieving the same level of accuracy by the expansion will require fewer coefficients on Γ_i than it would on Γ , and this property would reduce the accumulation of basis functions due to the splitting of Γ into smaller subintervals. Specific quantitative estimates along these lines will be a subject for future study.

3.6. Derivatives of the Chebyshev polynomials near the endpoints. The extension from the continuous boundary Γ to the discrete boundary γ via the Taylor formula (3.1) is done independently for each basis function (see section 3.2), and this process requires us to provide the basis functions themselves as well as their tangential derivatives up to the fourth order; see formulae (3.6) and (3.7).⁴ For the Chebyshev basis it is well known that the derivatives of the polynomials $T_n(x)$ near the endpoints $x = \pm 1$ are not singular, but their values become large. For example, the first derivative of the n th polynomial is

$$T'_n(x) = (\cos(n \arccos x))' = \frac{n \sin(n \arccos x)}{\sqrt{1-x^2}},$$

and taking successive derivatives will clearly retain the term $1-x^2$ in the denominator with increasingly higher exponents. Inevitably, whether by chance or by sufficient refinement of the grid, we will need to compute the values of these derivatives “close” to the endpoints $x = 1$ and/or $x = -1$. Specifically, this happens when the foot of the normal dropped from a given node of γ to Γ (see section 3.1) appears to be close to one of the points that partition Γ into segments ($\theta = 0$ or $\theta = a$ in section 3.5). In this case, the overall accuracy may deteriorate via the loss of significant digits. We have, in fact, computationally observed such a loss of accuracy.

To avoid this undesirable phenomenon, we employ an approach that allows us to completely eliminate the need to compute the derivatives of the Chebyshev basis functions near the endpoints. The key idea of the approach is to use an extended interval for the Chebyshev basis. That is, instead of linearly transforming the intervals $\theta \in [0, a)$ and $\theta \in [a, 2\pi)$, i.e., the arcs of the circle Γ_1 and Γ_2 , to the interval $x \in [-1, 1)$ to form the Chebyshev expansion (see formulae (3.22)), we will instead linearly transform them to a smaller interval $x \in [-1 + \varepsilon, 1 - \varepsilon)$, where $\varepsilon > 0$:

$$(3.28a) \quad x = \left(\frac{\theta}{a} + \frac{\theta - a}{a} \right) (1 - \varepsilon)$$

and

$$(3.28b) \quad x = \left(\frac{\theta - a}{2\pi - a} + \frac{\theta - 2\pi}{2\pi - a} \right) (1 - \varepsilon).$$

In doing so, formulae (3.28a) and (3.28b) obviously provide a transformation between the full interval $x \in [-1, 1)$ and the two extended intervals of the variable θ :

$$(3.29a) \quad x \in [-1, 1) \longleftrightarrow \theta \in [-a\sigma, a + a\sigma)$$

⁴Additionally, for a higher-order scheme one needs to use a higher-order Taylor formula [35], requiring even higher-degree tangential derivatives to be supplied.

and

$$(3.29b) \quad x \in [-1, 1] \longleftrightarrow \theta \in [a - (2\pi - a)\sigma, 2\pi + (2\pi - a)\sigma],$$

respectively, where $\sigma = \frac{1}{2} \frac{\varepsilon}{1-\varepsilon} > 0$.

We then extend all the functions that define the problem, $\alpha^{(1)}(\theta)$, $\beta^{(1)}(\theta)$, $\phi^{(1)}(\theta)$ and $\alpha^{(2)}(\theta)$, $\beta^{(2)}(\theta)$, $\phi^{(2)}(\theta)$ (see formula (3.21)) smoothly but otherwise arbitrarily from their respective intervals $[0, a)$ and $[a, 2\pi)$ to the extended intervals (3.29a) and (3.29b), so that they can subsequently be represented as functions of x using a standard Chebyshev series on $[-1, 1]$. We also formally assume that the unknown functions u and $\frac{\partial u}{\partial n}$ are defined on the same extended intervals (3.29) so that we can identically reproduce all the arguments of section 3.5 and obtain the same supplemental equations (3.25). At the same time, to build the extension (3.1), (3.2) of a given basis function from Γ to γ we will need to know this function only on the corresponding arc $\Gamma_1 \Leftrightarrow \theta \in [0, a)$ or $\Gamma_2 \Leftrightarrow \theta \in [a, 2\pi)$ or, equivalently, on the reduced interval $[-1 + \varepsilon, 1 - \varepsilon]$, because it is those actual arcs where the normals from γ can meet Γ , whereas the “tails” $x \in [-1, -1 + \varepsilon]$ and $x \in [1 - \varepsilon, 1]$ are artificial. Hence, the extension operator $\mathbf{E}x$ will never require any information from these tails and there will be no need to compute the derivatives of the basis functions near the endpoints (i.e., closer than ε to endpoints). It is only the reduced interval $[-1 + \varepsilon, 1 - \varepsilon]$ that will eventually contribute to system (3.12) which represents the discrete BEP (2.8).

Let $f(x)$ be a smooth function defined on $[-1 + \varepsilon, 1 - \varepsilon]$, where $0 < \varepsilon < 1$, and let $\tilde{f}(x)$ be any smooth extension of $f(x)$ to the full interval $[-1, 1]$ (i.e., $\tilde{f}(x) \equiv f(x)$ for $|x| \leq 1 - \varepsilon$, and $\tilde{f}(x)$ is smooth on all of $[-1, 1]$). Then, we expand \tilde{f} in the Chebyshev basis (see (A.1)):

$$(3.30) \quad \tilde{f}(x) = \sum_{n=0}^{\infty} \hat{f}_n T_n(x), \quad x \in [-1, 1].$$

Since series (3.30) converges uniformly, then, clearly, the same convergence takes place on any subinterval, which implies, in particular, that

$$(3.31) \quad f(x) = \sum_{n=0}^{\infty} \hat{f}_n T_n(x), \quad x \in [-1 + \varepsilon, 1 - \varepsilon].$$

We emphasize that even though the coefficients \hat{f}_n of the series (3.30) depend on what particular extension of $f(x)$ from $[-1 + \varepsilon, 1 - \varepsilon]$ to $[-1, 1]$ we choose, this dependence manifests itself only through the fact that the sum of the series (3.30) will vary on the tails $1 - \varepsilon < |x| \leq 1$. At the same time, on the central subinterval $|x| \leq 1 - \varepsilon$ the sum of the series (3.31) remains the same, i.e., equal to the original $f(x)$, regardless of the specific behavior of $\tilde{f}(x)$ for $1 - \varepsilon < |x| \leq 1$.

Accordingly, the Chebyshev coefficients $\hat{\alpha}_n^{(i)}$, $\hat{\beta}_n^{(i)}$, and $\hat{\phi}_n^{(i)}$, where $i = 1, 2$, will depend on the respective extensions of the functions $\alpha^{(i)}$, $\beta^{(i)}$, and $\phi^{(i)}$ from $[-1 + \varepsilon, 1 - \varepsilon]$ to $[-1, 1]$. As these coefficients provide the data for the supplemental equations (3.25), the solution $\mathbf{c} = [\mathbf{c}^{(0,1)}, \mathbf{c}^{(1,1)}, \mathbf{c}^{(0,2)}, \mathbf{c}^{(1,2)}]^T$ of the overall system (3.12), (3.25), where the matrix \mathbf{Q} is given by (3.24), will also be affected by what extension of $\alpha^{(i)}$, $\beta^{(i)}$, and $\phi^{(i)}$, $i = 1, 2$, has been chosen. However, the resulting variation of \mathbf{c} will, again, correspond only to the variation of u and $\frac{\partial u}{\partial n}$ on the extension tails $1 - \varepsilon < |x| \leq 1$ of both arcs, Γ_1 and Γ_2 , and will not affect the solution on the interior subinterval $|x| \leq 1 - \varepsilon$, i.e., on the actual arcs themselves.

In other words, what we essentially do is enforce the boundary conditions (3.21) on the extended intervals (3.29), while the differential equation (2.1a) is still enforced on the original boundary $\Gamma = \Gamma_1 \cup \Gamma_2$ through the discrete BEP (2.8). In doing so, extension (3.29) keeps the original boundary data, and hence the definition of the problem, unaffected. At the same time, the redundancy that we build into the treatment of the boundary conditions by using the extensions allows us to circumvent the difficulties in computing the derivatives of the Chebyshev basis functions near the endpoints. The practical choice of the tolerance ε is discussed in section 4.1.

The only remaining question is how to actually obtain the extended function \tilde{f} for a given f . Recall that the behavior of the tails of \tilde{f} (i.e., the part for which $|x| > 1 - \varepsilon$) will not affect the convergence of the series (3.31) to $f(x)$ for $x \in [-1 + \varepsilon, 1 - \varepsilon]$. Thus we are free to choose any smooth, bounded extension of f . In the case that we are given an analytic formula for $f(x)$, $x \in [-1 + \varepsilon, 1 - \varepsilon]$, which also defines a smooth function on $[-1, 1]$, we can simply use the same formula on the larger interval. Otherwise, we can employ a polynomial extension of order J :

$$(3.32) \quad \tilde{f}(x) = \begin{cases} f(x), & x \in [-1 + \varepsilon, 1 - \varepsilon], \\ \sum_{j=0}^J \frac{1}{j!} \frac{d^j f(-1+\varepsilon)}{dx^j} (x + 1 - \varepsilon)^j, & x < -1 + \varepsilon, \\ \sum_{j=0}^J \frac{1}{j!} \frac{d^j f(1-\varepsilon)}{dx^j} (x - 1 + \varepsilon)^j, & x > 1 - \varepsilon. \end{cases}$$

Formula (3.32) guarantees that the extended function \tilde{f} is smooth and hence bounded on $[-1, 1]$.

In sections 4.4 and 4.6, we compare the performance of the algorithm with and without use of the extended Chebyshev intervals. Our computations convincingly corroborate that the proposed approach completely eliminates the adverse numerical effect of having large derivatives near the endpoints. Therefore, we did not feel it necessary to look for any alternatives. Yet one may, of course, use other strategies as well. For example, another well-known orthogonal system that guarantees rapid convergence of the expansion for smooth nonperiodic functions is the Legendre polynomials [20, Appendix B.1]. The behavior of the derivatives of the original Legendre polynomials near the endpoints is similar to that of the Chebyshev polynomials. However, the associated Legendre functions vanish near the endpoints along with a certain number of derivatives [3, section 18.11]. Hence, they can potentially be used, although this approach will require a further inquiry. An alternative approach is to use a mapping which alleviates the denseness of the Chebyshev nodes near the endpoints; see [55].

3.7. Structure of the algorithm. From the discussion in sections 3.1 through 3.5 we see that the entire computational procedure can fundamentally be split into two parts. The first part involves selecting the basis on the boundary Γ (see (3.8) or (3.23)); extending the individual basis functions from Γ to γ with the help of the operator $\mathbf{E}x$ (see (3.1), (3.2)); applying the discrete projection (2.7) by solving the AP (2.5); and eventually obtaining the linear system (3.12) based on the discrete BEP (2.8). (The matrix \mathbf{Q} is given by (3.13) or (3.24).) This part does not involve the boundary condition (2.1b) in any way and hence does not change when this boundary condition changes. In particular, one and the same system (3.12), (3.13) will work for any boundary condition of type (3.14), whether it be a pure Dirichlet, pure Neumann, or general Robin boundary condition with smooth α , β , and ϕ .

The second part of the algorithm accounts for the boundary conditions via the supplemental equations (3.19) or (3.25). It is this part only that changes when the boundary condition changes, while otherwise the algorithm stays intact. Of course, if

the boundary condition involves a discontinuity (see section 3.5), and the boundary is accordingly partitioned into segments, then the matrix \mathbf{Q} of (3.24) also needs to be recomputed when the partition $\Gamma = \Gamma_1 \cup \Gamma_2$ changes, i.e., when the locations of the discontinuities move. However, for a given fixed partition the corresponding system (3.12), (3.24) will be appropriate for any boundary condition of type (3.21), i.e., any combination of the Dirichlet, Neumann, and Robin boundary conditions.

A natural example where the foregoing split of the algorithm into two relatively independent parts may be useful is electromagnetic scattering off conducting materials coated with dielectrics. It is known that the dielectric coating on the surface of a conductor can have a considerable effect on electromagnetic scattering, and even when the shape of the scatterer stays the same and only the coating changes the radar cross section can still vary substantially; see, e.g., [53]. In turn, various types of coating (pure dielectric, lossy dielectric, dielectric with magnetic losses, etc.) can be modeled by the Leontovich [27] or, equivalently, impedance [54, 50] boundary conditions. In the framework of the second-order governing equations (Helmholtz type) those become Robin boundary conditions. The proposed approach will therefore enable an efficient numerical simulation of electromagnetic scattering and radar cross section for a fixed conducting shape that may be fully or partially coated by different types of dielectrics.

4. Numerical results.

4.1. Parameters of the computational setting. For all the test cases, the boundary Γ is a circle of radius 1 centered at the origin, and the auxiliary domain is a square of side length 2.2 also centered at the origin. The simulations are conducted using the fourth-order accurate compact finite difference scheme (2.3) on a series of Cartesian grids containing 64, 128, 256, 512, 1024, and 2048 cells uniformly spaced in each direction. Scheme (2.3) is supplemented by the Sommerfeld-type boundary conditions (2.5c)–(2.5d) at the left and right edges of the auxiliary square and a Dirichlet condition (2.5b) at its top and bottom edges.

Following the approach developed in [35] (see also section 3.3), the number of basis functions N used to expand $\boldsymbol{\xi}_\Gamma$ by formula (3.9) is chosen grid-independent.⁵ Specifically, it is taken as a number sufficiently large to represent the given boundary functions α , β , and ϕ (see formulae (3.14) and (3.21)) to a prescribed tolerance, 10^{-10} . This boundary tolerance is selected a priori so that it would exceed any accuracy that we expect to obtain on all grids. Given that the boundary coefficients and data are (piecewise) smooth, the corresponding Fourier or Chebyshev expansions converge fast (on each interval of smoothness), and the resulting number N appears relatively small. Altogether, this is a robust and universal strategy that allows us to choose the boundary representation once and for all and then control the final accuracy exclusively by adjusting the grid size. Furthermore, keeping the same number of basis functions for all grids is convenient for the analysis of the computational complexity, as it allows for an unbiased measure of the scaling of the problem relative to the grid size; see sections 4.2–4.4.

We realize that this strategy may still be excessive, because the boundary expansion needs to be only as accurate as the finite difference scheme on a particular grid.

⁵While we always use one and the same notation N , it represents $2N + 1$ functions for the Fourier basis, $e^{-iN\theta}, \dots, 1, \dots, e^{iN\theta}$, and N functions for the Chebyshev basis, $T_0(x), \dots, T_{N-1}(x)$. Moreover, as $\boldsymbol{\xi}_\Gamma$ is a two-component vector function, we need a separate system of basis functions for each component; see (3.9). This makes the overall dimension equal to $2(2N + 1)$ for the Fourier basis and $2N$ for the Chebyshev basis.

In other words, for coarser grids one can take fewer basis functions (3.8). Moreover, taking too many basis functions on coarse grids may result in a loss of accuracy. Indeed, if the dimension of the basis on the boundary is higher than the number of nodes $|\gamma|$, then the least squares problems derived from (3.12) are no longer overdetermined. We have observed this phenomenon in sections 4.2–4.4. Therefore, in section 4.5 we demonstrate that reducing the number of basis functions on these coarse grids restores the accuracy of the least squares solution and the expected fourth-order convergence rate is achieved.

For the examples of sections 4.2–4.6, we take $k = 10$ in the Helmholtz equation (2.1a) and consider a known smooth exact solution, $u = e^{ikx}$. Subsequently, the error is computed in the maximum norm on the set of nodes \mathbb{N}^+ (see Figure 2.3), and the convergence rate is determined by taking the binary logarithm of the ratio of the errors on successively doubled grids. For the experiments of section 4.7 the exact solution is not known, and the convergence is assessed by evaluating the maximum norm of the difference between the numerical solutions obtained on pairs of consecutive grids.

For the examples of sections 4.3–4.7 that involve the Chebyshev basis, the circle is decomposed into two arcs which meet, by design, at the points of discontinuity of the Robin coefficients or at the point of discontinuity in the Dirichlet or Neumann data (section 4.7). In doing so, the trace of the solution on each arc of the circle is represented separately by a set of Chebyshev basis functions.⁶ To avoid computing the derivatives of the Chebyshev functions near the endpoints, we implement a Chebyshev basis on the extended interval, as described in section 3.6. In doing so, the “gap” ε is estimated as follows. The “worst term” in the extension operator (3.1) applied to a given a Chebyshev basis function on Γ will be that with the highest normal derivative, because it translates into the highest tangential derivative according to (3.6)–(3.7). The highest tangential derivative is of order four and it contains the problematic term $(1 - x^2)^{7/2}$ in the denominator⁷ (see section 3.6). Given that the machine precision is on the order of 10^{-16} , we seek ε such that $(1 - x^2)^{7/2} \leq 10^{-10}$ when $|x| > 1 - \varepsilon$ to provide a rough estimate. This yields $\varepsilon > 0.0007$, and thus we have conservatively chosen $\varepsilon = 0.001$ for our computations. This ensures that no calculations of the derivatives of the Chebyshev basis functions will occur within the problematic region near the endpoints $x = \pm 1$. By comparing the results of simulations in section 4.4 (Chebyshev basis on an extended interval with $\varepsilon = 0.001$) with those in section 4.6 (regular Chebyshev system, i.e., $\varepsilon = 0$), we demonstrate that the approach of section 3.6 indeed provides a very efficient remedy for the “near-singular” behavior of the Chebyshev derivatives at the endpoints.

The computer implementation of the entire algorithm is done in MATLAB. To assess the computational complexity, we measure the run time for different parts of the code. The overall time is dominated by applying the projection operator (2.7) to the extended basis functions to form the matrix \mathbf{Q} of (3.13), since this requires the expensive step of solving the discrete AP (2.5), i.e., inverting the finite difference operator $\mathbf{L}^{(h)}$. By representing $\mathbf{L}^{(h)}$ as a matrix, this is accomplished using MATLAB’s built-in sparse direct solver UMFPACK (see [12] and also <http://www.suitesparse.com>) that

⁶Those bases could have different dimensions. The larger of the two arcs may require more basis functions than the smaller, but we have, for simplicity, used the same number of basis functions N on each arc. In this case, the overall number of basis functions is $4N$, i.e., $2N$ per one arc.

⁷Since Chebyshev functions are polynomials, neither they nor their derivatives can become singular. In other words, the apparent zero in the denominator is canceled by the same order of zero in the numerator. However, this l’Hôpital-type indeterminacy makes the numerical computation of the derivatives at the endpoints difficult.

employs the LUPQR factorization. The factorization itself is performed only once and then applied to multiple right-hand sides consisting of the extended basis functions. Hence, the cost of building the matrix \mathbf{Q} is split into two parts: the cost of LU factorization and the cost of multiple backward substitutions.

Once the matrix \mathbf{Q} has been formed, the next step is to take into account the boundary condition via the supplemental equations (3.19) or (3.25) and solve the resulting overdetermined linear system in the sense of the least squares. This is done by means of the QR factorization, and we measure the corresponding CPU times for various cases that we investigate. Finally, once all the coefficients of the expansion (3.9) have been determined, one needs to solve the discrete AP (2.5) one more time to obtain the solution $\mathbf{P}_{N+}\xi_\gamma$, where $\xi_\gamma = \mathbf{E}\mathbf{x}\xi_\Gamma$ is the reconstructed density. This amounts to performing one additional backward substitution, as the LUPQR factorization of the matrix $\mathbf{L}^{(h)}$ stays the same. Altogether, the cost of QR, as well as that of the final solution, is much smaller than the cost of building the matrix \mathbf{Q} . This implies that if \mathbf{Q} is available, changing the boundary condition and solving a new problem can be done very economically.

4.2. Smooth periodic Robin boundary condition via Fourier basis. Consider the Robin boundary condition (3.14) with the smooth, periodic coefficients $\alpha(\theta) = e^{\cos\theta}$ and $\beta(\theta) = e^{\sin\theta}$. We “reverse engineer” the data $\phi(\theta)$ for the Robin condition (3.14) by evaluating the known exact solution $u = e^{ikx}$, $k = 10$, and its normal derivative on the boundary. The results presented in Table 4.1 fully corroborate the theoretical design rate of grid convergence for the proposed methodology (fourth-order). We also see that the dominant part of the computational cost is indeed the formation of the matrix \mathbf{Q} of (3.13). In sections 4.3 and 4.4, we show how one can efficiently solve a range of various boundary value problems using one and the same matrix \mathbf{Q} .

4.3. Mixed Dirichlet/Neumann boundary conditions using Chebyshev basis. In this section, as well as in section 4.4, we consider the same partition of the boundary into two arcs: $\Gamma = R_1 \cup R_2$, where $R_1 = \{0 \leq \theta < 2\pi/3\}$ and $R_2 = \{2\pi/3 \leq \theta < 2\pi\}$. Following section 3.5, we assign a separate Chebyshev basis to R_1 and R_2 . The dimension of the basis is chosen the same for both arcs: $N = 67$. The combination of the partition $\Gamma = R_1 \cup R_2$ and the basis defines the matrix \mathbf{Q} for each grid (via the extension $\mathbf{E}\mathbf{x}$ of each basis function and the application of the projection (2.7)). In Table 4.2, we present the CPU times needed to build this matrix \mathbf{Q} on every grid that we use. In doing so, we distinguish between the time for LUPQR factorization (done once per grid) and the time for $4N$ backward substitutions (see footnote 6).

We emphasize that one and the same matrix \mathbf{Q} will be used to solve all boundary value problems in this section, as well as in section 4.4. We also note that instead of

TABLE 4.1

Grid convergence and execution times for the smooth periodic Robin boundary condition (3.14). $N = 32$ in formula (3.9), and $\psi_n^{(0)}(\theta) = \psi_n^{(1)}(\theta) = e^{in\theta}$ in formula (3.8).

Grid	Error	Conv. rate	Build \mathbf{Q} time	QR time	$\mathbf{P}_{N+}\xi_\gamma$ time
64×64	$9.44 \cdot 10^{-3}$	-	0.66	0.095	0.0024
128×128	$5.92 \cdot 10^{-4}$	4.00	2.47	0.13	0.012
256×256	$3.69 \cdot 10^{-5}$	4.00	11.13	0.21	0.081
512×512	$2.29 \cdot 10^{-6}$	4.01	55.07	0.38	0.30
1024×1024	$1.44 \cdot 10^{-7}$	3.99	228.35	0.76	1.42
2048×2048	$8.25 \cdot 10^{-9}$	4.12	1193.07	2.52	6.68

TABLE 4.2
Time to build the matrix \mathbf{Q} of (3.13).

Grid	LUPQR time	$4N$ back solves time
64×64	0.06	1.16
128×128	0.27	4.28
256×256	1.59	20.38
512×512	8.17	91.03
1024×1024	46.5 1	471.11
2048×2048	1516.08	2563.69

Gaussian elimination, we could have used a more efficient direct solver for the AP, based on the separation of variables and FFT. Nonetheless, we chose to stay with LU decomposition because it permits an easy extension to variable coefficients.

Let us now specify a mix of the Dirichlet and Neumann conditions on separate parts of the boundary Γ by choosing the Robin coefficients in formula (3.21) to be either 1 or 0. Specifically, we let $\alpha^{(1)}(\theta) = 1$ and $\beta^{(1)}(\theta) = 0$ for $\theta \in R_1$ and $\alpha^{(2)}(\theta) = 0$ and $\beta^{(2)}(\theta) = 1$ for $\theta \in R_2$. As in section 3.5, we reduce the general linear system (3.25) to its simplified form (3.26)–(3.27). The exact solution is taken as $u = e^{ikx}$ with $k = 10$ and is used along with its normal derivate to supply the boundary data $\phi^{(1)}(\theta)$ and $\phi^{(2)}(\theta)$. Table 4.3 shows the grid convergence results and CPU times for QR and for the final solution. As the overall solution remains smooth, the method yields fourth-order accuracy even though the boundary conditions are mixed.

Next, we change the boundary data $\phi^{(1)}(\theta)$ and $\phi^{(2)}(\theta)$ by choosing a different exact solution: $u = e^{iky}$ with $k = 10$. The matrix \mathbf{Q} remains unaffected, and the corresponding convergence and timing results are presented in Table 4.4. By comparing the CPU times in Tables 4.3 and 4.4 with those in Table 4.2 we see that once the matrix \mathbf{Q} is available, changing the boundary condition and solving a new boundary value problem becomes a relatively inexpensive task.

TABLE 4.3

Grid convergence and execution times for the mixed Dirichlet/Neumann boundary conditions for $u = e^{ikx}$ with $k = 10$. $\psi_n^{(0)} = \psi_n^{(1)} = T_n$ in formula (3.8), and the dimension of the Chebyshev basis on each arc is $N = 67$.

Grid	Error	Conv. rate	QR time	$\mathbf{P}_{N+\xi_\gamma}$ time
64×64	$5.47 \cdot 10^2$	-	0.011	0.0030
128×128	$3.52 \cdot 10^{-1}$	10.60	0.017	0.013
256×256	$1.61 \cdot 10^{-5}$	14.41	0.035	0.065
512×512	$8.10 \cdot 10^{-7}$	4.32	0.070	0.32
1024×1024	$5.26 \cdot 10^{-8}$	3.95	0.17	1.44
2048×2048	$3.05 \cdot 10^{-9}$	4.11	0.46	7.43

TABLE 4.4

Grid convergence and execution times for the mixed Dirichlet/Neumann boundary conditions for $u = e^{iky}$ with $k = 10$. $\psi_n^{(0)} = \psi_n^{(1)} = T_n$ in formula (3.8), and the dimension of the Chebyshev basis on each arc is $N = 67$.

Grid	Error	Conv. rate	QR time	$\mathbf{P}_{N+\xi_\gamma}$ time
64×64	$2.26 \cdot 10^3$	-	0.011	0.0094
128×128	$2.27 \cdot 10^{-1}$	13.05	0.072	0.012
256×256	$9.71 \cdot 10^{-6}$	14.74	0.17	0.071
512×512	$3.56 \cdot 10^{-7}$	4.77	0.12	0.31
1024×1024	$2.26 \cdot 10^{-8}$	3.98	0.16	1.43
2048×2048	$1.29 \cdot 10^{-9}$	4.15	0.43	7.50

4.4. Piecewise smooth Robin boundary condition via Chebyshev basis.

In this section, we continue to use the same pair of Chebyshev bases and the same matrix \mathbf{Q} as in Table 4.2 but apply these to the Robin boundary condition (3.21) with piecewise smooth coefficients. First, we take $\alpha^{(1)}(\theta) = e^{\cos \theta}$, $\beta^{(1)}(\theta) = \arctan \theta + 1$ for $\theta \in R_1$, and $\alpha^{(2)}(\theta) = e^{2 \sin \theta}$, $\beta^{(2)}(\theta) = 1$ for $\theta \in R_2$. It is easily verified that these coefficients exhibit both jump discontinuities and discontinuities in the first derivative at $\theta = 0$ and $\theta = 2\pi/3$. The boundary data $\phi^{(1)}(\theta)$ and $\phi^{(2)}(\theta)$ are again generated by the smooth exact solution $u = e^{ikx}$ with $k = 10$, and the linear system (3.12) for the Chebyshev coefficients of u and $\frac{\partial u}{\partial n}$ is supplemented by (3.25). The data in Table 4.5 corroborate the fourth-order rate of grid convergence.

Again, as in section 4.3, changing the boundary condition does not require a recomputation of \mathbf{Q} and amounts to only doing another least squares solve and another final solve; see Table 4.6. Note also that the QR times for the piecewise smooth Robin cases (Tables 4.5 and 4.6) are longer than those for the mixed Dirichlet/Neumann cases (Tables 4.3 and 4.4). This is because in the mixed Dirichlet/Neumann case we are solving a reduced system (3.12), (3.26), (3.27), whereas in the piecewise Robin case we are solving a full system (3.12) supplemented by (3.25), which implies a larger dimension.

Altogether, we see that the method provides the design fourth-order accuracy for all the test cases that we have investigated (Tables 4.3–4.6) and also that, once the matrix \mathbf{Q} has been precomputed (Table 4.2), taking a different boundary condition from a rather broad class can be done at a low computational cost.

4.5. Reducing the number of basis functions for coarser grids. In this example, we repeat the first case of section 4.4 with the same parameters except that we alter the number of basis functions for each grid. This alleviates the loss

TABLE 4.5

Grid convergence and execution times for the piecewise smooth Robin boundary condition (3.21) with $\alpha^{(1)}(\theta) = e^{\cos \theta}$, $\beta^{(1)}(\theta) = \arctan \theta + 1$ and $\alpha^{(2)}(\theta) = e^{2 \sin \theta}$, $\beta^{(2)}(\theta) = 1$. The exact solution is $u = e^{ikx}$, $k = 10$. $\psi_n^{(0)} = \psi_n^{(1)} = T_n$ in formula (3.8). The dimension of the Chebyshev basis on each arc is $N = 67$.

Grid	Error	Conv. rate	QR time	$\mathbf{P}_{N+\xi_\gamma}$ time
64×64	$2.63 \cdot 10^3$	-	0.49	0.0023
128×128	$3.69 \cdot 10^{-1}$	12.80	0.56	0.012
256×256	$1.42 \cdot 10^{-5}$	14.67	0.80	0.076
512×512	$8.54 \cdot 10^{-7}$	4.05	1.47	0.31
1024×1024	$5.35 \cdot 10^{-8}$	4.00	2.72	1.44
2048×2048	$3.07 \cdot 10^{-9}$	4.12	5.39	7.52

TABLE 4.6

Grid convergence and execution times for the piecewise smooth Robin boundary condition (3.21) with $\alpha^{(1)}(\theta) = e^{\sin \theta}$, $\beta^{(1)}(\theta) = (\theta + 3)^2$ and $\alpha^{(2)}(\theta) = \log(\theta + 3)$, $\beta^{(2)}(\theta) = \sqrt{\theta + 3}$. The exact solution is $u = e^{ikx}$, $k = 10$. $\psi_n^{(0)} = \psi_n^{(1)} = T_n$ in formula (3.8). The dimension of the Chebyshev basis on each arc is $N = 67$.

Grid	Error	Conv. rate	QR time	$\mathbf{P}_{N+\xi_\gamma}$ time
64×64	$1.95 \cdot 10^3$	-	0.58	0.0024
128×128	$4.65 \cdot 10^{-1}$	12.03	0.68	0.012
256×256	$1.81 \cdot 10^{-5}$	14.64	0.89	0.076
512×512	$1.05 \cdot 10^{-6}$	4.11	1.58	0.31
1024×1024	$6.58 \cdot 10^{-8}$	4.00	2.85	1.44
2048×2048	$3.07 \cdot 10^{-9}$	4.15	5.57	7.29

of accuracy on the coarser grids (see the errors for the 64×64 and 128×128 grids in Tables 4.3–4.6) and also reduces the overall execution times. Recall that the loss of accuracy on coarser grids occurs because the vertical dimension $|\gamma|$ of the matrix \mathbf{Q} of (3.13) may be smaller than its horizontal dimension if too many basis functions are taken. In fact, it may occur even in the case where the system is still formally overdetermined, i.e., its vertical dimension exceeds the horizontal dimension, but only slightly. A possible explanation is that when the grid γ is coarse, the feet of the normals dropped from γ to Γ that are used for building the extension $\mathbf{E}\mathbf{x}$ by formulae (3.1)–(3.2) do not provide a sufficient resolution for the oscillations of the Chebyshev basis functions, which become increasingly oscillatory with N . This effectively implies that two different continuous basis functions may become practically indistinguishable when extended from Γ to γ , which renders the matrix \mathbf{Q} almost rank-deficient. Every time the grid is refined the vertical dimension $|\gamma|$ roughly doubles, the resolution on Γ increases accordingly, and the foregoing adverse phenomenon quickly vanishes. But in the beginning of our sequence of grids, it manifests itself by a higher-than-expected convergence rate as the matrix \mathbf{Q} becomes “taller” but maintains the same width.

We alleviate this issue by choosing N as follows. From our previous computations (Table 4.5) we know that the error for the final grid will be approximately 3×10^{-9} . Hence, we replace our initial tolerance that N should approximate α, β , and ϕ within 10^{-10} by 3×10^{-9} , and this determines the value of N for the finest grid. For this problem this yields $N = 62$ as the number of basis functions needed for the finest grid. To maintain a fourth-order convergence rate, the ratio of the errors on successive grids should be 16. Therefore, for each coarsening of the grid we multiply the tolerance for that grid by 16. Using this rule of thumb, we estimate a smaller but sufficient number of basis functions. In doing so, the matrix \mathbf{Q} obtained in section 4.3 is not recomputed; we simply drop the columns that correspond to the basis functions that are not included in the reduced set. If, on the other hand, we were to adopt this strategy of choosing N from the very beginning, then we would have attained an additional time saving compared to Table 4.2. The results of the computations with the reduced set of basis functions are presented in Table 4.7.

The errors for the coarse grids with 64×64 and 128×128 nodes in Table 4.7 are several orders of magnitude smaller than their counterparts in Table 4.5, with the errors for the remaining finer grids being almost identical. Since there are no other distinctions between this section and the first example of section 4.4, we have demonstrated that taking unnecessarily many basis functions on Γ relative to the grid dimension can cause a loss of accuracy and that grid refinement removes this loss of accuracy. Note also that the QR times in Table 4.7 are somewhat smaller than those in Table 4.5. This is because the horizontal dimensions of the corresponding matrices \mathbf{Q} are smaller.

TABLE 4.7

Grid convergence and execution times for the piecewise smooth Robin boundary condition (3.21) of Table 4.5 with all the same parameters but fewer basis functions on coarser grids.

Grid	N	Error	Conv. rate	QR time	$\mathbf{P}_{N+\xi_\gamma}$ time
64×64	39	$2.20 \cdot 10^{-2}$	-	0.13	0.0032
128×128	44	$2.17 \cdot 10^{-4}$	6.66	0.21	0.013
256×256	49	$1.31 \cdot 10^{-5}$	4.06	0.40	0.066
512×512	54	$7.59 \cdot 10^{-7}$	4.10	0.84	0.30
1024×1024	58	$5.19 \cdot 10^{-8}$	3.87	2.06	1.41
2048×2048	62	$3.10 \cdot 10^{-9}$	4.06	4.88	7.12

We emphasize that the approach for choosing N that we have used in this section is somewhat artificial because it requires an a priori estimate of the error on the finest grid. It, however, allows us to demonstrate that smaller dimensions of the basis can, in principle, be chosen if necessary.

4.6. Loss of accuracy due to nontreatment of Chebyshev endpoints. In section 3.6, we noted that the position of the grid nodes γ relative to the continuous boundary Γ may require computing the Taylor extension (3.1) of the Chebyshev basis functions very close to (or even precisely at) the Chebyshev endpoints, resulting in a loss of accuracy. To demonstrate this phenomenon, as well as the effectiveness of our method in eliminating it, we recompute the first example of section 4.4 without the use of any technique to correct the poorly conditioned computation of Chebyshev derivatives near the endpoints, allowing us to observe the disruption of convergence. We again divide the circle into two arcs, $R_1 = [0, 2\pi/3)$ and $R_2 = [2\pi/3, 2\pi)$, and set the same Robin boundary condition (3.21) with the coefficients $\alpha^{(1)}(\theta) = e^{\cos \theta}$, $\beta^{(1)}(\theta) = \arctan(\theta) + 1$ for $\theta \in R_1$ and $\alpha^{(2)}(\theta) = e^{2\sin \theta}$, $\beta^{(2)}(\theta) = 1$ for $\theta \in R_2$. The boundary data $\phi^{(1)}(\theta)$ and $\phi^{(2)}(\theta)$ are still supplied from the exact solution $u = e^{ikx}$ with $k = 10$. The results are presented in Table 4.8.

Comparing the results in Table 4.8 to those in Table 4.5, we observe an immediate decline in the convergence rate for all the grids from 256×256 on, culminating in a complete loss of convergence for the finest grid. Since the example of section 4.4 has precisely the same parameters as those of the current example, except that in section 4.4 we had used an extended interval for the Chebyshev basis, we conclude that the approach of section 3.6 indeed removes the difficulties that would otherwise arise due to computing the derivatives of the Chebyshev functions near the endpoints.

4.7. Boundary data that lead to solutions with singularities. In each of the following experiments, we use the same partition of the circle as before: $R_1 = [0, 2\pi/3)$ and $R_2 = [2\pi/3, 2\pi)$. However, we no longer consider a given smooth exact solution. We rather specify the boundary data independently on each of the two arcs, R_1 or R_2 , allowing for discontinuities at the points $\theta = 0$ and $\theta = 2\pi/3$. This gives rise to near-boundary singularities in the solution. We also choose $k = 5$ in the Helmholtz equation (2.1a).

On each of the two arcs, R_1 or R_2 , we set either a Dirichlet or a Neumann boundary condition. The data we specify on R_2 always correspond to a plane wave with $k = 5$ traveling in the x direction. For this wave, $u(x, y) = e^{ikx}$, the Dirichlet data are $u = e^{ikr \cos \theta}$ and the Neumann data are $\frac{\partial u}{\partial n} = ik \cos \theta e^{ikr \cos \theta}$, where $r = 1$ because Γ is a circle of radius 1. The data on R_1 are intentionally specified with a mismatch, and we consider mismatches of three different types: in the amplitude

TABLE 4.8

Computation of the same case as that in Table 4.5 but with no special treatment of the Chebyshev endpoints. $\psi_n^{(0)} = \psi_n^{(1)} = T_n$ in formula (3.8), and the dimension of the Chebyshev basis on each arc is $N = 67$.

Grid	Error	Conv. rate
64×64	$2.53 \cdot 10^3$	-
128×128	$3.49 \cdot 10^{-1}$	12.82
256×256	$2.90 \cdot 10^{-4}$	10.23
512×512	$1.37 \cdot 10^{-4}$	1.08
1024×1024	$9.73 \cdot 10^{-6}$	3.82
2048×2048	$2.56 \cdot 10^{-5}$	-1.39

of the wave, in its traveling direction, and in its phase. In the case of a mismatch in the amplitude, the data specified on R_1 correspond to the wave $u(x, y) = 2e^{ikx}$, which yields $u = 2e^{ikr \cos \theta}$ and $\frac{\partial u}{\partial n} = 2ik \cos \theta e^{ikr \cos \theta}$ on Γ in the Dirichlet and the Neumann case, respectively. In the case of a mismatch in the traveling direction, we provide two examples: one with the wave that travels in the y direction and the other with the wave that travels at an angle of 1 radian with respect to the positive x axis. For the wave traveling in the y direction, $u(x, y) = e^{iky}$, on the arc R_1 we have $u = e^{ikr \sin \theta}$ and $\frac{\partial u}{\partial n} = ik \sin \theta e^{ikr \sin \theta}$, $r = 1$, for the Dirichlet and Neumann data, respectively. For the wave traveling at an angle of 1 radian, we have $u = e^{ikr \cos(\theta+1)}$ and $\frac{\partial u}{\partial n} = ik \cos(\theta+1) e^{ikr \cos(\theta+1)}$. Finally, for the experiments in which the waves are out of phase, we have chosen a phase shift of 0.7 radians, yielding $u = e^{i(kr \cos(\theta)+0.7)}$ and $\frac{\partial u}{\partial n} = ik \cos(\theta) e^{i(kr \cos(\theta)+0.7)}$ on R_1 . The numerical results for each type of the mismatch subject to different boundary conditions (Dirichlet, Neumann, and mixed) are summarized in Tables 4.9, 4.10, 4.11, and 4.12. For comparison, we have the continuous case in Table 4.13, i.e., the case with no mismatch, in which the plane wave $u(x, y) = e^{ikx}$ with $k = 5$ supplies the data for both R_1 and R_2 .

TABLE 4.9

Grid convergence for boundary data with amplitude mismatch and either Dirichlet, Neumann, or mixed boundary conditions.

Grid	Dirichlet on R_1 and R_2		Neumann on R_1 and R_2		Dirichlet on R_1 Neumann on R_2	
	Error	Rate	Error	Rate	Error	Rate
128×128	$2.60 \cdot 10^{-2}$	-	$2.64 \cdot 10^{-2}$	-	$1.88 \cdot 10^{-2}$	-
256×256	$1.35 \cdot 10^{-2}$	0.95	$1.61 \cdot 10^{-2}$	0.71	$1.08 \cdot 10^{-2}$	0.80
512×512	$9.91 \cdot 10^{-3}$	0.44	$1.00 \cdot 10^{-2}$	0.69	$5.67 \cdot 10^{-3}$	0.93
1024×1024	$4.80 \cdot 10^{-3}$	1.05	$1.73 \cdot 10^{-3}$	2.53	$2.84 \cdot 10^{-3}$	1.00
2048×2048	$2.87 \cdot 10^{-3}$	0.74	$3.62 \cdot 10^{-3}$	-1.06	$1.68 \cdot 10^{-3}$	0.75

TABLE 4.10

Grid convergence for boundary data with direction mismatch (plane wave traveling in the y direction) and either Dirichlet, Neumann, or mixed boundary conditions.

Grid	Dirichlet on R_1 and R_2		Neumann on R_1 and R_2		Dirichlet on R_1 Neumann on R_2	
	Error	Rate	Error	Rate	Error	Rate
128×128	$1.42 \cdot 10^{-2}$	-	$6.23 \cdot 10^{-2}$	-	$2.32 \cdot 10^{-2}$	-
256×256	$1.67 \cdot 10^{-2}$	-0.24	$2.85 \cdot 10^{-2}$	1.13	$1.44 \cdot 10^{-2}$	0.69
512×512	$1.04 \cdot 10^{-2}$	0.69	$1.67 \cdot 10^{-2}$	0.77	$6.11 \cdot 10^{-3}$	1.24
1024×1024	$5.33 \cdot 10^{-3}$	0.96	$4.68 \cdot 10^{-3}$	1.84	$6.08 \cdot 10^{-3}$	0.01
2048×2048	$2.86 \cdot 10^{-3}$	0.90	$6.92 \cdot 10^{-3}$	-0.56	$2.70 \cdot 10^{-3}$	1.17

TABLE 4.11

Grid convergence for boundary data with direction mismatch (plane wave traveling at 1 radian with respect to the positive x direction) and either Dirichlet, Neumann, or mixed boundary conditions.

Grid	Dirichlet on R_1 and R_2		Neumann on R_1 and R_2		Dirichlet on R_1 Neumann on R_2	
	Error	Rate	Error	Rate	Error	Rate
128×128	$4.86 \cdot 10^{-2}$	-	$8.22 \cdot 10^{-2}$	-	$2.94 \cdot 10^{-2}$	-
256×256	$2.49 \cdot 10^{-2}$	0.96	$2.70 \cdot 10^{-2}$	1.61	$1.75 \cdot 10^{-2}$	0.75
512×512	$1.90 \cdot 10^{-2}$	0.39	$1.49 \cdot 10^{-2}$	0.86	$8.25 \cdot 10^{-3}$	1.08
1024×1024	$9.25 \cdot 10^{-3}$	1.04	$7.05 \cdot 10^{-3}$	1.08	$6.61 \cdot 10^{-3}$	0.32
2048×2048	$5.17 \cdot 10^{-3}$	0.84	$7.81 \cdot 10^{-3}$	-0.15	$3.08 \cdot 10^{-3}$	1.10

TABLE 4.12

Grid convergence for boundary data with phase mismatch (waves out of phase by 0.7 radians) and either Dirichlet, Neumann, or mixed boundary conditions.

Grid	Dirichlet on R_1 and R_2		Neumann on R_1 and R_2		Dirichlet on R_1 Neumann on R_2	
	Error	Rate	Error	Rate	Error	Rate
128×128	$1.78 \cdot 10^{-2}$	-	$1.81 \cdot 10^{-2}$	-	$1.27 \cdot 10^{-2}$	-
256×256	$9.24 \cdot 10^{-3}$	0.95	$1.10 \cdot 10^{-2}$	0.71	$7.41 \cdot 10^{-3}$	0.78
512×512	$6.80 \cdot 10^{-3}$	0.44	$6.86 \cdot 10^{-3}$	0.69	$3.89 \cdot 10^{-3}$	0.93
1024×1024	$3.29 \cdot 10^{-3}$	1.05	$1.19 \cdot 10^{-3}$	2.53	$1.94 \cdot 10^{-3}$	1.00
2048×2048	$1.97 \cdot 10^{-3}$	0.74	$2.48 \cdot 10^{-3}$	-1.06	$1.15 \cdot 10^{-3}$	0.75

TABLE 4.13

Grid convergence for boundary data with no mismatch and either Dirichlet, Neumann, or mixed boundary conditions.

Grid	Dirichlet on R_1 and R_2		Neumann on R_1 and R_2		Dirichlet on R_1 Neumann on R_2	
	Error	Rate	Error	Rate	Error	Rate
128 × 128	$6.62 \cdot 10^{-5}$	-	$1.35 \cdot 10^{-4}$	-	$6.17 \cdot 10^{-5}$	-
256 × 256	$4.18 \cdot 10^{-6}$	3.99	$1.66 \cdot 10^{-5}$	3.02	$4.02 \cdot 10^{-6}$	3.94
512 × 512	$2.62 \cdot 10^{-7}$	3.99	$1.14 \cdot 10^{-7}$	7.19	$2.53 \cdot 10^{-7}$	3.99
1024 × 1024	$1.65 \cdot 10^{-8}$	3.99	$7.10 \cdot 10^{-9}$	4.00	$1.57 \cdot 10^{-8}$	4.01
2048 × 2048	$1.30 \cdot 10^{-9}$	3.66	$5.64 \cdot 10^{-10}$	3.65	$1.24 \cdot 10^{-9}$	3.66

As the exact solutions to these problems (except the formulation with no mismatch) are not available, we cannot evaluate the error by comparing the numerical solution to the actual solution on the grid. Instead, we introduce a grid-based metric, which compares the numerical solutions on successive grids that have common nodes. Specifically, we structure our Cartesian grids so that each refinement retains all the nodes of the previous grid, and then compute the maximum norm of the difference between the two successive solutions on the nodes of the coarser grid. Since this measure involves a pair of grids, Tables 4.9–4.13 display the finer of the pair for each resulting error (i.e., the coarsest grid on which we compute is 64×64). Additionally, we found that the error spikes at the nodes of the discrete boundary γ which are closest to the discontinuities at $\theta = 2\pi/3$ and $\theta = 0$ and that the maximum norm when these points are included does not exhibit significant convergence. Therefore, as an additional modification, we compute the maximum error strictly on the interior of the disk or, more precisely, inside the circle of radius 0.8. We also note that changing the maximum norm to l_2 norm makes no substantial difference in the observed convergence.

From Tables 4.9–4.13 we see that the singularities substantially slow the convergence. In all the cases with mismatches, the rate of grid convergence is at most first-order. This behavior is expected though, because the scheme essentially loses its consistency near the singularity.

5. Conclusions. We have investigated theoretically and demonstrated experimentally the capability of the method of difference potentials to handle complex boundary conditions, such as variable coefficient Robin, mixed, and discontinuous. The governing Helmholtz equation was approximated on a regular Cartesian grid by an economical fourth-order accurate compact finite difference scheme. For a number of test cases that involve a nonconforming circular boundary and various boundary conditions, we have been able to recover the design fourth-order accuracy of the

scheme provided that the overall solution was sufficiently smooth. The accuracy was not adversely affected by either staircasing [9, 22] or the nonstandard nature of the boundary conditions. When the overall solution is not smooth and has a singularity at the location where the coefficients and/or data in the boundary conditions are discontinuous, then the convergence of the method slows down, as expected, because the finite difference scheme loses its consistency.

The approach that we use to reduce the original problem from its domain to the boundary is based on Calderon's operators. It automatically guarantees the well-posedness of the resulting boundary formulation as long as the original problem is well-posed, regardless of the type of the boundary condition. Moreover, it is very important that when changing the boundary condition only a particular component of the overall numerical algorithm changes, whereas most of it remains unaffected (see section 3.7). Accordingly, the computational cost associated with solving a new problem for a new boundary condition is small. In contrast, a change of the boundary condition in the classical method of boundary integral equations often requires a complete change of the algorithm. We note that there is another group of techniques based on finite differences/volumes as opposed to integral equations and designed to handle nonaligned boundaries/interfaces: immersed boundary [38], immersed interface [29], ghost fluid [16, 30], and embedded boundary [23, 11] methods. To the best of our knowledge, there are no reported uses in the literature of those methods for anything but simple Dirichlet, Neumann, or interface conditions (continuity of the solution and its normal flux), changing the boundary condition requires major changes to the algorithm [11], and extension to higher than second-order accuracy is not straightforward.

If the boundary condition involves products of functions, such as the general variable coefficient Robin boundary condition (3.14), then the method developed in this paper relies on expressing the Fourier coefficients of a product as a convolution of the coefficients of individual factors and on a similar relation for the Chebyshev coefficients. In fact, the proposed method will work for any system of basis functions on the boundary Γ for which a relation of this type can be obtained. An alternative to convolution-type formulae may be provided by collocation techniques; see, e.g., [3, 20].

Partitioning the boundary into segments helps for several difficulties and not only for handling sophisticated boundary conditions, such as those in sections 4.3 and 4.4. It is also useful when, for example, the resolution at the boundary needs to be increased locally, for example, in those parts of Γ that have high curvature. It may also help when the boundary is defined as a composition of segments with an independent parametrization rather than as one curve with a continuous parametrization. In particular, in future work we will investigate the method for domains with corners.

An immediate logical extension of the current work is to modify the proposed methodology so that it would provide a faster convergence rate in the case of solutions with singularities. As we have noted, the near-boundary singularities of the solution due to discontinuities in the boundary conditions result in a substantial slowdown of convergence. The solution may also develop a singularity due to the geometric irregularities of the boundary itself, e.g., corners. In many cases, though, the type of the singularity is known, at least to the leading order; see, e.g., [37, 60, 28, 61, 17, 41, 42, 10, 32, 34, 33], as well as the monograph [18]. Then, the singular component can be subtracted, and the problem can be solved with respect to the remaining part of the solution. The latter should have higher regularity and thus enable an improved convergence of the numerical approximation. Combined with the method of difference potentials, this approach was previously implemented for the Laplace equation in [25] and the Chaplygin equation in [24].

A more comprehensive extension will be to three dimensions (3D), which will not only, and even not necessarily, involve the treatment of nonstandard boundary conditions in the initial stages. Rather, it will require building and testing all the basic components of an algorithm based on the method of difference potentials. Previously, the method of difference potentials was used in 3D to construct the artificial boundary conditions for fluid flow computations [56]. While all the key ideas of the method remain the same, its implementation for solving 3D boundary value problems will imply a number of substantial changes throughout the entire procedure.

Many issues that require special attention in 3D are related to the geometry. For example, it is a more challenging problem in 3D to determine whether a given point (i.e., a grid node) lies inside or outside a given closed surface than an analogous 2D problem of determining whether a given point is inside or outside a given closed curve. An efficient solution to this problem determines the efficiency of partitioning of the set \mathbb{M}_0 into \mathbb{M}^+ and \mathbb{M}^- ; see section 2.3 and Figure 2.3. Finding a convenient representation for the surface Γ itself may not be as easy (unless it is a simple analytical shape, e.g., a sphere) and may require using multiple patches, such as in [7, 6]. Furthermore, choosing a good basis (3.8) on the two-dimensional surface Γ that would enable an efficient (i.e., low-dimensional) representation (3.9) may be more difficult than constructing a basis on the one-dimensional curve Γ . Construction of the extension operator $\mathbf{E}\mathbf{x}$ (see section 3.1) is likely to require special surface-oriented coordinates [21]. Finally, unlike in 2D, preconditioned iterative solvers provide the only realistic venue for solving the AP in the case of variable coefficients. An iterative solver that can be easily parallelized and that has already been applied successfully to the Helmholtz equation discretized by compact high-order schemes is described in [58], and a class of efficient complex-shifted Helmholtz preconditioners that can be inverted by multigrid is described in [15, 14]. For constant coefficient interior problems, solution by the separation of variables method will remain the most efficient approach in 3D. However, for exterior problems one may use convolution with the discrete fundamental solution (see [26, Appendix C]), accelerated by the fast multipole method. The advantage of this approach is that it automatically takes into account the proper behavior of the solution in the far field.

Achieving a higher than fourth-order accuracy may be another objective to pursue. In [35], we have reported some preliminary computations for the constant coefficient Helmholtz equation using a sixth-order accurate compact scheme of [52]. A more advanced sixth-order scheme that can handle variable coefficients has been introduced and tested in [58]. Increasing the order of accuracy even further naturally leads to a question of whether a spectral approximation can be used. On one hand, the AP is always formulated on a simple domain, such as a rectangle. This is done primarily for making its numerical solution easy and efficient, and of course, such an AP can be as easily solved by a spectral method as by high-order finite differences. On the other hand, in the core of the method of difference potentials is the reduction of the governing equation from the domain to the boundary in the form of the BEP (2.8). This requires that the equation be approximated on a local stencil, so that the grid boundary γ is located near Γ . For spectral methods this is not intuitively possible because formally one can interpret a spectral discretization as having a stencil that occupies the entire domain. Altogether, we leave the question of further improving the approximation accuracy of the method of difference potentials for future investigation.

Another issue worth investigating is the relationship between boundary equations with projections and Schur complements. In the current version of the method

of difference potentials the unknowns that we are solving for are not associated with the discretization grid at all. They are rather the coefficients $\mathbf{c} = [c_{-N}^{(0)}, \dots, c_N^{(0)}, c_{-N}^{(1)}, \dots, c_N^{(1)}]^T$ of the expansion (3.9) of the trace of the solution ξ_Γ at the continuous boundary Γ . Therefore, the resulting linear system (3.12), along with any supplemental equations that account for the boundary conditions, e.g., (3.19), may not be obtained directly by elimination from the system of difference equations that is formulated on the grid \mathbb{N}_0 . On the other hand, there are other versions of the method of difference potentials [45], in particular, a purely discrete one, for which the unknowns are components of the discrete trace ξ_γ . Then, the question arises of whether the BEP (2.8) can be obtained with the help of a Schur complement from the finite difference AP (2.5). The first intuitive answer to this question is that there is no direct equivalence, simply because the AP (2.5) has a unique solution, whereas the homogeneous BEP (2.8) has multiple solutions. Nonetheless, there may be deeper similarities that will be interesting to explore.

Appendix A. Chebyshev coefficients for the product of two functions.

The expansion of a smooth bounded function $f(x)$, $x \in [-1, 1]$, in the Chebyshev basis $\{T_n(x)\}_{n=0}^\infty$ is given by

$$(A.1) \quad f(x) = \sum_{n=0}^\infty \hat{f}_n T_n(x), \quad \text{where} \quad \hat{f}_n = \begin{cases} \frac{1}{2} \int_{-1}^1 \omega(x) f(x) T_n(x) dx, & n = 0, \\ \int_{-1}^1 \omega(x) f(x) T_n(x) dx, & n > 0. \end{cases}$$

Let $f(x)$ and $g(x)$ be smooth on $x \in [-1, 1]$. Then, according to the definition of the Chebyshev coefficients (see (A.1)), we have

$$(A.2a) \quad \begin{aligned} (\widehat{fg})_0 &= \frac{1}{2} \int_{-1}^1 \omega(x) f(x) g(x) T_0(x) dx = \frac{1}{2} \int_{-1}^1 \omega(x) f(x) \left[\sum_{m=0}^\infty \hat{g}_m T_m \right] dx \\ &= \frac{1}{2} \sum_{m=0}^\infty \hat{g}_m \int_{-1}^1 \omega(x) f(x) T_m(x) dx = \frac{1}{2} \left(\hat{g}_0 (2\hat{f}_0) + \sum_{m=1}^\infty \hat{g}_m \hat{f}_m \right) \\ &= \hat{g}_0 \hat{f}_0 + \frac{1}{2} \sum_{m=1}^\infty \hat{g}_m \hat{f}_m. \end{aligned}$$

For $n > 0$, we take into account that

$$T_n(x) T_m(x) = \begin{cases} \frac{1}{2} (T_{m+n}(x) + T_{m-n}(x)), & m \geq n, \\ \frac{1}{2} (T_{m+n}(x) + T_{n-m}(x)), & m < n, \end{cases}$$

and using (A.1) we obtain

$$(A.2b) \quad \begin{aligned} (\widehat{fg})_n &= \int_{-1}^1 \omega(x) f(x) g(x) T_n(x) dx = \int_{-1}^1 \omega(x) f(x) \left[\sum_{m=0}^\infty \hat{g}_m T_m \right] T_n(x) dx \\ &= \sum_{m=0}^\infty \hat{g}_m \int_{-1}^1 \omega(x) f(x) T_m(x) T_n(x) dx \\ &= \sum_{m=0}^{n-1} \hat{g}_m \cdot \frac{1}{2} \left[\left(\int_{-1}^1 \omega(x) f(x) T_{m+n}(x) dx \right) + \left(\int_{-1}^1 \omega(x) f(x) T_{n-m}(x) dx \right) \right] \end{aligned}$$

$$\begin{aligned}
& + \sum_{m=n}^{\infty} \hat{g}_m \cdot \frac{1}{2} \left[\left(\int_{-1}^1 \omega(x) f(x) T_{m+n}(x) dx \right) + \left(\int_{-1}^1 \omega(x) f(x) T_{m-n}(x) dx \right) \right] \\
& = \frac{1}{2} \left(\sum_{m=0}^{n-1} \hat{g}_m (\hat{f}_{m+n} + \hat{f}_{n-m}) + \hat{g}_n (\hat{f}_{2n} + 2\hat{f}_0) + \sum_{m=n+1}^{\infty} \hat{g}_m (\hat{f}_{m+n} + \hat{f}_{m-n}) \right).
\end{aligned}$$

We expect formula (A.2b) to be symmetric with respect to f and g . While this is not immediate in the formula itself, it can easily be shown. To see the symmetry of (A.2b) with respect to f and g , first rearrange the summations by the \hat{f} subscripts (i.e., one summation for $m+n$ terms, one for $n-m$, and one for $m-n$). Then, after appropriate substitutions in the summation indices (respectively, $j = m+n$, $j = n-m$, and $j = m-n$), the form of (A.2b) with interchanged g and f terms can be obtained by regrouping the summation terms.

Acknowledgment. We would like to thank two anonymous referees for their most helpful comments.

REFERENCES

- [1] I. M. BABUŠKA AND S. A. SAUTER, *Is the pollution effect of the FEM avoidable for the Helmholtz equation considering high wave numbers?*, SIAM Rev., 42 (2000), pp. 451–484.
- [2] A. BAYLISS, C. I. GOLDSTEIN, AND E. TURKEL, *An iterative method for the Helmholtz equation*, J. Comput. Phys., 49 (1983), pp. 443–457.
- [3] J. P. BOYD, *Chebyshev and Fourier Spectral Methods*, 2nd ed., Dover, Mineola, NY, 2001.
- [4] S. BRITT, S. TSYNKOV, AND E. TURKEL, *A compact fourth order scheme for the Helmholtz equation in polar coordinates*, J. Sci. Comput., 45 (2010), pp. 26–47.
- [5] S. BRITT, S. TSYNKOV, AND E. TURKEL, *Numerical simulation of time-harmonic waves in inhomogeneous media using compact high order schemes*, Commun. Comput. Phys., 9 (2011), pp. 520–541.
- [6] O. P. BRUNO AND L. A. KUNYANSKY, *A fast, high-order algorithm for the solution of surface scattering problems: Basic implementation, tests, and applications*, J. Comput. Phys., 169 (2001), pp. 80–110.
- [7] O. P. BRUNO AND L. A. KUNYANSKY, *Surface scattering in three dimensions: An accelerated high-order solver*, Roy. Soc. London Proc. Ser. A Math. Phys. Engrg. Sci., 457 (2001), pp. 2921–2934.
- [8] A. P. CALDERON, *Boundary-value problems for elliptic equations*, in Proceedings of the Soviet-American Conference on Partial Differential Equations in Novosibirsk, Moscow, 1963, Fizmatgiz, pp. 303–304.
- [9] A. C. CANGELLARIS AND D. B. WRIGHT, *Analysis of the numerical error caused by the stair-stepped approximation of a conducting boundary in FDTD simulations of electromagnetic phenomena*, IEEE Trans. Antennas and Propagation, 39 (1991), pp. 1518–1525.
- [10] M. COSTABEL AND M. DAUGE, *Construction of corner singularities for Agmon-Douglis-Nirenberg elliptic systems*, Math. Nachr., 162 (1993), pp. 209–237.
- [11] R. K. CROCKETT, P. COLELLA, AND D. T. GRAVES, *A Cartesian grid embedded boundary method for solving the Poisson and heat equations with discontinuous coefficients in three dimensions*, J. Comput. Phys., 230 (2011), pp. 2451–2469.
- [12] T. A. DAVIS AND Y. HU, *The University of Florida sparse matrix collection*, ACM Trans. Math. Softw., 38 (2011).
- [13] A. DERAEMAEKER, I. M. BABUŠKA, AND P. BOUILLARD, *Dispersion and pollution of the FEM solution for the Helmholtz equation in one, two and three dimensions*, Internat. J. Numer. Methods Engrg., 46 (1999), pp. 471–499.
- [14] Y. ERLANGGA AND E. TURKEL, *Iterative schemes for high order compact discretizations to the exterior Helmholtz equation*, Math. Model. Numer. Anal., 46 (2012), pp. 647–660.
- [15] Y. A. ERLANGGA, C. VUIK, AND C. W. OOSTERLEE, *On a class of preconditioners for solving the Helmholtz equation*, Appl. Numer. Math., 50 (2004), pp. 409–425.
- [16] R. P. FEDKIW, T. ASLAM, AND S. XU, *The ghost fluid method for deflagration and detonation discontinuities*, J. Comput. Phys., 154 (1999), pp. 393–427.

- [17] L. FOX AND R. SANKAR, *Boundary singularities in linear elliptic differential equations*, J. Inst. Math. Appl., 5 (1969), pp. 340–350.
- [18] P. GRISVARD, *Elliptic Problems in Nonsmooth Domains*, Monogr. Stud. Math. 24, Pitman, Boston, 1985.
- [19] I. HARARI AND E. TURKEL, *Accurate finite difference methods for time-harmonic wave propagation*, J. Comput. Phys., 119 (1995), pp. 252–270.
- [20] J. S. HESTHAVEN, S. GOTTLIEB, AND D. GOTTLIEB, *Spectral Methods for Time-Dependent Problems*, Cambridge Monogr. Appl. Comput. Math., 21, Cambridge University Press, Cambridge, UK, 2007.
- [21] E. H. HIRSCHL AND W. KORDULLA, *Shear Flow in Surface-Oriented Coordinates*, Notes Numer. Fluid Mech. 4, Vieweg, Braunschweig, 1981.
- [22] R. HOLLAND, *Pitfalls of staircase meshing*, IEEE Trans. Electromagnetic Compatibility, 35 (1993), pp. 434–439.
- [23] H. JOHANSEN AND P. COLELLA, *A Cartesian grid embedded boundary method for Poisson's equation on irregular domains*, J. Comput. Phys., 147 (1998), pp. 60–85.
- [24] D. S. KAMENETSKII, *A Numerical Method for Solving a Singular Boundary Value Problem for the Chaplygin Equation in the Hodograph Plane* (in Russian), Preprint 60, Keldysh Institute for Applied Mathematics, Russian Academy of Sciences, Moscow, Russia, 1992.
- [25] D. S. KAMENETSKII AND V. S. RYABEN'KII, *Solution of Boundary Value Problems for the Laplace Equation in a Domain with a Cut by the Method of Difference Potentials* (in Russian), Preprint 33, Keldysh Institute for Applied Mathematics, USSR Academy of Sciences, Moscow, USSR, 1990.
- [26] E. KANSA, U. SHUMLAK, AND S. TSYNKOV, *Discrete Calderon's projections on parallelepipeds and their application to computing exterior magnetic fields for FRC plasmas*, J. Comput. Phys., 234 (2013), pp. 172–198.
- [27] M. L. LEVIN, S. M. RYTOV, AND V. D. SHAFRANOV, *M. A. Leontovich's researches in electrodynamics*, Sov. Phys. Usp., 26 (1983), pp. 353–355.
- [28] R. S. LEHMAN, *Developments at an analytic corner of solutions of elliptic partial differential equations*, J. Math. Mech., 8 (1959), pp. 727–760.
- [29] Z. LI AND K. ITO, *The Immersed Interface Method*, Frontiers in Appl. Math. 33, SIAM, Philadelphia, 2006.
- [30] X.-D. LIU, R. P. FEDKIW, AND M. KANG, *A boundary condition capturing method for Poisson's equation on irregular domains*, J. Comput. Phys., 160 (2000), pp. 151–178.
- [31] J. LONČARIĆ, V. S. RYABEN'KII, AND S. V. TSYNKOV, *Active shielding and control of noise*, SIAM J. Appl. Math., 62 (2001), pp. 563–596.
- [32] V. MANTIĆ, F. PARÍS, AND J. BERGER, *Singularities in 2D anisotropic potential problems in multi-material corners: Real variable approach*, Internat. J. Solids Structures, 40 (2003), pp. 5197–5218.
- [33] L. MARIN, *Treatment of singularities in the method of fundamental solutions for two-dimensional Helmholtz-type equations*, Appl. Math. Model., 34 (2010), pp. 1615–1633.
- [34] L. MARIN, D. LESNIC, AND V. MANTIĆ, *Treatment of singularities in Helmholtz-type equations using the boundary element method*, J. Sound Vibration, 278 (2004), pp. 39–62.
- [35] M. MEDVINSKY, S. TSYNKOV, AND E. TURKEL, *The method of difference potentials for the Helmholtz equation using compact high order schemes*, J. Sci. Comput., 53 (2012), pp. 150–193.
- [36] M. MEDVINSKY, S. TSYNKOV, AND E. TURKEL, *High order numerical simulation of the transmission and scattering of waves using the method of difference potentials*, J. Comput. Phys., 243 (2013), pp. 305–322.
- [37] F. OBERHETTINGER, *Diffraction of waves by a wedge*, Comm. Pure Appl. Math., 7 (1954), pp. 551–563.
- [38] C. S. PESKIN, *The immersed boundary method*, Acta Numer., 11 (2002), pp. 479–517.
- [39] A. A. REZNIK, *Approximation of surface potentials of elliptic operators by difference potentials*, Dokl. Akad. Nauk SSSR, 263 (1982), pp. 1318–1321.
- [40] A. A. REZNIK, *Approximation of the Surface Potentials of Elliptic Operators by Difference Potentials and Solution of Boundary-Value Problems* (in Russian), Ph.D. thesis, Moscow Institute for Physics and Technology, Moscow, USSR, 1983.
- [41] R. J. RIDDELL, JR., *Boundary-distribution solution of the Helmholtz equation for a region with corners*, J. Comput. Phys., 31 (1979), pp. 21–41.
- [42] R. J. RIDDELL, JR., *Numerical solution of the Helmholtz equation for two-dimensional polygonal regions*, J. Comput. Phys., 31 (1979), pp. 42–59.
- [43] V. S. RYABEN'KII, *Boundary equations with projections*, Russian Math. Surveys, 40 (1985), pp. 147–183.

- [44] V. S. RYABEN'KII, *Difference potentials method and its applications*, Math. Nachr., 177 (1996), pp. 251–264.
- [45] V. S. RYABEN'KII, *Method of Difference Potentials and Its Applications*, Springer Ser. Comput. Math. 30, Springer-Verlag, Berlin, 2002.
- [46] V. S. RYABEN'KII, *On the method of difference potentials*, J. Sci. Comput., 28 (2006), pp. 467–478.
- [47] V. S. RYABEN'KII, *Difference potentials analogous to Cauchy integrals*, Russian Math. Surveys, 67 (2012), pp. 541–567.
- [48] V. S. RYABEN'KII AND S. V. TSYNKOV, *A Theoretical Introduction to Numerical Analysis*, Chapman & Hall/CRC, Boca Raton, FL, 2007.
- [49] R. T. SEELEY, *Singular integrals and boundary value problems*, Amer. J. Math., 88 (1966), pp. 781–809.
- [50] T. B. A. SENIOR AND JOHN L. VOLAKIS, *Approximate Boundary Conditions in Electromagnetics*, IEEE Electromagnetic Waves Series 41, IEEE, London, 1995.
- [51] I. SINGER AND E. TURKEL, *High-order finite difference methods for the Helmholtz equation*, Comput. Methods Appl. Mech. Engrg., 163 (1998), pp. 343–358.
- [52] I. SINGER AND E. TURKEL, *Sixth-order accurate finite difference schemes for the Helmholtz equation*, J. Comput. Acoust., 14 (2006), pp. 339–351.
- [53] H. C. STRIFORS AND G. C. GAUNAURD, *Scattering of electromagnetic pulses by simple-shaped targets with radar cross section modified by a dielectric coating*, IEEE Trans. Antennas and Propagation, 46 (1998), pp. 1252–1262.
- [54] H. H. SYED AND J. L. VOLAKIS, *Electromagnetic Scattering by Coated Convex Surfaces and Wedges Simulated by Approximate Boundary Conditions*, Contractor report 190540, NASA, 1992.
- [55] H. TAL-EZER, *Non-periodic trigonometric approximation*, J. Sci. Comput., submitted.
- [56] S. V. TSYNKOV, *External boundary conditions for three-dimensional problems of computational aerodynamics*, SIAM J. Sci. Comput., 21 (1999), pp. 166–206.
- [57] S. V. TSYNKOV, *On the definition of surface potentials for finite-difference operators*, J. Sci. Comput., 18 (2003), pp. 155–189.
- [58] E. TURKEL, D. GORDON, R. GORDON, AND S. TSYNKOV, *Compact 2D and 3D sixth order schemes for the Helmholtz equation with variable wave number*, J. Comput. Phys., 232 (2013), pp. 272–287.
- [59] S. V. UTYUZHNIKOV, *Generalized Calderón-Ryaben'kii's potentials*, IMA J. Appl. Math., 74 (2009), pp. 128–148.
- [60] W. WASOW, *Asymptotic development of the solution of Dirichlet's problem at analytic corners*, Duke Math. J., 24 (1957), pp. 47–56.
- [61] N. M. WIGLEY, *Asymptotic expansions at a corner of solutions of mixed boundary value problems*, J. Math. Mech., 13 (1964), pp. 549–576.