Numer. Math.
DOI 10.1007/s00211-006-0040-3

Numerische Mathematik

Error analysis of an enhanced DtN-FE method for exterior scattering problems

David P. Nicholls · Nilima Nigam

Received: 23 August 2005 / Revised: 13 June 2006

Abstract In this work we analyze the convergence of the high-order Enhanced DtN-FEM algorithm, described in our previous work (Nicholls and Nigam, J. Comput. Phys. 194:278–303, 2004), for solving exterior acoustic scattering problems in $\mathbf{R}^2$. This algorithm consists of using an exact Dirichlet-to-Neumann (DtN) map on a hypersurface enclosing the scatterer, where the hypersurface is a perturbation of a circle, and, in practice, the perturbation can be very large. Our theoretical work had shown the DtN map was analytic as a function of this perturbation. In the present work, we carefully analyze the error introduced by virtue of using this algorithm. Specifically, we give a full account of the error introduced by truncating the DtN map at a finite order in the perturbation expansion, and study the well-posedness of the associated formulation. During computation, the Fourier series of the Dirichlet data on the artificial boundary must be truncated. To deal with the ensuing loss of uniqueness of solutions, we propose a modified DtN map, and prove well-posedness of the resulting problem. We quantify the spectral error introduced due to this truncation of the data. The key tools in the analysis include a new theorem on the analyticity of the DtN map in a suitable Sobolev space, and another on the perturbation of non-self-adjoint Fredholm operators.

D. P. Nicholls
Department of Mathematics, Statistics, and Computer Science,
University of Illinois at Chicago, 851 South Morgan Street,
Chicago, IL 60607, USA
e-mail: nicholls@math.uic.edu

N. Nigam (✉)
Department of Mathematics and Statistics,
McGill University, 805 Sherbrooke West,
Montréal, QC H3A 2K6, Canada
e-mail: nigam@math.mcgill.ca

1 Introduction

The propagation of acoustic and electromagnetic waves arises in a wide variety of applications, for example non-destructive testing, spectroscopy, remote sensing, and radar imaging. In such settings, robust and high-accuracy numerical approximation is an invaluable tool for design and simulation. A fundamental aspect of many of these problems is that they are most naturally stated on an unbounded domain, and conditions at spatial infinity must be imposed to specify a unique solution [14]. Such domains and conditions pose severe challenges for numerical simulations. For computational strategies involving volumetric discretizations the domain needs to be truncated to one of finite extent before numerical simulation. Typically, finite elements are used to discretize the region between the scatterer and an artificial boundary, and the question arises as to what boundary conditions to prescribe at this artificial boundary. Clearly, to have any hope of stating a well-posed problem, the solution on the exterior of this boundary must be resolved to some extent. This can be achieved by several means: Boundary element methods, infinite elements, absorbing boundary conditions, or, as we advocate, transparent boundary conditions. An ideal condition on this artificial surface will not only preserve the well-posedness of the original problem, but will also be highly accurate and computationally efficient. In this paper we present the rigorous error analysis of such a method, the "Enhanced DtN-FE" method, which we proposed in [43]. The technique involves constructing efficient and exact boundary conditions in the frequency domain (i.e., for time-harmonic incident radiation). We note that our present analysis is performed for the two-dimensional case, however, most of the mathematical techniques can be extended to the three-dimensional situation (provided that the artificial boundary is slightly smoother, see Sect. 3).

If the artificial boundary is denoted $\mathcal{B}$, an exact (Neumann) boundary condition on it can be specified via the Dirichlet-to-Neumann (DtN) map (see Sect. 2 for more details). This is a map between the Dirichlet data on a given surface to Neumann data, and is thus a natural candidate for a suitable boundary condition. The map is well-studied, and is also referred to as the Dirichlet–Neumann operator or the Stekhlov–Poincaré map. Such maps are defined for quite general $\mathcal{B}$, are non-local in nature, and lead to well-posed problems. They can be coupled to different volumetric discretizations, including finite difference and finite volume methods. The numerical implementation of these maps presents an interesting algorithmic challenge, most notably in trying to simultaneously achieve accuracy and efficiency of implementation. The explicit form of the fundamental solution of the governing equations (the Helmholtz equation) suggests a boundary integral approach and several have been implemented [36,39,21,18,33]. We point out the recent developments of Bruno and Kunyansky [10,11], and Ganesh and Graham [24] which have dealt with many of the restrictions of these methods in three dimensions, particularly the treatment of singular kernels and the lack of high-order quadrature rules.

On the other hand, if the surface $\mathcal{B}$ is such that its infinite complement is separable (e.g., if $\mathcal{B}$ is circular or elliptical) then the DtN map can be computed *explicitly* using separation of variables [22,34,41]. This observation leads to the DtN-FE method of Feng and Yu [22,50–53,49,54,38] (where it was called the Natural Boundary Element-Finite Element (NBE-FE) or Natural Boundary Reduction-Finite Element (NBR-FE) method), Han and Wu [34], Keller and Givoli [41,27,28,26], and Keller and Grote [31] (please see [47] for a brief history of the DtN-FE method). However, this approach is limited in two ways: The requirement that the artificial boundary be of a quite simple shape, and the non-local nature of the DtN map. The latter shortcoming results in dense sub-matrices being introduced into otherwise sparse linear systems. This concern has been examined by several authors [19,6] (see also [26]), and many *local* approximate boundary conditions have been devised, including the perfectly matched layer [8,1,7]. Of course, as local approximations of global operators, these boundary conditions have limitations, not the least of which result in questions regarding their stability. The DtN-FE method, which makes no such approximation avoids any spurious reflections which may arise from an artificial boundary.

In a recent paper [43] we addressed the difficulty associated with the very specific *shape* required of $\mathcal{B}$. We described an algorithm for two-dimensional time-harmonic scattering which permits quite general artificial boundaries of the form

$$\mathcal{B} = \{r = a + \delta f(\theta) \mid 0 \leq \theta < 2\pi\},$$

which still utilize the *exact* DtN map. Considering the form of $\mathcal{B}$ given above, a perturbative approach to evaluating the DtN map seems natural and is quite powerful. The authors showed that not only is the DtN map analytic in the parameter δ (if f is sufficiently smooth), but it can also be analytically continued to any *real* value of the parameter δ up to physical obstruction [43]. A numerical implementation of this "Enhanced DtN-FE" method validated this theoretical work by preserving the order of convergence of the underlying piecewise linear FE method. In addition, we demonstrated that numerical analytic continuation (via Padé approximation) permitted the use of artificial boundaries which were *large* departures from a circle.

The objective of the current paper is to give a rigorous numerical analysis of this method in the spirit of Harari and Hughes [35] and Demkowicz and Ihlenburg [17] for the original DtN-FE method. We consider the effects of truncating the perturbative approximation of the DtN map, and rigorously prove well-posedness and convergence of the resulting formulations. We then examine the effect of truncating the Fourier series of the Dirichlet data on $\mathcal{B}$. This will clearly lead to stability problems. To alleviate these problems, we modify our DtN map in a manner analogous to that described in [31]. The resulting formulation is well-posed, as we shall show. Moreover, the modified DtN map is exact for the untruncated modes, and we can precisely describe the error introduced by the truncation procedure.

The organization of the paper is as follows: In Sect. 2 we review the governing equations of bounded-obstacle scattering and show how the DtN map can be used to specify transparent boundary conditions at a general artificial boundary $\mathcal{B}$. In Sect. 3 we recall known analyticity properties of DtN maps, and in Sect. 3.1 we give a new theorem on analyticity of these maps for Dirichlet data in the weak Sobolev class $H^{1/2}$. This result is necessary to establish the well-posedness of the variational formulations of Sect. 4; these results are proven in Sect. 5 for the full problem (using an elementary argument), and in Sect. 6 for the problem where the Taylor series of the DtN map is truncated after a finite number of terms. In Sect. 7 we introduce a modified DtN map suitable for use when the Fourier series of the boundary data is truncated. We prove the variational formulation involving this new DtN map is well-posed, and perform an error analysis. We make some concluding remarks in Sect. 8.

2 Governing equations

As we mentioned in the Introduction, the topic of this paper is the scattering of time-harmonic acoustic (or electromagnetic) radiation by a bounded obstacle. It is well-known that in two-dimensional problems the incident, scattered, and total acoustic (electromagnetic) fields all satisfy the Helmholtz equation [14]. Given the time-harmonic incident field

$$\tilde{v}_i(x) = \mathrm{e}^{-\mathrm{i}\omega t} v_i(x),$$

the scattered field will also be time-harmonic, and we are interested in determining the (reduced) scattered field, $v = v_s$. The physics of the obstacle, Σ, determines a condition on v at the boundary of the scatterer, $\Gamma := \partial\Sigma$, which need only be Lipschitz. We select a sound-hard (perfectly conducting in TM polarization in electromagnetics) obstacle for definiteness. However, other boundary conditions can be accommodated in a straightforward manner. Finally, to specify a unique solution we impose the Sommerfeld radiation condition which requires that scattered waves be outgoing. Together, these equations are

$$\Delta v + k^2 v = 0, \qquad x \in \mathbf{R}^2 \setminus \bar{\Sigma} \tag{1a}$$

$$\partial_n v = g, \qquad x \in \Gamma \tag{1b}$$

$$\lim_{r\to\infty} \sqrt{r}\,(\partial_r v - \mathrm{i}kv) = 0, \tag{1c}$$

where $g := -\partial_n v_i$ at Γ, and we use the standard notation for the Laplacian:

$$\Delta := \partial_r^2 + \frac{1}{r}\partial_r + \frac{1}{r^2}\partial_\theta^2.$$

In addition, $k^2 = (2\pi/\lambda)^2$, λ is the wavelength of radiation, and n is the unit normal pointing exterior to Σ.

Clearly, the problem (1) is not suitable for discretization by finite elements until the infinite computational domain is truncated. For this we introduce an

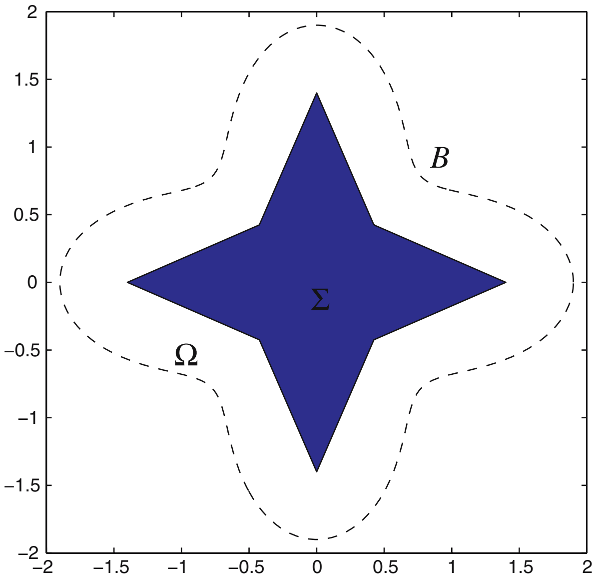

Fig. 1 Σ is the obstacle, $\mathcal{B}$ is the enclosing artificial boundary, and Ω is the "annulus" between them

artificial boundary $\mathcal{B}$ properly enclosing $\bar{\Sigma}$, and discretize the annular domain, Ω, between them, see Fig. 1. This introduces a natural domain decomposition which leads to a system of equations which are coupled across $\mathcal{B}$ and are equivalent to (1):

$$\Delta u + k^2 u = 0, \qquad x \in \Omega \tag{2a}$$

$$\partial_n u = g \qquad x \in \Gamma \tag{2b}$$

$$\partial_N u = \partial_N w \qquad x \in \mathcal{B} \tag{2c}$$

$$u = w \qquad x \in \mathcal{B} \tag{2d}$$

$$\Delta w + k^2 w = 0, \qquad x \in \mathrm{Ext}(\mathcal{B}) \tag{2e}$$

$$\lim_{r\to\infty} \sqrt{r}\,(\partial_r w - \mathrm{i}kw) = 0, \tag{2f}$$

where N is a normal vector to $\mathcal{B}$ directed towards Σ (see [43] for a particular choice and its importance). Gathering (2d)–(2f), we note that the resulting problem

$$\Delta w + k^2 w = 0, \qquad x \in \mathrm{Ext}(\mathcal{B}) \tag{3a}$$

$$w = u \qquad x \in \mathcal{B} \tag{3b}$$

$$\lim_{r\to\infty} \sqrt{r}\,(\partial_r w - \mathrm{i}kw) = 0, \tag{3c}$$

is, given the trace of u on $\mathcal{B}$, uniquely solvable. From this unique solution we can produce Neumann data,

$$\nu := \nabla w|_{\mathcal{B}} \cdot N,$$

and this procedure of producing Neumann data from Dirichlet data is known as the Dirichlet-to-Neumann (DtN) map (or, alternatively, the Dirichlet–Neumann operator or Steklov–Poincaré map). We introduce the notation

$$G(\mathcal{B})[u] := \nabla w|_{\mathcal{B}} \cdot N,$$

and note that G maps the Sobolev class $H^{s+1}(\mathcal{B})$ to $H^s(\mathcal{B})$, for any $s \geq 0$ (see, e.g., [44]). With this notation, (2a)–(2c) can be written as

$$\Delta u + k^2 u = 0, \qquad x \in \Omega, \tag{4a}$$
$$\partial_n u = g \qquad x \in \Gamma, \tag{4b}$$
$$\partial_N u = G(\mathcal{B})[u] \qquad x \in \mathcal{B}, \tag{4c}$$

which is completely equivalent to (1). The right-hand side of (4c) is meant to be understood as the DtN map applied to the trace of u at the boundary.

For future reference we note that for scattering problems (unlike problems from potential theory), the DtN map is not self-adjoint. This fact has real consequences for the current study as it prevents us from using the classical perturbation theory for self-adjoint operators on Banach spaces [40]. To realize this property consider a general Fourier multiplier, $m(D)$, defined by

$$m(D)[\xi] := \sum_{p=-\infty}^{\infty} m(p)\hat{\xi}_p e^{ip\theta},$$

where $\hat{\xi}_p$ is the pth Fourier coefficient of ξ. The natural inner product for 2π-periodic functions ξ and ψ is

$$\begin{aligned}
(\xi, \psi) &= \int_0^{2\pi} \xi(\theta)\bar{\psi}(\theta)\, d\theta, \\
&= \int_0^{2\pi} \sum_{p=-\infty}^{\infty} \hat{\xi}_p e^{ip\theta} \sum_{q=-\infty}^{\infty} \bar{\hat{\psi}}_q e^{-iq\theta}\, d\theta, \\
&= 2\pi \sum_{p=-\infty}^{\infty} \hat{\xi}_p \bar{\hat{\psi}}_p.
\end{aligned}$$

We now compute

$$(m(D)[\xi], \psi) = 2\pi \sum_{p=-\infty}^{\infty} m(p)\hat{\xi}_p \bar{\hat{\psi}}_p,$$

$$(\xi, m(D)[\psi]) = 2\pi \sum_{p=-\infty}^{\infty} \overline{m(p)}\hat{\xi}_p \bar{\hat{\psi}}_p,$$

which clearly indicates that $m(D)$ is self-adjoint only if $m(p)$ is real-valued for all integer p. We recall from [43] that the DtN map for scattering on a circular boundary $\mathcal{B} = \{r = a\}$ is specified by the Fourier multiplier

$$m(p) = -k \frac{d_z H_p^{(1)}(ka)}{H_p^{(1)}(ka)}, \tag{5}$$

where $H_p^{(1)}$ is the pth Hankel function of the first kind, and d_z denotes differentiation with respect to the argument of the Hankel function. This is clearly not real-valued and thus the DtN map is not self-adjoint even in this simple case.

The DtN map G incorporates information about the outgoing nature of the solution w. If this map can be accurately computed, solving problem (4) provides the solution of the original problem (1) in the near field, in the sense that v and u agree on Ω. Therefore, we can simply solve (4) for u. If required, we can then use the traces of u and its normal derivative $\partial_n u$ to compute v in the unbounded complement of $\mathcal{B}$, using a representation formula,

$$v(x) = \int_{\mathcal{B}} u(y)(\partial_{n(y)}\Phi(x, y)) - \Phi(x, y)(\partial_{n(y)}u(y))\mathrm{d}y,$$

for $x \in \mathrm{Ext}(\mathcal{B})$. Here,

$$\Phi(x, y) := \frac{\mathrm{i}}{4} H_0^{(1)}(k\,|x - y|)$$

is the fundamental solution for the Helmholtz equation.

3 Dirichlet-to-Neumann maps

As with all DtN-FE methods, the finite element approximation of solutions of (4) is standard save the treatment of the non-local pseudodifferential operator, G, appearing in (4c). In our paper [43], we showed how this operator could be computed for quite general $\mathcal{B}$ (deformations of a circle) using a Boundary Perturbation method. With this freedom one may fit the artificial boundary much more closely to the scatterer than if $\mathcal{B}$ is required to be a circle (cf.

[34,41,31,32]). The benefits of this are not only a smaller computational domain, but also better uniqueness and stability properties (see [35]).

The Boundary Perturbation approach that we advocate for the computation of the DtN map requires rigorous justification not only to give us confidence that our method will converge rapidly, but also to permit the numerical analysis which we provide in the ensuing sections. Such theories have been provided by several authors beginning with the work of Coifman and Meyer [13], based upon the results of Calderón [12], which state that, in a Cartesian geometry, the DtN map for a two-dimensional domain is analytic with respect to boundary deformations provided that the deformation is Lipschitz continuous. This result was extended by Craig et al. [16] and Craig and Nicholls [15] for three and general m dimensions, respectively, provided that the deformation is C^1. Using a completely different, and much more direct, technique Nicholls and Reitich showed that, in m dimensions, the DtN map is analytic as a function of boundary deformation provided that the profile is $C^{3/2+\alpha}$ for any $\alpha > 0$ [44], and *jointly* analytic in both parametric and spatial variables provided that the deformation itself is analytic [46]. While the results in [44] were not optimal in terms of boundary smoothness, they did point the way to a new, stable, high-order numerical scheme for the computation of DtN maps [45]. Subsequently, Hu and Nicholls have shown how these Nicholls–Reitich recursions *can* be used to realize analyticity within the class of Lipschitz perturbations [37], essentially recovering the original result of Coifman and Meyer.

All of this work was conducted for geometries most conveniently expressed in Cartesian coordinates. Moreover, since the wavenumber in these cases was $k = 0$, difficulties involving eigenvalues or exceptional values were not an issue. One component of our recent work [43] was to show that this analysis could be extended to the case of polar coordinates and $k > 0$. In fact, we showed that if G_0 denotes the DtN map on a circle, and if

$$\mathcal{B} := \{(r, \theta) \mid r = a + \delta f(\theta)\},$$

then G is an analytic perturbation of G_0. More precisely,

Theorem 1 (Nicholls and Nigam [43]) *Given an integer $s \geq 0$, if $f \in C^{s+2}([0, 2\pi])$ and $\xi \in H^{s+3/2}([0, 2\pi])$ then the series*

$$G(\delta f)[\xi] = \sum_{n=0}^{\infty} G_n(f)[\xi]\, \delta^n \tag{6}$$

converges strongly as an operator from $H^{s+3/2}([0, 2\pi])$ to $H^{s+1/2}([0, 2\pi])$. In other words there exist constants $\tilde{K}_1$ and C, depending on the smoothness s, such that

$$\|G_n(f)[\xi]\|_{H^{s+1/2}} \leq \tilde{K}_1\, \|\xi\|_{H^{s+3/2}}\, \tilde{B}^n$$

for any $\tilde{B} > C\, |f|_{C^{s+2}}$.

The main goal of this paper is to rigorously analyze the error introduced in the Enhanced DtN-FE method by approximating the DtN map with a truncation of the expansion (6) to a finite number of terms N, i.e.,

$$G(\delta f)[\xi] \approx G^N(\delta f)[\xi] := \sum_{n=0}^{N} G_n(f)[\xi]\, \delta^n. \tag{7}$$

We accomplish this by establishing the well-posedness of the weak formulations of not only (4), but also (4) with G replaced by G^N. Clearly, the correct functional setting for this problem, particularly if the scatterer is only Lipschitz continuous, is to seek solutions u in $H^1(\Omega)$. This results in Dirichlet traces which sit in the Sobolev space $H^{1/2}(\mathcal{B})$ and therefore the estimates given by Theorem 1 are insufficient. However, a modification of the proof given in [43], which we elucidate in the next subsection, allows us to prove the following theorem which can be utilized.

Theorem 2 *Given an integer $s \geq 0$, if $f \in H^{s+5}([0, 2\pi])$ and $\xi \in H^{s+1/2}([0, 2\pi])$ then the series (6) converges strongly as an operator from $H^{s+1/2}([0, 2\pi])$ to $H^{s-1/2}([0, 2\pi])$. In other words there exist constants C and K_1 (depending on s) such that*

$$\|G_n(f)[\xi]\|_{H^{s-1/2}} \leq K_1\, \|\xi\|_{H^{s+1/2}}\, B^n$$

for any $B > C\, \|f\|_{H^{s+5}}$.

3.1 Analyticity of the DtN map for weak Dirichlet data

In this subsection we show how the analyticity proof for DtN maps given in [43] can be extended to include quite rough Dirichlet data in the class $H^s([0, 2\pi])$, $s \geq 1/2$. We note that the *boundedness* of the DtN map has been previously established in such weak spaces, see e.g., [4,5]. The cost of this extension is a stricter smoothness requirement on the boundary perturbation f (we will require $f \in H^5$ rather than $f \in$ Lip when $\xi \in H^1([0, 2\pi])$, [13]). In fact, the difference between the proof presented in [43] and the abridged one given here is that the "algebra estimates" [44] of Lemma 3 are replaced by those of Lemma 4 [23].

Lemma 3 *For any integer $s \geq 0$, any $\varepsilon > 0$, and any set $U \subset \mathbf{R}^m$, if $f, u, g, \mu : U \to \mathbf{C}$, $f \in C^s(U)$, $u \in H^s(U)$, $g \in C^{s+1/2+\varepsilon}(U)$, $\mu \in H^{s+1/2}(U)$, then*

$$\|fu\|_{H^s} \leq \tilde{M}(m, s, U)\, |f|_{C^s}\, \|u\|_{H^s}\,,$$
$$\|g\mu\|_{H^{s+1/2}} \leq \tilde{M}(m, s, U)\, |g|_{C^{s+1/2+\varepsilon}}\, \|\mu\|_{H^{s+1/2}}\,,$$

for some constant $\tilde{M}$.

Lemma 4 *For any $s \in \mathbf{R}$ and any set $U \subset \mathbf{R}^m$, if $\varphi, \psi : U \to \mathbf{C}$, $\varphi \in H^{|s|+m+2}(U)$, and $\psi \in H^s(U)$, then*

$$\|\varphi\psi\|_{H^s} \leq M(m,s,U)\,\|\varphi\|_{H^{|s|+m+2}}\,\|\psi\|_{H^s}\,, \tag{8}$$

for some constant M.

Remark 5 In fact, the result found in [23] (Proposition 6.16) states that, for any $s \in \mathbf{R}$,

$$\|\varphi\psi\|_{H^s} \leq |\varphi|_{L^\infty}\,\|\psi\|_{H^s} + C(m,s,U)\,\|\varphi\|_{H^{|s|+m+2}}\,\|\psi\|_{H^{s-1}}\,. \tag{9}$$

By Sobolev embedding results [2,25], if $\varphi \in H^t$ and $t > m/2$ then $\varphi \in L^\infty$. In this way (9) implies (8).

Recall that we have now specialized to the case of artificial boundaries of the form

$$\mathcal{B} = \{(r,\theta) \mid r = a + \delta f(\theta)\},$$

and that the Helmholtz problem which defines the DtN map, $G[\xi]$, is

$$\begin{aligned} &\Delta w + k^2 w = 0, && r > a + \delta f \\ &w(a+\delta f(\theta),\theta) = \xi(\theta) && \\ &\lim_{r\to\infty} \sqrt{r}\,(\partial_r w - \mathrm{i}kw) = 0. && \end{aligned}$$

To simplify the analysis, in [43] we introduced a second, exterior artificial boundary at $r = b$ ($b > a + \delta\,|f|_{L^\infty}$), and a second transparent boundary condition via the (DtN) operator T

$$T(b)[\mu] := T(b)\left[\sum_{p=-\infty}^{\infty} \hat{\mu}_p \mathrm{e}^{\mathrm{i}p\theta}\right] = \sum_{p=-\infty}^{\infty} k\frac{d_z H_p^{(1)}(kb)}{H_p^{(1)}(kb)}\hat{\mu}_p \mathrm{e}^{\mathrm{i}p\theta},$$

cf. (5), with the orientation of the normal reversed. In the same way that we restated (1) as (4), we can equivalently pose (3) as

$$\Delta w + k^2 w = 0, \qquad \text{in } A_{a+\delta f,b} \tag{10a}$$

$$w(a+\delta f(\theta),\theta) = \xi(\theta) \tag{10b}$$

$$\partial_r w = Tw \qquad \text{at } r = b, \tag{10c}$$

where $A_{a+\delta f,b} := \{(r,\theta) \mid a + \delta f(\theta) < r < b\}$.

The Transformed Field Expansions (TFE) method [44–46] has proven quite successful in establishing analyticity properties for boundary value and free boundary problems, and we shall use it again in this setting. This TFE method proceeds by effecting a "domain flattening" change of variables which, in this geometry, is

$$r' = \frac{(a-b)r + \delta b f(\theta)}{a-b+\delta f(\theta)}, \quad \theta' = \theta. \tag{11}$$

Notice that this transformation maps the perturbed annulus $A_{a+\delta f,b}$ to the annulus $A_{a,b}$. Defining

$$U(r',\theta') = w\left(\frac{(a-b+\delta f(\theta'))r' - \delta b f(\theta')}{a-b}, \theta'\right),$$

the change of variables (11) transforms (10), upon dropping primes, into

$$\Delta U + k^2 U = F(r,\theta;U,f) \quad \text{in } A_{a,b} \tag{12a}$$
$$U(a,\theta) = \xi(\theta) \tag{12b}$$
$$\partial_r U = TU + h(\theta) \quad \text{at } r=b. \tag{12c}$$

The precise forms of F and h, both of which are $\mathcal{O}(\delta)$, are given in [43]. Following the TFE philosophy, we expand the transformed field in a Taylor series

$$U(r,\theta,\delta) = \sum_{n=0}^{\infty} U_n(r,\theta)\,\delta^n,$$

and derive equations for the U_n:

$$\Delta U_n + k^2 U_n = (1-\delta_{n,0})F_n(r,\theta;U_l,f) \quad \text{in } A_{a,b} \tag{13a}$$
$$U_n(a,\theta) = \delta_{n,0}\,\xi(\theta) \tag{13b}$$
$$\partial_r U_n = TU_n + (1-\delta_{n,0})h_n(\theta) \quad \text{at } r=b, \tag{13c}$$

where $\delta_{n,l}$ is the Kronecker delta. Again, the precise form of F_n is given in [43], however, let us point out that

$$F_n = \frac{1}{(a-b)^2}\left[\partial_r F_n^{(1)} + \partial_\theta F_n^{(2)} + F_n^{(3)}\right],$$

where no more than one derivative acts upon a U_l $(0 \le l \le n-1)$ appearing in the $F_n^{(j)}$. A representative term is

$$F_n^{(1)}(r,\theta) = -(\partial_\theta f(\theta))^2(b-r)^2\partial_r U_{n-2}(r,\theta) + \cdots, \tag{14}$$

while the h_n are simply

$$h_n(\theta) = \frac{f(\theta)}{a-b}(TU_{n-1})(b,\theta). \tag{15}$$

In addition, the DtN map, G, must be stated in transformed coordinates and then expanded in a Taylor series:

$$G(\delta f)[\xi] = \sum_{n=0}^{\infty} G_n(f)[\xi]\,\delta^n.$$

The G_n were fully derived in [43], where we showed that

$$G_n(f)[\xi] = -\frac{1}{a}(\partial_\theta f(\theta))^2 \partial_r U_{n-2}(a,\theta) + \cdots. \tag{16}$$

Remark 6 As we shall see in the following calculations, we must estimate terms of the form:

$$\|(\partial_\theta f)V\|_{H^s(A_{a,b})}, \quad \|(\partial_\theta f)\mu\|_{H^{s-1/2}([0,2\pi])},$$

for integer $s \geq 0$. To accomplish this we use Lemma 4 with $U = A_{a,b}$ $(m = 2)$ and $U = [0, 2\pi]$ $(m = 1)$, respectively. We can now bound these terms

$$\begin{aligned}
\|(\partial_\theta f)V\|_{H^s(A_{a,b})} &\leq M\,\|(\partial_\theta f)\|_{H^{s+2+2}(A_{a,b})}\,\|V\|_{H^s(A_{a,b})} \\
&\leq M\,\|f\|_{H^{s+5}([0,2\pi])}\,\|V\|_{H^s(A_{a,b})}, \\
\|(\partial_\theta f)\mu\|_{H^{s-1/2}([0,2\pi])} &\leq M\,\|(\partial_\theta f)\|_{H^{|s-1/2|+1+2}([0,2\pi])}\,\|\mu\|_{H^s([0,2\pi])} \\
&\leq M\,\|f\|_{H^{s+4}([0,2\pi])}\,\|\mu\|_{H^s([0,2\pi])}.
\end{aligned}$$

Thus, if we require $f \in H^{s+5}([0, 2\pi])$ all of these terms are bounded; in the special case $s = 0$ (which we will consider exclusively in later sections) this, of course, implies $f \in H^5$. We note that if $f \in H^5([0, 2\pi])$ then $f \in C^4([0, 2\pi])$ by standard Sobolev embeddings [2,25].

With our transformation and notation in place we can finally state our new result which will establish Theorem 2.

Theorem 7 *Given any integer $s \geq 0$, if $f \in H^{s+5}([0, 2\pi])$ and $\xi \in H^{s+1/2}([0, 2\pi])$ there exist constants C_0 and K_0, and a unique solution $U_n \in H^{s+1}(A_{a,b})$ of (13) such that*

$$\|U_n\|_{H^{s+1}} \leq K_0\,\|\xi\|_{H^{s+1/2}}\,B^n \tag{17}$$

for any $B > 2K_0 C_0\,\|f\|_{H^{s+5}}$

To prove this we inductively estimate the problem (13) using an elliptic lemma (Lemma 8) and a recursive lemma (Lemma 9). To begin we state, without proof, the following well-known elliptic estimate [25,42,43].

Lemma 8 *For any integer $s \geq 0$ there exists a constant K_0 such that for any $F \in H^{s-1}(A_{a,b})$, $\xi \in H^{s+1/2}([0, 2\pi])$, $h \in H^{s-1/2}([0, 2\pi])$, the solution $W \in H^{s+1}(A_{a,b})$ of*

$$\begin{aligned} &\Delta W + k^2 W = F(r, \theta) && \text{in } A_{a,b} \\ &W(a, \theta) = \xi(\theta) && \\ &\partial_r W = TW + h(\theta) && \text{at } r = b \\ &W(r, \theta + 2j\pi) = W(r, \theta) && \forall j \in \mathbf{Z}, \end{aligned}$$

satisfies

$$\|W\|_{H^{s+1}} \leq K_0 \left[\|F\|_{H^{s-1}} + \|\xi\|_{H^{s+1/2}} + \|h\|_{H^{s-1/2}} \right].$$

To control the right-hand side of (13) we prove the following.

Lemma 9 *Let $s \geq 0$ be an integer and $f \in H^{s+5}([0, 2\pi])$. Assume that, for $U_n \in H^{s+1}(A_{a,b})$,*

$$\|U_n\|_{H^{s+1}} \leq K_2 B^n$$

for all $n < N$ and constants K_2 and B. If

$$B > \|f\|_{H^{s+5}},$$

then $F_N \in H^{s-1}(A_{a,b})$, $h_N \in H^{s-1/2}([0, 2\pi])$, and there exists a C_0 such that

$$\begin{aligned} \|F_N\|_{H^{s-1}} &\leq K_2 \|f\|_{H^{s+5}} C_0 B^{N-1}, \\ \|h_N\|_{H^{s-1/2}} &\leq K_2 \|f\|_{H^{s+5}} C_0 B^{N-1}. \end{aligned}$$

Proof Note that

$$\begin{aligned} \|F_N\|_{H^{s-1}} &\leq \frac{1}{(a-b)^2} \left\{ \left\| \partial_r F_N^{(1)} \right\|_{H^{s-1}} + \left\| \partial_\theta F_N^{(2)} \right\|_{H^{s-1}} + \left\| F_N^{(3)} \right\|_{H^{s-1}} \right\} \\ &\leq \frac{1}{(a-b)^2} \left\{ \left\| F_N^{(1)} \right\|_{H^s} + \left\| F_N^{(2)} \right\|_{H^s} + \left\| F_N^{(3)} \right\|_{H^s} \right\}, \end{aligned}$$

and consider the representative term (14)

$$\begin{aligned} \left\| F_N^{(1)} \right\|_{H^s} &\leq \left\| -(\partial_\theta f(\theta))^2 (b-r)^2 \partial_r U_{N-2}(r, \theta) \right\|_{H^s} + \cdots \\ &\leq M^2 \|\partial_\theta f\|_{H^{s+4}}^2 R^2 \|U_{N-2}\|_{H^{s+1}} + \cdots \\ &\leq K_2 \|f\|_{H^{s+5}} (C_0/3) B^{N-1} \end{aligned}$$

for some C_0, where we have used the inductive hypothesis, Lemma 4, and

$$\|(b-r)W\|_{H^s} \leq R(s)\,\|W\|_{H^s}$$

for some constant $R = R(s)$. We are finished provided that $B > \|f\|_{H^{s+5}}$. Regarding h_N we simply compute

$$\begin{aligned}
\|h_N\|_{H^{s-1/2}} &\leq \left\| \frac{f}{a-b} TU_{N-1} \right\|_{H^{s-1/2}} \\
&\leq \frac{M}{|a-b|} \|f\|_{H^{|s-1/2|+3}} \left\| TU_{N-1} \right\|_{H^{s-1/2}} \\
&\leq \frac{M}{|a-b|} \|f\|_{H^{|s-1/2|+3}}\, C_T \left\| U_{N-1} \right\|_{H^{s+1/2}} \\
&\leq \frac{MC_T}{|a-b|} \|f\|_{H^{|s-1/2|+3}}\, C_{\mathrm{tr}} \left\| U_{N-1} \right\|_{H^{s+1}(A_{a,b})} \\
&\leq K_2 \|f\|_{H^{|s-1/2|+3}}\, C_0 B^{N-1} \leq K_2 \|f\|_{H^{s+5}}\, C_0 B^{N-1},
\end{aligned}$$

where C_T and C_{tr} are bounding constants for the DtN map T and trace operator, respectively.

We can now present the proof of Theorem 7.

Proof (Theorem 7) The estimate on U_0 follows directly from Lemma 8 with F and h identically zero. Letting $K_2 := K_0\,\|\xi\|_{H^{s+1/2}}$ we now assume that (17) holds for all $n < N$. Applying Lemma 8 implies that

$$\|U_N\|_{H^{s+1}} \leq K_0 \left[\|F_N\|_{H^{s-1}} + \|h_N\|_{H^{s-1/2}} \right].$$

Lemma 9 implies that

$$\|U_N\|_{H^{s+1}} \leq K_2 \left(2K_0 C_0 \|f\|_{H^{s+5}} \right) B^{N-1},$$

and (17) is verified provided $B > 2K_0 C_0 \|f\|_{H^{s+5}}$.

Finally, we establish Theorem 2.

Proof (Theorem 2) We merely consider the representative term (16)

$$\begin{aligned}
\|G_n(f)[\xi]\|_{H^{s-1/2}} &\leq \left\| -\frac{1}{a} (\partial_\theta f)^2 \partial_r U_{n-2}(a,\theta) \right\|_{H^{s-1/2}} + \cdots \\
&\leq \frac{M^2}{a} \|\partial_\theta f\|^2_{H^{|s-1/2|+3}} \|\partial_r U_{n-2}(a,\theta)\|_{H^{s-1/2}} + \cdots \\
&\leq \frac{M^2}{a} \|f\|^2_{H^{|s-1/2|+4}}\, C_{\mathrm{tr}} \|U_{n-2}\|_{H^{s+1}(A_{a,b})} + \cdots \\
&\leq \frac{M^2}{a} \|f\|^2_{H^{s+4}}\, C_{\mathrm{tr}} K_0 \|\xi\|_{H^{s+1/2}} B^{n-2} + \cdots \\
&\leq K_1 \|\xi\|_{H^{s+1/2}} B^n,
\end{aligned}$$

where we have used $B > 2K_0C_0 \|f\|_{H^{s+5}}$ and chosen K_1 to be a constant depending on M, a, C_{tr}, K_0, and C_0, but not ξ or f.

4 Variational formulation

We are now in a position to examine the effect of using the truncated DtN map G^N in place of G. To facilitate our analysis we gather here some convenient notation that occurs frequently. We denote by V the Hilbert space $H^1(\Omega)$, equipped with the usual norm:

$$V := \{u \in L^2(\Omega) \mid \nabla u \in L^2(\Omega)\}, \quad \|u\|_V^2 := \int_\Omega |u|^2 + |\nabla u|^2 \, \mathrm{d}V.$$

We denote by V' the dual space of V, and use $\langle\langle\cdot,\cdot\rangle\rangle$ for the L^2-duality pairing between V and V'. The norm $\|\cdot\|_{L(V,V')}$ describes the operator norm for linear operators between V and V'.

The underlying assumption on Ω is that it has a boundary smooth enough so that the trace operator $[\cdot] : V \to H^{1/2}(\mathcal{B})$ is well-defined for elements of V. We are thus able to define the Sobolev space

$$H^{1/2}(\mathcal{B}) := \left\{\mu \in L^2(\mathcal{B}) \mid \exists\, v \in V \text{ such that } [v] = \mu \text{ on } \Gamma\right\}.$$

We denote by $\langle\cdot,\cdot\rangle$ the L^2-duality pairing between $H^{1/2}(\mathcal{B})$ and $H^{-1/2}(\mathcal{B})$, and we reserve $(\cdot,\cdot)$ for the $L^2(\Omega)$ inner product,

$$(u, v) := \int_\Omega u\, \bar{v} \mathrm{d}V.$$

The notation $\mathcal{A}(\cdot,\cdot)$ (sometimes with subscripts) is used for the sesquilinear form $\mathcal{A}$ on $V \times V$ defined by

$$\mathcal{A}(w, v) := \int_\Omega \nabla w \cdot \nabla \bar{v} - k^2 w\, \bar{v} \mathrm{d}V + \langle G[w], v\rangle.$$

Given this notation we now give three variational problems which approximate our original problem (4) to varying degrees, and lead us to the total error of our numerical procedure. The first is the weak formulation of the full problem (4):

$$\text{Find } u \in V \text{ such that for all } v \in V,$$

$$\mathcal{A}(u, v) = \int_\Gamma g\, \bar{v} \mathrm{d}s. \tag{18}$$

The well-posedness of Eq. (18) will be studied in Sect. 5. The second variational problem, which we analyze in Sect. 6, is the weak formulation of (4) where the DtN map G is replaced by the truncation G^N, (7):

$$\begin{aligned}&\text{Find } u^N \in V \text{ such that for all } v \in V,\\ &\mathcal{A}^N(u^N, v) = \int_\Gamma g\,\bar{v}\mathrm{d}s,\end{aligned} \tag{19}$$

where

$$\mathcal{A}^N(w,v) := \int_\Omega \nabla w \cdot \nabla \bar{v} - k^2 w\,\bar{v}\mathrm{d}V + \left\langle G^N[w], v\right\rangle.$$

To motivate the final formulation, we recall that it is most convenient in computations of the DtN map to utilize the Fourier series representation of the boundary trace ξ of the scattered field u at $\mathcal{B}$, [43]. Of course this Fourier series will also be truncated after a finite number of terms, say m, and we represent this procedure by the operator

$$G^{N,m}[\xi] = G^{N,m}(\delta f)[\xi] := \sum_{n=0}^{N} G_n(f)\left[\sum_{p=-m}^{m} \hat{\xi}_p \mathrm{e}^{\mathrm{i}p\theta}\right]\delta^n.$$

Also useful in this formulation is the space of all functions in V which have Dirichlet trace at $\mathcal{B}$ with only $(2m+1)$ Fourier coefficients,

$$B^m := \left\{ v \in V \;\middle|\; v|_{\mathcal{B}} \in H^{1/2}(\mathcal{B}),\, v|_{\mathcal{B}} = \sum_{p=-m}^{m} \hat{\xi}_p \mathrm{e}^{\mathrm{i}p\theta}\;,\; 0 \le \theta < 2\pi. \right\}. \tag{20}$$

In Sect. 7 we consider the variational formulation:

$$\begin{aligned}&\text{Find } u^{N,m} \in B^m \text{ such that for all } v \in V,\\ &\mathcal{A}^{N,m}(u^{N,m}, v) = \int_\Gamma g\,\bar{v}\mathrm{d}s,\end{aligned} \tag{21}$$

where

$$\mathcal{A}^{N,m}(w,v) := \int_\Omega \nabla w \cdot \nabla \bar{v} - k^2 w\,\bar{v}\mathrm{d}V + \left\langle G^{N,m}[w], v\right\rangle.$$

As noted by several authors, (e.g., [31,35]), when the artificial boundary is a circle, the truncation in Fourier space leads to a loss of uniqueness. To deal with this, Grote and Keller proposed a modification to the DtN map which remains exact for the first m modes, but stabilizes the formulation. We shall introduce and analyze a similar modification in Sect. 7.

If $u^{N,m,h}$ is the finite element solution of (21) then the total error for our scheme is

$$\left\|u - u^{N,m,h}\right\|_V \leq \left\|u - u^N\right\|_V + \left\|u^N - u^{N,m}\right\|_V + \left\|u^{N,m} - u^{N,m,h}\right\|_V . \quad (22)$$

This error is composed of a consistency error, a spectral approximation error, and a finite element approximation error. In Sect. 5 we show that the full variational problem (18) is indeed well-posed, and subsequently in Sect. 6, we estimate the consistency error $\|u - u^N\|_V$ introduced by using a finite number of terms in the power series of the DtN map. In order to achieve this estimate we first need to show that $\mathcal{A}^N$ satisfies a discrete inf-sup condition, which, in turn, requires information concerning the well-posedness of the variational problem (19). Thereafter in Sect. 7 we study the spectral approximation error $\|u^N - u^{N,m}\|_V$ induced by truncating the Fourier series of the boundary trace in DtN map computations (and using a modified DtN map). The final, finite element approximation error $\|u^{N,m} - u^{N,m,h}\|_V$ can be estimated using standard techniques.

5 Well-posedness of the full variational problem

To begin our analysis of the full variational problem, (18), we recall the results of Demkowicz and Ihlenburg [17], and Harari and Hughes [35] in the case $\mathcal{B} = \{r = a\}$. Defining the linear operator B_0 via

$$\langle\langle B_0 u, v\rangle\rangle_0 := \int_{\Omega_0} \nabla u \cdot \nabla \bar{v} - k^2 u\, \bar{v} \mathrm{d}V + \int_{r=a} G_0[u]\, \bar{v} \mathrm{d}s,$$

where Ω_0 is the annulus between Γ and $\{r = a\}$, they showed that the variational problem:

$$\langle\langle B_0 u, v\rangle\rangle_0 = \int_{\Gamma} g\, \bar{v} \mathrm{d}V \quad \forall v \in H^1(\Omega_0)$$

is well-posed. Here $G_0 = G(\{r = a\})$ is the DtN map associated with a circular artificial boundary. They showed that the linear operator $B_0 : H^1(\Omega_0) \to (H^1(\Omega_0))'$ was Fredholm, and that the variational problem had a unique solution in $H^1(\Omega_0)$. (We have denoted by $\langle\langle\cdot,\cdot\rangle\rangle_0$ the duality pairing between $H^1(\Omega_0)$ and its dual.) The proof relies on the spectral characterization of G_0. In particular, one can write

$$\langle G_0[\xi], v\rangle_0 := \int_{r=a} G_0[\xi]\,\bar{v}\mathrm{d}s = \int_0^{2\pi} G_0[\xi]\,\bar{v}\,a\,\mathrm{d}\theta = \sum_{p=-\infty}^{\infty} a\lambda_p \hat{\xi}_p \hat{\bar{v}}_p, \tag{23a}$$

$$\lambda_p = -k\frac{d_z H_p^{(1)}(ka)}{H_p^{(1)}(ka)}. \tag{23b}$$

Well-posedness follows by showing that the $\operatorname{Im}\left\{\lambda_p\right\} < 0$ are bounded, and that $\operatorname{Re}\left\{\lambda_p\right\} \geq 1/a > 0$.

Of course, since the problem (18) is completely equivalent to the variational form of the *full* exterior scattering problem we could appeal to classical results (see e.g., [14]) to deduce existence and uniqueness of solutions. However, as no such analogy is available for the truncated problem (19) a much more direct method is required to establish well-posedness of (19). To accomplish this we give a complete proof of the well-posedness of (18), closely following that of Demkowicz and Ihlenburg [17] and Harari and Hughes [35], where the DtN map appears *explicitly*. By doing this we are able to determine necessary modifications which deliver a well-posedness proof for (19). Furthermore, we point out that due to the non-self-adjoint nature of the DtN map (see Sect. 2) we cannot appeal to the classical perturbation theory for self-adjoint operators [40]. Indeed, we have found the following, rather explicit, method of proof attractive in its simplicity.

The argument will proceed as follows: We first show that (18) has at most one solution (see Theorem 10) which follows closely the arguments in [17]. Defining the linear operators $S^N : V \to V'$ and $S : V \to V'$ via

$$\left\langle\left\langle S^N w, v\right\rangle\right\rangle := \mathcal{A}^N(w,v), \quad \langle\langle Sw, v\rangle\rangle := \mathcal{A}(w,v), \tag{24}$$

we next show that the linear operator $S^0 : V \to V'$ is a Fredholm operator (see Theorem 12). We then compare the operator S to S^0, and use a perturbation argument to show that S is also Fredholm. This, along with the uniqueness result, provides the required existence result.

Theorem 10 *If* $f \in H^5([0,2\pi])$ *the variational problem* (18) *has at most one solution.*

Proof Regarding the smoothness of f, for future use we choose it to fit into the analyticity theory of the DtN map; see Remark 6. The proof follows an argument similar to the one in [35], where $\mathcal{B}$ was taken to be a circle. Suppose $e \in V$ satisfies the homogeneous problem, then $\operatorname{Im}\{\mathcal{A}(e,e)\} = 0$. Since k is real, this means there is zero energy flux through $\mathcal{B}$:

$$0 = \int_{\mathcal{B}} \operatorname{Im}\{G[e]\,\bar{e}\}\,\mathrm{d}s = \frac{1}{2\mathrm{i}}\int_{\mathcal{B}} (\partial_n e)\,\bar{e} - e\,(\partial_n \bar{e})\mathrm{d}s.$$

We now consider the region between $\mathcal{B}$ and a large circle of radius R, B_R, containing Int($\mathcal{B}$) to obtain (in the limit as $R \to 0$)

$$
\begin{aligned}
0 &= \int_{\mathcal{B}} (\partial_n e)\, \bar{e} - e\, (\partial_n \bar{e}) \mathrm{d}s \\
&= \int_{\mathrm{Ann}(\mathcal{B}, B_R)} (\Delta e)\, \bar{e} - e\, (\Delta \bar{e}) \mathrm{d}V - \lim_{R \to 0} \int_{\partial B_R} (\partial_r e)\, \bar{e} - e\, (\partial_r \bar{e}) \mathrm{d}s.
\end{aligned}
$$

Since both e and $\bar{e}$ solve the Helmholtz equation in the exterior of $\mathcal{B}$, we obtain

$$
0 = \lim_{R \to 0} \int_{\partial B_R} (\partial_r e)\, \bar{e} - e\, (\partial_r \bar{e}) \mathrm{d}s.
$$

We use the radiation condition and Rellich's Lemma (see e.g., [48]) to conclude that $e = 0$ in the exterior of $\mathcal{B}$. By continuity, both e and $\partial_n e = \partial_N e = G[e]$ vanish on $\mathcal{B}$.

It remains to show that $e = 0$ in the interior of $\mathcal{B}$; to this end define

$$
\tilde{\mathcal{B}} := \{(r, \theta) \mid r = a + \delta\, |f|_{L^\infty}\},
$$

a circle slightly bigger than $\mathcal{B}$. We have just shown that e and $\partial_n e$ are zero on $\mathcal{B}$, and, by analytic continuation (valid since f is assumed to be in $H^5([0, 2\pi])$), we can assume they vanish on $\mathcal{B}$. The eigenvalue problem for the Laplacian in the ball contained within $\tilde{\mathcal{B}}$ with zero Dirichlet and Neumann data has no finite eigenvalues. Since reducing the region to the interior of $\mathcal{B}$ would only increase such eigenvalues, we can conclude that there are no non-trivial solutions of the homogeneous problem, establishing uniqueness.

Remark 11 Note that the theorem remains true even if we prescribe Dirichlet data on Γ, since such data would only modify the eigenvalues obtained in the Neumann case.

To build a perturbation argument for the well-posedness of (18) we now carefully examine the variational problem with truncated DtN map characterized by the operator S^0. At this juncture we will use the analyticity properties of G and consequently we specialize to boundary perturbations f smooth enough to satisfy the hypotheses of Theorem 2, i.e., for $\xi \in H^{1/2}$ we need $f \in H^5$.

Theorem 12 *If $f \in H^5([0, 2\pi])$ then there exists a $\delta_0 > 0$ such that if $0 < \delta < \delta_0$ then $S^0 = A + C$ where the linear operators A and C are respectively invertible and compact as maps from V to V'. Hence, S^0 is a Fredholm operator.*

Proof It is clear that $\mathcal{A}^0(u, v)$ is a continuous sesquilinear form on $V \times V$. We now define the sesquilinear forms a, d on $V \times V$

$$a(u,v) := \int_\Omega \nabla u \cdot \nabla \bar{v} + u\, \bar{v} \mathrm{d}V + \mathrm{Re}\left\{ \int_\mathcal{B} G_0[u]\, \bar{v} \mathrm{d}s \right\}, \tag{25a}$$

and

$$d(u,v) := -\int_\Omega (k^2+1) u\, \bar{v} \mathrm{d}V + \mathrm{Im}\left\{ \int_\mathcal{B} G_0[u]\, \bar{v} \mathrm{d}s \right\}; \tag{25b}$$

clearly $\mathcal{A}^0(u, v) = a(u, v) + d(u, v)$.

By inspection a is continuous; for coercivity we note that

$$a(u,u) = \|u\|_V^2 + \mathrm{Re}\left\{ \int_\mathcal{B} G_0[u]\, \bar{u} \mathrm{d}s \right\}.$$

If we can show $\mathrm{Re}\left\{\int_\mathcal{B} G_0[\xi]\, \bar{\xi}\ ds\right\} \geq 0$ for all $\xi \in H^{1/2}(\mathcal{B})$, the coercivity of $a(u, v)$ is established, since the trace of $u \in V$ lies in $H^{1/2}(\mathcal{B})$. We describe the arc-length parameter ds on $\mathcal{B}$ as

$$(ds)^2 = \left[(a+\delta f)^2 + (\delta f')^2\right] (d\theta)^2,$$

so that

$$\begin{aligned}
\mathrm{Re}\left\{ \int_\mathcal{B} G_0[\xi]\, \bar{\xi}\ ds \right\} &= \mathrm{Re}\left\{ \int_0^{2\pi} G_0[\xi]\, \bar{\xi} \left[(a+\delta f)^2 + (\delta f')^2\right]^{1/2} \mathrm{d}\theta \right\} \\
&= \mathrm{Re}\left\{ \int_0^{2\pi} G_0[\xi]\, \bar{\xi}\, a \,\mathrm{d}\theta \right\} + \mathrm{Re}\left\{ \int_0^{2\pi} G_0[\xi]\, \bar{\xi} \right. \\
&\quad \left. \times \left(\left[(a+\delta f)^2 + (\delta f')^2\right]^{1/2} - a \right) \mathrm{d}\theta \right\} \\
&\geq (1 - \bar{c}(\delta)) \mathrm{Re}\left\{ \int_0^{2\pi} G_0[\xi]\, \bar{\xi}\, a \,\mathrm{d}\theta \right\} \\
&= (1 - \bar{c}(\delta)) \sum_{p=-\infty}^{\infty} \mathrm{Re}\left\{\lambda_p\right\} \left|\hat{\xi}_p\right|^2 \geq 0
\end{aligned} \tag{26}$$

for δ sufficiently small, where we used $\mathrm{Re}\left\{\lambda_p\right\} \geq 0$, from (23). Here

$$(1-\bar{c}(\delta)) = 1 - \max_{\theta\in[0,2\pi]} \left| \left[\left(1+\delta\frac{f(\theta)}{a}\right)^2 + \delta^2 \frac{f'(\theta)^2}{a^2} \right]^{1/2} - 1 \right| \geq 0,$$

provided δ is chosen small enough; let δ_0 be the largest such perturbation. The relation (26) establishes the coercivity of the sesquilinear form $a(u,v)$, which in turn can be used to define an invertible operator $A : V \to V'$:

$$\langle\langle Au, v\rangle\rangle := a(u,v) \quad \forall u, v \in V.$$

We now turn our attention to $d(u,v)$; the continuity of the first term in (25b) is clear, while the continuity of the second follows by the calculation:

$$\mathrm{Im}\left\{ \int_0^{2\pi} G_0[\xi]\,\bar{\sigma}\; a\mathrm{d}\theta \right\} = \sum_{p=-\infty}^{\infty} \mathrm{Im}\left\{\lambda_p\right\} \hat{\xi}_p \bar{\hat{\sigma}}_p.$$

This defines a continuous sesquilinear map on $L^2([0,2\pi]) \times L^2([0,2\pi])$, since the $\mathrm{Im}\left\{\lambda_p\right\}$ are bounded for all p, see (23). Therefore,

$$\begin{aligned}\left| \mathrm{Im}\left\{ \int_{\mathcal{B}} G_0[\xi]\,\bar{\sigma}\,\mathrm{d}s \right\} \right| &= \left| \mathrm{Im}\left\{ \int_0^{2\pi} G_0[\xi]\,\bar{\sigma} \left[(a+\delta f)^2 \right.\right.\right. \\ &\qquad \left.\left.\left. +(\delta f')^2 \right]^{1/2} \mathrm{d}\theta \right\} \right| \\ &\leq \bar{C}(\delta) \left| \mathrm{Im}\left\{ \int_0^{2\pi} G_0[\xi]\,\bar{\sigma}\; a\,\mathrm{d}\theta \right\} \right|\end{aligned}$$

for

$$\bar{C}(\delta) = \max_{\theta\in[0,2\pi]} \left\{ \left(1+\delta\frac{f(\theta)}{a}\right)^2 + \delta^2\frac{f'(\theta)^2}{a^2} \right\},$$

and hence, d is a continuous map. The embeddings of $H^1(\Omega)$ in $L^2(\Omega)$, and $H^{1/2}(\mathcal{B})$ in $L^2(\mathcal{B})$ are compact, and therefore the sesquilinear form $d(u,v)$ can be used to define a compact operator $C : V \to V'$. Consequently, the variational problem (Eq. (19) with $N = 0$)

$$\text{Find } u^0 \in V \text{ such that for all } v \in V,$$
$$\mathcal{A}^0(u^0, v) = \int_{\Gamma} g\,\bar{v}\mathrm{d}s,$$

can then be written in operator notation as $(A + C)u^0 = F$, for some $F \in V'$, or $(I + A^{-1}C)u^0 = A^{-1}F$, where $A^{-1}C$ is a compact map from $V \to V$. This proves the assertion that S^0 is a Fredholm operator.

We are now in a position to investigate the well-posedness of (18) (involving the *full* DtN map) by comparing its induced linear operator S to S^0 via a perturbation argument.

Theorem 13 *If $f \in H^5([0, 2\pi])$ then, for a given fixed perturbation $\delta > 0$, S has a bounded inverse on V' provided $0 < \delta < \delta_0$ is sufficiently small. Further, there exists an inf-sup constant $\gamma > 0$ for the sesquilinear form $\mathcal{A}(\cdot, \cdot)$:*

$$\gamma \|u\|_V \leq \sup_{v \in V\setminus\{0\}} \frac{|\mathcal{A}(u, v)|}{\|v\|_V} \quad \forall u \in V.$$

Proof We proceed by showing that S is a Fredholm operator. Then, using the uniqueness result in Theorem 10, we will have obtained the desired invertibility result and, consequently, the inf-sup constant γ. Now, for all u, v in V, the difference between the operators S and S^0 is given by

$$\left\langle\!\left\langle (S - S^0)u, v \right\rangle\!\right\rangle = \int_{\mathcal{B}} (G - G_0)[u]\,\bar{v}\mathrm{d}s.$$

We can estimate this difference in operator norm $\|S - S^0\|_{L(V,V')}$ in terms of $\|G - G_0\|_{L(H^{1/2}(\mathcal{B}),H^{-1/2}(\mathcal{B}))}$, which can be made small by Theorem 2. Using a perturbation argument (see, e.g., [3]), we obtain that S is also a Fredholm operator.

6 Well-posedness of the truncated variational problem

The well-posedness of the full problem, (18), discussed in the previous section, provides the basis for the study of the well-posedness of (19) where the DtN map is truncated to a finite number of terms. In other words, we wish to examine the well-posedness of the weak form, (19), of the problem

$$\Delta u^N + k^2 u^N = 0, \qquad x \in \Omega, \tag{27a}$$
$$\partial_n u^N = g, \qquad x \in \Gamma, \tag{27b}$$
$$\partial_N u^N = G^N[u^N], \qquad x \in \mathcal{B}. \tag{27c}$$

The study of well-posedness for the weak form of problem (27) proceeds by showing the associated linear operators are close (in some suitable norm) to those obtained in the previous section. We shall then use a perturbation argument. This leads us to conclude the existence of inf-sup constants γ^N associated with the variational problems (19). We then show that these inf-sup constants are bounded away from 0, enabling us to estimate the error term $\|u-u^N\|_V$ in (22).

Theorem 14 *If $f \in H^5([0,2\pi])$ then for a fixed δ, $0<\delta<\delta_0$, there exists a positive integer $N_0 = N_0(\delta)$ such that S^N has a bounded inverse on V' for all $N > N_0$. Furthermore, for each of these N there exists a positive inf-sup constant γ^N:*

$$\gamma^N \|u\|_V \leq \sup_{v\in V\backslash\{0\}} \frac{\left|\mathcal{A}^N(u,v)\right|}{\|v\|_V} \quad \forall u \in V.$$

Proof From the previous theorem, $S : V \to (V)'$ has a bounded inverse. Also,

$$\left\langle\!\left\langle (S-S^N)u, v \right\rangle\!\right\rangle = \mathcal{A}(u,v) - \mathcal{A}^N(u,v) = \langle (G-G_N)[u], v\rangle$$

which implies that

$$\left\|S-S^N\right\|_{L(V,V')} \leq C\left\|G-G^N\right\|_{L(H^{1/2},H^{-1/2})} \leq \frac{1}{\left\|S^{-1}\right\|_{L(V,V')}}$$

provided N is chosen large enough; here C incorporates the trace constant.

Since $V = H^1(\Omega)$ is complete, by Theorem 2.3.5 of [3], we conclude that $(S^N)^{-1}$ is a linear bijection between V' and V, and

$$\left\|(S_N)^{-1}\right\|_{L(V,V')} \leq \frac{\left\|S^{-1}\right\|_{L(V,V')}}{1-\left\|S^{-1}\right\|_{L(V,V')}\left\|S-S^N\right\|_{L(V,V')}}.$$

In other words, S^N is an invertible operator on V, and therefore the variational problem (19) has a unique solution $u^N \in V$. Then, from Theorem 3.6 in [9], $\mathcal{A}^N$ is a continuous sesquilinear form on $V\times V$ which satisfies the inf-sup condition, yielding the existence of γ^N as claimed.

Remark 15 Note that we cannot yet exclude the possibility that $\gamma^N \to 0$ as $N\to\infty$. This, of course, would prevent us from concluding that the consistency error $\|u-u^N\|_V \to 0$ as $N\to\infty$. The next theorem shows that, indeed, the inf-sup constants, γ^N, are bounded strictly away from zero.

Theorem 16 *The inf-sup constants γ^N defined in Theorem 14 are bounded strictly away from zero. Specifically, if $\gamma > 0$ is the inf-sup constant of Theorem 13, then*

$$\gamma^N \geq \frac{\gamma}{2} > 0 \quad \forall N > N_0.$$

Proof Clearly, $\gamma^N > 0$ for all $N \geq N_0$. Now, for $u, v \in V$, we have

$$\mathcal{A}(u,v) = \mathcal{A}(u,v) - \mathcal{A}^N(u,v) + \mathcal{A}^N(u,v)$$

which implies that

$$\frac{|\mathcal{A}(u,v)|}{\|v\|_V} \leq C \left\|(G - G^N)\right\|_{L(H^{1/2},H^{-1/2})} \|u\|_V + \frac{\left|\mathcal{A}^N(u,v)\right|}{\|v\|_V},$$

where $C > 0$ involves the trace constant. Thus,

$$\sup_{v \in V \setminus \{0\}} \frac{|\mathcal{A}(u,v)|}{\|v\|_V} - C \left\|(G - G^N)\right\|_{L(H^{1/2},H^{-1/2})} \|u\|_V \leq \sup_{v \in V \setminus \{0\}} \frac{|\mathcal{A}_N(u,v)|}{\|v\|_V}.$$

If γ is the inf-sup constant for $\mathcal{A}(\cdot,\cdot)$, we can choose, using Theorem 2, $N_1 \geq N_0$ such that for all $N > N_1$,

$$\gamma - C \left\|(G - G^N)\right\|_{L(H^{1/2},H^{-1/2})} \geq \frac{\gamma}{2}.$$

Then, for all $N \geq N_1$,

$$\frac{\gamma}{2} \leq \inf_{u \in V \setminus \{0\}} \sup_{v \in V \setminus \{0\}} \frac{\left|\mathcal{A}^N(u,v)\right|}{\|v\|_V \|u\|_V}.$$

This shows that for $N > N_1$, $\gamma_N \geq \gamma/2 > 0$.

To establish the convergence of our Enhanced DtN-FE method, note that

$$\begin{aligned} \left|\mathcal{A}^N(u^N - u, v)\right| &\leq \left|\mathcal{A}^N(u^N, v) - \mathcal{A}(u,v)\right| + \left|\mathcal{A}(u,v) - \mathcal{A}^N(u,v)\right| \\ &= \left|\mathcal{A}(u,v) - \mathcal{A}^N(u,v)\right| = \left|\left\langle (G - G^N)u, v\right\rangle\right|. \end{aligned}$$

Using the inf-sup constant γ_N for $\mathcal{A}^N$, we get

$$\begin{aligned} \left\|u - u^N\right\|_V &\leq \frac{C}{\gamma^N} \left\|G - G^N\right\|_{L(H^{1/2},H^{-1/2})} \|u\|_V \\ &\leq \frac{2C}{\gamma} \left\|G - G^N\right\|_{L(H^{1/2},H^{-1/2})} \|u\|_V, \end{aligned}$$

where C involves the trace constant. Since $\|G - G^N\|_{L(H^{1/2},H^{-1/2})} \to 0$ as $N \to \infty$, we have convergence.

7 Spectral approximation error

In the previous section, we successfully estimated the key new source of error in our approximation scheme, introduced by replacing the DtN map G by G^N. We now turn to the spectral approximation error, introduced by the practical numerical necessity of using only a finite number of terms m in the Fourier series expansion of the Dirichlet data ξ provided to the DtN map.

To be more precise, recall that $u^N \in V$ is the unique solution of (19):

$$\text{Find } u^N \in V \text{ such that for all } v \in V,$$
$$\mathcal{A}^N(u^N, v) = \int_\Gamma g\, \bar{v} \mathrm{d}s. \tag{28}$$

The spectral truncation error is the difference $\|u^N - u^{N,m}\|_V$ where $u^{N,m} \in B^m$ solves (21):

$$\mathcal{A}^{N,m}(u^{N,m}, v) = \int_\Gamma g\, \bar{v} \quad \forall v \in V. \tag{29}$$

We remind the reader that

$$\mathcal{A}^{N,m}(w, v) := \int_\Omega \nabla w \cdot \nabla \bar{v} - k^2 w\, \bar{v} \mathrm{d}V + \left\langle G^{N,m}[w], v \right\rangle.$$

In our analysis it is convenient to define the projection $P^m : H^s(\mathcal{B}) \to H^s(\mathcal{B})$, $s \geq 0$, via

$$P^m \mu := \sum_{p=-m}^{m} \hat{\mu}_p \mathrm{e}^{\mathrm{i}p\theta},$$

where $\hat{\mu}_p$ is the pth Fourier coefficient of μ. With this we note that

$$G^{N,m}[\mu] = G^N[P^m \mu].$$

While (29) is straightforward to implement numerically (see [43]) it suffers from a loss of stability. To see this note that the map $G^{N,m}$ is "blind" to Fourier modes of frequency higher than m which results in a lack of uniqueness. For example, when the artificial boundary $\mathcal{B}$ is a circle one could add any solution w of the Neumann eigenvalue problem

$$\Delta w + k^2 w = 0, \quad \text{in } \Omega_0,$$
$$\partial_n w = 0 \quad \text{on } \partial\Omega_0,$$

to our discrete solution and still obtain a valid solution. This results in a loss of stability if k is near an interior Neumann eigenvalue.

Two possible approaches to address this issue in the setting of circular artificial boundaries are used. Harari and Hughes [35] have suggested the heuristic of choosing $m > ka$, where a is the radius of the artificial boundary. Grote and Keller [31] advocate modifying the DtN map, in our notation $G^{0,m}$. This latter approach, which we will pursue in this section, begins with any operator $\mathcal{G} : H^{1/2}(\mathcal{B}_0) \to H^{-1/2}(\mathcal{B}_0)$ such that

$$\operatorname{Im}\left\{\int_{\mathcal{B}} \bar{\mu}\mathcal{G}\mu \; ds\right\} < 0 \quad \forall \mu \neq 0. \tag{30}$$

Defining

$$H^{0,m} := G^{0,m} - \mathcal{G}P^m + \mathcal{G},$$

Grote and Keller showed that using the operator $H^{0,m}$ in place of the DtN map $G^{0,m}$ leads to a well-posed problem. We point out that $H^{0,m} = G^{0,m}$ when applied to $\mu = \sum_{|p|<m} \hat{\mu}_p \mathrm{e}^{\mathrm{i}p\theta}$. This is a desirable feature of the modified map, since it remains exact for all modes up to the m-th one.

In the present context of perturbed artificial boundaries we will use

$$H^{N,m}\mu := G^N P^m \mu + \mathcal{G}(I - P^m)\mu,$$

defined for any $\mu \in H^{1/2}(\mathcal{B})$, where

$$\mathcal{G}\mu := -\mathrm{i}k\mu.$$

The map $\mathcal{G}$ is essentially the Sommerfeld radiation condition, used in the near-field. It is an easy-to-implement condition which prevents the loss of stability in (29) since

$$\operatorname{Im}\left\{\int_{\mathcal{B}} \bar{\mu} H^{N,m} \mu \mathrm{d}s\right\} < 0 \quad \forall \mu \neq 0,$$

cf. (30).

We now examine the solvability of (29) with $G^{N,m}$ replaced by $H^{N,m}$.

Theorem 17 *Let $S^{N,m} : V \to V'$ be defined by*

$$\left\langle\!\left\langle S^{N,m} w, v \right\rangle\!\right\rangle := \int_\Omega \nabla w \cdot \nabla \bar{v} - k^2 w \, \bar{v} \mathrm{d}V + \left\langle H^{N,m}[w], v \right\rangle \quad \forall w, v \in V.$$

If $f \in H^5([0, 2\pi])$ then there exist positive integers $\bar{N}$ and $\bar{m}$ such that for all $N > \bar{N}$, $m > \bar{m}$, $S^{N,m}$ is invertible as a map from V to V' with bounded inverse. Hence, the discrete variational problem (29) is uniquely solvable.

Proof We will show that $S^{N,m}$ is close, in operator norm, to the invertible operator S^N (see Theorem 14) which, we recall, has a bounded inverse. For every $w, v \in V$,

$$\begin{aligned} \left\langle\!\left\langle (S^{N,m} - S^N) w, v \right\rangle\!\right\rangle &= \left\langle (H^{N,m} - G^N)[w], v \right\rangle \\ &= \left\langle (\mathcal{G} - G^N)(I - P^m)[w], v \right\rangle. \end{aligned}$$

This allows us to estimate

$$\begin{aligned} \left| \left\langle\!\left\langle (S^{N,m} - S^N) w, v \right\rangle\!\right\rangle \right| &\le \left\| (\mathcal{G} - G^N)(I - P^m) w \right\|_{H^{-1/2}(\mathcal{B})} \|v\|_{H^{1/2}(\mathcal{B})} \\ &\le C \left\| (\mathcal{G} - G^N)(I - P^m) w \right\|_{H^{1/2}(\mathcal{B})} \|v\|_V \\ &\le C_N \left\| I - P^m \right\|_{L(H^{1/2}(\mathcal{B}), H^{1/2}(\mathcal{B}))} \|w\|_V \, \|v\|_V \, . \end{aligned}$$

Here, the constant $C_N > 0$ includes dependence on the trace constant and on the sum of the operator norms of $\mathcal{G}$ and G^N. Since these are bounded, and

$$\left\| I - P^m \right\|_{L(H^{1/2}(\mathcal{B}), H^{1/2}(\mathcal{B}))} \to 0$$

as $m \to \infty$, this allows us to conclude that

$$\left\| S^{N,m} - S^N \right\|_{L(V, V')} \to 0$$

as $m \to \infty$. If we choose N large enough so that S^N is invertible, the perturbation result Theorem 2.3.5 in [3] provides the desired invertibility result.

The theorem above allows us to deduce the existence of a family of inf-sup constants $\gamma_{N,m}$ such that

$$\gamma_{N,m} \|w\|_V \le \sup_{\|v\|_V = 1} \left| \mathcal{A}^{N,m}(w, v) \right|,$$

where $H^{N,m}$ has replaced $G^{N,m}$ in the definition of $\mathcal{A}^{N,m}$. However, the theorem does not guarantee that these constants will not tend to zero as $N, m \to \infty$.

Using an argument analogous to that given for Theorem 16 we will show that this cannot happen. First we note that

$$\begin{aligned}\mathcal{A}^N(w,v) &= \mathcal{A}^{N,m}(w,v) + \mathcal{A}^N(w,v) - \mathcal{A}^{N,m}(w,v)\\ &= \mathcal{A}^{N,m}(w,v) + \left\langle (G^N - \mathcal{G})(I - P^m)[w], v \right\rangle\end{aligned}$$

This implies that

$$\frac{\left|\mathcal{A}^{N,m}(w,v)\right|}{\|v\|_V} + C_N \left\|(I - P^m)\right\|_{L(H^{1/2}(\mathcal{B}),H^{1/2}(\mathcal{B}))} \|w\|_V \|v\|_V \geq \frac{\left|\mathcal{A}^N(w,v)\right|}{\|v\|_V}.$$

and hence

$$\begin{aligned}&\sup_{\|v\|_V=1} \left|\mathcal{A}^N(w,v)\right| - C_N \left\|(I - P^m)\right\|_{L(H^{1/2}(\mathcal{B}),H^{-1/2}(\mathcal{B}))} \|w\|_V\\ &\quad\leq \sup_{\|v\|_V=1} \left|\mathcal{A}^{N,m}(w,v)\right|,\end{aligned}$$

where $C_N > 0$ is the same constant as in the previous proof. Recalling that Theorem 16 gives

$$(\gamma/2)\,\|w\|_V \leq \sup_{\|v\|_V=1} \left|\mathcal{A}^N(w,v)\right| \quad \forall w \in V,$$

if we choose m sufficiently large then

$$(\gamma/4)\,\|w\|_V \leq \sup_{\|v\|_V=1} \left|\mathcal{A}^{N,m}(w,v)\right| \quad \forall w \in V. \tag{31}$$

We can now precisely estimate the discretization error.

Theorem 18 *Suppose $f \in H^5([0,2\pi])$, $N > \bar{N}$ and $m > \bar{m}$, and that $u^N \in V$ solves (7) and $u^{N,m} \in V$ solves (29) (with $G^{N,m}$ replaced by $H^{N,m}$). Then there is a constant $C > 0$ such that*

$$\left\|u^N - u^{N,m}\right\|_V \leq C \left\|(I - P^m)[u^N]\right\|_{H^{1/2}(\mathcal{B})}$$

where $[u^N]$ is the trace of u^N on $\mathcal{B}$. Furthermore, if $[u^N] \in H^{1/2+\alpha}(\mathcal{B})$, for any $\alpha \geq 0$, the error decays as $|m|^{-1/2-\alpha}$.

Proof From (31) we have

$$
\begin{aligned}
(\gamma/4)\left\|u^N-u^{N,m}\right\|_V &\le \sup_{\|v\|_V=1}\left|\mathcal{A}^{N,m}(u^N-u^{N,m},v)\right| \\
&\le \sup_{\|v\|_V=1}\Big|\mathcal{A}^{N,m}(u^N,v)-\mathcal{A}^N(u^N,v) \\
&\quad -\mathcal{A}^{N,m}(u^{N,m},v)+\mathcal{A}^N(u^N,v)\Big| \\
&\le \sup_{\|v\|_V=1}\left|\left\langle (H^{N,m}-G^N)[u^N],v\right\rangle\right| \\
&\le \left\|G^N-\mathcal{G}\right\|_{L(H^{1/2}(\mathcal{B}),H^{-1/2}(\mathcal{B}))} \\
&\quad \times \left\|(I-P^m)[u^N]\right\|_{H^{1/2}(\mathcal{B})}
\end{aligned}
$$

which gives the required error estimate. Using Lemma 19, if $[u^N]\in H^{1/2+\alpha}(\mathcal{B})$, $\alpha\ge 0$, we can estimate

$$
\left\|(I-P^m)[u^N]\right\|_{H^{1/2}(\mathcal{B})}\le C\,|m|^{-1/2-\alpha}\,.
$$

Finally, the standard spectral approximation result we need is presented below without proof [29].

Lemma 19 *If $f\in H^s([0,2\pi])$ and $s\ge t$ then*

$$
\left\|(I-P^m)f\right\|_{H^t([0,2\pi])}\le C\,|m|^{-1/2-(s-t)}\,.
$$

8 Conclusion

In this paper we have presented a numerical analysis of the “Enhanced DtN-FE” method devised by the authors in a previous publication [43]. This algorithm uses a perturbative approach to compute the DtN map on a domain shaped by the deformation of a circle. If the full DtN map is utilized the exact solution of the full problem is recovered (without spurious reflections), furthermore, if a truncated Taylor series approximation of the DtN map is used the problem is still well-posed, and the solution converges exponentially (as a function of the number of Taylor series terms retained) to the exact solution. We provide a stabilization of our map in the setting where the boundary data is truncated in Fourier space.

In this analysis we have focused on the error introduced by truncation of the DtN map in perturbation parameter. Consequently, once this analysis is in place, the error introduced by a particular choice of finite elements can be studied using standard techniques. This suggests that our method can be used in conjunction with *any* variational algorithm in the truncated region, e.g., a spectral method. This feature of our method is intentional, and our design philosophy has been to create a flexible boundary condition which can be implemented as a "black-box." Since the boundary condition we introduce is high-order, the accuracy of the overall computation becomes essentially limited by that of the numerical discretization in the bounded region. In a forthcoming paper we shall describe efficient implementations of these DtN maps, including preconditioning strategies and preprocessing steps which significantly enhance the speed and accuracy of this DtN-FE method.

Acknowledgements DPN gratefully acknowledges support from NSF through grants No. DMS–0139822, and DMS–0406007. NN gratefully acknowledges support from the National Science and Engineering Research Council and the FQRNT. We would like to thank the anonymous referees for their careful and helpful reading of this manuscript. We would also like to thank the McGill Applied Mathematics Working Seminar for numerous invigorating and challenging discussions on this work.

References

1. Abarbanel, S., Gottlieb, D., Hesthaven, J.S.: Long time behavior of the perfectly matched layer equations in computational electromagnetics. In: Proceedings of the Fifth International Conference on Spectral and High Order Methods (ICOSAHOM-01) (Uppsala), vol. 17, pp. 405–422 (2002)
2. Adams, R.A.: Sobolev spaces. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London (1975). Pure and Applied Mathematics, vol. 65
3. Atkinson, K., Han, W.: Theoretical Numerical Analysis. Texts in Applied Mathematics, vol. 39. Springer, Berlin Heidelberg New York (2001). A functional analysis framework
4. Barrenechea, G.R., Barrientos, M.A., Gatica, G.N.: On the numerical analysis of finite element and Dirichlet-to-Neumann methods for nonlinear exterior transmission problems. Numer. Funct. Anal. Optim. **19**(7–8), 705–735 (1998)
5. Barrenechea, G.R., Gatica, G.N., Hsiao, G.C.: Weak solvability of interior transmission problems via mixed finite elements and Dirichlet-to-Neumann mappings. J. Comput. Appl. Math., **100**(2), 145–160 (1998)
6. Alvin Bayliss, Max Gunzburger, Eli Turkel: Boundary conditions for the numerical solution of elliptic equations in exterior regions. SIAM J. Appl. Math. **42**(2), 430–451 (1982)
7. Eliane Bécache, Patrick Joly. On the analysis of Bérenger's perfectly matched layers for Maxwell's equations. M2AN Math. Model. Numer. Anal. **36**(1), 87–119 (2002)
8. Berenger, J.-P.: A perfectly matched layer for the absorption of electromagnetic waves. J. Comput. Phys. **114**(2), 185–200 (1994)
9. Braess, D.: Finite Elements, 2nd edn. Cambridge University Press, Cambridge, (2001). Theory, fast solvers, and applications in solid mechanics, Translated from the 1992 German edition by Larry L. Schumaker.
10. Bruno, O.P., Kunyansky, L.A.: A fast, high-order algorithm for the solution of surface scattering problems: basic implementation, tests, and applications. J. Comput. Phys. **169**(1), 80–110 (2001)
11. Bruno, O.P., Kunyansky, L.A.: Surface scattering in three dimensions: an accelerated high-order solver. Proc. R. Soc. Lond. Proc. A **457** (2001)
12. Calderón, A.P.: Cauchy integrals on Lipschitz curves and related operators. Proc. Nat. Acad. Sci. USA **75**, 1324–1327 (1977)
13. Coifman, R., Meyer, Y.: Nonlinear harmonic analysis and analytic dependence. In: Pseudo-differential Operators and Applications (Notre Dame, Ind., 1984), pp. 71–78. Amer. Math. Soc. (1985)
14. Colton, D., Kress, R.: Inverse acoustic and electromagnetic scattering theory. Springer, Berlin Heidelberg New York (1998)
15. Craig, W., Nicholls, D.P.: Traveling two and three dimensional capillary gravity water waves. SIAM J. Math. Anal. **32**(2), 323–359 (2000)

16. Craig, W., Schanz, U., Sulem, C.: The modulation regime of three-dimensional water waves and the Davey–Stewartson system. Ann. Inst. Henri Poincaré **14**, 615–667 (1997)
17. Demkowicz, L., Ihlenburg, F.: Analysis of a coupled finite-infinite element method for exterior Helmholtz problems. Numer. Math. **88**(1), 43–73 (2001)
18. Djellouli, R., Farhat, C., Macedo, A., Tezaur, R.: Finite element solution of two-dimensional acoustic scattering problems using arbitrarily shaped convex artificial boundaries. J. Comput. Acoust. **8**(1), 81–99 (2000). Finite elements for wave problems (Trieste, 1999)
19. Bjorn Engquist and Andrew Majda. Absorbing boundary conditions for the numerical simulation of waves. Math. Comput. **31**(139), 629–651 (1977)
20. Evans, L.C.: Partial Differential Equations. American Mathematical Society, Providence, RI (1998)
21. Feng, K.: Finite element method and natural boundary reduction. In Proceedings of the International Congress of Mathematicians, vol. 1, 2 (Warsaw, 1983), pp. 1439–1453. Warsaw (1984)
22. Feng, K., Yu, D.H.: Canonical integral equations of elliptic boundary value problems and their numerical solutions. In: Proceedings of China–France Symposium on the Finite Element Method, pp. 211–252. Science Press, Beijing (1982)
23. Folland, G.B.: Introduction to Partial Differential Equations. Princeton University Press, Princeton, NJ (1976). Preliminary informal notes of university courses and seminars in mathematics, Mathematical Notes
24. Ganesh, M., Graham, I.G.: A high-order algorithm for obstacle scattering in three dimensions. J. Comput. Phys. **198**(1), 211–242 (2004)
25. Gilbarg, D., Trudinger, N.S.: Elliptic Partial Differential Equations of Second Order, 2nd edn. Springer, Berlin Heidelberg New York (1983)
26. Givoli, D.: Recent advances in the DtN FE method. Arch. Comput. Methods Eng. **6**(2), 71–116 (1999)
27. Givoli, D.: Nonreflecting boundary conditions. J. Comput. Phys. **94**(1), 1–29 (1991)
28. Givoli, D.: Numerical Methods for Problems in Infinite Domains. Studies in Applied Mechanics, vol. 33. Elsevier, Amsterdam (1992)
29. Gottlieb, D., Orszag, S.A.: Numerical Analysis of Spectral Methods: Theory and Applications. Society for Industrial and Applied Mathematics, Philadelphia, PA (1977). CBMS-NSF Regional Conference Series in Applied Mathematics, No. 26
30. Grote, M.J., Keller, J.B.: Exact nonreflecting boundary conditions for the time dependent wave equation. SIAM J. Appl. Math. **55**(2), 280–297 (1995)
31. Grote, M.J., Keller, J.B.: On nonreflecting boundary conditions. J. Comput. Phys. **122**(2), 231–243 (1995)
32. Grote, M.J., Keller, J.B.: Nonreflecting boundary conditions for Maxwell's equations. J. Comput. Phys. **139**(2), 327–342 (1998)
33. Guidotti, P.: A new first kind boundary integral formulation for the Dirichlet-to-Neumann map in 2D. J. Comput. Phys. **190**(1), 325–345 (2003)
34. Han, H.D., Wu, X.N.: Approximation of infinite boundary condition and its application to finite element methods. J. Comput. Math. **3**(2), 179–192 (1985)
35. Harari, I., Hughes, T.J.R.: Analysis of continuous formulations underlying the computation of time-harmonic acoustics in exterior domains. Comput. Methods Appl. Mech. Eng. **97**(1), 103–124 (1992)
36. Hsiao, G., Wendland, W.: A finite element method for some integral equations of the first kind. J. Math. Anal. Appl. **58** (1977)
37. Hu, B., Nicholls, D.P.: Analyticity of Dirichlet–Neumann operators on Hölder and Lipschitz domains. SIAM J. Math. Anal. **37**(1), 302–320 (2005)
38. Jia, Z.P., Wu, J.M., Yu, D.H.: The coupling natural boundary–finite element method for solving 3–D exterior Helmholtz problem. Math. Numer. Sinica **23**(3), 357–382 (2001)
39. Johnson, C., Nédélec, J.-C.: On the coupling of boundary integral and finite element methods. Math. Comput. **35**(152), 1063–1079 (1980)
40. Kato, T.: Perturbation Theory for Linear Operators. Springer, Berlin Heidelberg New York (1976)
41. Keller, J.B., Givoli, D.: Exact nonreflecting boundary conditions. J. Comput. Phys. **82**(1), 172–192 (1989)

42. Ladyzhenskaya, O.A., Ural'tseva, N.N.: Linear and Quasilinear Elliptic Equations. Academic Press, New York (1968)
43. Nicholls, D.P., Nigam, N.: Exact non-reflecting boundary conditions on general domains. J. Comput. Phys. **194**(1), 278–303 (2004)
44. Nicholls, D.P., Reitich, F.: A new approach to analyticity of Dirichlet-Neumann operators. Proc. Roy. Soc. Edinburgh Sect. A **131**(6), 1411–1433 (2001)
45. Nicholls, D.P., Reitich, F.: Stability of high-order perturbative methods for the computation of Dirichlet-Neumann operators. J. Comput. Phys. **170**(1), 276–298 (2001)
46. Nicholls, D.P., Reitich, F.: Analytic continuation of Dirichlet-Neumann operators. Numer. Math. **94**(1), 107–146 (2003)
47. Ushijima, T.: An FEM–CSM combined method for planar exterior Laplace problems. Japan J. Indust. Appl. Math. **18**, 359–382 (2001)
48. Wilcox, C.H.: Scattering Theory for the d'Alembert Equation in Exterior Domains. Lecture Notes in Mathematics, vol. 442 Springer, Berlin Heidelberg New York (1975)
49. Wu, J.M., Yu, D.H.: The natural integral equation of 3–D exterior Helmholtz problem and its numerical solution. Chinese J. Comput. Phys. **16**(5), 449–456 (1999)
50. Yu, D.H.: Coupling canonical boundary element method with FEM to solve harmonic problem over cracked domain. J. Comput. Math. **1**(3), 195–202 (1983)
51. Yu, D.H.: Approximation of boundary conditions at infinity for a harmonic equation. J. Comput. Math. **3**(3), 219–227 (1985)
52. Yu, D.H.: Mathematical Theory of Natural Boundary Element Method. Science Press, Beijing (1993)
53. Yu, D.H.: The Natural Boundary Integral Method and its Application. Science Press & Kluwer Academic Publishers (2002)
54. Yu, D.H., Jia, Z.P.: The non–overlapping DDM based on natural boundary reduction for 2–D exterior Helmholtz problem. Math Numer. Sinica **22**(2), 227–240 (2000)