SIAM J. NUMER. ANAL.
Vol. 47, No. 4, pp. 2708–2734

A RIGOROUS NUMERICAL ANALYSIS OF THE TRANSFORMED FIELD EXPANSION METHOD*

DAVID P. NICHOLLS† AND JIE SHEN‡

Abstract. Boundary perturbation methods, in which the deviation of the problem geometry from a simple one is taken as the small quantity, have received considerable attention in recent years due to an enhanced understanding of their convergence properties. One approach to deriving numerical methods based upon these ideas leads to Bruno and Reitich's generalization [*Proc. Roy. Soc. Edinburgh Sect. A*, 122 (1992), pp. 317–340] of Rayleigh and Rice's classical algorithm giving the "method of variation of boundaries" which is very fast and accurate within its domain of applicability. Treating problems outside this domain (e.g., boundary perturbations which are large and/or rough) led Nicholls and Reitich [*Proc. Roy. Soc. Edinburgh Sect. A*, 131 (2001), pp. 1411–1433] to design the "transformed field expansions" (TFE) method, and the rigorous numerical analysis of these recursions is the subject of the current work. This analysis is based upon analyticity estimates for the TFE expansions coupled to the convergence of Fourier–Legendre Galerkin methods. This powerful and flexible analysis is extended to a wide range of problems including those governed by Laplace and Helmholtz equations, and the equations of traveling free-surface ideal fluid flow.

Key words. water wave equation, Helmholtz equation, scattering, error analysis, spectral-Galerkin, perturbation

AMS subject classifications. 65N35, 65N22, 65F05, 35J05

DOI. 10.1137/080741914

1. Introduction.

Perturbation techniques have been a crucial tool for scientists and engineers for hundreds of years, many classical and contemporary books, e.g., [25, 4], have been devoted to their explication. One very important subclass of such methods are those designed for problems where the perturbation occurs in its geometry; of particular interest for us are partial differential equations posed on complicated domains which are "nearly" simple (e.g., rectangular, circular, and spherical). Such considerations go back at least as far as Stokes [51] in the context of free-surface ideal fluid mechanics (water waves) and Rayleigh [43] regarding the scattering of linear acoustic waves by an irregular obstacle, for example.

Quite recently, this class of "boundary perturbation" (BP) methods has been carefully reexamined for use as highly accurate computational methodologies for the simulation of engineering problems. For instance, for the problem of electromagnetic or acoustic scattering by irregular obstacles, Bruno and Reitich [7, 8, 9, 10, 11, 12] generalized the method of Rayleigh [43] and Rice [46] to arbitrarily high order and demonstrated that this method can deliver highly accurate answers with computational complexities which match the current state-of-the-art. These methods were largely ignored for many years due to an incomplete understanding of their convergence properties. However, these issues were resolved by Bruno and Reitich [6] (and later clarified by one of the authors and Reitich [31]) using the theory of complex

*Received by the editors November 26, 2008; accepted for publication (in revised form) May 22, 2009; published electronically July 31, 2009.
http://www.siam.org/journals/sinum/47-4/74191.html
†Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, Chicago, IL 60607 (nicholls@math.uic.edu). The work of this author was partially supported by NSF grants DMS–0537511 and DMS–0810958.
‡Department of Mathematics, Purdue University, West Lafayette, IN 47907 (shen@math.purdue.edu). The work of this author was partially supported by NSF grant DMS-0610646 and AFOSR grant FA9550-08-1-0416.

analysis and showed that the scattered field depends *analytically* upon boundary perturbations thus enabling the success of the BP method they outlined.

One shortcoming of the "method of variation of boundaries" (later named the method of "field expansions") of Bruno and Reitich is its numerical ill-conditioning, particularly for large and/or rough obstacles [32] due to subtle, but quite strong, cancellations present in the underlying recursions. This property not only renders the scheme unstable for extreme geometries but also prevents a *direct* proof of their convergence. This difficulty was completely resolved in [31, 32] with the development of the method of "transformed field expansions" (TFE). This algorithm, derived in much the same way as the field expression approach save that a domain–flattening transformation is affected before the perturbation expansion is made, is not only a stable, high-order numerical scheme but also can also be used to *directly* establish the analyticity of the field. It is upon this TFE method that our current numerical analysis is focused. While many of the requisite pieces are in place in the literature, the process of assembling them to conclude the rigorous convergence of this numerical method is far from trivial. In fact, we are able to show the power and flexibility of our analysis by establishing the convergence of these TFE recursions not only as applied to the Laplace equation [31, 32, 33, 39] but also in regard to the Helmholtz equation (which arises in electromagnetic and acoustic scattering) [29, 34, 35, 38, 30, 19] and the water wave problem [36, 37].

The organization of the paper is as follows: In section 2 we set up the general framework in which our numerical analysis is applicable. This includes the cases of Laplace's equation and the Helmholtz equation mentioned above. In section 3 we provide complete details of our results for the case of Laplace's equation on a perturbed domain (discussing the TFE recursions in section 3.1, the Fourier–Legendre method in section 3.2, and the combined error estimates in section 3.3). We provide similar results for the TFE method as applied to water wave problems in section 4 and Helmholtz's equation in section 5. We give concluding remarks in section 6.

2. General framework.

In this section we outline the generic problem to which the TFE method can be applied and for which our subsequent numerical analysis is relevant. To begin we consider a quite general elliptic boundary value problem of the form

$$L[V(x)] = G(x), \qquad x \in \Omega, \tag{2.1a}$$

$$B[V(x)|_{x\in\partial\Omega}] = \zeta(x), \tag{2.1b}$$

where $\Omega \subset \mathbf{R}^d$ is the domain of definition of our problem with boundary $\partial\Omega$, L is an elliptic differential operator (e.g., $L[\cdot] = \operatorname{div}[A\nabla\cdot]$ if A is symmetric and positive definite), and B is a boundary operator determining a unique solution for (2.1). If Ω has a complicated shape, then the design of robust, highly accurate numerical algorithms becomes much more involved, and there are many approaches that one may take. One very successful method involves mapping the complex domain Ω to a simpler (perhaps separable) one, say, Ω_0, and we suppose that this can be accomplished via the mapping $x' = \phi(x)$. Supposing further that the triple $\{V, G, \zeta\}$ is mapped by ϕ to $\{U, F, \xi\}$, the problem (2.1) becomes

$$\mathcal{L}[U(x')] = F(x'; U, \phi), \qquad x' \in \Omega_0, \tag{2.2a}$$

$$\mathcal{B}[U(x')|_{x'\in\partial\Omega_0}] = J(x'; \xi, U, \phi). \tag{2.2b}$$

Of course the properties of $\mathcal{L}$ and $\mathcal{B}$ (and thus the well posedness of (2.2)) depend strongly upon the mapping ϕ, and an important hypothesis for our approach is that Ω is "close" to Ω_0 in a manner made precise below.

The TFE approach works under the assumption that the original domain can be written as a small perturbation of the simpler domain $\Omega = \Omega_{\varepsilon f}$, where ε measures the size of the deformation and the smooth function f measures the shape. It can typically be shown that the smallness assumption on ε can be dropped [33], but we do not pursue this aspect in this work. With this parameter in hand we postulate the regular perturbation expansion

$$U = U(x'; \varepsilon) = \sum_{n=0}^{\infty} U_n(x') \varepsilon^n \tag{2.3}$$

and insert this into (2.2). Noting that the transformation ϕ has moved the boundary perturbation from the domain shape to the right-hand side of (2.2) we can easily deduce that, at every perturbation order, we have

$$\mathcal{L}[U_n(x')] = F_n(x'; f, U_0, \ldots, U_{n-1}), \qquad x' \in \Omega_0, \tag{2.4a}$$

$$\mathcal{B}[U_n(x')|_{x' \in \partial\Omega_0}] = J_n(x'; \xi, f, U_0, \ldots, U_{n-1}); \tag{2.4b}$$

these are the TFE recursions.

Remark 2.1. Several important remarks are in order.

1. If we take the boundary perturbation around $\varepsilon = 0$, as we have done above, then the linear operators satisfy $\mathcal{L} = L$ and $\mathcal{B} = B$.
2. The right-hand sides $\{F_n, J_n\}$, while being quite nonlinear, depend only upon information that we already possess, $\{U_0, \ldots, U_{n-1}\}$.
3. At order zero, $F_0 \equiv 0$ and $J_0 = \xi$, which is the "linearization" of the original problem about the simpler geometry.

The key ingredients for our numerical analysis are that the series (2.3) in fact converges strongly in an appropriate function space and that solutions of the generic elliptic problem (2.4) on the simplified domain Ω_0 can be analyzed numerically. Regarding the former, for a second order elliptic problem, we require the following proposition.

Proposition 2.1. *Given an integer $r \geq 2$, if $f \in H^r(\partial\Omega_0)$ and $\xi \in H^{r-1/2}(\partial\Omega_0)$, then there exists a unique solution U of* (2.2) *such that the expansion* (2.3) *converges strongly. More precisely, there exists C_1, $C_2 > 0$ such that*

$$\|U_n\|_{H^r(\Omega_0)} \leq C_1 \|\xi\|_{H^{r-1/2}(\partial\Omega_0)} B^n \tag{2.5}$$

for any constant $B > C_2 \|f\|_{H^r(\partial\Omega_0)}$.

For the numerical analysis, consider the prototype elliptic problem

$$\mathcal{L}[W(x')] = F(x'), \qquad x' \in \Omega_0, \tag{2.6a}$$

$$\mathcal{B}[W(x')|_{x' \in \partial\Omega_0}] = J(x'), \tag{2.6b}$$

which is approximated by a Galerkin method (parametrized by a mesh-size $\bar{h}$) delivering an approximate solution $W^{\bar{h}}$. We require the following result to be true.

Proposition 2.2. *Consider the unique solution W of* (2.6) *and its numerical approximation $W^{\bar{h}}$; then we have the estimate*

$$\left\| W - W^{\bar{h}} \right\|_E \leq C_0 \bar{h}^{r - s(E)} \|W\|_{H^r}, \tag{2.7}$$

where $\|\cdot\|_E$ is the natural energy norm associated with (2.6)*, $s(E)$ is a constant depending on the norm $\|\cdot\|_E$, and $r \geq 2 \geq s(E)$ is a constant depending on the specific numerical method and the regularity of W.*

With these two results in hand we make the following approximation of the solution of (2.2):

$$U^{N,\bar{h}}(x,y;\varepsilon) := \sum_{n=0}^{N} U_n^{\bar{h}}(x,y)\varepsilon^n,$$

where $U_n^{\bar{h}}$ is an approximate solution of (2.4) and, for future reference,

$$U^N(x,y;\varepsilon) := \sum_{n=0}^{N} U_n(x,y)\varepsilon^n.$$

Under these assumptions, we have the following convergence result.

THEOREM 2.1. *Assuming that the hypotheses of Propositions* 2.1 *and* 2.2 *are valid for a given integer $r \geq 2$, then*

$$\left\|U - U^{N,\bar{h}}\right\|_E \lesssim (B\varepsilon)^{N+1} + \|\xi\|_{H^{r-1/2}(\partial\Omega)} \bar{h}^{r-s(E)}$$

for any constant $B > C_2 \|f\|_{H^r(\partial\Omega_0)}$ such that $B\varepsilon < 1$.

Proof. Using the fact that $\|U_n\|_E \leq \|U_n\|_{H^2(\Omega_0)}$, we derive from (2.5) and (2.7) that

$$\begin{aligned}
\left\|U - U^{N,\bar{h}}\right\|_E &\leq \left\|U - U^N\right\|_E + \left\|U^N - U^{N,\bar{h}}\right\|_E \\
&\leq \sum_{n=N+1}^{\infty} \|U_n\|_E \varepsilon^n + \sum_{n=0}^{N} \left\|U_n - U_n^{\bar{h}}\right\|_E \varepsilon^n \\
&\leq \sum_{n=N+1}^{\infty} \|U_n\|_{H^2(\Omega_0)} \varepsilon^n + C_0 \bar{h}^{r-s(E)} \sum_{n=0}^{N} \|U_n\|_{H^r(\Omega_0)} \varepsilon^n \\
&\leq C_1 \|\xi\|_{H^{3/2}(\partial\Omega_0)} \sum_{n=N+1}^{\infty} (B\varepsilon)^n + C_0 C_1 \|\xi\|_{H^{r-1/2}(\partial\Omega_0)} \bar{h}^{r-s(E)} \sum_{n=0}^{N} (B\varepsilon)^n \\
&\leq \frac{C_1}{1 - B\varepsilon} \|\xi\|_{H^{3/2}(\partial\Omega_0)} (B\varepsilon)^{N+1} + C_0 C_1 \|\xi\|_{H^{r-1/2}(\partial\Omega_0)} \bar{h}^{r-s(E)} \frac{1 - (B\varepsilon)^{N+1}}{1 - B\varepsilon} \\
&\leq C_1 \frac{(B\varepsilon)^{N+1}}{1 - B\varepsilon} \|\xi\|_{H^{3/2}(\partial\Omega_0)} + \frac{C_0 C_1}{1 - B\varepsilon} \|\xi\|_{H^{r-1/2}(\partial\Omega_0)} \bar{h}^{r-s(E)},
\end{aligned}$$

provided that $B\varepsilon < 1$. ∎

Remark 2.2. The problems (2.2) that we are interested in for this paper usually involve an artificial boundary where the transparent boundary condition, namely, the Dirichlet-to-Neumann mapping, is used. It is therefore clear, from a practical point of view, that we should place the artificial boundary in such a way that the computational domain Ω_0 is as small as possible. The effect of the artificial boundary location on the convergence rate is manifested through the norms $\|U_n\|_{H^r(\Omega_0)}$ in (2.5), more precisely the constant C_1 in the above proof (see Remark 4.2 in [50] for a concrete example).

3. A surface formulation of water waves and Laplace's equation. To begin our more specific discussions, let us discuss Laplace's equation as it arises in the surface formulation of the water wave problem due to Zakharov [52] and Craig and Sulem [17]. For this, the relevant (ocean) geometry is

$$S_{h,g} := \{(x,y) \in \mathbf{R}^{d-1} \times \mathbf{R} \mid -h < y < g(x)\}, \quad d = 2,3$$

(cf. Ω from section 2), where h may be infinity, g is meant to represent the ocean surface, and the governing partial differential equation is Laplace's equation

$$\Delta V(x,y) = 0, \quad (x,y) \in S_{h,g}, \tag{3.1}$$

for the velocity potential V [27]. The problem of free-surface ideal fluid flows further demands the Dirichlet condition at the surface

$$V(x,g(x)) = \zeta(x) \tag{3.2}$$

for some given data ζ, and for lateral boundary conditions we choose the classical assumption of periodicity. To specify this we define a lattice of periodicity Γ and a fundamental period cell $P(\Gamma)$. This lattice generates a conjugate lattice Γ' of wave numbers so that any function periodic with respect to Γ can be represented by its Fourier series:

$$\psi(x) = \sum_{p \in \Gamma'} \hat{\psi}_p e^{ip\cdot x}.$$

For instance, in the case of single-variable 2π-periodic functions we have $\Gamma = (2\pi)\mathbf{Z}$, $P(\Gamma) = [0,2\pi]$, and $\Gamma' = \mathbf{Z}$. With this notation in hand we require that V be periodic in the x variable:

$$V(x+\gamma,y) = V(x,y) \quad \forall\, \gamma \in \Gamma,\ x \in P(\Gamma). \tag{3.3}$$

To specify a unique solution one must enforce a condition at $y = -h$, and for this we again follow the guidance of the ideal fluid problem and require

$$\partial_y V(x,-h) = 0. \tag{3.4}$$

We can greatly reduce the size of the problem domain by introducing a "transparent boundary condition" [36]: Consider a plane $y = -a$ such that $-h < -a < -|g|_{L^\infty}$, and define the Fourier multiplier

$$T[\psi] := \partial_y W(x,-a) = \sum_{p\in\Gamma'} \hat{\psi}_p \,|p| \tanh(|p|\,(h-a)) e^{ip\cdot x} =: |D| \tanh((h-a)\,|D|)\psi(x). \tag{3.5}$$

The operator T is typically called a Dirichlet–Neumann operator (DNO) or Dirichlet-to-Neumann map, and (3.1)–(3.4) are equivalent to

$$\Delta V(x,y) = 0, \qquad (x,y) \in S_{a,g}, \tag{3.6a}$$

$$V(x,g(x)) = \zeta(x), \tag{3.6b}$$

$$\partial_y V(x,-a) - T[V(x,-a)] = 0, \tag{3.6c}$$

$$V(x+\gamma,y) = V(x,y); \tag{3.6d}$$

cf. [36]. Notice that we can accommodate the case $h = \infty$, which requires that $\partial_y V \to 0$ as $y \to -\infty$, with the operator $T = |D|$:

$$T[\psi] = |D|\,[\psi] := \sum_{p\in\Gamma'} |p|\,\hat{\psi}_p e^{ip\cdot x}.$$

To close the system of equations for the water wave problem one must pose boundary conditions at the unknown fluid interface $y = g$. This involves both the tangential and normal derivatives of the velocity potential V at the surface which, in turn, involves the solution of (3.6). We defer until section 4 a precise statement of these conditions and focus now upon the crucial operation of producing *normal* derivatives (tangential derivatives are easily computed via $\partial_x \zeta(x)$), i.e., the *surface* DNO:

$$G(g)[\zeta] := \nabla V|_{y=g(x)} \cdot N_g, \tag{3.7}$$

where $N_g = (-\nabla_x g(x), 1)^T$ is a normal to the fluid domain. Once this operator is implemented the simulation of the full water wave problem is fairly straightforward [17].

Remark 3.1. It should be pointed out that while both T and G are DNO, the trivial boundary shape associated with T renders this operator a simple Fourier multiplier. By contrast, the complicated geometry associated with G produces a quite general, order one, pseudodifferential operator which is a multiple composition of Fourier multipliers and multiplications (e.g., by the boundary shape g). For instance, in the case $h = \infty$,

$$G(g)[\zeta] = |D|\,\zeta - \partial_x[g\partial_x\zeta] - |D|\,[g\,|D|\,\zeta] + \cdots.$$

Please see [17, 31] for more details of the "operator expansion" representation of G.

3.1. Transformed field expansions. Our boundary perturbation method for the simulation of the problem (3.6) is the so-called method of TFE [31, 32, 33]. To begin, we make a change of variables [41, 13, 31] which "flattens" the boundary:

$$x' = x, \quad y' = a\left(\frac{y - g(x)}{a + g(x)}\right). \tag{3.8}$$

Notice that $S_{a,g} = \{-a < y < g\}$ maps to $S_{a,0} = \{-a < y' < 0\}$ under this change of variables. We now define a *transformed* field

$$U(x', y') := V\left(x', \frac{(a + g(x'))y'}{a} + g(x')\right), \tag{3.9}$$

and a simple calculation shows that (3.6) transforms (upon dropping primes) to

$$\Delta U(x, y) = F(x, y; U, g), \qquad (x, y) \in S_{a,0}, \tag{3.10a}$$
$$U(x, 0) = \xi(x), \tag{3.10b}$$
$$\partial_{y'} U(x, -a) - T[U(x, -a)] = J(x; U, g), \tag{3.10c}$$
$$U(x + \gamma, y) = U(x, y), \tag{3.10d}$$

where $\xi(x) = \zeta(x)$ and F and J are reported in, e.g., [39].

With our boundary perturbation philosophy in mind we set $g(x) = \varepsilon f(x)$ and make the expansion

$$U(x,y;\varepsilon) = \sum_{n=0}^{\infty} U_n(x,y)\varepsilon^n, \tag{3.11}$$

which can be shown to be strongly convergent if f has two continuous derivatives [31, 33]. Inserting (3.11) into (3.10) results in the following set of problems to solve:

$$\Delta U_n(x,y) = F_n(x,y), \qquad (x,y) \in S_{a,0}, \tag{3.12a}$$
$$U_n(x,0) = \delta_{n,0}\xi(x), \tag{3.12b}$$
$$\partial_y U_n(x,-a) - T[U_n(x,-a)] = J_n(x), \tag{3.12c}$$
$$U_n(x+\gamma,y) = U_n(x,y), \tag{3.12d}$$

where $\delta_{m,n}$ is the Kronecker delta,

$$\begin{aligned} a^2 F_n = & -\mathrm{div}_x\left[2a\ f\ \nabla_x U_{n-1}\right] + \mathrm{div}_x\left[a(a+y)(\nabla_x f)\partial_y U_{n-1}\right] \\ & + \partial_y\left[a(a+y)(\nabla_x f)\cdot\nabla_x U_{n-1}\right] - \mathrm{div}_x\left[f^2\ \nabla_x U_{n-2}\right] \\ & + \mathrm{div}_x\left[(a+y)\ f\ (\nabla_x f)\partial_y U_{n-2}\right] + \partial_y\left[(a+y)\ f\ (\nabla_x f)\cdot\nabla_x U_{n-2}\right] \\ & - \partial_y\left[(a+y)^2\ \left|\nabla_x f\right|^2 \partial_y U_{n-2}\right] + a\ (\nabla_x f)\cdot\nabla_x U_{n-1} \\ & + f\ (\nabla_x f)\cdot\nabla_x U_{n-2} - (a+y)\left|\nabla_x f\right|^2 \partial_y U_{n-2}, \end{aligned} \tag{3.12e}$$

and

$$aJ_n = fT[U_{n-1}(x,-a)]. \tag{3.12f}$$

In the previous work of Nicholls and Reitich [31, 33], the analyticity of the expansion (3.11) was shown in function spaces which gave nearly optimal results with regard to smoothness ($\xi \in H^{3/2}$ and $f \in C^2$) while remaining quite straightforward technically. In a later publication, Hu and Nicholls [20] showed that the smoothness requirements on f could be eased to $C^{1+\alpha}$ and even Lipschitz, provided that the Dirichlet data (and thus the field U) sits in a suitable, though complicated, function space. For the present purposes it is necessary that both the Dirichlet data ξ and the deformation f lie in L^2-based Sobolev classes H^s. We now establish this new result; however, as it is quite close in spirit to the existing ones outlined above we present only an abbreviated proof.

To begin, we recall the algebra property of the Sobolev spaces H^s [1].

LEMMA 3.1. *If $s > d/2$ and $f, g \in H^s(P(\Gamma))$, $U \in H^s(S_{a,0})$, then for some positive constant $M = M(d,s)$ both of the following hold:*

1. *The product $fg \in H^s(P(\Gamma))$ and*

$$\|fg\|_{H^s(P(\Gamma))} \leq M \|f\|_{H^s(P(\Gamma))} \|g\|_{H^s(P(\Gamma))}.$$

2. *The product $fU \in H^s(S_{a,0})$ and*

$$\|fU\|_{H^s(S_{a,0})} \leq M \|f\|_{H^s(P(\Gamma))} \|U\|_{H^s(S_{a,0})}.$$

Notice that in the first case, $s > (d-1)/2$ would suffice since $P(\Gamma) \subset \mathbf{R}^{d-1}$.

Next, we need the following well-known elliptic regularity result [26, 18].

LEMMA 3.2. *For any integer $s \geq 0$, if $F \in H^s(S_{a,0})$, $\xi \in H^{s+3/2}(P(\Gamma))$, and $J \in H^{s+1/2}(P(\Gamma))$, then there exists a unique solution of the elliptic boundary value problem*

$$\begin{aligned}
&\Delta W(x,y) = F(x,y), && (x,y) \in S_{a,0},\\
&W(x,0) = \xi(x), &&\\
&\partial_y W(x,-a) - T[W(x,-a)] = J(x), &&\\
&W(x+\gamma, y) = W(x,y) &&
\end{aligned}$$

such that the estimate

$$\|W\|_{H^{s+2}(S_{a,0})} \leq C_e \left\{ \|F\|_{H^s(S_{a,0})} + \|\xi\|_{H^{s+3/2}(P(\Gamma))} + \|J\|_{H^{s+1/2}(P(\Gamma))} \right\} \tag{3.13}$$

holds for some positive constant C_e.

For our inductive method of proof we require a recursive estimate based upon the problem (3.12).

LEMMA 3.3. *Given an integer $s \geq 0$ and constants $B, C > 0$, if*

$$\|U_n\|_{H^{s+2}(S_{a,0})} \leq CB^n \qquad \forall n < N,$$

then there exists $C_3 > 0$ such that the functions F_N and J_N satisfy

$$\begin{aligned}
&\|F_N\|_{H^s(S_{a,0})} \leq C_3 \left\{ \|f\|_{H^{s+2}(P(\Gamma))} B^{N-1} + \|f\|^2_{H^{s+2}(P(\Gamma))} B^{N-2} \right\},\\
&\|J_N\|_{H^{s+1/2}(P(\Gamma))} \leq C_3 \|f\|_{H^{s+2}(P(\Gamma))} B^{N-1}.
\end{aligned}$$

Proof. Let us start with F_N and consider the representative term in (3.12e):

$$R_N := \frac{1}{a^2} \mathrm{div}_x \left[(a+y)\ f\ (\nabla_x f) \partial_y U_{N-2} \right].$$

We begin with

$$\begin{aligned}
\|R_N\|_{H^s} &\leq \frac{1}{a^2} \|(a+y)\ f\ (\nabla_x f)\partial_y U_{N-2}\|_{H^{s+1}(S_{a,0})}\\
&\leq \frac{Y}{a^2} \|f\ (\nabla_x f)\partial_y U_{N-2}\|_{H^{s+1}(S_{a,0})}\\
&\leq \frac{YM}{a^2} \|f\|_{H^{s+1}(P(\Gamma))} M \|f\|_{H^{s+2}(P(\Gamma))} \|\partial_y U_{N-2}\|_{H^{s+1}(S_{a,0})}\\
&\leq \frac{YM}{a^2} \|f\|_{H^{s+1}(P(\Gamma))} M \|f\|_{H^{s+2}(P(\Gamma))} \|U_{N-2}\|_{H^{s+2}(S_{a,0})}\\
&\leq \frac{YM^2}{a^2} \|f\|^2_{H^{s+2}(P(\Gamma))} CB^{N-2},
\end{aligned}$$

where we have used the inductive hypothesis Lemma 3.1 and the constant Y defined as the smallest value such that

$$\|(a+y)F\|_{H^s(S_{a,0})} \leq Y \|F\|_{H^s(S_{a,0})}.$$

It is clear that other terms in F_N can be estimated in the same fashion.

Regarding J_N, (3.12f), we may use Lemma 3.1 and classical trace theorems (e.g., Adams [1, Chapter 7, Theorem 7.53]) to show that

$$\begin{aligned}\|J_N\|_{H^{s+1/2}(P(\Gamma))} &\le \frac{1}{a}\|fT[U_{N-1}(\cdot,-a)]\|_{H^{s+1/2}(P(\Gamma))}\\ &\le \frac{M}{a}\|f\|_{H^{s+1/2}(P(\Gamma))}\|T[U_{N-1}(\cdot,-a)]\|_{H^{s+1/2}(P(\Gamma))}\\ &\le \frac{M}{a}\|f\|_{H^{s+1/2}(P(\Gamma))}\|U_{N-1}(\cdot,-a)\|_{H^{s+3/2}(P(\Gamma))}\\ &\le \frac{M}{a}\|f\|_{H^{s+2}(P(\Gamma))}\|U_{N-1}\|_{H^{s+2}(S_{a,0})}\\ &\le \frac{M}{a}\|f\|_{H^{s+2}(P(\Gamma))}CB^{N-1}. \qquad \square\end{aligned}$$

We are now in a position to establish the analyticity result corresponding to Proposition 2.1.

THEOREM 3.1. *Given an integer* $s \ge 0$*, if* $\xi \in H^{s+3/2}(P(\Gamma))$ *and* $f \in H^{s+2}(P(\Gamma))$*, then the expansion* (3.11) *converges strongly; i.e. there exist constants* $C_1, C_2 > 0$ *such that*

$$\|U_n\|_{H^{s+2}(S_{a,0})} \le C_1 \|\xi\|_{H^{s+3/2}(P(\Gamma))} B^n \tag{3.14}$$

for any constant $B > C_2 \|f\|_{H^{s+2}(P(\Gamma))}$.

Proof. As in [31] we work by induction in the perturbation order n. Each of the U_n satisfies (3.12), and at order zero the problem is particularly simple as the right-hand sides F_0 and J_0 vanish. Here we may use the elliptic regularity result Lemma 3.2 to show that

$$\|U_0\|_{H^{s+2}(S_{a,0})} \le C_e \|\xi\|_{H^{s+3/2}(P(\Gamma))},$$

and we define $C_1 := C_e$. We now assume the estimate (3.14) for every order $n < N$. Using Lemma 3.2 followed by Lemma 3.3, we deduce that

$$\begin{aligned}\|U_N\|_{H^{s+2}(S_{a,0})} &\le C_e \left\{ \|F_N\|_{H^s(S_{a,0})} + \|J_N\|_{H^{s+1/2}(P(\Gamma))} \right\}\\ &\le 2C_eC_3 \left\{ \|f\|_{H^{s+2}(P(\Gamma))} B^{N-1} + \|f\|^2_{H^{s+2}(P(\Gamma))} B^{N-2} \right\} \le C_1 B^N,\end{aligned}$$

provided that

$$B \ge \max\left\{4C_3C_1, 2\sqrt{C_3C_1}\right\} \|f\|_{H^{s+2}(P(\Gamma))}. \qquad \square$$

Remark 3.2. As we noted before, the analogue of Theorem 3.1 holds with weaker assumptions on the boundary shape, namely, $f \in C^{s+3/2+\delta}(P(\Gamma))$ [31], $f \in C^{1+\alpha}(P(\Gamma))$, or even $f \in \mathrm{Lip}(P(\Gamma))$ [20]. However, as we shall see, it is far more convenient for all of the relevant functions to belong to one of the L^2-based Sobolev spaces H^r.

3.2. A Fourier–Legendre Galerkin approximation. Our numerical approach is to simulate the solution of (3.12) by a discrete approximation $U_n^{\bar{h}}$, for every perturbation order n, and then produce the approximation

$$U^{N,\bar{h}}(x,y) := \sum_{n=0}^{N} U_n^{\bar{h}}(x,y)\varepsilon^n. \tag{3.15}$$

To find the $U_n^{\bar{h}}$ we must solve the prototype problem; cf. (3.12):

$$\Delta u(x,y) = f(x,y), \qquad (x,y) \in S_{a,0}, \tag{3.16a}$$

$$u(x,0) = b_1(x), \tag{3.16b}$$

$$\partial_y u(x,-a) - T[u(x,-a)] = b_2(x), \tag{3.16c}$$

$$u(x+\gamma, y) = u(x,y). \tag{3.16d}$$

In light of the periodic boundary conditions, we begin by expanding u in a Fourier series

$$u(x,y) = \sum_{p\in\Gamma'} \hat{u}_p(y) e^{ip\cdot x}$$

and likewise for f, b_1, and b_2. Inserting these expansions into (3.16), we find that $\hat{u}_p(y)$ satisfies the two-point boundary value problem

$$\partial_y^2 \hat{u}_p(y) - |p|^2 \hat{u}_p(y) = \hat{f}_p(y), \qquad -a < y < 0, \tag{3.17a}$$

$$\hat{u}_p(0) = \hat{b}_{1,p}, \tag{3.17b}$$

$$\partial_y \hat{u}_p(-a) - |p| \tanh((h-a)|p|) \hat{u}_p(-a) = \hat{b}_{2,p}. \tag{3.17c}$$

Denoting ${}^bH^1(-a,0) = \{v \in H^1(-a,0) \mid v(0) = b\}$, the weak form of (3.17) reads as follows:

Find $\hat{u}_p \in {}^{\hat{b}_{1,p}}H^1(-a,0)$ such that

$$\begin{aligned} &\int_{-a}^0 \partial_y \hat{u}_p \partial_y \phi \,\mathrm{d}y + |p| \tanh((h-a)|p|) \hat{u}_p(-a)\phi(-a) + |p|^2 \int_{-a}^0 \hat{u}_p \phi \,\mathrm{d}y \\ &\qquad = -\int_{-a}^0 \hat{f}_p \phi \,\mathrm{d}y - \hat{b}_{2,p}\phi(-a) \quad \forall \phi \in {}^0H^1(-a,0). \end{aligned} \tag{3.18}$$

Let us denote

$${}^bP_M = \operatorname{span}\{v \in P_M \mid v(0) = b\},$$

where P_M is the space of polynomials of degree less than or equal to M. The Legendre–Galerkin method for (3.17) is as follows:

Find $\hat{u}_{p,M} \in {}^{\hat{b}_{1,p}}P_M$ such that

$$\begin{aligned} &\int_{-a}^0 \partial_y \hat{u}_{p,M} \partial_y \phi \,\mathrm{d}y + |p| \tanh((h-a)|p|) \hat{u}_{p,M}(-a)\phi(-a) + |p|^2 \int_{-a}^0 \hat{u}_{p,M} \phi \,\mathrm{d}y \\ &\qquad = -\int_{-a}^0 \hat{f}_p \phi \,\mathrm{d}y - \hat{b}_{2,p}\phi(-a) \quad \forall \phi \in {}^0P_M. \end{aligned} \tag{3.19}$$

In order to derive an error estimate for this method with explicit dependence on p, we need the following result (cf. [5, 3]).

LEMMA 3.4. *There exists an operator* $\tilde{\Pi}_M \colon H^1(-a,0) \to P_M$ *satisfying* $\tilde{\Pi}_M u(-a) = u(-a)$ *and* $\tilde{\Pi}_M u(0) = u(0)$ *such that*

$$\|\partial_y(u - \tilde{\Pi}_M u)\|_{H^s(-a,0)} \lesssim M^{s-r} \|u\|_{H^r(-a,0)} \quad \forall u \in H^r(-a,0), \quad 0 \le s \le 1 \le r. \tag{3.20}$$

We can then prove the following.

THEOREM 3.2. *If $\hat{u}_p \in H^r(-a,0)$ with $r \geq 1$, we have*

$$\|\hat{u}_p - \hat{u}_{p,M}\|_{H^1(-a,0)} + |p|\,\|\hat{u}_p - \hat{u}_{p,M}\|_{H^0(-a,0)} \lesssim M^{1-r}(\|\hat{u}_p\|_{H^r(-a,0)} + |p|\,\|\hat{u}_p\|_{H^{r-1}(-a,0)}).$$

Proof. Subtracting (3.19) from (3.18), we obtain the following error equation:

$$\int_{-a}^{0} \partial_y(u - \hat{u}_{p,M})\partial_y \phi \, \mathrm{d}y + |p|^2 \int_{-a}^{0} (u - \hat{u}_{p,M})\phi \, \mathrm{d}y = 0 \quad \forall \phi \in {}^0P_M. \tag{3.21}$$

Let us denote

$$e_{p,M} = u - \hat{u}_{p,M} = (u - \tilde{\Pi}_N u) + (\tilde{\Pi}_N u - \hat{u}_{p,M}) := \tilde{e}_{p,M} + \hat{e}_{p,M}.$$

Since $\hat{e}_{p,M} \in {}^0P_M$, taking $\phi = \hat{e}_{p,M}$ in (3.21), we derive from the Schwartz inequality that

(3.22)

$$\begin{aligned}
\|\partial_y \hat{e}_{p,M}\|^2_{H^0(-a,0)} + |p|^2 \|\hat{e}_{p,M}\|^2_{H^0(-a,0)} &= -\int_{-a}^{0} \partial_y \tilde{e}_{p,M} \partial_y \hat{e}_{p,M} \, \mathrm{d}y \\
&\quad - |p|^2 \int_{-a}^{0} \tilde{e}_{p,M} \hat{e}_{p,M} \, \mathrm{d}y \\
&\leq \frac{1}{2}\left(\|\partial_y \hat{e}_{p,M}\|^2_{H^0(-a,0)} + |p|^2 \|\hat{e}_{p,M}\|^2_{H^0(-a,0)} \right) \\
&\quad + \frac{1}{2}\left(\|\partial_y \tilde{e}_{p,M}\|^2_{H^0(-a,0)} + |p|^2 \|\tilde{e}_{p,M}\|^2_{H^0(-a,0)} \right),
\end{aligned}$$

which, along with Lemma 3.4, implies that

$$\begin{aligned}
\|\partial_y \hat{e}_{p,M}\|^2_{H^0(-a,0)} + |p|^2 \|\hat{e}_{p,M}\|^2_{H^0(-a,0)} &\leq \left(\|\partial_y \tilde{e}_{p,M}\|^2_{H^0(-a,0)} + |p|^2 \|\tilde{e}_{p,M}\|^2_{H^0(-a,0)} \right) \\
&\lesssim M^{2(1-r)} \left(\|u\|^2_{H^r(-a,0)} + |p|^2 \|u\|^2_{H^{r-1}(-a,0)} \right).
\end{aligned}$$

We conclude by invoking the triangle inequality. ∎

Defining

$$\Gamma'_P := \{p \in \Gamma' \mid |p| < P\},$$

we now let

$$u^{P,M}(x,y) := \sum_{p \in \Gamma'_P} \hat{u}_{p,M}(y) e^{ip\cdot x}$$

be the Fourier–Legendre approximation of the solution u of (3.16). Then, we have the following error estimate.

THEOREM 3.3. *If $u \in H^r(S_{a,0})$ for any integer $r \geq 2$, then*

$$\left\| u - u^{P,M} \right\|_{H^1(S_{a,0})} \lesssim \left(P^{1-r} + M^{1-r} \right) \|u\|_{H^r(S_{a,0})}.$$

Proof. Since

$$u(x,y) - u^{P,M}(x,y) = \sum_{p \in \Gamma'_P} (\hat{u}_p(y) - \hat{u}_{p,M}(y))\, e^{ip\cdot x} + \sum_{p \in \Gamma' \setminus \Gamma'_P} \hat{u}_p(y) e^{ip\cdot x},$$

we have

$$\left\|\nabla(u-u^{P,M})\right\|_{H^0(-a,0)}^2 \lesssim \sum_{p\in\Gamma'_P}\left(\|\partial_y\hat{u}_p-\partial_y\hat{u}_{p,N}\|_{H^0(-a,0)}^2+|p|^2\|\hat{u}_p-\hat{u}_{p,N}\|_{H^0(-a,0)}^2\right) + \sum_{p\in\Gamma'\setminus\Gamma'_P}\left(\|\partial_y\hat{u}_p\|_{H^0(-a,0)}^2+|p|^2\|\hat{u}_p\|_{H^0(-a,0)}^2\right) := S_1^2+S_2^2.$$

We derive from Theorem 3.2 that

$$S_1^2 \lesssim M^{2(1-r)}\sum_{p\in\Gamma'_P}\|\hat{u}_p\|_{H^r(-a,0)}^2 \lesssim M^{2(1-r)}\|u\|_{H^r(S_{a,0})}^2.$$

On the other hand,

$$S_2^2 \lesssim P^{2(1-r)}\sum_{p\in\Gamma'\setminus\Gamma'_P}p^{2(r-1)}\left(\|\partial_y\hat{u}_p\|_{H^0(-a,0)}^2+|p|^2\|\hat{u}_p\|_{H^0(-a,0)}^2\right) \lesssim P^{2(1-r)}\|u\|_{H^r(S_{a,0})}^2.$$

The above estimates establish our result. ∎

Remark 3.3. Regarding this numerical approach, we have the following comments:

1. The matrix system associated with (3.19) is usually full. However, if we replace ${}^0P_M$ by

$$X_M^{(p)} = \mathrm{span}\,\{v\in P_M \mid v(0)=0,\quad \partial_y v(-a)-|p|\tanh((h-a)\,|p|)v(-a)=0\}$$

   and construct a suitable basis for $X_M^{(p)}$ as suggested in [47], we can obtain a sparse matrix system.
2. A Chebyshev–tau method for (3.17) was used in [32]. However, the Chebyshev–tau method does not lead to optimal error estimates. In practice, it is more efficient to use a Chebyshev–Galerkin method [48] for (3.17). A result similar to Theorem 3.3 for the Chebyshev–Galerkin approach can be established, but the proof is much more involved due to the nonuniform weight associated with the Chebyshev polynomials.

3.3. Combined error estimates. Let $U_n^{P,M}$ be the Fourier–Legendre Galerkin approximation of U_n, the solution of (3.12). Then, Theorem 3.3 leads to the following corollary.

COROLLARY 3.1. *Let U_n be the solution of* (3.12). *If $U_n\in H^r(S_{a,0})$ for any integer $r\geq 2$, then*

$$\left\|U_n-U_n^{P,M}\right\|_{H^1(S_{a,0})} \lesssim \left(P^{1-r}+M^{1-r}\right)\|U_n\|_{H^r(S_{a,0})}.$$

This is essentially the analogue of Proposition 2.2 that we require. Now, set

$$U^{N,P,M}(x,y) := \sum_{n=0}^{N} U_n^{P,M}(x,y)\varepsilon^n$$

which is our approximation to the solution U of (3.10). With Theorem 3.1 we use the proof of Theorem 2.1 to realize the following convergence estimate.

THEOREM 3.4. *Let U be the solution of* (3.10). *Assume that for some integer $r\geq 2$, $f\in H^r(P(\Gamma))$ and $\xi\in H^{r-1/2}(P(\Gamma))$; then*

$$\left\|U-U^{N,P,M}\right\|_{H^1(S_{a,0})} \lesssim (B\varepsilon)^{N+1}+\left(P^{1-r}+M^{1-r}\right)\|\xi\|_{H^{r-1/2}(P(\Gamma))}$$

for any constant $B \geq C_2 \|f\|_{H^r(P(\Gamma))}$ *such that* $B\varepsilon < 1$*, where* C_2 *is the constant in Theorem* 3.1.

Remark 3.4. We point out that, once Theorem 3.4 is established for the convergence of $U^{N,P,M}$ to the field U, a similar result follows quite straightforwardly for the convergence of a corresponding numerical approximation of the DNO, say, $G^{N,P}$, to the DNO, G (see (3.7)). Since this adds no new difficulties for our method, we simply state the result here and omit the proof.

4. Water wave equations. Computing the free-surface evolution of an ideal fluid under the influence of gravity and capillarity (the water wave problem) is one of the most difficult problems in the field of computational fluid mechanics. Not only is the system of equations highly nonlinear without any dissipation mechanism but also the domain of the fluid is an unknown of the problem.

Recall from section 3 that, in specifying the model for free-surface potential flow, we denoted the (unknown) domain of the fluid by

$$S_{h,\eta} := \{(x, y) \in \mathbf{R}^{d-1} \times \mathbf{R} \mid -h < y < \eta(x, t)\},$$

where $y = \eta(x, t)$ is the free air-fluid interface (we have replaced g with the more standard notation η) and the bottom of the ocean is at $y = -h$. As we saw earlier, an artificial boundary can be introduced to the domain (at $y = -a$, $-h < -a < -|\eta|_{L^\infty}$) and a transparent boundary condition enforced there with the operator T from (3.5). With this, the well-known model for the evolution of ocean waves is potential flow, governed by the following system of equations [27]:

$$\Delta\varphi = 0, \qquad (x, y) \in S_{a,\eta}, \tag{4.1a}$$

$$\partial_y\varphi - T[\varphi] = 0 \qquad \text{at } y = -a, \tag{4.1b}$$

$$\partial_t\eta = \partial_y\varphi - (\nabla_x\eta) \cdot \nabla_x\varphi \qquad \text{at } y = \eta, \tag{4.1c}$$

$$\partial_t\varphi = -g\eta + \sigma\Delta_x\eta + \sigma\kappa(\eta) - \frac{1}{2}|\nabla\varphi|^2 \qquad \text{at } y = \eta, \tag{4.1d}$$

where $\varphi = \varphi(x, y, t)$ is the velocity potential (so that the velocity is given by $\vec{v} = \nabla\varphi$), g and σ are the constants of gravity and capillarity, respectively, and the modified curvature κ is given by

$$\kappa(\eta) := \operatorname{div}_x \left[\frac{\nabla_x\eta}{\sqrt{1 + |\nabla_x\eta|^2}}\right] - \Delta_x\eta.$$

Once again, we investigate solutions periodic with respect to a lattice Γ, defining a period cell $P(\Gamma)$ and generating a conjugate lattice Γ' of wave numbers (see section 3).

Among the many questions of interest regarding these equations, one that the authors have taken up is the existence of traveling wave solutions [15, 36] and their numerical solution [16, 37, 28]. It is not difficult to show that steady state solutions of (4.1) in a reference frame moving uniformly with velocity $c \in \mathbf{R}^{d-1}$ satisfy

$$\Delta\varphi = 0, \qquad (x, y) \in S_{a,\eta}, \tag{4.2a}$$

$$\partial_y\varphi - T[\varphi] = 0 \qquad \text{at } y = -a, \tag{4.2b}$$

$$c \cdot \nabla_x\eta = \partial_y\varphi - (\nabla_x\eta) \cdot \nabla_x\varphi \qquad \text{at } y = \eta, \tag{4.2c}$$

$$c \cdot \nabla_x\varphi = -g\eta + \sigma\Delta_x\eta + \sigma\kappa(\eta) - \frac{1}{2}|\nabla\varphi|^2 \qquad \text{at } y = \eta; \tag{4.2d}$$

see [36, 37]. Once again, the change of variables (3.8) is the starting point for the specification of a TFE method for this problem, and, upon defining

$$U(x',y') := \varphi\left(x', \frac{(a+\eta(x'))y'}{a} + \eta(x')\right),$$

we find that U, η, and c must satisfy (upon dropping primes)

$$\begin{aligned}
&(4.3a) && \Delta U = F(x,y;U,\eta), && (x,y)\in S_{a,0},\\
&(4.3b) && \partial_y U - T[U] = J(x;U,\eta) && \text{at } y=-a,\\
&(4.3c) && c\cdot\nabla_x\eta - \partial_y U = \tilde{Q}(x;U,\eta) && \text{at } y=0,\\
&(4.3d) && c\cdot\nabla_x U + g\eta - \sigma\Delta_x\eta = \tilde{R}(x;U,\eta) && \text{at } y=0,
\end{aligned}$$

where the precise forms for F and J are given in (3.10) while those for $\tilde{Q}$ and $\tilde{R}$ are reported in [36, 37]. If we now suppose that the expansions

$$U(x,y;\varepsilon) = \sum_{n=1}^{\infty} U_n(x,y)\varepsilon^n, \quad \eta(x;\varepsilon) = \sum_{n=1}^{\infty}\eta_n(x)\varepsilon^n, \quad c(\varepsilon) = c_0 + \sum_{n=1}^{\infty} c_n\varepsilon^n$$

are valid, then it is easy to show that the $\{U_n, \eta_n, c_{n-1}\}$ $(n\geq 1)$ satisfy

$$\begin{aligned}
&(4.4a) && \Delta U_n = F_n(x,y), && (x,y)\in S_{a,0},\\
&(4.4b) && \partial_y U_n - T[U_n] = J_n(x) && \text{at } y=-a,\\
&(4.4c) && c_0\cdot\nabla_x\eta_n - \partial_y U_n = Q_n(x) - c_{n-1}\cdot\nabla_x\eta_1 && \text{at } y=0,\\
&(4.4d) && c_0\cdot\nabla_x U_n + g\eta_n - \sigma\Delta_x\eta_n = R_n(x) - c_{n-1}\cdot\nabla_x U_1 && \text{at } y=0,
\end{aligned}$$

where

$$Q_n = \tilde{Q}_n - \sum_{l=1}^{n-2} c_l\cdot\nabla_x\eta_{n-l}, \quad R_n = \tilde{R}_n - \sum_{l=1}^{n-2} c_l\cdot\nabla_x U_{n-l}.$$

We separate out the term involving c_{n-1} explicitly as this allows us to specify a solvability condition at every perturbation order [36] (see the developments following (4.12)).

To solve (4.4) for each n we take a rather different approach than that described for the problem in section 3. At order one we must find (U_1, η_1, c_0) satisfying

$$\begin{aligned}
&(4.5a) && \Delta U_1 = 0, && (x,y)\in S_{a,0},\\
&(4.5b) && \partial_y U_1 - T[U_1] = 0 && \text{at } y=-a,\\
&(4.5c) && c_0\cdot\nabla_x\eta_1 - \partial_y U_1 = 0 && \text{at } y=0,\\
&(4.5d) && c_0\cdot\nabla_x U_1 + g\eta_1 - \sigma\Delta_x\eta_1 = 0 && \text{at } y=0,
\end{aligned}$$

and it is not difficult to see that (4.5a) and (4.5b) have periodic solutions:

$$U_1(x,y) = \sum_{p\in\Gamma'} a_{1,p}\frac{\cosh(|p|\,(y+h))}{\cosh(|p|\,h)}e^{ip\cdot x}, \quad \eta_1(x) = \sum_{p\in\Gamma'} d_{1,p}e^{ip\cdot x}.$$

(4.5c) and (4.5d) demand that $M_p w_{1,p} = 0$ for every $p\in\Gamma'$ where

$$M_p := \begin{pmatrix} ic_0\cdot p & -|p|\tanh(h\,|p|) \\ g+\sigma\,|p|^2 & ic_0\cdot p\end{pmatrix}, \quad w_{1,p} := \begin{pmatrix} d_{1,p}\\ a_{1,p}\end{pmatrix}.$$

This equation will have only nontrivial solutions if the matrix M_p is singular, measured by the (opposite of the) determinant function

$$\Lambda(c_0, p) := (c_0 \cdot p)^2 - (g + \sigma |p|^2)\, |p| \tanh(h\, |p|).$$

Our strategy for finding interesting (nonzero) solutions is governed by the dimension of the problem ($d = 2, 3$) and the presence or absence of capillarity σ. We summarize as follows [15]:

1. If $d = 2$ and $\sigma = 0$, then we select a wave number $p_0 \in \Gamma' \subset \mathbf{R}$ and find the unique (up to sign) value $c_0 \in \mathbf{R}$ such that $\Lambda(c_0, p_0) = 0$.
2. If $d = 2$ and $\sigma > 0$, then, again, we select a wave number $p_0 \in \Gamma' \subset \mathbf{R}$ and find the unique (up to sign) value $c_0 \in \mathbf{R}$ such that $\Lambda(c_0, p_0) = 0$. However, in this case there may be a second wave number $p \neq 0, \pm p_0$ such that $\Lambda(c_0, p) = 0$ which impedes our algorithm as it is currently formulated. This "resonance" gives rise to the "Wilton ripples" [45, 44], and we do not consider such waves in this contribution.
3. If $d = 3$ and $\sigma > 0$, then we select two linearly independent wave numbers $p_1, p_2 \in \Gamma' \subset \mathbf{R}^2$ and find the unique (up to sign) value $c_0 \in \mathbf{R}^2$ such that $\Lambda(c_0, p_j) = 0$. As in the case $d = 2$, the Wilton ripple resonance may arise, but we assume that this is not the case.
4. If $d = 3$ and $\sigma = 0$, then for any given $c_0 \in \mathbf{R}^2$ it may be the case that $\Lambda(c_0, p) = 0$ for an *infinite* number of linearly independent $p \in \Gamma' \subset \mathbf{R}^2$. This instance of "small denominators" is outside the scope of our algorithm and is still an open problem (see the recent results of [42, 22, 21, 23, 24]).

For simplicity we focus upon the first case ($d = 2$, $\sigma = 0$) and note that in this case we can find solutions at order $n = 1$, for any $\tau \in \mathbf{R}$,

$$d_{1,p_0} = \tau\, |p_0| \tanh(h\, |p_0|), \tag{4.6a}$$
$$d_{1,-p_0} = \tau\, |p_0| \tanh(h\, |p_0|), \tag{4.6b}$$
$$a_{1,p_0} = \tau(i c_0 p_0), \tag{4.6c}$$
$$a_{1,-p_0} = -\tau(i c_0 p_0), \tag{4.6d}$$

and $d_{1,p} = a_{1,p} = 0$ for all $p \neq \pm p_0$. Having selected our solution at order $n = 1$, we can proceed to solve (4.4) for $n \geq 2$ as the discussion of the next paragraph shows. In fact, we have shown the following analyticity result in [36].

THEOREM 4.1 (Nicholls and Reitich [36]). *If $\sigma > 0$ (or $\sigma = 0$ for $d = 2$) and no Wilton ripple resonance is present, then there exists $C_1(r, q)$, $B > 0$ such that, for all integers $r > d/2$ and $q > d - 1$, the solutions $\{U_n, \eta_n, c_{n-1}\}$ of (4.4) satisfy, for $n \geq 1$,*

$$\|U_n\|_{H^{r+2}(S_{a,0})} \leq C_1(r,q) \frac{B^{n-1}}{n^q}, \ \|\eta_n\|_{H^{r+5/2}(P(\Gamma))} \leq C_1(r,q) \frac{B^{n-1}}{n^q}, \ |c_{n-1}| \leq C_1(r,q) \frac{B^{n-1}}{n^q}.$$

Now, for each $n \geq 2$ we have to solve the prototype problem (cf. (4.4))

$$\Delta u = f(x, y), \qquad (x, y) \in S_{a,0}, \tag{4.7a}$$
$$\partial_y u - T[u] = J(x) \qquad \text{at } y = -a, \tag{4.7b}$$
$$c_0 \cdot \nabla_x \eta - \partial_y u = Q(x) - c_{n-1} \cdot \nabla_x \eta_1 \qquad \text{at } y = 0, \tag{4.7c}$$
$$c_0 \cdot \nabla_x u + g\eta - \sigma \Delta_x \eta = R(x) - c_{n-1} \cdot \nabla_x U_1 \qquad \text{at } y = 0, \tag{4.7d}$$

and, to prescribe a precise numerical formulation, we use the periodic boundary condition in x to write

$$u(x,y) = \sum_{p\in\Gamma'} \hat{u}_p(y) e^{ip\cdot x}, \quad \eta(x) = \sum_{p\in\Gamma'} \hat{\eta}_p e^{ip\cdot x}.$$

Upon inserting this into (4.7) we realize

$$\partial_y^2 \hat{u}_p - |p|^2 \hat{u}_p = \hat{f}_p(y), \qquad -a < y < 0, \tag{4.8a}$$

$$\partial_y \hat{u}_p(-a) - |p| \tanh((h-a)|p|) \hat{u}_p(-a) = \hat{J}_p, \tag{4.8b}$$

$$(ic_0 \cdot p)\hat{\eta}_p - \partial_y \hat{u}_p(0) = \hat{Q}_p - \delta_{p,\pm p_0}(ic_{n-1}\cdot p) d_{1,p}, \tag{4.8c}$$

$$(g + \sigma |p|^2)\hat{\eta}_p + (ic_0\cdot p)\hat{u}_p(0) = \hat{R}_p - \delta_{p,\pm p_0}(ic_{n-1}\cdot p) a_{1,p}. \tag{4.8d}$$

Our treatment of this problem depends upon the value of p which can be broken into three distinct cases.

1. *Case* $p \neq 0, \pm p_0$
   We can eliminate η_p by multiplying (4.8c) by $-(g+\sigma|p|^2)$ and (4.8d) by $(ic_0\cdot p)$, and summing up to obtain

$$(g+\sigma|p|^2)\partial_y \hat{u}_p(0) + (ic_0\cdot p)^2 \hat{u}_p(0) = -(g+\sigma|p|^2)\hat{Q}_p + (ic_0\cdot p)\hat{R}_p, \tag{4.9}$$

   which leads to the following generic problem:

$$\partial_y^2 \hat{u}_p - |p|^2 \hat{u}_p = \hat{f}_p(y), \qquad -a < y < 0, \tag{4.10a}$$

$$\partial_y \hat{u}_p(-a) - |p| \tanh((h-a)|p|) \hat{u}_p(-a) = \hat{b}_{2,p}, \tag{4.10b}$$

$$\partial_y \hat{u}_p(0) - \frac{(c_0\cdot p)^2}{(g+\sigma|p|^2)} \hat{u}_p(0) = \hat{b}_{1,p}. \tag{4.10c}$$

   In this case, since $\Lambda(c_0, p) \neq 0$, the problem (4.10) is well-posed.
2. *Case* $p = 0$
   Notice that $\Lambda(c_0, 0) = 0$, so care is required. In this case (4.10) becomes

$$\partial_y^2 \hat{u}_0 = \hat{f}_0(y), \qquad -a < y < 0, \tag{4.11a}$$

$$\partial_y \hat{u}_0(-a) = \hat{b}_{2,0}, \tag{4.11b}$$

$$\partial_y \hat{u}_0(0) = \hat{b}_{1,0}, \tag{4.11c}$$

   which is not uniquely solvable. To render it so, we must use the uniqueness constraint

$$\int_{P(\Gamma)} \partial_y U(x, -h)\, \mathrm{d}x = 0,$$

   see [36], which specifies $\hat{u}_0(-a)$ so that the problem (4.11) becomes well-posed.
3. *Case* $p = \pm p_0$

Here the analogue of (4.10), in the representative case $p = p_0$, becomes

(4.12a)
$$\partial_y^2 \hat{u}_{p_0} - |p_0|^2 \hat{u}_{p_0} = \hat{f}_{p_0}(y), \qquad -a < y < 0,$$
(4.12b)
$$\partial_y \hat{u}_{p_0}(-a) - |p_0| \tanh((h-a)|p_0|)\hat{u}_{p_0}(-a) = \hat{b}_{2,p_0} - \mu \hat{c}_{2,p_0},$$
(4.12c)
$$\partial_y \hat{u}_{p_0}(0) - \frac{(c_0 \cdot p_0)^2}{(g + \sigma |p_0|^2)} \hat{u}_{p_0}(0) = \hat{b}_{1,p_0} - \mu \hat{c}_{1,p_0}.$$

From our choice of c_0 we know that this problem is not uniquely solvable and we must choose μ (meant to represent c_{n-1}) to force the right-hand side into the range of the linear operator on the left-hand side. The details of this choice are given in [36] and involve the representation of $\hat{u}_{p_0}(y)$ as an integral equation. Once this has been arranged, we must make a choice to find a unique solution, and here we use Stokes' choice that η_n be L^2-orthogonal to η_1. Since the Fourier transform of η_1 is supported only at wave numbers $p = \pm p_0$, this is easily accomplished by setting

$$\hat{\eta}_{n,\pm p_0} = 0, \quad n \geq 2.$$

In light of these considerations, from (4.8) we obtain the modified prototype problem (with μ specially chosen and $\hat{\eta}_{p_0} = 0$):

$$\text{(4.13a)} \quad \partial_y^2 \hat{u}_{p_0} - |p_0|^2 \hat{u}_{p_0} = \hat{f}_{p_0}, \qquad -a < y < 0,$$
$$\text{(4.13b)} \quad \partial_y \hat{u}_{p_0}(-a) - |p_0| \tanh((h-a)|p_0|)\hat{u}_{p_0}(-a) = \hat{b}_{2,p_0},$$
$$\text{(4.13c)} \quad (ic_0 \cdot p_0)\hat{u}_{p_0}(0) = \hat{b}_{1,p_0} - \mu \hat{c}_{1,p_0},$$

which is obviously well-posed.

Now, for the three prototype problems (4.10), (4.11), and (4.13), we can write down their weak formulations as in (3.18) and define $\hat{u}_{p,M}$ as their Legendre–Galerkin approximations. Once $\hat{u}_{p,M}$ is determined, we can then define $\eta_{p,M}$, the approximation of η_p, by (4.8d), namely,

$$\text{(4.14)} \qquad \hat{\eta}_{p,M} = \frac{1}{(g + \sigma |p|^2)} \left(-(ic_0 \cdot p)\hat{u}_{p,M}(0) + \hat{R}_p - \delta_{p,\pm p_0}(ic_{n-1} \cdot p) a_{1,p} \right).$$

The pair $\{\hat{u}_{p,M}, \hat{\eta}_{p,M}\}$ will be called the Legendre–Galerkin approximation for $\{\hat{u}_p, \hat{\eta}_p\}$, the solution of (4.8).

Then, we have the following two theorems.

Theorem 4.2. *Let $\{\hat{u}_p, \hat{\eta}_p\}$ be the solution of* (4.8) *and $\{\hat{u}_{p,M}, \hat{\eta}_{p,M}\}$ be its Legendre–Galerkin approximation. If $\hat{u}_p \in H^r(-a, 0)$ with $r \geq 1$, we have*

$$\|\hat{u}_p - \hat{u}_{p,M}\|_{H^1(-a,0)} + |p| \|\hat{u}_p - \hat{u}_{p,M}\|_{H^0(-a,0)}$$
$$\text{(4.15a)} \qquad \lesssim M^{1-r}(\|\hat{u}_p\|_{H^r(-a,0)} + |p| \|\hat{u}_p\|_{H^{r-1}(-a,0)}),$$
$$\text{(4.15b)} \qquad |\hat{\eta}_p - \hat{\eta}_{p,M}| \lesssim M^{1-r}(\|\hat{u}_p\|_{H^r(-a,0)} + |p| \|\hat{u}_p\|_{H^{r-1}(-a,0)}).$$

Proof. Since the problem (4.13) takes the same form as (3.17), Theorem 4.2 can be applied to (4.13). For the two other prototype problems (4.10) and (4.11), we can slightly modify the analysis for (3.17) to show that Theorem 4.2 is also valid for (4.10) and (4.11). Hence, we have proved (4.15a) and have only to show (4.15b). We derive from (4.8d) and (4.14) that

$$|\hat{\eta}_p - \hat{\eta}_{p,M}| = \frac{|c_o p|}{|g + \sigma |p|^2|} |\hat{u}_p(0) - \hat{u}_{p,M}(0)| \lesssim |p| \|\hat{u}_p - \hat{u}_{p,M}\|_{H^1(-a,0)}.$$

Hence, (4.15b) follows from the above and (4.15a). ∎

Denoting

$$u^{P,M}(x,y) := \sum_{p \in \Gamma'_P} \hat{u}_{p,M}(y) e^{ip\cdot x}, \quad \eta^{P,M}(x) := \sum_{p \in \Gamma'_P} \hat{\eta}_{p,M} e^{ip\cdot x}$$

as the Fourier–Legendre approximations of the solutions $\{u, \eta\}$ of (4.7), we can prove the following error estimate by using the same argument as in the proof of Theorem 3.3.

THEOREM 4.3. *Let $\{u, \eta\}$ be the solution of* (4.7) *and $\{u^{P,M}, \eta^{P,M}\}$ be its Fourier–Legendre Galerkin approximation. Then, if $u \in H^r(S_{a,0})$ and $\eta \in H^r(P(\Gamma))$, we have*

$$\begin{aligned} \left\| u - u^{P,M} \right\|_{H^1(S_{a,0})} &\lesssim \left(P^{1-r} + M^{1-r}\right) \|u\|_{H^r(S_{a,0})}, \\ \left\| \eta - \eta^{P,M} \right\|_{H^1(P(\Gamma))} &\lesssim \left(P^{1-r} + M^{1-r}\right) \left(\|u\|_{H^r(S_{a,0})} + \|\eta\|_{H^r(P(\Gamma))} \right). \end{aligned}$$

Theorem 4.3 immediately leads to the following corollary.

COROLLARY 4.1. *Let $\{U_n^{P,M}, \eta^{P,M}\}$ be the Fourier–Legendre Galerkin approximation of $\{U_n, \eta_n\}$ from* (4.4)*. Then, if $U_n \in H^r(S_{a,0})$ and $\eta_n \in H^r(P(\Gamma))$, we have*

$$\begin{aligned} \left\| U_n - U_n^{P,M} \right\|_{H^1(S_{a,0})} &\lesssim \left(P^{1-r} + M^{1-r}\right) \|U_n\|_{H^r(S_{a,0})}, \\ \left\| \eta_n - \eta_n^{P} \right\|_{H^1(P(\Gamma))} &\lesssim \left(P^{1-r} + M^{1-r}\right) \left(\|\eta_n\|_{H^r(P(\Gamma))} + \|U_n\|_{H^r(S_{a,0})} \right). \end{aligned}$$

This is, once again, the analogue of Proposition 2.2. Finally, we set

$$U^{N,P,M}(x,y) := \sum_{n=0}^{N} U_n^{P,M}(x,y)\varepsilon^n, \quad \eta^{N,P}(x) := \sum_{n=0}^{N} \eta_n^P(x,y)\varepsilon^n,$$

which are our approximations to the solutions $\{U, \eta\}$ of (4.3). With Theorem 4.1 we can easily modify the proof of Theorem 2.1 to realize the following error estimate.

THEOREM 4.4. *Under the assumptions of Theorem* 4.1*, for all $r \geq 2$ and $q > d-1$, we have*

$$\left\| \eta - \eta^{N,P} \right\|_{H^1(P(\Gamma))} \lesssim C_1(r,q) \left(\frac{(B\varepsilon)^{N+1}}{(N+1)^q} + P^{1-r} + M^{1-r} \right), \tag{4.16a}$$

$$\left\| U - U^{N,M,P} \right\|_{H^1(S_{a,0})} \lesssim C_1(r,q) \left(\frac{(B\varepsilon)^{N+1}}{(N+1)^q} + P^{1-r} + M^{1-r} \right) \tag{4.16b}$$

for all ε such that $B\varepsilon < 1$, where B and $C_1(r,q)$ are given in Theorem 4.1.

5. The Helmholtz equation. In this section we outline how the results of the previous sections can be extended to problems of linear acoustic scattering in two and three dimensions. More specifically, we consider the problem of time-harmonic linear acoustic waves scattered by periodic surfaces (see section 5.1) and bounded obstacles (see section 5.2). For each of these problems we state the TFE recursions and then, based upon analyticity and numerical analysis results, conclude the convergence of the resulting numerical scheme.

5.1. Periodic rough surface scattering. Consider the scattering of two-dimensional time-harmonic acoustic waves from a γ-periodic impenetrable rough surface $y = g(x)$, $g(x + \gamma) = g(x)$, which defines the exterior domain

$$\Omega_\infty := \{(x, y) \in \mathbf{R} \times \mathbf{R} \mid y > g(x)\}.$$

If an incident plane wave has the form

$$V_i(x, y) = e^{i(\alpha x - \beta y)},$$

then it is well-known [40] that the problem is governed by Helmholtz's equation with quasi-periodic lateral boundary conditions:

$$\Delta V + k^2 V = 0, \qquad (x, y) \in \Omega_\infty, \tag{5.1a}$$

$$V(x + \gamma, y) = e^{i\alpha\gamma} V(x, y), \tag{5.1b}$$

where V is the scattered field and $k^2 = \alpha^2 + \beta^2$. Depending upon the properties of the grating, one of several surface boundary conditions is suitable. A representative choice is a pressure release condition at the surface of the scatterer:

$$V(x, g(x)) = -V_i(x, g(x)) =: \zeta(x). \tag{5.1c}$$

To specify a unique solution, a condition is required as $y \to \infty$, and an "upward propagating condition" provides this for (5.1). Again, this type of condition can be enforced *exactly* in the near-field via a transparent boundary condition, in this case

$$\partial_y V(x, a) - S[V(x, a)] = 0 \tag{5.1d}$$

for any $a > |g|_{L^\infty}$ which defines the domain

$$\Omega_{g,a} := \{(x, y) \in \mathbf{R} \times \mathbf{R} \mid g(x) < y < a\}.$$

Here

$$S[\psi(x)] := \sum_{p=-\infty}^{\infty} (i\beta_p) \hat{\psi}_p e^{i\alpha_p x},$$

where we define

$$\alpha_p := \alpha + (2\pi/\gamma)p, \quad \beta_p := \sqrt{k^2 - \alpha_p^2}$$

(see [34, 35]). Note that we always choose β_p such that $\operatorname{Im}\{\beta_p\} \geq 0$.

The TFE approach to (5.1) is very similar to that presented in section 3.1 for Laplace's equation. If we modify the change of variables slightly from (3.8) to

$$x' = x, \quad y' = a \left(\frac{y - g(x)}{a - g(x)} \right),$$

we see that $\Omega_{g,a}$ is mapped to $\Omega_{0,a}$. Defining the *transformed* field

$$U(x',y') := V\left(x', \frac{(a-g(x'))y'}{a} + g(x')\right),$$

one can easily see [34, 35] that (5.1a)–(5.1d) changes to (upon dropping primes)

$$\Delta U + k^2 U = F(x,y;U,g), \qquad (x,y) \in \Omega_{0,a}, \tag{5.2a}$$
$$U(x,0) = \xi(x), \tag{5.2b}$$
$$\partial_y U(x,a) - S[U(x,a)] = J(x;U,g), \tag{5.2c}$$
$$U(x+\gamma,y) = e^{i\alpha\gamma}U(x,y), \tag{5.2d}$$

where the F and J are identical to those in (3.10) save the addition of terms involving k^2 to F, and the replacement of T by S in J (please see [35] for complete details).

Again, our boundary perturbation approach uses the strongly convergent expansions

$$U(x,y;\varepsilon) = \sum_{n=0}^{\infty} U_n(x,y)\varepsilon^n \tag{5.3}$$

to approximate U by

$$U^N(x,y;\varepsilon) = \sum_{n=0}^{N} U_n(x,y)\varepsilon^n.$$

As in section 3.1, it is straightforward to find equations for the U_n:

$$\Delta U_n + k^2 U_n = F_n(x,y), \qquad (x,y) \in \Omega_{0,a}, \tag{5.4a}$$
$$U_n(x,0) = \delta_{n,0}\xi(x), \tag{5.4b}$$
$$\partial_y U_n(x,a) - S[U_n(x,a)] = J_n(x), \tag{5.4c}$$
$$U_n(x+\gamma,y) = e^{i\alpha\gamma}U_n(x,y). \tag{5.4d}$$

Again, the forms for $\{F_n, J_n\}$ are similar to those given in section 3.1.

As with Laplace's equation, a variety of analyticity results exist which justify the expansions of U in ε [36, 37] (provided that k is not a Wood's anomaly; see [40]). However, for our numerical analysis we need that both the deformation f and Dirichlet data ξ lie in the Sobolev spaces H^s. This is, once again, a new result; however, the proof of the result is nearly identical to that of Theorem 3.1, so we present it without proof.

THEOREM 5.1. *Provided that k is not a Wood's anomaly, for any integer $s \geq 0$, if $\xi \in H^{s+3/2}(P(\Gamma))$ and $f \in H^{s+2}(P(\Gamma))$, then the expansion* (5.3) *converges strongly; i.e. there exist constants $C_1, C_2 > 0$ such that*

$$\|U_n\|_{H^{s+2}(\Omega_{0,a})} \leq C_1 \|\xi\|_{H^{s+3/2}(P(\Gamma))} B^n \tag{5.5}$$

for some $B > C_2 \|f\|_{H^{s+2}(P(\Gamma))}$.

At this point we can fully describe our numerical approach. In the same way that we outlined in section 3.2 for Laplace's equation, we make an approximation $U^{N,\bar{h}}$, cf. (3.15), to U. Again, the $U_n^{\bar{h}}$ satisfy a prototype problem, cf. (3.16), of Helmholtz type:

$$\Delta u(x,y) + k^2 u(x,y) = f(x,y), \qquad (x,y) \in \Omega_{0,a}, \tag{5.6a}$$

$$u(x,0) = b_1(x), \tag{5.6b}$$

$$\partial_y u(x,a) - S[u(x,a)] = b_2(x), \tag{5.6c}$$

$$u(x+\gamma, y) = e^{i\alpha\gamma} u(x,y). \tag{5.6d}$$

The quasi-periodic boundary conditions suggest the (generalized) Fourier expansion

$$u(x,y) = \sum_{p=-\infty}^{\infty} \hat{u}_p(y) e^{i\alpha_p x}$$

and likewise for f, b_1, and b_2. Inserting this expansion into (5.6), we find that $\hat{u}_p(y)$ satisfies the two-point boundary value problem

$$\partial_y^2 \hat{u}_p(y) + (k^2 - \alpha_p^2)\hat{u}_p(y) = \hat{f}_p(y), \qquad 0 < y < a, \tag{5.7a}$$

$$\hat{u}_p(0) = \hat{b}_{1,p}, \tag{5.7b}$$

$$\partial_y \hat{u}_p(a) - i\beta_p \hat{u}_p(a) = \hat{b}_{2,p}. \tag{5.7c}$$

We define

$${}^b\mathbb{P}_M = \text{span}\left\{ v \in \mathbb{P}_M \mid v(0) = b \right\},$$

where $\mathbb{P}_M$ is the space of all complex valued polynomials of degree less than or equal to M. Then, the Legendre–Galerkin approximation is as follows:

Find $\hat{u}_{p,M} \in {}^{\hat{b}_{1,p}}\mathbb{P}_M$ such that

$$\begin{aligned} -\int_0^a \partial_y \hat{u}_{p,M} \partial_y \bar{\phi}\, \mathrm{d}y + i\beta_p u_{p,M}(a)\bar{\phi}(a) + (k^2 - \alpha_p^2) \int_0^a \hat{u}_{p,M}\bar{\phi}\, \mathrm{d}y \\ = \int_0^a q_p \bar{\phi}\, \mathrm{d}y + \hat{b}_{2,p}\phi(a) \quad \forall \phi \in {}^0\mathbb{P}_M, \end{aligned} \tag{5.8}$$

where $\bar{\phi}$ is the complex conjugate of ϕ. With regards to this method we have the following error estimate (cf. Theorem 3.2 in [49]).

THEOREM 5.2. *If* $\hat{u}_p \in H^r(0,a)$ *with* $r \geq 1$, *we have*

$$\|\partial_y(\hat{u}_p - \hat{u}_{p,M})\|_{H^0(0,a)} + k\,\|\hat{u}_p - \hat{u}_{p,M}\|_{H^0(0,a)} \lesssim (1 + k^2 M^{-1}) M^{1-r}\, \|\hat{u}_p\|_{H^r(0,a)}.$$

Remark 5.1. Comparing to the estimate in Theorem 3.2 for (3.17), the above result contains an extra term $(1 + k^2 M^{-1})$ which is the pollution error associated with the Helmholtz equation (cf. [2]).

Letting

$$u^{P,M}(x,y) := \sum_{p=-P}^{P} \hat{u}_{p,M}(y) e^{i\alpha_p x}$$

be the Fourier–Legendre approximation of the solution u of (5.6) and using the same argument as in the proof of Theorem 3.3, we can prove the following estimate.

THEOREM 5.3. *If $u \in H^r(\Omega_{0,a})$ with an integer $r \geq 1$, then*

$$\left\|\nabla(u - u^{P,M})\right\|_{H^0(\Omega_{0,a})} + k\left\|u - u^{P,M}\right\|_{H^0(\Omega_{0,a})} \lesssim \left(P^{1-r} + (1 + k^2M^{-1})M^{1-r}\right)\|u\|_{H^r(\Omega_{0,a})}.$$

Finally, we can set

$$U^{N,P,M}(x,y) := \sum_{n=0}^{N} U_n^{P,M}(x,y)\varepsilon^n$$

which is our approximation to the solution U of (5.1). Using Theorem 5.1 and applying the proof of Theorem 2.1, we have the next estimate.

THEOREM 5.4. *For a given integer $r \geq 2$, if $f \in H^r(\Omega_{0,a})$ and $\xi \in H^{r-1/2}(P(\Gamma))$, then we have*

$$\begin{aligned}\left\|\nabla(U - U^{N,P,M})\right\|_{H^0(\Omega_{0,a})} &+ k\left\|U - U^{N,P,M}\right\|_{H^0(\Omega_{0,a})}\\ &\lesssim (B\varepsilon)^{N+1} + \left(P^{1-r} + (1 + k^2M^{-1}M^{1-r})\right)\|\xi\|_{H^{r-1/2}(P(\Gamma))}\end{aligned}$$

for any constant $B \geq C_2\|f\|_{H^r(\Omega)}$ such that $B\varepsilon < 1$, where C_2 is the constant in Theorem 5.1.

Remark 5.2. Before proceeding, we note that the analysis is exactly the same if the scattering surface is two-dimensional (so that the problem domain is a subset of $\mathbf{R}^3$) and biperiodic. The only modifications of importance are that the incident radiation has the form

$$V_i(x,y,z) = e^{i(\alpha_1 x_1 + \alpha_2 x_2 - \beta y)},$$

$k^2 = \alpha_1^2 + \alpha_2^2 + \beta^2$, and the x-derivatives must be vectorized. Otherwise, all of the other formulas and estimates remain identical, including Theorem 5.4

5.2. Bounded-obstacle scattering. Before leaving this section on the TFE algorithm as applied to scattering configurations, we point out that this algorithm has been adapted to the problem of scattering by a bounded obstacle [38, 19]. Here we consider a (two-dimensional) domain shaped as

$$\Omega_\infty := \{(r,\theta) \in (0,\infty) \times [0, 2\pi) \mid r > a + g(\theta)\}$$

incident by the two-dimensional field

$$V_i = e^{i(\alpha x - \beta y)}.$$

Again, Helmholtz's equation governs the scattered field [14], coupled to a Dirichlet boundary condition at the obstacle (for a pressure release surface) and periodic boundary conditions:

$$\Delta V + k^2 V = 0, \qquad (r,\theta) \in \Omega_\infty, \tag{5.9a}$$

$$V(a + g(\theta), \theta) = -\left.V_i\right|_{r=a+g} =: \zeta(\theta), \tag{5.9b}$$

$$V(r, \theta + 2\pi) = V(r,\theta). \tag{5.9c}$$

Here a Sommerfeld radiation condition must be enforced to ensure a unique solution, and, for this, we, once again, have a transparent boundary condition [29, 38]

$$\partial_r V(b,\theta) - R[V(b,\theta)] = 0 \tag{5.9d}$$

for any $b > a + |g|_{L^\infty}$ which defines the domain

$$\Omega_{a+g,b} := \{(r,\theta) \in (0,\infty) \times [0,2\pi) \mid a + g(\theta) < r < b\}.$$

Here

$$R[\psi(\theta)] := \sum_{p=-\infty}^{\infty} \frac{d_z H_p^{(1)}(kb)}{H_p^{(1)}(kb)} \hat{\psi}_p e^{ip\theta},$$

where $H_p^{(1)}(z)$ is the pth Hankel function of the first kind and $d_z H_p^{(1)}(z)$ is its first derivative (see [29, 38]).

The TFE approach to (5.9) is very similar to what is presented above; if we now consider the change of variables

$$r' = \frac{(b-a)r - bg(\theta)}{(b-a) - g(\theta)}, \quad \theta' = \theta,$$

then $\Omega_{a+g,b}$ is mapped to $\Omega_{a,b}$. Defining the *transformed* field

$$U(r',\theta') := V\left(r' + \frac{g(\theta')(b-r')}{b-a}, \theta'\right),$$

one can deduce [29, 38] that (5.9) changes to (upon dropping primes)

$$\Delta U + k^2 U = F(r,\theta;U,g), \qquad (r,\theta) \in \Omega_{a,b}, \tag{5.10a}$$
$$U(a,\theta) = \xi(\theta), \tag{5.10b}$$
$$\partial_r U(b,\theta) - R[U(b,\theta)] = J(\theta;U,g), \tag{5.10c}$$
$$U(r,\theta+2\pi) = U(r,\theta), \tag{5.10d}$$

where the F and J are presented in detail in [38].

Using the expansion

$$U(r,\theta;\varepsilon) = \sum_{n=0}^{\infty} U_n(r,\theta)\varepsilon^n, \tag{5.11}$$

our boundary perturbation method seeks to approximate U by

$$U^N(r,\theta;\varepsilon) = \sum_{n=0}^{N} U_n(r,\theta)\varepsilon^n.$$

As above, equations for the U_n are given by

$$\Delta U_n + k^2 U_n = F_n(r,\theta), \qquad (r,\theta) \in \Omega_{a,b}, \tag{5.12a}$$
$$U_n(a,\theta) = \delta_{n,0}\xi(\theta), \tag{5.12b}$$
$$\partial_r U_n(b,\theta) - R[U_n(b,\theta)] = J_n(\theta), \tag{5.12c}$$
$$U_n(r,\theta+2\pi) = U_n(x,\theta), \tag{5.12d}$$

and the precise forms for $\{F_n, J_n,\}$ are listed in [38].

Again, analyticity results exist which justify the expansion of U in ε [29], and again they do not feature function spaces which are convenient for our analysis. However, these results do inspire the following result which is what we require; again, the proof is quite similar to that presented for Theorem 3.1.

THEOREM 5.5. *For any integer $s \geq 0$, if $\xi \in H^{s+3/2}(0,2\pi)$ and $f \in H^{s+2}(0,2\pi)$, then the expansion* (5.11) *converges strongly; i.e. there exist constants $C_1, C_2 > 0$ such that*

$$\|U_n\|_{H^{s+2}(\Omega_{a,b})} \leq C_1 \|\xi\|_{H^{s+3/2}(0,2\pi)} B^n \tag{5.13}$$

for some $B > C_2 \|f\|_{H^{s+2}(0,2\pi)}$.

We now discuss how the developments of section 3.2 can be adapted to the problem of bounded-obstacle scattering. The prototype problem of Helmholtz type in this setting is

$$\Delta u + k^2 u = f(r,\theta), \qquad (r,\theta) \in \Omega_{a,b}, \tag{5.14a}$$

$$u(a,\theta) = b_1(\theta), \tag{5.14b}$$

$$\partial_r u(b,\theta) - R[u(b,\theta)] = b_2(\theta), \tag{5.14c}$$

$$u(r,\theta + 2\pi) = u(r,\theta). \tag{5.14d}$$

The periodic boundary conditions suggest

$$u(r,\theta) = \sum_{p=-\infty}^{\infty} \hat{u}_p(r) e^{ip\theta}$$

and likewise for f, b_1, and b_2. Inserting these expansions into (5.14), we find that $\hat{u}_p(r)$ satisfies the two-point boundary value problem

$$[r\partial_r]^2 \hat{u}_p(r) + (r^2k^2 - p^2)\hat{u}_p(r) = \hat{f}_p(r), \qquad a < r < b, \tag{5.15a}$$

$$\hat{u}_p(a) = \hat{b}_{1,p}, \tag{5.15b}$$

$$\partial_r \hat{u}_p(b) - \left(\frac{d_z H_p^{(1)}(kb)}{H_p^{(1)}(kb)} \right) \hat{u}_p(b) = \hat{b}_{2,p}. \tag{5.15c}$$

The Legendre–Galerkin approximation for (5.15) and the Fourier Legendre–Galerkin approximation for (5.14) have been studied recently in [50]. We now summarize the corresponding results below. Let $\hat{u}_{p,M}$ be the Legendre–Galerkin approximation of $\hat{u}_p$ in (5.15); then

$$u^{P,M}(r,\theta) := \sum_{|p|\leq P} \hat{u}_{p,M}(r) e^{ip\theta}$$

is the natural Fourier–Legendre approximation of the solution u of (5.14). Then, we have the following error estimate (cf. Theorem 4.3 in [50]).

THEOREM 5.6. *If $u \in H^r(\Omega_{0,a})$ with an integer $r \geq 1$, then*

$$\left\|\nabla(u - u^{P,M})\right\|_{H^0(\Omega_{a,b})} + k\left\|u - u^{P,M}\right\|_{H^0(\Omega_{a,b})} \lesssim \left((1 + KP^{-1})P^{1-r} + k^{\frac{1}{3}}(1 + k^2M^{-1})M^{1-r} \right) \|u\|_{H^r(\Omega_{a,b})}.$$

Setting

$$U^{N,P,M}(r,\theta) := \sum_{n=0}^{N} U_n^{P,M}(r,\theta)\varepsilon^n,$$

using Theorem 5.5, and applying the proof of Theorem 2.1, we have the final estimate.

THEOREM 5.7. *Given an integer $r \geq 2$, if $f \in H^r(0,2\pi)$ and $\xi \in H^{r-1/2}(0,2\pi)$, then*

$$\left\|U - U^{N,P,M}\right\|_{H^1(\Omega_{a,b})} + k\left\|U - U^{N,P,M}\right\|_{H^0(\Omega_{a,b})} \\ \lesssim (B\varepsilon)^{N+1} + \left((1+KP^{-1})P^{1-r} + k^{\frac{1}{3}}(1+k^2M^{-1})M^{1-r}\right)\|\xi\|_{H^{r-1/2}(0,2\pi)}$$

for any constant $B \geq C_2\|f\|_{H^r(0,2\pi)}$ such that $B\varepsilon < 1$, where C_2 is the constant in Theorem 5.5.

Remark 5.3. For bounded-obstacle scattering in three dimensions, the results corresponding to Theorems 5.5 and 5.6 were proved in [30] and [50], respectively. Therefore, a result similar to Theorem 5.7 can be established. For a detailed description of the algorithm and numerical implementation, we refer the interested reader to [19].

6. Conclusions. In this paper we have provided a rigorous numerical analysis for a large class of boundary value and free boundary problems which can be treated with the stable and high-order TFE method. The examples have ranged from acoustics to fluid mechanics, but the particular details of these applications are largely immaterial, and it is easy to see that these methods (and our numerical analysis) can be extended to a much wider class of problems in a relatively straightforward manner. Our approach to establishing the convergence and accuracy of the TFE methodology is to combine analyticity theorems extended to the most convenient Sobolev spaces (presented here for the first time) with results on Legendre–Galerkin methods applied to elliptic boundary value problems. The flexibility of our approach (decoupling the boundary perturbation expansion from the Galerkin approximation) indicates that similar results can be established for other discretization schemes (e.g., Chebyshev–Galerkin methods or finite-element discretizations).

REFERENCES

[1] R. A. Adams, *Sobolev Spaces*, Pure Appl. Math. 65, Academic Press, New York-London, 1975.
[2] I. Babuška and S. A. Sauter, *Is the pollution effect of the FEM avoidable for the Helmholtz equation considering high wave numbers?*, SIAM J. Numer. Anal., 34 (1997), pp. 2392–2423.
[3] G. Ben-Yu and W. Li-Lian, *Jacobi interpolation approximations and their applications to singular differential equations*, Adv. Comput. Math., 14 (2001), pp. 227–276.
[4] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers.* I, Asymptotic Methods and Perturbation Theory, Springer-Verlag, New York, 1999.
[5] C. Bernardi and Y. Maday, *Approximations Spectrales de Problèmes aux Limites Elliptiques*, Springer-Verlag, Paris, 1992.
[6] O. P. Bruno and F. Reitich, *Solution of a boundary value problem for the Helmholtz equation via variation of the boundary into the complex domain*, Proc. Roy. Soc. Edinburgh Sect. A, 122 (1992), pp. 317–340.
[7] O. P. Bruno and F. Reitich, *Numerical solution of diffraction problems: A method of variation of boundaries*, J. Opt. Soc. Amer. A, 10 (1993), pp. 1168–1175.
[8] O. P. Bruno and F. Reitich, *Numerical solution of diffraction problems: A method of variation of boundaries. II. Finitely conducting gratings, Padé approximants, and singularities*, J. Opt. Soc. Amer. A, 10 (1993), pp. 2307–2316.
[9] O. P. Bruno and F. Reitich, *Numerical solution of diffraction problems: A method of variation of boundaries. III. Doubly periodic gratings*, J. Opt. Soc. Amer. A, 10 (1993), pp. 2551–2562.
[10] O. P. Bruno and F. Reitich, *Calculation of electromagnetic scattering via boundary variations and analytic continuation*, Appl. Comput. Electromagn. Soc. J., 11 (1996), pp. 17–31.

[11] O. P. Bruno and F. Reitich, *Boundary-variation solutions for bounded-obstacle scattering problems in three dimensions*, J. Acoust. Soc. Amer., 104 (1998), pp. 2579–2583.
[12] O. P. Bruno and F. Reitich, *High-order boundary perturbation methods*, in Mathematical Modeling in Optical Science, Frontiers Appl. Math. 22, G. Bao, L. Cowsar, and W. Masters, eds., SIAM, Philadelphia, PA, 2001, pp. 71–109.
[13] J. Chandezon, D. Maystre, and G. Raoult, *A new theoretical method for diffraction gratings and its numerical application*, J. Opt., 11 (1980), pp. 235–241.
[14] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, 2nd ed., Springer-Verlag, Berlin, 1998.
[15] W. Craig and D. P. Nicholls, *Traveling two and three dimensional capillary gravity water waves*, SIAM J. Math. Anal., 32 (2000), pp. 323–359.
[16] W. Craig and D. P. Nicholls, *Traveling gravity water waves in two and three dimensions*, Eur. J. Mech. B Fluids, 21 (2002), pp. 615–641.
[17] W. Craig and C. Sulem, *Numerical simulation of gravity waves*, J. Comput. Phys., 108 (1993), pp. 73–83.
[18] L. C. Evans, *Partial Differential Equations*, Grad. Stud. Math. 19, American Mathematical Society, Providence, RI, 1998.
[19] Q. Fang, D. P. Nicholls, and J. Shen, *A stable, high-order method for three-dimensional bounded-obstacle scattering*, J. Comput. Phys., 224 (2007), pp. 1145–1169.
[20] B. Hu and D. P. Nicholls, *Analyticity of Dirichlet–Neumann operators on Hölder and Lipschitz domains*, SIAM J. Math. Anal., 37 (2005), pp. 302–320.
[21] G. Iooss and P. Plotnikov, *Existence of multimodal standing gravity waves*, J. Math. Fluid Mech., 7 (2005), pp. S349–S364.
[22] G. Iooss and P. Plotnikov, *Multimodal standing gravity waves: A completely resonant system*, J. Math. Fluid Mech., 7 (2005), pp. S110–S126.
[23] G. Iooss, P. Plotnikov, and J. Toland, *Standing waves on infinite depth*, C. R. Math. Acad. Sci. Paris, 338 (2004), pp. 425–431.
[24] G. Iooss, P. I. Plotnikov, and J. F. Toland, *Standing waves on an infinitely deep perfect fluid under gravity*, Arch. Ration. Mech. Anal., 177 (2005), pp. 367–478.
[25] T. Kato, *Perturbation Theory for Linear Operators*, Classics Math., Springer-Verlag, Berlin, 1995.
[26] O. A. Ladyzhenskaya and N. N. Ural'tseva, *Linear and quasilinear elliptic equations*, Academic Press, New York, 1968.
[27] H. Lamb, *Hydrodynamics*, 6th ed., Cambridge University Press, Cambridge, 1993.
[28] D. P. Nicholls, *Boundary perturbation methods for water waves*, GAMM Mitt. Ges. Angew. Math. Mech., 30 (2007), pp. 44–74.
[29] D. P. Nicholls and N. Nigam, *Exact non-reflecting boundary conditions on general domains*, J. Comput. Phys., 194 (2004), pp. 278–303.
[30] D. P. Nicholls and N. Nigam, *Error analysis of a coupled finite element/DtN map algorithm on general domains*, Numer. Math., 105 (2006), pp. 267–298.
[31] D. P. Nicholls and F. Reitich, *A new approach to analyticity of Dirichlet-Neumann operators*, Proc. Roy. Soc. Edinburgh Sect. A, 131 (2001), pp. 1411–1433.
[32] D. P. Nicholls and F. Reitich, *Stability of high-order perturbative methods for the computation of Dirichlet-Neumann operators*, J. Comput. Phys., 170 (2001), pp. 276–298.
[33] D. P. Nicholls and F. Reitich, *Analytic continuation of Dirichlet-Neumann operators*, Numer. Math., 94 (2003), pp. 107–146.
[34] D. P. Nicholls and F. Reitich, *Shape deformations in rough surface scattering: Cancellations, conditioning, and convergence*, J. Opt. Soc. Amer. A, 21 (2004), pp. 590–605.
[35] D. P. Nicholls and F. Reitich, *Shape deformations in rough surface scattering: Improved algorithms*, J. Opt. Soc. Amer. A, 21 (2004), pp. 606–621.
[36] D. P. Nicholls and F. Reitich, *On analyticity of traveling water waves*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 461 (2005), pp. 1283–1309.
[37] D. P. Nicholls and F. Reitich, *Rapid, stable, high-order computation of traveling water waves in three dimensions*, Eur. J. Mech. B Fluids, 25 (2006), pp. 406–424.
[38] D. P. Nicholls and J. Shen, *A stable, high-order method for two-dimensional bounded-obstacle scattering*, SIAM J. Sci. Comput., 28 (2006), pp. 1398–1419.
[39] D. P. Nicholls and M. Taber, *Joint analyticity and analytic continuation for Dirichlet–Neumann operators on doubly perturbed domains*, J. Math. Fluid Mech., 10 (2008), pp. 238–271.
[40] R. Petit, ed., *Electromagnetic Theory of Gratings*, Springer-Verlag, Berlin, 1980.
[41] N. A. Phillips, *A coordinate system having some special advantages for numerical forecasting*, J. Atmospheric Sci., 14 (1957), pp. 184–185.

[42] P. Plotnikov, *Solvability of the problem of spatial gravitational waves on the surface of an ideal fluid*, Soviet Phys. Doklady, 25 (1980), pp. 170–171.
[43] L. Rayleigh, *On the dynamical theory of gratings*, Proc. R. Soc. London Ser. A Math. Phys. Eng. Sci., 79 (1907), pp. 399–416.
[44] J. Reeder and M. Shinbrot, *On Wilton ripples, I: Formal derivation of the phenomenon*, Wave Motion, 3 (1981), pp. 115–135.
[45] J. Reeder and M. Shinbrot, *On Wilton ripples, II: Rigorous results*, Arch. Rat. Mech. Anal., 77 (1981), pp. 321–347.
[46] S. O. Rice, *Reflection of electromagnetic waves from slightly rough surfaces*, Comm. Pure Appl. Math., 4 (1951), pp. 351–378.
[47] J. Shen, *Efficient spectral-Galerkin method I. Direct solvers of second- and fourth-order equations using Legendre polynomials*, SIAM J. Sci. Comput., 15 (1994), pp. 1489–1505.
[48] J. Shen, *Efficient spectral-Galerkin method II. Direct solvers of second- and fourth-order equations using Chebyshev polynomials*, SIAM J. Sci. Comput., 16 (1995), pp. 74–87.
[49] J. Shen and L.-L. Wang, *Spectral approximation of the Helmholtz equation with high wave numbers*, SIAM J. Numer. Anal., 43 (2005), pp. 623–644.
[50] J. Shen and L.-L. Wang, *Analysis of a spectral-Galerkin approximation to the Helmholtz equation in exterior domains*, SIAM J. Numer. Anal., 45 (2007), pp. 1954–1978.
[51] G. Stokes, *On the theory of oscillatory waves*, Trans. Cambridge Philos. Soc., 8 (1847), pp. 441–473.
[52] V. Zakharov, *Stability of periodic waves of finite amplitude on the surface of a deep fluid*, J. Appl. Mech. Tech. Phys., 9 (1968), pp. 190–194.