

High order methods to simulate scattering problems in locally perturbed periodic waveguides

Ruming Zhang

CRC Preprint 2021/25, May 2021

KARLSRUHE INSTITUTE OF TECHNOLOGY

CRC 1173



Wave
phenomena

Participating universities



Universität Stuttgart

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN



Funded by

DFG

High order methods to simulate scattering problems in locally perturbed periodic waveguides

Ruming Zhang*

Abstract

In this paper, two high order numerical methods, the CCI method and the decomposition method, are proposed to simulate wave propagation in locally perturbed periodic closed waveguides. As is well known the problem is not always uniquely solvable due to the existence of guided modes, the limiting absorption principle is a standard way to get the unique physical solution. Both methods are based on the Floquet-Bloch transform which transforms the original problem to an equivalent family of cell problems. The CCI method is based on a modification of integral contours of the inverse transform, and the decomposition method comes from an explicit definition of the radiation condition. Due to the local perturbation, the family of cell problems are coupled thus the whole system is actually defined in 3D. Based on different types of singularities, high order methods are developed for faster convergence rates. Finally we show the convergence results by both theoretical explanations and numerical examples.

Keywords: periodic waveguide, Floquet-Bloch transform, high order method, finite element method

1 Introduction

Periodic structures are widely used in applications such as photonic crystals, for details we refer to [13, 14, 35]. This topic also attracts the interests of many mathematicians and we refer to [3, 22, 23] for the studies from mathematical point of view. It is well known that this kind of problems is challenging due to existence of guided modes. To obtain the unique physical solution, the *Limiting Absorption Principle (LAP)* is a standard process. The LAP process is to define the physical solution by the limit of unique solutions with absorption, as the absorption parameter tends to zero. For simplicity, in this paper we call the solution from the LAP process an LAP solution. Significant progresses have been made in the past few years in the study of this kind of problems, from both theoretical and numerical point of view. For example, in [12] with an analysis on the resolvent of the differential operator, a radiation condition was given for LAP solution in a periodic half guide, and in [10] the authors gave the radiation condition as well as semi-analytic representations for LAP solutions in the full guide. On the other hand, with the singular perturbation theory (see [1]), the radiation condition is given by authors in [19]. With this method, radiation conditions for LAP solutions are also developed for periodic layers in 2D spaces and periodic open tubes in 3D spaces, see [17–19]. Besides the LAP, a Kondrat'ev's weighted spaces based method was adopted by S. A. Nazarov in [29] and further works were carried out by him and his collaborators in [30–34]. On the other hand, numerical methods are also developed to compute the LAP solutions. For example, in [15] an algorithm was proposed to compute exact Dirichlet-to-Neumann maps from the LAP process and this method was extended to periodic structures with local perturbation [6–9]. A doubling recursive procedure based method with an extrapolation technique was also developed to compute the DtN maps, see [4, 5, 36]. With a decomposition of Bloch waves, a numerical method was proposed for waveguides with different refractive indexes on both directions in [2].

In this paper, we aim to develop two high order numerical methods to simulate wave traveling in locally perturbed periodic waveguides. For both methods, the Floquet-Bloch transform is the key to transform

*Institute of Applied and Numerical mathematics, Karlsruhe Institute of Technology, Karlsruhe, Germany ; ruming.zhang@kit.edu.

the problem defined in 2D unbounded domain to an equivalent coupled family of cell problems. The idea comes from some older papers of the author and A. Lechleiter for (locally perturbed) periodic surfaces see [24, 26–28, 38]. However, the problems discussed in the above papers are relatively easier since there are no guided modes and the problems are always uniquely solvable. For the problems discussed in this paper, a radiation condition is introduced to describe the LAP solutions. The first (complex contour integral/CCI) method is based on the author’s previous work on purely periodic waveguides from both theoretical (see [40]) and numerical (see [39]) point of view. The second (decomposition) method comes from an explicit characterisation of the radiation condition proposed in [17]. For both methods, analytic formulations for LAP solutions are given as extensions of periodic cases. Due to the local perturbation, the whole system is coupled thus the formulations are actually defined in 3D. This makes the computational complexity much larger than periodic problems thus high order methods are necessary. Based on different types of singularities, we develop different high order algorithms for both cases. Finally, we show that the numerical converges super-algebraically.

The rest of this paper is organized as follows. In the second section, the mathematical model is given and two explicit formulations via the Floquet-Bloch transform are given in the third section. In Section 4, different numerical schemes are developed, and numerical examples are shown in Section 5.

2 Mathematical model

Let the closed waveguide $\Omega := \mathbb{R} \times (0, 1)$ with the upper and lower boundaries $\Sigma_- := \mathbb{R} \times \{0\}$, $\Sigma_+ := \mathbb{R} \times \{1\}$. The scattering problem in Ω is formulated by the following equations:

$$\Delta u + k^2(n + q)u = f \text{ in } \Omega; \quad \frac{\partial u}{\partial x_2} = 0 \text{ on } \Sigma_{\pm}. \quad (1)$$

Here n is 1-periodic in x_1 -direction, both f and q are compactly supported. Moreover both n and $n + q$ are strictly positive, i.e., there is a constant $c > 0$ such that

$$n(x) \geq c > 0; \quad n(x) + q(x) \geq c > 0 \quad \text{for all } x \in \Omega.$$

For convenience, we define the following periodicity cells and their boundaries:

$$\begin{aligned} \Omega_j &:= \left(j - \frac{1}{2}, j + \frac{1}{2}\right) \times (0, 1); & \Gamma_j &= \left\{j - \frac{1}{2}\right\} \times (0, 1); \\ \Sigma_-^j &= \left(j - \frac{1}{2}, j + \frac{1}{2}\right) \times \{0\}; & \Sigma_+^j &= \left(j - \frac{1}{2}, j + \frac{1}{2}\right) \times \{1\}. \end{aligned}$$

Thus $\partial\Omega_j = \Gamma_j \cup \Sigma_-^j \cup \Gamma_{j+1} \cup \Sigma_+^j$. For simplicity, we assume that both f and q are supported in Ω_0 . For visualization of the locally perturbed periodic waveguide we refer to Figure 1.

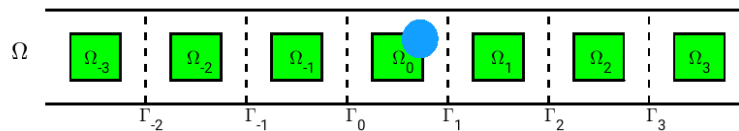


Figure 1: Periodic waveguide with local perturbation (blue disk).

As is well known, the problem (1) is not always uniquely solvable for positive valued wave number k . To carry out the LAP, we first consider the damped problem, by replacing k^2 with $k^2 + i\varepsilon$ where $\varepsilon > 0$. It is well known that the damped problem is uniquely solvable in $H^1(\Omega)$. Then the limit of solution when $\varepsilon \rightarrow 0^+$ is set to be the physical solution, which is called an LAP solution in this paper. From [10, 19, 40], the radiation condition for LAP solutions in periodic waveguides have been described in different forms, and the forms are actually equivalent.

Definition 1 (Radiation Condition). *Suppose for the positive valued periodic refractive index n and wavenumber $k > 0$, there are no standing waves. Then an LAP solution for (1) satisfies the following radiation conditions:*

$$\begin{aligned} u(x) &= u_+(x) + \sum_{\ell \in L_+} a_\ell^+ \varphi_\ell^+(x), \quad x_1 > 1/2; \\ u(x) &= u_-(x) + \sum_{\ell \in L_-} a_\ell^- \varphi_\ell^-(x), \quad x_1 < -1/2; \end{aligned}$$

where u_\pm decays exponentially when $x_1 \rightarrow \pm\infty$, φ_ℓ^+ (φ_ℓ^-) are propagating modes travelling to the right (left). L_+ and L_- are finite set of indexes for left or right propagating modes and $a_\ell^\pm \in \mathbb{C}$ are coefficients.

For definitions of standing waves and propagating modes we refer to Section 3.1 for details. The explicit formulation of LAP solutions plays the crucial role in development of numerical methods. In this paper, we show two different ways to develop different numerical schemes. For convenience, we first define a subset $H_{LAP}(\Omega) \subset H_{loc}^1(\Omega)$ which contains all the functions that satisfies the radiation condition.

Multiply the equation (1) with any compactly supported smooth function $\bar{\varphi}$ on both sides. From Green's identity, the following equation holds:

$$\int_{\Omega} [\nabla u \cdot \nabla \bar{\varphi} - k^2(n+q)u\bar{\varphi}] \, dx = - \int_{\Omega} f\bar{\varphi} \, dx.$$

From Riesz representation theorem, there is an operator $B : H_{loc}^1(\Omega) \rightarrow (H^1(\Omega))^*$ such that

$$\int_{\Omega} \nabla u \cdot \nabla \bar{\varphi} \, dx = \langle Bu, \varphi \rangle_{n+q, L^2},$$

where $\langle \cdot, \cdot \rangle_{n+q, L^2}$ is the weighted L^2 -norm in Ω with norm $n+q$. As $n+q$ is a strictly positive and bounded function in Ω , this norm is equivalent to the normally defined L^2 -norm. Then

$$\langle (B - k^2 I)u, \varphi \rangle_{n+q, L^2} = \int_{\Omega} [\nabla u \cdot \nabla \bar{\varphi} - k^2(n+q)u\bar{\varphi}] \, dx,$$

where I is the identity operator.

Assumption 2. *We assume that k^2 does not lie in the spectrum of the operator B in the space $H_{LAP}(\Omega)$.*

Remark 3. *For perfectly periodic waveguide, an important result is that Assumption 2 always holds (see [10, 19]). However, this case may happen for a locally perturbed periodic waveguide.*

Now we show an example of such a $k > 0$ that the lies in the spectrum of B for positive n and $n+q$. Suppose n and f are defined by:

$$n(x) = \begin{cases} 1, & |x - a_0| > 0.3; \\ 9, & 0.1 < |x - a_0| < 0.3; \\ 1 + 8\zeta(|x - a_0|; 0.1, 0.3), & \text{otherwise.} \end{cases} \quad ; \quad f(x) = \begin{cases} 0, & |x - a_0| > 0.3; \\ 0.5, & 0.1 < |x - a_0| < 0.3; \\ 0.5\zeta(|x - a_0|; 0.1, 0.3), & \text{otherwise;} \end{cases}$$

where $a_0 = (0, 0.5)^\top$, and $\zeta(t)$ is a C^8 -continuous function defined by

$$\zeta(t; a, b) = \begin{cases} 1, & t \leq a; \\ 0, & t \geq b; \\ 1 - \left[\int_{\tau=a}^b (\tau - a)^4 (\tau - b)^4 \, d\tau \right]^{-1} \left[\int_{\tau=a}^t (\tau - a)^4 (\tau - b)^4 \, d\tau \right], & a < t < b. \end{cases}$$

When $k^2 = 3.2$, the problem (1) with $q = 0$ is uniquely solvable in $H^1(\Omega)$. The solution u decays exponentially when $|x_1| \rightarrow \infty$ thus the radiation condition defined in Definition 13 is satisfied. The solution in Ω_0 is shown in (a) in Figure 2, which is strictly positive in $\bar{\Omega}_0$. Let $q = -k^{-2}f/u$ (see (b) in Figure 2), then it is the constructed local perturbation. As $n+q$ is strictly positive (see (c) in Figure 2), u is the solution satisfies (1) with $f = 0$ and satisfies the radiation condition. Thus $k^2 = 3.2$ is an eigenvalue of B in $H_{LAP}(\Omega)$.

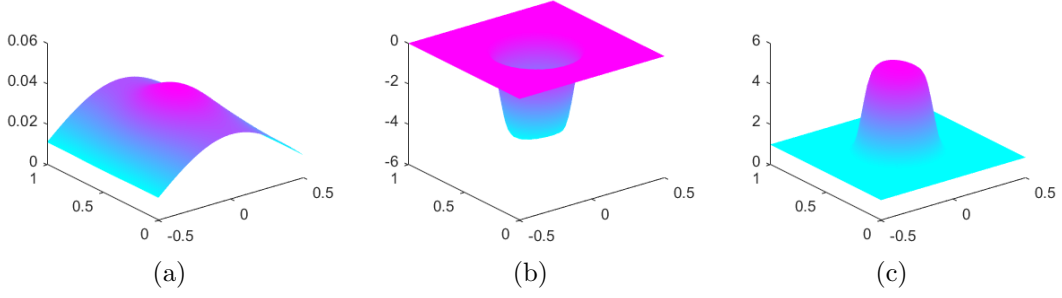


Figure 2: (a): numerical solution in Ω_0 ; (b): the constructed local perturbation q ; (c): the function $n + q$.

3 Explicit formulations of LAP solutions

In this section, we formulate the LAP solutions by two different methods introduced in [39] and [19], respectively. Note that the equation (1) can be rewritten as:

$$\Delta u + k^2 n u = r \text{ in } \Omega, \quad \frac{\partial u}{\partial x_2} = 0 \text{ on } \partial \Sigma_{\pm} \quad (2)$$

where $r = f - k^2 q u$ is compactly supported in Ω_0 . In this section, we extend the formulations of LAP solutions with purely periodic refractive indexes to periodic ones with local perturbations, and also prove the unique solvability of the formulations.

3.1 The quasi-periodic problems

From the Floquet theory, the quasi-periodic problems (especially corresponding eigenvalues and eigenfunctions) particularly interesting as they are related to the propagating modes. A function φ is α -quasi-periodic in Ω_0 if it satisfies $\varphi(x_1 + 1, x_2) = e^{i\alpha} \varphi(x)$ when $x \in \Omega$. The strong formulation for the quasi-periodic problem is, to find an α -quasi-periodic solution such that:

$$\Delta w + k^2 n w = r \text{ in } \Omega_0; \quad \frac{\partial w}{\partial x_2} = 0 \text{ on } \Sigma_{\pm}^0. \quad (3)$$

It is more convenient to consider the family of α -dependent problems in one fixed space, thus let $v := e^{-i\alpha x_1} w$, then v is 0-quasi-periodic, i.e., it is periodic in Ω_0 . The strong formulation for v is:

$$\Delta v + 2i\alpha \frac{\partial v}{\partial x_1} + (k^2 n - \alpha^2) v = e^{-i\alpha x_1} r(x) \text{ in } \Omega_0; \quad \frac{\partial v}{\partial x_2} = 0 \text{ on } \Sigma_{\pm}^0. \quad (4)$$

For each α , the weak formulation of the periodic problems is, to find $v \in H_{per}^1(\Omega_0)$ such that

$$\int_{\Omega_0} \left[\nabla v \cdot \nabla \bar{\psi} - i\alpha \left(\frac{\partial v}{\partial x_1} \bar{\psi} - v \frac{\partial \bar{\psi}}{\partial x_1} \right) - (k^2 n - \alpha^2) v \bar{\psi} \right] dx = - \int_{\Omega_0} e^{-i\alpha x_1} r \bar{\psi} dx \quad (5)$$

for any $\psi \in H_{per}^1(\Omega_0)$. From Riesz representation theorem, there is an operator $A(\alpha, k) : H_{per}^1(\Omega_0) \rightarrow H_{per}^1(\Omega_0)$ such that

$$\langle A(\alpha, k) w, \psi \rangle = \int_{\Omega_0} \left[\nabla w \cdot \nabla \bar{\psi} - i\alpha \left(\frac{\partial w}{\partial x_1} \bar{\psi} - w \frac{\partial \bar{\psi}}{\partial x_1} \right) - (k^2 n - \alpha^2) w \bar{\psi} \right] dx,$$

where $\langle \cdot, \cdot \rangle$ is the inner product in the space $H_{per}^1(\Omega_0)$. It is obvious that $A(\alpha, k)$ is a Fredholm operator. When $k > 0$ and $\alpha \in \mathbb{R}$, $A(\alpha, k)$ is a self-adjoint operator. There are operators A_1, A_2, A_3, A_4 which do not depend on α and k with obvious definitions such that

$$A(\alpha, k) = A_1 + \alpha A_2 + \alpha^2 A_3 + k^2 A_4. \quad (6)$$

Thus $A(\alpha, k)$ depends analytically on both α and k . From [19], for fixed k^2 , there are only finitely many α 's in $[-\pi, \pi]$ such that $A(\alpha, k)$ is not invertible, which are called exceptional values. The set of exceptional values is denoted by $S(k)$. They are solutions of the quadratic eigenvalue problems for fixed k , thus can be computed in a standard way (for details we refer to Section 4.2). In the following, we introduce some definitions and notations concerning exceptional values without proofs. For details we refer to [4, 5, 10].

For any $\alpha \in [-\pi, \pi]$, there is a family of analytic functions $\mu_i(\alpha)$ defined in $[-\pi, \pi]$ such that:

$$(A_1 + \alpha A_2 + \alpha^2 A_3)\varphi = -\mu_i(\alpha)A_4.$$

For 2D periodic waveguide, none of the analytic functions is constant (see [10]). For any index i , the graph of the function μ_i , i.e., $\{(\alpha, \mu_i(\alpha)) : \alpha \in [-\pi, \pi]\}$, is a dispersion curve. All the dispersion curves compose dispersion diagrams (see Figure 3).

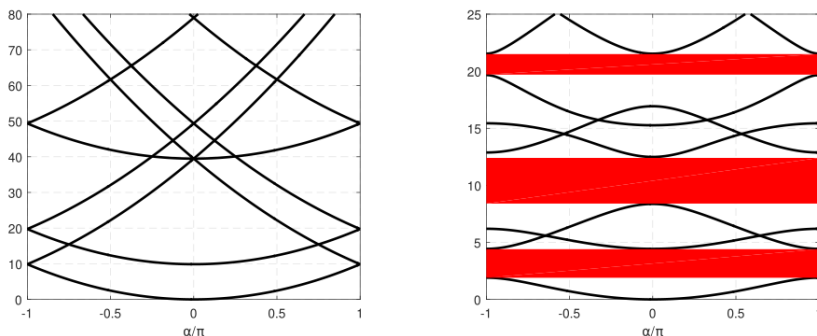


Figure 3: Dispersion diagrams for different n

For any fixed $k > 0$, there is a finite set I (may be empty, for example when k^2 lies in the red bands on the right picture of Figure 3) such that $S(k) = \{\alpha \in [-\pi, \pi] : \exists i \in I, \text{ s.t.}, \mu_i(\alpha) = k^2\}$. Corresponding to each dispersion curve $\mu_i(\alpha)$, there is also a family of eigenfunctions $\{\varphi_i(\alpha, x) : i \in I\}$ which also depend analytically on α . When $\mu'_i(\alpha) > 0$ ($\mu'_i(\alpha) < 0$), then the corresponding eigenfunction $\varphi_i(\alpha, \cdot)$ propagates to the right (left); when $\mu'_i(\alpha) = 0$, then $\varphi_i(\alpha, \cdot)$ is a standing wave. For fixed $k > 0$, $S(k)$ is divided into the following subsets:

$$\begin{aligned} S_-(k) &:= \{\alpha \in [0, 2\pi] : \exists i \in I, \text{ s.t.}, \mu_i(\alpha) = k^2, \mu'_i(\alpha) < 0\}; \\ S_+(k) &:= \{\alpha \in [0, 2\pi] : \exists i \in I, \text{ s.t.}, \mu_i(\alpha) = k^2, \mu'_i(\alpha) > 0\}; \\ S_0(k) &:= \{\alpha \in [0, 2\pi] : \exists i \in I, \text{ s.t.}, \mu_i(\alpha) = k^2, \mu'_i(\alpha) = 0\}. \end{aligned}$$

Note that when $S_0(k) \neq \emptyset$, the LAP does not work, so we have to make the following assumption.

Assumption 4. *In this paper, we assume that $S_0(k) = \emptyset$.*

In [10], it has been proved that the positive valued k 's such that Assumption 4 holds is only a discrete set, thus this assumption is reasonable.

At the end of this subsection, we apply the Floquet-Bloch transform to (1) (for details we refer to the appendix). Let $v(\alpha, x) := \mathcal{J}u$ where $\alpha \in [-\pi, \pi]$. With formal calculation, $v(\alpha, \cdot)$ satisfies (4) and the solution u is given by the inverse transform of v :

$$u(x) = \int_{-\pi}^{\pi} e^{i\alpha x_1} v(\alpha, x) d\alpha, \quad x \in \Omega_0. \quad (7)$$

From above arguments, (5) is uniquely solvable in $H_{per}^1(\Omega_0)$ for all $\alpha \in [-\pi, \pi] \setminus S(k)$. From analytic Fredholm theory, the function v is also extended to $\alpha \in \mathbb{C}$ (see Theorem 4 in [39]).

Theorem 5. *The Floquet-Bloch transformed field $v(\alpha, \cdot) = (\mathcal{J}u)(\alpha, \cdot)$ is extended to an analytic function for $\alpha \in \mathbb{C} \setminus \mathbb{F}$ (where \mathbb{F} is a discrete set) and meromorphic function for $\alpha \in \mathbb{C}$. Moreover, $e^{i\alpha x_1} v(\alpha, \cdot)$ is periodic with respect to the real part of α , i.e.,*

$$e^{i(\alpha+2\pi)x_1} v(\alpha + 2\pi, \cdot) = e^{i\alpha x_1} v(\alpha, \cdot), \quad \text{for almost all } \alpha \in \mathbb{C}.$$

Note that when $S(k) \neq \emptyset$, $v(\alpha, \cdot)$ does not exist when $\alpha \in S(k)$ thus the inverse transform (7) is not well defined. The formulation (7) no longer works thus we need some modifications to give exact formulations for LAP solutions.

3.2 Complex contour integral method

In this subsection, we introduce the first method – the complex contour integral (CCI) method. To guarantee that the CCI method works, we have to make the following assumption.

Assumption 6. *Assume that $k > 0$ such that $S_-(k) \cap S_+(k) = \emptyset$.*

From [39], the positive valued k 's such that Assumption 6 is satisfied is only a discrete set.

Remark 7. *Assumption 6 is not necessary for the LAP and the CCI method has been extended to the case without it in [39]. Concerning the length of this paper, we keep this assumption to only focus on the most important part.*

Assume that Assumption 4 and 6 are satisfied. The main idea of the CCI method is to modify the integral contour $[-\pi, \pi]$ in (7) such that the points in $S(k)$ are avoided. As the set $S(k)$ is finite, from the symmetry between $S_+(k)$ and $S_-(k)$ (see [39]), let

$$S_+(k) = \{\widehat{\alpha}_1^+, \dots, \widehat{\alpha}_N^+\} \text{ and } S_-(k) = \{\widehat{\alpha}_1^-, \dots, \widehat{\alpha}_N^-\} \quad \text{where } \widehat{\alpha}_j^+ = -\widehat{\alpha}_j^-.$$

First let the disk with center α and radius $\delta > 0$ be denoted by $B(\alpha, \delta)$, and define

$$D_\delta := [(-\pi, \pi) \times (0, +\infty)] \cup \left[\bigcup_{j=1}^N B(\widehat{\alpha}_j^+, \delta) \right] \setminus \left[\bigcup_{j=1}^N \overline{B(\widehat{\alpha}_j^-, \delta)} \right].$$

Define the new integral contour by:

$$\Lambda = \partial D_\delta \setminus [\{-\pi, \pi\} \times \mathbb{R}].$$

With Assumption 4 and 6, we choose $\delta > 0$ such that the following conditions are satisfied:

- first, any two disks do not have nonempty intersection;
- second, the closure of each disk only contains one exceptional value $\widehat{\alpha}_j^\pm$.

Theorem 8 (Theorem 7, [39]). *Assumption 4 and 6 hold. Then the LAP solution for (2) is given by*

$$u(x) := \int_{\Lambda} e^{i\alpha x_1} v(\alpha, x) d\alpha, \quad x \in \Omega. \quad (8)$$

where $v(\alpha, \cdot) \in H_{per}^1(\Omega_0)$ solves (5) for any fixed $\alpha \in \Lambda$.

From [40], the solution represented by (8) satisfies the radiation condition defined in Definition 1. Since $v(\alpha, \cdot)$ satisfies (5), assume $\psi = \psi(\alpha, x)$ and integral the equation on both sides with respect to α on the contour Λ , we arrive at the variational form of the problem. The weak formulation is to find $v \in X := L^2(\Lambda; H_{per}^1(\Omega_0))$ such that

$$\int_{\Lambda} \langle A(\alpha, k)v(\alpha, \cdot), \psi(\alpha, \cdot) \rangle d\alpha = - \int_{\Omega_0} e^{-i\alpha x_1} r(x) \overline{\left(\int_{\Lambda} \psi(\alpha, x) d\alpha \right)} dx$$

for all $\psi \in X$. From Riesz representation theorem, there is an operator \mathcal{A} defined in X such that

$$\langle \mathcal{A}w, \psi \rangle_X = \int_{\Lambda} \langle A(\alpha, k)w(\alpha, \cdot), \psi(\alpha, \cdot) \rangle d\alpha, \quad \forall w, \psi \in X.$$

As any $\alpha \in \Lambda$ is not an exceptional value, $A(\alpha, k)$ is always invertible. Thus \mathcal{A} is invertible in X .

As $r = f - k^2qu$, the variational formulation is written in the form:

$$\langle \mathcal{A}v, \psi \rangle_X - k^2 \int_{\Lambda} \int_{\Omega_0} e^{-i\alpha x_1} q(x) u(x) \overline{\psi(\alpha, x)} dx d\alpha = - \int_{\Lambda} \int_{\Omega_0} e^{-i\alpha x_1} f(x) \overline{\psi(\alpha, x)} dx d\alpha. \quad (9)$$

For simplicity, define the following operators. Let

$$(\mathcal{K}v)(x) := \int_{\Lambda} e^{i\alpha x_1} v(\alpha, x) d\alpha, \quad x \in \Omega_0.$$

From Riesz representation theorem, there is an operators \mathcal{T} in $H_{per}^1(\Omega_0)$ such that

$$\langle \mathcal{T}\varphi, \psi \rangle = \int_{\Omega_0} q(x)\varphi(x)\overline{\psi(x)} dx, \quad \text{for any } \varphi, \psi \in H_{per}^1(\Omega_0).$$

As $L^2(\Omega_0)$ is compact embedded in $H^1(\Omega_0)$, the operator \mathcal{T} is also compact. Similarly, define \mathcal{L} by:

$$\langle \mathcal{L}\varphi, \psi \rangle = \int_{\Omega_0} \varphi(x)\overline{\psi(x)} dx, \quad \text{for any } \varphi \in L^2(\Omega_0), \psi \in H_{per}^1(\Omega_0).$$

Then \mathcal{L} is bounded from $L^2(\Omega_0)$ to $H_{per}^1(\Omega_0)$. Thus (9) is equivalent to

$$(\mathcal{A} - k^2\mathcal{K}^*\mathcal{T}\mathcal{K})w = -\mathcal{K}^*\mathcal{L}f. \quad (10)$$

Now we study the property of the operator \mathcal{K} . We begin with a classical Minkowski integral inequality.

Lemma 9 ([11], Theorem 202). *Suppose (S_1, μ_1) and (S_2, μ_2) are two measure spaces and $F : S_1 \times S_2 \rightarrow \mathbb{R}$ is measurable. Then the following inequality holds for any $p \geq 1$*

$$\left[\int_{S_2} \left| \int_{S_1} F(y, z) d\mu_1(y) \right|^p d\mu_2(z) \right]^{1/p} \leq \int_{S_1} \left(\int_{S_2} |F(y, z)|^p d\mu_2(z) \right)^{1/p} d\mu_1(y). \quad (11)$$

Lemma 10. *The operator \mathcal{K} is bounded from $L^2(\Lambda; H^m(\Omega_0))$ to $H^m(\Omega_0)$ for any fixed non-negative integer m and $j \in \mathbb{Z}$. Especially, it is bounded from X to $H^1(\Omega_0)$.*

Proof. First consider the case when $m = 0$. Given any $w \in C^\infty(\Lambda \times \Omega_0)$, then $\mathcal{K}w$ is well defined and uniformly bounded for any $x \in \Omega_0$. Then from Lemma 9 (with $p = 1$) and the Cauchy-Schwartz inequality,

$$\begin{aligned} \|\mathcal{K}w\|_{L^2(\Omega_0)} &= \left[\int_{\Omega_0} \left| \int_{\Lambda} e^{i\alpha x_1} w(\alpha, x) d\alpha \right|^2 dx \right]^{1/2} \leq \int_{\Lambda} \left(\int_{\Omega_0} |e^{i\alpha x_1} w(\alpha, x)|^2 dx \right)^{1/2} d\alpha \\ &\leq \left(\int_{\Lambda} \int_{\Omega_0} |e^{i\alpha x_1} w(\alpha, x)|^2 dx d\alpha \right)^{1/2} \left(\int_{\Lambda} 1 d\alpha \right)^{1/2} \leq C\|w\|_{L^2(\Lambda; L^2(\Omega_0))}. \end{aligned}$$

From the density of $C^\infty(\Lambda \times \Omega_0)$ in $L^2(\Lambda; L^2(\Omega_0))$, the above inequality holds for all $w \in L^2(\Lambda; L^2(\Omega_0))$. Thus \mathcal{K} is bounded from $L^2(\Lambda; L^2(\Omega_j))$ to $L^2(\Omega_j)$. For $m \geq 1$, the proof is similar thus is omitted. So \mathcal{K} is bounded from $L^2(\Lambda; H^m(\Omega_0))$ to $H^m(\Omega_0)$ and particularly it is bounded from X to $H^1(\Omega_0)$. \square

As \mathcal{A} is invertible, \mathcal{K} and \mathcal{T} are bounded and \mathcal{T} is compact, $\mathcal{A} - k^2\mathcal{K}^*\mathcal{T}\mathcal{K}$ is a Fredholm operator. Thus it is invertible if and only if it is an injection. Then we study the well-posedness of the problem (10) in the following theorem.

Theorem 11. *With Assumption 2, the operator $\mathcal{A} - k^2\mathcal{K}^*\mathcal{TK}$ is invertible. That is, given any compactly supported $f \in L^2(\Omega_0)$, the problem (10) has a unique solution in X .*

Proof. Suppose w is a solution of (10) with $f = 0$, i.e., $(\mathcal{A} - k^2\mathcal{K}^*\mathcal{TK})w = 0$. Then $u = \mathcal{K}w$ lies in $H^1(\Omega_0)$. From the arguments in [39], $u = \mathcal{K}w$ is the LAP solution with source term $-k^2qu$ thus it is the solution of (1) with $f = 0$ and satisfies the radiation condition defined in Definition 1. From Assumption 2, $u = 0$. Thus $w = 0$, which implies that $\mathcal{A} - k^2\mathcal{K}^*\mathcal{TK}$ is injective thus is invertible. So the problem (10) is well-posed in the space X . \square

We have further regularity results for the solution w .

Corollary 12. *With Assumption 2, given any compactly supported $f \in L^2(\Omega_0)$. The solution w depends piecewise smoothly on $\alpha \in \Lambda$ and for fixed α , $w(\alpha, \cdot) \in H_{per}^2(\Omega_0)$.*

Proof. As for any $\alpha \in \Lambda$, $A(\alpha, k)$ is invertible and Λ is a piecewise smooth curve, $A(\alpha, k)$ depends piecewise smoothly on α from the perturbation theory. For each fixed $\alpha \in \Lambda$, $w(\alpha, \cdot) \in H_{per}^2(\Omega_0)$ from interior regularity. \square

3.3 Decomposition method

In [17, 18], an explicit formulation of the radiation condition is given for periodic open tubes in 3D. However, since the method is easily transferred to this case, we introduce the decomposition method based on the definition without proofs. For simplicity, let

$$S(k) := \left\{ \widehat{\beta}_j : j \in J \right\} \text{ where } J \text{ is a finite set.}$$

Note that this set is exactly the same as in the previous subsection, and we change the notation just to avoid confusions. For any fixed j , the Fredholm operator $A(\widehat{\beta}_j, k)$ is not an injection, and the null space $\widehat{Y}_j := \mathcal{N}(A(\widehat{\beta}_j, k))$ is finite dimensional with dimension m_j . There is an orthonormal eigensystem in \widehat{Y}_j

$$\left\{ (\lambda_{\ell,j}, \widehat{\varphi}_{\ell,j}) : \ell = 1, 2, \dots, m_j \right\} \quad (12)$$

such that

$$\int_{\Omega_0} \left[-i \frac{\partial \widehat{\varphi}_{\ell,j}}{\partial x_1} + \widehat{\beta}_j \widehat{\varphi}_{\ell,j} \right] \overline{\psi} \, dx = \lambda_{\ell,j} k \int_{\Omega_0} n \widehat{\varphi}_{\ell,j} \overline{\psi} \, dx \text{ for all } \psi \in \widehat{Y}_j \quad (13)$$

with normalization

$$2k \int_{\Omega_0} n \widehat{\varphi}_{\ell,j} \overline{\widehat{\varphi}_{\ell',j}} \, dx = \delta_{\ell,\ell'} \text{ for } \ell, \ell' = 1, 2, \dots, m_j.$$

Let $\varphi_{\ell,j} := e^{i\widehat{\beta}_j x_1} \widehat{\varphi}_{\ell,j}$, then $\varphi_{\ell,j}$ is $\widehat{\beta}_j$ -quasi-periodic eigenfunction to the Helmholtz equation

$$\Delta \varphi_{\ell,j} + k^2 n \varphi_{\ell,j} = 0 \text{ in } \Omega_0; \quad \varphi_{\ell,j} = 0 \text{ on } \Sigma_{\pm}^0.$$

For each $j \in J$, let the space spanned by $\{\varphi_{\ell,j}, \ell = 1, 2, \dots, m_j\}$ be denoted by Y_j .

With Assumption 4, $\lambda_{\ell,j} \neq 0$ for all $j \in J$ and $\ell = 1, 2, \dots, m_j$. In [19], a positive wave number k such that this condition is satisfied is also called regular. With above notations, we recall the radiation condition which is equivalent, but more explicit, compared to that defined in Definition 1.

Definition 13 (Theorem 3.7, [17]). *Suppose Assumption 2 and 4 hold. The LAP solution u has a decomposition $u = u^{(1)} + u^{(2)}$ where $u^{(1)} \in H^1(\Omega)$ and $u^{(2)}$ has the form*

$$u^{(2)}(x) = \psi^+(x_1) \sum_{j \in \mathcal{J}} \sum_{\lambda_{\ell,j} > 0} f_{\ell,j}^+ \varphi_{\ell,j}(x) + \psi^-(x_1) \sum_{j \in \mathcal{J}} \sum_{\lambda_{\ell,j} < 0} f_{\ell,j}^- \varphi_{\ell,j}(x)$$

for some $f_{\ell,j}^{\pm} \in \mathbb{C}$ defined by (see [18])

$$f_{\ell,j}^{\pm} = f_{\ell,j} = \frac{i}{|\lambda_{\ell,j}|} \int_{\Omega_0} r \overline{\varphi_{\ell,j}} \, dx. \quad (14)$$

and

$$\psi^+(x_1) \rightarrow 1 \text{ as } x_1 \rightarrow \infty, \psi^+(x_1) \rightarrow 0 \text{ as } x_1 \rightarrow -\infty; \quad \psi^-(x_1) = \psi^+(-x_1).$$

With an assumption that $\text{supp}(q) \subset (-1/2 + \delta, 1/2 - \delta) \times (0, 1)$ for a $\delta \in (0, 0.5)$, we can choose proper ψ^+ and ψ^- such that

$$\psi^+(x_1) = 1 \text{ for } x_1 \geq 1 - \frac{3\delta}{4} \text{ and } \psi^+(x_1) = 0 \text{ for } x_1 \leq 1 - \frac{\delta}{4}.$$

Then the property of $\psi^-(x_1) := \psi^+(-x_1)$ comes directly. Thus $\psi^\pm(x_1)q(x) = 0$ for all $x \in \Omega$.

Remark 14. *The choice of $\psi^\pm(x_1)$ is of course multiple. We require that $\psi^\pm(x_1)q(x) = 0$ just for a simplified formulation. Without the assumption on the support of q , we can set for example $\psi^+(x_1) = 1$ for $x_1 \geq 3/2$ and $\psi^+(x_1) = 0$ for $x_1 \leq 1/2$. Thus $\psi^-(x_1) = \psi^+(-x_1)$ satisfies $\psi^-(x_1) = 1$ for $x_1 \leq -3/2$ and $\psi^-(x_1) = 0$ for $x_1 \geq -1/2$. This setting is the one adopted in the numerical examples.*

With the particular ψ^\pm , $u^{(1)} \in H^1(\Omega)$ is actually decaying exponentially when $|x_1| \rightarrow \infty$. This also coincides with the radiation condition in Definition 1.

With the representation of $u^{(2)}$, we easily arrive at the following equation for $u^{(1)}$:

$$\Delta u^{(1)} + k^2 n u^{(1)} = \mathcal{M}(r) g_{\ell,j}(x), \quad (15)$$

where

$$\mathcal{M}(g) := g - \sum_{j \in J} \sum_{j=1}^m \left[\frac{i}{|\lambda_{\ell,j}|} \int_{\Omega_0} g \overline{\varphi_{\ell,j}} dx \right] g_{\ell,j}(x), \text{ for } g \in L^2(\Omega_0).$$

The functions $g_{\ell,j}$ are defined by:

$$g_{\ell,j}(x) = 2(\psi^\pm(x_1))' \frac{\partial \varphi_{\ell,j}}{\partial x_1}(x) + (\psi^\pm(x_1))'' \varphi_{\ell,j}(x) \text{ when } \pm \lambda_{\ell,j} > 0.$$

From the property of ψ^\pm , $\text{supp}(g_{\ell,j}) \subset \Omega_0$ for all j and ℓ . It is easily checked that \mathcal{M} is a bounded linear operator in $L^2(\Omega_0)$ and $\text{ran}(\mathcal{M})$ is orthogonal to the kernel of the differential operator $\Delta + k^2 n$.

Let $v(\beta, x) := (\mathcal{J}u^{(1)})(\beta, x)$. Since $u^{(1)} \in H^1(\Omega)$ decays exponentially as $|x_1| \rightarrow \infty$, $v(\beta, \cdot) \in H_{per}^1(\Omega_0)$ exists for all $\beta \in [-\pi, \pi]$ and depends analytically on β in the interval (and also extended analytically to a small neighbourhood of $[-\pi, \pi]$). For any fixed β , it is the unique solution of (5) with r be replaced by $\mathcal{M}r$. Similar to the arguments for the CCI method, the weak formulation for the problem is to find $v \in L^2((-\pi, \pi); H_{per}^1(\Omega_0))$ such that

$$\begin{aligned} \int_{-\pi}^{\pi} \langle A(\beta, k)v(\beta, \cdot), \varphi(\beta, \cdot) \rangle d\beta - k^2 \int_{\Omega_0} \mathcal{M}(qu^{(1)}) \overline{\left(\int_{-\pi}^{\pi} e^{i\beta x_1} \varphi(\beta, x) d\beta \right)} dx \\ = - \int_{\Omega_0} \mathcal{M}(f) \overline{\left(\int_{-\pi}^{\pi} e^{i\beta x_1} \varphi(\beta, x) d\beta \right)} dx. \end{aligned} \quad (16)$$

Then with the definition of \mathcal{T} and \mathcal{L} in the previous section,

$$\begin{aligned} \int_{\Omega_0} \mathcal{M}(qu^{(1)}) \overline{\left(\int_{-\pi}^{\pi} e^{i\beta x_1} \varphi(\beta, x) d\beta \right)} dx &= \int_{\Omega_0} qu^{(1)} \overline{\mathcal{M}^* \mathcal{J}^{-1} \varphi} dx = \langle \mathcal{T}u^{(1)}, \mathcal{M}^* \mathcal{J}^{-1} \varphi \rangle \\ &= \langle \mathcal{T} \mathcal{J}^{-1} v, \mathcal{M}^* \mathcal{J}^{-1} \varphi \rangle = \langle \mathcal{J} \mathcal{M} \mathcal{T} \mathcal{J}^{-1} v, \varphi \rangle; \end{aligned}$$

and

$$\int_{\Omega_0} \mathcal{M}(f) \overline{\left(\int_{-\pi}^{\pi} e^{i\beta x_1} \varphi(\beta, x) d\beta \right)} dx = \int_{\Omega_0} f \overline{\mathcal{M}^* \mathcal{J}^{-1} \varphi} dx = \langle \mathcal{L}f, \mathcal{M}^* \mathcal{J}^{-1} \varphi \rangle = \langle \mathcal{J} \mathcal{M} \mathcal{L}f, \varphi \rangle.$$

Note that \mathcal{M}^* is the adjoint operator of \mathcal{M} with respect to the L^2 -inner product thus is bounded.

Define the operator \mathcal{B} by:

$$\langle \mathcal{B}v, \varphi \rangle = \int_{-\pi}^{\pi} \langle A(\beta, k)v(\beta, x), \varphi(\beta, x) \rangle d\beta, \quad \text{for all } v, \varphi \in L^2((-\pi, \pi); H_{per}^1(\Omega_0)).$$

As $A(\beta, k)$ is self-adjoint for real valued β , \mathcal{B} is also self-adjoint. Thus the space $L^2((-\pi, \pi); H_{per}^1(\Omega_0))$ has the following decomposition:

$$L^2((-\pi, \pi); H_{per}^1(\Omega_0)) = \text{ran}(\mathcal{B}) \oplus \ker(\mathcal{B}) \text{ and } \text{ran}(\mathcal{B}) \perp \ker(\mathcal{B}).$$

Moreover, $\ker(\mathcal{B}) = \bigoplus_{j \in J} \widehat{Y}_j$. Thus \mathcal{B} is an isomorphism from $\text{ran}(\mathcal{B})$ to itself. For simplicity, let $Y := \text{ran}(\mathcal{B})$. The variational form (16) is now equivalent to

$$(\mathcal{B} - k^2 \mathcal{J} \mathcal{M} \mathcal{T} \mathcal{J}^{-1})v = -\mathcal{J} \mathcal{M} \mathcal{L} f. \quad (17)$$

Moreover, as $\text{ran}(\mathcal{M})$ is orthogonal to the kernel of $\Delta + k^2 n \cdot$, $\text{ran}(\mathcal{J} \mathcal{M})$ is orthogonal to $\ker(\mathcal{B})$, which implies that $\text{ran}(\mathcal{J} \mathcal{M}) \subset Y$. Thus the operator $\mathcal{B} - k^2 \mathcal{J} \mathcal{M} \mathcal{T} \mathcal{J}^{-1}$ is from Y to Y .

As \mathcal{T} is compact, the operator $\mathcal{J} \mathcal{M} \mathcal{T} \mathcal{J}^{-1}$ is a compact operator with range in Y . As \mathcal{B} is invertible in Y , $\mathcal{B} - k^2 \mathcal{J} \mathcal{M} \mathcal{T} \mathcal{J}^{-1}$ is a Fredholm operator. Thus we arrive at the well-posed result of (17).

Theorem 15. *Given any compactly supported $f \in L^2(\Omega_0)$ and Assumption 2, 4 hold. The operator $\mathcal{B} - k^2 \mathcal{J} \mathcal{M} \mathcal{T} \mathcal{J}^{-1}$ is invertible in Y , i.e., the problem (17) is uniquely solvable in Y .*

When u is the unique solution of (17), from previous process it satisfies the radiation condition introduced in Definition 13. Thus it is the LAP solution for the problem (1).

At the end of this section, we have to mention particularly the difficulty to solve the problem (17). Since $A(\beta, k)$ is not invertible in $H_{per}^1(\Omega_0)$ when $\beta = \widehat{\beta}_j$ and it is complicated to define the space Y from numerical point of view, to solve the problem directly may cause problems. To this end, we modify the formulation to avoid the singularities. Note that $v(\beta, \cdot)$ depends analytically on β in $(-\pi - \delta, \pi + \delta) \times (-2\sigma, 2\sigma)$ for sufficiently small $\delta, \sigma > 0$. From Cauchy integral formula,

$$\oint_{C_\sigma} e^{i\beta x_1} v(\beta, x) d\beta = 0,$$

where C_σ is the boundary of the rectangle $[-\pi, \pi] \times [0, \sigma]$ which is oriented counter-clockwise. On the other hand, $e^{i\beta x_1} v(\beta, \cdot)$ is 2π -periodic with respect to the real part of β , this implies that the integrals on the left- and right edges cancel. Thus

$$\int_{-\pi}^{\pi} e^{i\beta x_1} v(\beta, x) d\beta = \int_{-\pi}^{\pi} e^{i(\beta+i\sigma)x_1} v(\beta+i\sigma, x) d\beta$$

where $v(\beta+i\sigma, \cdot)$ solves (5) with parameter $\beta+i\sigma$. In the numerical schemes, to avoid the singularities we always replace β by $\beta+i\sigma$ in the variational formulation (16).

4 Numerical schemes

In this section, we introduce two numerical schemes based on the two representations. For both methods, the periodic problems (5) are computed for several times. So we briefly recall the finite element method for these problems at the very beginning.

Let \mathcal{M}_h be a family of regular curved and quasi-uniform meshes which cover the domain Ω_0 . We also assume that there the number and heights of nodal points on the left and right boundaries of Ω_0 are the same, such that we can define functions that can be extended to periodic ones on the meshes. By omitting nodal points on the right boundary, we assume that the points x_j ($j = 1, 2, \dots, M'$) be all the nodal points, where M' is a positive integer. Let $M > M'$ be a positive integer and x_j ($j = M'+1, \dots, M$) are points on the right boundary. Let $\{\zeta_j : j = 1, 2, \dots, M'\}$ be the family of piecewise linear and globally continuous basis functions that vanish on Σ_{\pm}^0 defined on the meshes \mathcal{M}_h , and $\zeta_j(x_\ell) = \delta_{j,\ell}$ where $j, \ell = 1, 2, \dots, M'$

and where $\delta_{j,\ell}$ is the Kronecker delta function. Define the function ζ_j on Γ_1 by the value on Γ_0 , then it is also extended periodically to Ω as a 1-periodic function. Then we define the finite dimensional subspace:

$$V_h := \text{span}\left\{\zeta_j : j = 1, 2, \dots, M'\right\} \subset H_{per}^1(\Omega_0).$$

Then the solutions $w(\alpha, \cdot)$ and v_α are approximated in the finite dimensional subspace V_h .

In the next subsections, we introduce the two numerical methods based on formulations of the LAP solution by the complex contour integral and decomposition.

4.1 The CCI method

First we recall the variational form for the CCI method, i.e., to find $w \in X$ such that

$$\begin{aligned} \int_{\Lambda} \langle A(\alpha, k)w(\alpha, \cdot), \psi(\alpha, \cdot) \rangle d\alpha - k^2 \int_{\Lambda} \int_{\Omega_0} e^{-i\alpha x_1} q(x) u(x) \overline{\psi(\alpha, x)} dx d\alpha \\ = - \int_{\Lambda} \int_{\Omega_0} e^{-i\alpha x_1} f(x) \overline{\psi(\alpha, x)} dx d\alpha \end{aligned} \quad (18)$$

$$u(x) = \int_{\Lambda} e^{i\alpha x_1} w(\alpha, x) d\alpha. \quad (19)$$

In this subsection, we focus on the discretization on the curve Λ . The modified integral contour Λ is composed with finite number of intervals and semicircles, thus it is piecewise smooth with finite number of corners. The first step is to parameterize Λ piecewisely:

$$\Lambda = \cup_{j=1}^Q \overline{\left\{s_j(t) = s_j^1(t) + is_j^2(t) : t \in I_j\right\}},$$

where I_j are closed bounded intervals, s_j^1, s_j^2 are real analytic functions in a neighbourhood of the interval I_j in \mathbb{R} . With a proper parameterization, let the intervals I_1, \dots, I_Q be

$$I_1 = (0, a_1), I_2 = (a_1, a_2), \dots, I_j = (a_{j-1}, a_j), \dots, I_Q = (a_{Q-1}, a_Q) = (a_{Q-1}, 1),$$

then

$$\Lambda = \left\{s(t) = s_1(t) + is_2(t) : t \in [0, 1]\right\}$$

where $s_1 = s_j^1$ and $s_2 = s_j^2$ in each I_j , respectively. Since the trapezoidal rule converges much faster for smooth periodic functions than nonperiodic smooth functions (see [37]), we require

$$s_j^{(m)}(a_{j-1}) = s_j^{(m)}(a_j) = 0, \quad \forall j = 1, 2, \dots, Q \text{ and } m = 1, 2.$$

Then $s'(t)$ is smooth in $[0, 1]$ and it can be extended to a periodic smooth function in \mathbb{R} . For details of this technique we refer to Section 9.6 in [20].

Replacing α by $s(t)$ in the equations (18)-(19) and let $\tilde{w}(t, x) := w(s(t), x)s'(t)$, we arrive at the following equivalent equations:

$$\begin{aligned} \int_0^1 \langle A(s(t), k)\tilde{w}(t, \cdot), \psi(t, \cdot) \rangle dt - k^2 \int_0^1 \int_{\Omega_0} e^{-is(t)x_1} q(x) u(x) \overline{\psi(t, x)} s'(t) dx dt \\ = - \int_0^1 \int_{\Omega_0} e^{-is(t)x_1} f(x) \overline{\psi(t, x)} s'(t) dx dt, \end{aligned} \quad (20)$$

$$u(x) = \int_0^1 e^{is(t)x_1} \tilde{w}(t, x) dt. \quad (21)$$

The variational problem is to seek for $\tilde{w} \in \tilde{X} := L^2((0, 1); H_{per}^1(\Omega_0))$ such that (20)-(21) hold for all test function $\psi \in \tilde{x}$. Moreover, from Corollary 12, $\tilde{w} \in C_{per}^\infty([0, 1]; H_{per}^2(\Omega_0))$.

Now we discretize the system (20)-(21). As the finite element discretization in the domain Ω_0 has already been introduced at the beginning of this section, we only introduce the trigonometric interpolation in the domain $[0, 1]$. Let N be a positive integer and the uniformly distributed nodal points be

$$t_0 = 0; \quad t_j = \frac{j}{N} \quad \forall j = 1, 2, \dots, N.$$

Suppose N is an even number, then the basic function:

$$\xi_\ell(t) := \frac{1}{N} \sum_{m=-N/2+1}^{N/2} \exp(i2\pi m(t - t_\ell)), \quad \ell = 1, 2, \dots, N.$$

It is well know that

$$\xi_\ell(t_{\ell'}) = \delta_{\ell, \ell'} \quad \text{and} \quad \int_0^1 \xi_\ell(t) \xi_{\ell'}(t) dt = \frac{1}{N} \delta_{\ell, \ell'}.$$

Now we are prepared to discretize the system (20)-(21) in the finite dimensional space:

$$\tilde{X}_{N,h} := \underbrace{V_h \oplus V_h \oplus \dots \oplus V_h}_{N \text{ spaces}}.$$

Let $\tilde{w}_{N,h} \in \tilde{X}_{N,h}$ be the approximation of \tilde{w} in the form of:

$$\tilde{w}_{N,h}(t, x) = \sum_{\ell=1}^N \sum_{j=1}^M \hat{w}_{\ell,j} \xi_\ell(t) \zeta_j(x).$$

With the test function $\psi(t, x) = \xi_{\ell'}(t) \zeta_{j'}(x)$, (20)-(21) is discretized as:

$$\begin{aligned} \frac{1}{N} \delta_{\ell, \ell'} \sum_{j=1}^{M'} \hat{w}_{\ell,j} \langle A(t_\ell, k) \zeta_j, \zeta_{j'} \rangle - \frac{k^2}{N} \sum_{j=1}^{M'} \hat{u}_j \left(\int_{\Omega_0} e^{-is(t_{\ell'})x_1} q(x) \zeta_j(x) \zeta_{j'}(x) dx \right) s'(t_{\ell'}) \\ = -\frac{1}{N} \left(\int_{\Omega_0} e^{-is(t_{\ell'})x_1} f(x) \zeta_{j'}(x) dx \right) s'(t_{\ell'}), \end{aligned} \quad (22)$$

together with the coefficients of \hat{u}_j :

$$\hat{u}_j = \frac{1}{N} \sum_{\ell=1}^N e^{is(t_\ell)x_1} \hat{w}_{\ell,j}. \quad (23)$$

With (23) we also get the approximation of u , i.e., $u_{N,h}$, at the same time.

Finally the system (22)-(23) is summarized as the following matrix:

$$\begin{pmatrix} A_1 & 0 & \dots & 0 & C_1 \\ 0 & A_2 & \dots & 0 & C_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & A_N & C_N \\ B_1 & B_2 & \dots & B_N & I \end{pmatrix} \begin{pmatrix} W_1 \\ W_2 \\ \vdots \\ W_N \\ U \end{pmatrix} = \begin{pmatrix} F_1 \\ F_2 \\ \vdots \\ F_N \\ 0 \end{pmatrix}, \quad (24)$$

where $W_j = (\hat{w}_{1,j}, \dots, \hat{w}_{M,j})^\top$ and $U = (\hat{u}_1, \dots, \hat{u}_M)^\top$. A_j comes from the first term of (22), C_j comes from the second term of (22), B_j comes from (23), F_j comes from the right hand side. The size of this system is $(N+1)M \times (N+1)M$. Note that U is treated as an additional unknown vector just for a simpler representation of the linear system.

4.2 The decomposition method

We introduce the decomposition method in this subsection. First we recall the variational form from (16) with the definition of \mathcal{M} (note that β is already replaced by $\beta + i\sigma$):

$$\begin{aligned} \int_{-\pi}^{\pi} \langle A(\beta + i\sigma, k)v(\beta, \cdot), \psi(\beta, \cdot) \rangle d\beta - k^2 \int_{-\pi}^{\pi} \int_{\Omega_0} e^{-i(\beta+i\sigma)x_1} q(x) u^{(1)}(x) \overline{\psi(\beta, x)} dx d\beta \\ + \sum_{j \in J} \sum_{\ell=1}^{m_j} C_{\ell,j}(\psi) = - \int_{-\pi}^{\pi} \int_{\Omega_0} e^{-i(\beta+i\sigma)x_1} f_0(x) \overline{\psi(\beta, x)} dx d\beta; \end{aligned} \quad (25)$$

$$u^{(1)}(x) = \int_{-\pi}^{\pi} e^{i(\beta+i\sigma)x_1} v(\beta + i\sigma, x) d\beta; \quad (26)$$

$$C_{\ell,j}(\psi) = \frac{ik^2}{|\lambda_{\ell,j}|} \left[\int_{\Omega_0} q(x) u^{(1)}(x) \overline{\varphi_{\ell,j}} dx \right] \int_{-\pi}^{\pi} \int_{\Omega_0} e^{-i(\beta+i\sigma)x_1} q(x) g_{\ell,j}(x) \overline{\psi(\beta, x)} dx d\beta \quad (27)$$

$$f_0(x) = f - \sum_{j \in J} \sum_{\ell=1}^{m_j} \frac{i}{|\lambda_{\ell,j}|} \left[\int_{\Omega_0} f(x) \overline{\varphi_{\ell,j}} dx \right] g_{\ell,j}(x); \quad (28)$$

The first task is to approximate the orthonormal eigensystem $\{(\widehat{\beta}_j, \widehat{\varphi}_{\ell,j}) : \ell = 1, 2, \dots, m_j\} \subset \widehat{Y}_j$ and the corresponding $\lambda_{\ell,j}$ for all $j \in J$ numerically. This implementation is carried out by the following two steps. The first step is to find out all the exceptional values $\widehat{\beta}_j \in (-\pi, \pi]$ and the corresponding eigenspace $\widehat{Y}_j \subset H_{per}^1(\Omega_0)$ such that $A(\widehat{\beta}_j, k)\widehat{\varphi} = 0$ for all $\widehat{\varphi} \in \widehat{Y}_j$. Recall that in (6), $A(\alpha, k) = A_1 + k^2 A_4 + \alpha A_2 + \alpha^2 A_3$, it is to solve the quadratic eigenvalue problem. Define

$$B_1 := \begin{pmatrix} A_1 + k^2 A_4 & 0 \\ 0 & I \end{pmatrix}, \quad B_2 = \begin{pmatrix} -A_2 & -A_3 \\ I & 0 \end{pmatrix}, \quad W = \begin{pmatrix} \widehat{\varphi}_{\ell,j} \\ \widehat{\beta}_j \widehat{\varphi}_{\ell,j} \end{pmatrix}$$

then $(\widehat{\beta}_j, W)$ is the solution of the following generalized eigenvalue problem:

$$B_1 W = \widehat{\beta}_j B_2 W. \quad (29)$$

By solving this problem, we find all the eigenvalues and eigenfunctions

$$\left\{ (\widehat{\beta}_j, \widehat{\varphi}_{\ell,j}) : j \in J, \ell = 1, 2, \dots, m_j \right\}.$$

Then $\widehat{Y}_j = \text{span}\{\widehat{\varphi}_{\ell,j} : \ell = 1, 2, \dots, m_j\}$. The discretization of the operators is also based on the finite element method. Let the corresponding matrices be denoted by $A_1^h, A_2^h, A_3^h, A_4^h, B_1^h, B_2^h$, by solving (29), the numerical results are denoted by

$$\left\{ (\widehat{\beta}_j^h, \widehat{\varphi}_{\ell,j}^h) : j \in J, \ell = 1, 2, \dots, m_j \right\}.$$

It is known that since the discretized form B_1^h and B_2^h converges to exact operators B_1 and B_2 as $h \rightarrow 0^+$,

$$\lim_{h \rightarrow 0^+} \widehat{\beta}_j^h = \widehat{\beta}_j, \quad \lim_{h \rightarrow 0^+} \widehat{\varphi}_{\ell,j}^h = \widehat{\varphi}_{\ell,j}$$

hold for all $j \in J$ and $\ell = 1, 2, \dots, m_j$ (see Chapter 4, Section 3.5, [16]). However, the convergence rate of $|\widehat{\beta}_j^h - \widehat{\beta}_j|$ and $\|\widehat{\varphi}_{\ell,j}^h - \widehat{\varphi}_{\ell,j}\|_{L^2(\Omega_0)}$ with respect to the meshsize h is not clear. For simplicity, let $E(h)$ be a function depends on $h \in (0, h_0)$ for a sufficiently small h_0 such that

$$|\widehat{\beta}_j^h - \widehat{\beta}_j|, \|\widehat{\varphi}_{\ell,j}^h - \widehat{\varphi}_{\ell,j}\|_{H_{per}^1(\Omega_0)} \leq E(h).$$

From the convergence result, it is reasonable to require that $E(h) \rightarrow 0$ when $h \rightarrow 0^+$.

With the eigenfunctions $\{\tilde{\varphi}_{\ell,j}^h : \ell = 1, 2, \dots, m_j\}$ for fixed $j \in J$, we go to the second step, i.e., to construct the orthonormal eigensystem (12) with condition (13). The eigenfunctions are in the form of

$$\tilde{\varphi}_{\ell,j}^h(x) = \sum_{\ell'=1}^{m_j} c_{\ell,\ell'}^j \tilde{\varphi}_{\ell',j}^h(x) \quad (30)$$

where the coefficients $c_{\ell,\ell'}^j \in \mathbb{C}$. Thus the problem is then to find out the approximated eigenvalues $\lambda_{\ell,j}^h$ and coefficients $c_{\ell,\ell'}^j$ for all $\ell, \ell' = 1, 2, \dots, m_j$. From (13), for fixed $\ell = 1, 2, \dots, m_j$,

$$\sum_{\ell'=1}^{m_j} c_{\ell,\ell'}^{j,h} \left(\int_{\Omega_0} \left[-i \frac{\partial}{\partial x_1} \tilde{\varphi}_{\ell',j}^h + \hat{\beta}_j \tilde{\varphi}_{\ell',j}^h \right] \overline{\tilde{\varphi}_{\ell',j}^h} dx \right) = \lambda_{\ell,j}^h \sum_{\ell'=1}^{m_j} c_{\ell,\ell'}^{j,h} \left(k \int_{\Omega_0} \tilde{\varphi}_{\ell',j}^h \overline{\tilde{\varphi}_{\ell',j}^h} dx \right)$$

holds for any $\ell'' = 1, 2, \dots, m_j$. Let

$$a_{\ell,\ell'}^j = \int_{\Omega_0} \left[-i \frac{\partial}{\partial x_1} \tilde{\varphi}_{\ell',j}^h + \hat{\beta}_j \tilde{\varphi}_{\ell',j}^h \right] \overline{\tilde{\varphi}_{\ell,j}^h} dx; \quad b_{\ell,\ell'}^j = k \int_{\Omega_0} \tilde{\varphi}_{\ell',j}^h \overline{\tilde{\varphi}_{\ell,j}^h} dx,$$

then

$$\begin{pmatrix} a_{1,1}^j & a_{1,2}^j & \cdots & a_{1,m_j}^j \\ a_{2,1}^j & a_{2,2}^j & \cdots & a_{2,m_j}^j \\ \vdots & \vdots & \cdots & \vdots \\ a_{m_j,1}^j & a_{m_j,2}^j & \cdots & a_{m_j,m_j}^j \end{pmatrix} \begin{pmatrix} c_{\ell,1}^{j,h} \\ c_{\ell,2}^{j,h} \\ \vdots \\ c_{\ell,m_j}^{j,h} \end{pmatrix} = \lambda_{\ell,j}^h \begin{pmatrix} b_{1,1}^j & b_{1,2}^j & \cdots & b_{1,m_j}^j \\ b_{2,1}^j & b_{2,2}^j & \cdots & b_{2,m_j}^j \\ \vdots & \vdots & \cdots & \vdots \\ b_{m_j,1}^j & b_{m_j,2}^j & \cdots & b_{m_j,m_j}^j \end{pmatrix} \begin{pmatrix} c_{\ell,1}^{j,h} \\ c_{\ell,2}^{j,h} \\ \vdots \\ c_{\ell,m_j}^{j,h} \end{pmatrix}.$$

By solving this generalized eigenvalue problem, we get all the coefficients $c_{\ell,\ell'}^{j,h}$ and eigenvalues $\lambda_{\ell,j}^h$. Then the function $\tilde{\varphi}_{\ell,j}^h$ is obtained directly by (30). Finally, we normalize the functions $\tilde{\varphi}_{\ell,j}^h$ such that

$$2k \int_{\Omega_0} n(x) \tilde{\varphi}_{\ell,j}^h(x) \overline{\tilde{\varphi}_{\ell,j}^h(x)} dx = 1.$$

Let $\varphi_{\ell,j}^h(x) := e^{i\hat{\beta}_j x_1} \tilde{\varphi}_{\ell,j}^h(x)$, then the system $\{(\lambda_{\ell,j}^h, \varphi_{\ell,j}^h) : \ell = 1, 2, \dots, m_j\}$ for fixed $j \in J$ is the numerical approximation of (12). Here we still let $E(h)$ be the error bound, i.e.,

$$|\lambda_{\ell,j}^h - \lambda_{\ell,j}|, \|\varphi_{\ell,j}^h - \varphi_{\ell,j}\|_{H_{per}^1(\Omega_0)} \leq E(h).$$

Now we replace $\lambda_{\ell,j}$ and $\varphi_{\ell,j}$ by $\lambda_{\ell,j}^h$ and $\varphi_{\ell,j}^h$ in (25)-(28), an approximated formulation is obtained:

$$\begin{aligned} & \int_{-\pi}^{\pi} \langle A(\beta + i\sigma, k) v^h(\beta + i\sigma, \cdot), \psi(\beta, \cdot) \rangle d\beta - k^2 \int_{-\pi}^{\pi} \int_{\Omega_0} e^{-i(\beta+i\sigma)x_1} q(x) u_h^{(1)}(x) \overline{\psi(\beta, x)} dx d\beta \\ & + \sum_{j \in J} \sum_{\ell=1}^{m_j} C_{\ell,j}^h(\psi) = - \int_{-\pi}^{\pi} \int_{\Omega_0} e^{-i(\beta+i\sigma)x_1} f_0^h(x) \overline{\psi(\beta, x)} dx d\beta; \end{aligned} \quad (31)$$

$$u_h^{(1)}(x) = \int_{-\pi}^{\pi} e^{i(\beta+i\sigma)x_1} v^h(\beta + i\sigma, x) d\beta; \quad (32)$$

$$C_{\ell,j}^h(\psi) = \frac{ik^2}{|\lambda_{\ell,j}^h|} \left[\int_{\Omega_0} q(x) u_h^{(1)}(x) \overline{\varphi_{\ell,j}^h} dx \right] \int_{-\pi}^{\pi} \int_{\Omega_0} e^{-i(\beta+i\sigma)x_1} q(x) g_{\ell,j}^h(x) \overline{\psi(\beta, x)} dx d\beta \quad (33)$$

$$f_0^h(x) = f - \sum_{j \in J} \sum_{\ell=1}^{m_j} \frac{i}{|\lambda_{\ell,j}^h|} \left[\int_{\Omega_0} f(x) \overline{\varphi_{\ell,j}^h} dx \right] g_{\ell,j}^h(x). \quad (34)$$

Here $g_{\ell,j}^h$ is defined in the same way as g_{ℓ} by replacing $\varphi_{\ell,j}$ with $\varphi_{\ell,j}^h$. Thus the error $\|g_{\ell,j}^h - g_{\ell,j}\|_{L^2(\Omega_0)} \leq E(h)$. Since $\lim_{h \rightarrow 0^+} E(h) = 0$, when h is sufficiently small, the above variational formulation is a small perturbation of (25)-(28). From perturbation theory the following estimation is obvious

$$\|u_h^{(1)} - u^{(1)}\|_{H^1(\Omega_0)} \leq E(h).$$

Now we are prepared to discretize the system (31)-(34). Let g be a complex valued smooth function defined in $[0, 1]$ such that $g(t) = \text{Re}(g(t)) + i\sigma$ and

$$\text{Re}(g(0)) = -\pi \text{ and } \text{Re}(g(1)) = \pi, \quad g^{(m)}(0) = g^{(m)}(1) = 0 \quad \text{for all } m = 1, 2, \dots$$

Replace $\beta + i\sigma$ by $g(t)$ in (31)-(33), and define $\tilde{v}(t, x) := v^h(g(t), x)g'(t)$, then $\tilde{v} \in C_{per}^\infty([0, 1]; H_{per}^2(\Omega_0))$. We arrive at the variational form for \tilde{v} , i.e., to find $\tilde{v} \in L^2((0, 1); H_{per}^1(\Omega_0))$ such that

$$\begin{aligned} \int_0^1 \langle A(g(t), k)\tilde{v}(t, \cdot), \psi(t, \cdot) \rangle dt - k^2 \int_0^1 \int_{\Omega_0} e^{-ig(t)x_1} q(x) u_h^{(1)}(x) \overline{\psi(t, x)} g'(t) dx dt \\ + \sum_{j \in J} \sum_{\ell=1}^{m_j} C_{\ell,j}^h(\psi) = - \int_0^1 \int_{\Omega_0} e^{-ig(t)x_1} f_0^h(x) \overline{\psi(t, x)} g'(t) dx dt; \end{aligned} \quad (35)$$

$$u_h^{(1)}(x) = \int_0^1 e^{ig(t)x_1} \tilde{v}(t, x) dt; \quad (36)$$

$$C_{\ell,j}^h(\psi) = \frac{ik^2}{|\lambda_{\ell,j}|} \left[\int_{\Omega_0} q(x) u_h^{(1)}(x) \overline{\varphi_{\ell,j}^h} dx \right] \int_0^1 \int_{\Omega_0} e^{-ig(t)x_1} q(x) g_{\ell,j}^h(x) \overline{\psi(t, x)} g'(t) dx dt \quad (37)$$

hold for any $\psi \in L^2((0, 1); H_{per}^1(\Omega_0))$, where f_0^h is defined in (34).

We use the same discretization as in the CCI method and consider the discretized problem in the space $\tilde{X}_{N,h}$, where $\tilde{v}_{N,h} \subset \tilde{X}_{N,h}$ is given by:

$$\tilde{v}_{N,h} = \sum_{\ell=1}^N \sum_{j=1}^{M'} \hat{v}_{\ell,j} \xi_\ell(t) \zeta_j(x).$$

With the test function $\psi(t, x) = \xi_{\ell'}(t) \zeta_{j'}(x)$, we get the discretized form of the system (35)-(37):

$$\begin{aligned} \frac{1}{N} \delta_{\ell,\ell'} \sum_{j=1}^M \hat{v}_{\ell,j} \langle A(t_\ell, k) \zeta_j, \zeta_{j'} \rangle - \frac{k^2}{N} \sum_{j=1}^M \hat{v}_j \left(\int_{\Omega_0} e^{-ig(t_{\ell'})x_1} q(x) \zeta_j(x) \zeta_{j'}'(x) dx \right) g'(t_{\ell'}) \\ + \sum_{j'' \in J} \sum_{\ell''=1}^{m_j} C_{\ell'',j''}^{N,h} \left[\int_{\Omega_0} e^{-ig(t_{\ell'})x_1} q(x) g_{\ell'',j''}^h(x) \zeta_{j'}(x) dx \right] g'(t_{\ell'}) \\ = - \frac{1}{N} \left(\int_{\Omega_0} e^{-ig(t_{\ell'})x_1} f_0^h(x) \zeta_{j'}(x) dx \right) g'(t_{\ell'}) \end{aligned} \quad (38)$$

together with the coefficients \tilde{v}_j and C :

$$\hat{v}_j = \frac{1}{N} \sum_{\ell=1}^N e^{ig(t_\ell)x_1} \hat{v}_{\ell,j}; \quad (39)$$

$$C_{\ell,j}^{N,h} = \frac{ik^2}{|\lambda_{\ell,j}|} \frac{1}{N} \left[\int_{\Omega_0} q(x) u_h^{(1)}(x) \overline{\varphi_{\ell,j}^h} dx \right]. \quad (40)$$

Similar to the CCI method, we approximate $u_h^{(1)}$ by $u_{N,h}^{(1)}$ with coefficients defined by (39) at the same time.

Finally the system (38)-(40) is summarized as the following matrix:

$$\begin{pmatrix} \tilde{A}_1 & 0 & \cdots & 0 & S_{1,1}^h & S_{2,1}^h & \cdots & S_{I,1}^h & \tilde{C}_1 \\ 0 & \tilde{A}_2 & \cdots & 0 & S_{1,2}^h & S_{2,2}^h & \cdots & S_{I,2}^h & \tilde{C}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \tilde{A}_N & S_{1,N}^h & S_{2,N}^h & \cdots & S_{I,N}^h & \tilde{C}_N \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & D_1 \\ 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & D_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & D_I \\ \tilde{B}_1 & \tilde{B}_2 & \cdots & \tilde{B}_N & 0 & 0 & \cdots & 0 & I \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \\ \vdots \\ V_N \\ C_{1,N,h}^{N,h} \\ C_{2,N,h}^{N,h} \\ \vdots \\ C_{I,N,h}^{N,h} \\ U^{(1)} \end{pmatrix} = \begin{pmatrix} \tilde{F}_1 \\ \tilde{F}_2 \\ \vdots \\ \tilde{F}_N \\ \hat{F}_1 \\ \hat{F}_2 \\ \vdots \\ \hat{F}_I \\ 0 \end{pmatrix}, \quad (41)$$

where \tilde{A}_j comes from the first term of (38), \tilde{C}_j comes from the second term of (38), $S_{\ell,j}^h$ comes from the third term of (38), \tilde{B}_j comes from (39), \tilde{F}_j comes from the right hand side, D_j comes from the second term in (40) and \hat{F}_j come from the right hand side of (38), $C_j^{N,h}$ are reordered with the indexes $j = 1, 2, \dots, I$. Note here $I = \sum_{j \in J} m_j$. The size of this system is $(NM + M + I) \times (NM + M + I)$.

4.3 Error estimations

In the last part of this section, we present the error estimations for both of the algorithms. Since both the solution $\tilde{w}(t, x)$ of (20)-(21) and the solution $\tilde{v}(t, x)$ of (35)-(37) lie in the space $C_{per}^\infty([0, 1]; H_{per}^2(\Omega_0))$, we can use the error estimation introduced in [38]. First result is the error estimation of the interpolation in the finite dimensional space $\tilde{X}_{N,h}$.

Theorem 16 (Theorem 32, [38]). *Suppose $v \in C_{per}^\infty([0, 1]; H_{per}^2(\Omega_0))$ and $v_{N,h}$ is the interpolation of v in the subspace $\tilde{X}_{N,h}$. Then*

$$\min_{v_{N,h} \in \tilde{X}_{N,h}} \|v - v_{N,h}\|_{L^2([0,1]; H_{per}^1(\Omega_0))} \leq C(N^{-n} + h) \|v\|_{C_{per}^\infty([0,1]; H_{per}^2(\Omega_0))},$$

where n can be any positive integer and C depends on n .

This implies that the interpolation decays super algebraically with respect to the parameter $1/N$ but linearly with respect to h . With this result, we get the error estimations for finite element solutions of the variational problems (20)-(21) and (35)-(37).

Theorem 17 (Theorem 34, [38]). *Let \tilde{w} be the exact solution of (20)-(21) and $\tilde{w}_{N,h}$ be the finite element solution of the corresponding discretized form (22)-(23). Let \tilde{v} be the exact solution of (35)-(37) and $\tilde{v}_{N,h}$ be the finite element solution of (38)-(40). For sufficiently small $h > 0$ and sufficiently large positive integer N , we have the following error estimations:*

$$\|\tilde{w} - \tilde{w}_{N,h}\|_{L^2([0,1] \times \Omega_0)} \leq Ch(N^{-n} + h) \|\tilde{w}\|_{C_{per}^\infty([0,1]; H_{per}^2(\Omega_0))}; \quad (42)$$

$$\|\tilde{v} - \tilde{v}_{N,h}\|_{L^2([0,1] \times \Omega_0)} \leq Ch(N^{-n} + h) \|\tilde{v}\|_{C_{per}^\infty([0,1]; H_{per}^2(\Omega_0))} \quad (43)$$

where C is a parameter depends on n .

For the both methods, we get the original solution from (23), and the error is also easily obtained:

$$\|u - u_{N,h}\|_{L^2(\Omega_0)} \leq Ch(N^{-n} + h) \|\tilde{v}\|_{C_{per}^\infty([0,1]; H_{per}^2(\Omega_0))}; \quad (44)$$

$$\|u_h^{(1)} - u_{N,h}^{(1)}\|_{L^2(\Omega_0)} \leq Ch(N^{-n} + h) \|\tilde{v}\|_{C_{per}^\infty([0,1]; H_{per}^2(\Omega_0))}. \quad (45)$$

For the CCI method, the final error estimation is already obtained by (44). For the decomposition method, since $\|u_h^{(1)} - u_h^{(1)}\|_{L^2(\Omega_0)} \leq E(h)$,

$$\|u_h^{(1)} - u_{N,h}^{(1)}\|_{L^2(\Omega_0)} \leq Ch(N^{-n} + h) \|\tilde{v}\|_{C_{per}^\infty([0,1]; H_{per}^2(\Omega_0))} + E(h).$$

From Definition 13 and (26), the original solution u is now approximated by

$$\tilde{u}_{N,h} := u_{N,h}^{(1)} + \psi^+(x_1) \sum_{j \in \mathcal{J}} f_{\ell,j}^h \varphi_{\ell,j}^h(x)$$

where

$$f_{\ell,j}^h = \frac{i}{|\lambda_{\ell,j}^h|} \int_{\Omega_0} \left[-f + k^2 q u_{N,h}^{(1)} \right] \overline{\varphi_{\ell,j}^h} dx.$$

As the second term also has an error bounded by $Ch(N^{-n} + h) \|\tilde{v}\|_{C_{per}^\infty([0,1]; H_{per}^2(\Omega_0))} + E(h)$, finally we get the error estimate for the decomposition method:

$$\|u - \tilde{u}_{N,h}\|_{L^2(\Omega_0)} \leq Ch(N^{-n} + h) \|\tilde{v}\|_{C_{per}^\infty([0,1]; H_{per}^2(\Omega_0))} + E(h). \quad (46)$$

5 Numerical examples

In this section, we use some numerical examples to show the efficiency of both of the methods. For both cases, we apply the same finite element discretization for quasi-periodic problem (5), and the meshsize h is chosen as 0.005, 0.01, 0.02, 0.04. The parameter N is chosen as 4, 8, 16, 32, 64.

We show numerical results for two different examples. For both examples, f and q are the same. f is already be defined in Remark 2, and q is defined by:

$$q(x) = \begin{cases} 0, & |x - b_0| > 0.15; \\ 2, & 0.1 < |x - b_0| < 0.15; \\ 2\zeta(|x - b_0|; 0.1, 0.15), & \text{otherwise.} \end{cases}$$

where $b_0 = (0.2, 0.2)^\top$. In Example 1, the periodic refractive index $n = n_1$ is defined also in Remark 2, while in Example 2,

$$n(x) = n_2(x) = 3 + \sin(4\pi x_1).$$

The wave number is chosen as $\sqrt{17}$ in Example 1 and $\sqrt{12}$ in Example 2. We plot the dispersion diagrams for different examples in Figure 4. From the diagrams we get the set of exceptional values for two examples. For Example 1,

$$S(k) = \{-0.9577, 0.9577\} \text{ and } S_-(k) = \{-0.9577\}, S_+(k) = \{0.9577\};$$

while for Example 2,

$$S(k) = \{-1.0326, 1.0326\} \text{ and } S_-(k) = \{1.0326\}, S_+(k) = \{-1.0326\}.$$

Note that since the above results are numerical, they are not exact. Based on the exceptional values, we also show the integral contour Λ in Figure 5.

5.1 The convergence of exceptional values

First, we show the convergence of the approximated exceptional values with respect to the meshsize h . As the convergence of eigenfunctions and $\lambda_{\ell,j}^h$ are similar, they are omitted here. For $h = 0.04, 0.02, 0.01, 0.005$, we compute the exceptional values and the positive valued ones are listed in Table 5.1.

	$h = 0.04$	$h = 0.02$	$h = 0.01$	$h = 0.005$
Example 1	0.8982	0.9435	0.9549	0.9577
Example 2	1.0738	1.0387	1.0337	1.0326

Table 1: Positive valued exceptional values computed with different h 's.

If the values at $h = 0.005$ is treated as “exact”, then we plot the relative error with respect to the parameter h in logarithmic scales in the first picture in Figure 6. From the plot, the slope for the red curve is about 2.2 and for the blue one is about 2.6, which corresponds to (and even a little faster than) the convergence of the finite element method. Thus we can trust the exceptional values as well as the corresponding eigenfunctions computed by this method and the convergence rate is about $O(h^2)$.

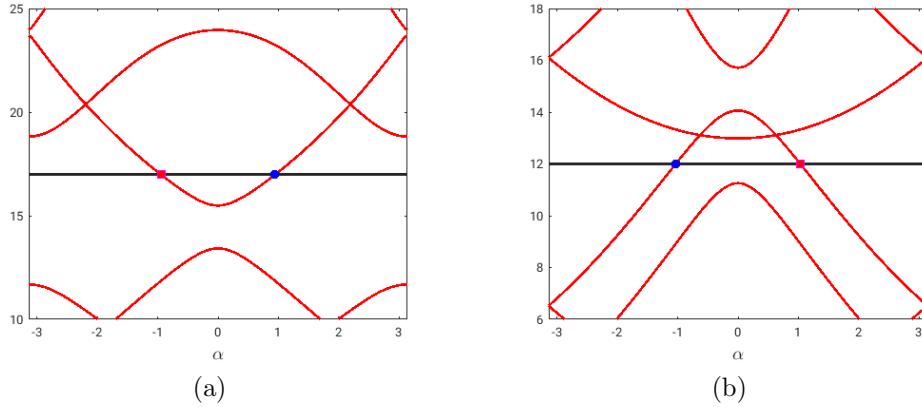


Figure 4: Dispersion diagrams: Example 1 in (a) and Example 2 in (b). Blue dots are points in $S_+(k)$ and purple squares are points in $S_-(k)$.

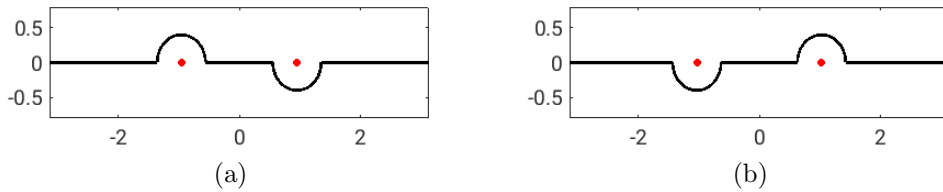


Figure 5: Integral contour Λ : Example 1 in (a) and Example 2 in (b). Red dots exceptional values.

5.2 Numerical results

In this section, we focus on the numerical results obtained by the proposed methods. For both examples, we use a completely different method introduced in [5] to produce results as “exact solutions”. In the computation we use Lagrangian element with meshsize 0.005 and an extrapolation technique with data points 0.001, 0.0005, 0.00025, and the solution is denoted by u_{exa} . Then the error is estimated as:

$$err_{N,h} = \frac{\|u_{N,h} - u_{exa}\|_{L^2(\Omega_0)}}{\|u_{exa}\|_{L^2(\Omega_0)}}.$$

For the decomposition method, the parameter σ is chosen to be 0.2. For the relative errors with different examples and methods we refer to Table 5.2-5.2. From the four tables, the relative errors decrease as h gets smaller and N gets larger, but the decreases stop at the level of 3×10^{-3} . This may due to the lack of accuracy of the “exact solutions”.

	$h = 0.04$	$h = 0.02$	$h = 0.01$	$h = 0.005$
$N = 16$	1.98E-1	1.85E-1	1.80E-1	1.78E-1
$N = 32$	8.87E-2	5.26E-2	4.30E-2	4.06E-2
$N = 64$	6.21E-2	1.96E-2	7.54E-3	4.60E-3
$N = 128$	6.12E-2	1.87E-2	6.60E-3	3.82E-3
$N = 256$	6.12E-2	1.87E-2	6.60E-3	3.82E-3

Table 2: Relative error for Example 1, CCI method.

Now let’s turn to the convergence rate. For both examples and methods, we study the dependence of the errors on h and N separately. To study the dependence on h , we fix a large N , i.e., $N = 256$ and let the solutions with $h = 0.005$ be the “exact” ones. Then we show the relative errors

$$err_{256,h} = \frac{\|u_{256,h} - u_{256,0.005}\|_{L^2(\Omega_0)}}{\|u_{256,0.005}\|_{L^2(\Omega_0)}}$$

	$h = 0.04$	$h = 0.02$	$h = 0.01$	$h = 0.005$
$N = 16$	7.89E-2	2.00E-2	5.42E-3	3.13E-3
$N = 32$	8.27E-2	2.05E-2	5.44E-3	3.09E-3
$N = 64$	8.38E-2	2.08E-2	5.44E-3	3.05E-3
$N = 128$	8.36E-2	2.08E-2	5.44E-3	3.05E-3
$N = 256$	8.36E-2	2.08E-2	5.44E-3	3.05E-3

Table 3: Relative error for Example 1, decomposition method.

	$h = 0.04$	$h = 0.02$	$h = 0.01$	$h = 0.005$
$N = 16$	6.87E-2	4.63E-2	4.28E-2	4.22E-2
$N = 32$	4.80E-2	1.51E-2	7.07E-3	5.78E-3
$N = 64$	4.71E-2	1.36E-2	4.67E-3	2.98E-3
$N = 128$	4.71E-2	1.36E-2	4.69E-3	3.02E-3
$N = 256$	4.71E-2	1.36E-2	4.69E-3	3.02E-3

Table 4: Relative error for Example 2, CCI method.

in Table 5.2. To study the dependence on N , we fix $h = 0.005$ and let the solutions with $N = 256$ be “exact” ones. Then we show the relative errors

$$err_{N,0.005} = \frac{\|u_{N,0.005} - u_{256,0.005}\|_{L^2(\Omega_0)}}{\|u_{256,0.005}\|_{L^2(\Omega_0)}}.$$

in Table 5.2. We also plot the data in logarithmic scales in Figure 6. (b) shows the dependence on h and (c) shows the dependence on N . (b) of Figure 6, the four curves are almost straight and the slopes are almost 2. This shows that the convergence rate with respect to h is about $O(h^2)$. In (c), the curves are no longer close to straight ones and the slopes gets faster as $\log N$ gets larger. This implies that the convergence rate is super-algebraic. Both results coincide with the error estimations given in Theorem 17.

At the end of this section, we also show the error between the two different methods with the same parameters. Since the super-algebraic convergence of both algorithms with respect to N is already shown, we fix $N = 256$ and compute the relative errors for different h 's:

$$err_{h,256} = \frac{\|u_{256,h}^{CCI} - u_{256,h}^D\|_{L^2(\Omega_0)}}{\|u_{256,h}^{CCI}\|_{L^2(\Omega_0)}}$$

where $u_{N,h}^{CCI}$ and $u_{N,h}^D$ are numerical results obtained from the CCI method and decomposition method, respectively. The relative errors are shown in Table 5.2. The data are also plotted in logarithmic scales in (d) in Figure 6. For Example 1, the slope of the curve is about 1.8 and for Example 2, the slope is about 1.9. The results also corresponds to the error analysis of the finite element method and the convergence of exceptional values shown in (a) in Figure 6. Moreover, since the solutions for both methods coincide with each other, we can trust both of the algorithms.

6 Conclusion

In the last section of this paper, we compare the two algorithms. Both methods have second order convergence with respect to the meshsize and super algebraically with respect to the number of nodal points on $[0, 1]$. The computational complexity is also similar with the same parameters. The CCI method requires the additional Assumption 6 which is not necessary for the LAP process and the decomposition method. Fortunately this only excludes a discrete subset of $(0, +\infty)$. On the other hand, the decomposition method involves two steps and has a relatively more complicated formulation compare to the CCI method. Since both algorithms converges supre-algebraically with respect to the parameter N , they are very efficient and we can choose different algorithm according to different settings and requirements.

	$h = 0.04$	$h = 0.02$	$h = 0.01$	$h = 0.005$
$N = 16$	9.01E-2	8.04E-2	7.92E-2	7.92E-2
$N = 32$	4.92E-2	3.25E-2	3.05E-2	3.03E-2
$N = 64$	3.71E-2	1.11E-2	4.83E-3	3.84E-3
$N = 128$	3.70E-2	1.08E-2	4.11E-3	2.92E-3
$N = 256$	3.70E-2	1.08E-2	4.12E-3	2.94E-3

Table 5: Relative error for Example 2, decomposition method.

	Example 1 (CCI)	Example 1 (D)	Example 2 (CCI)	Example 2 (D)
$h = 0.04$	5.96E-2	8.32E-2	4.54E-2	3.55E-2
$h = 0.02$	1.60E-2	2.00E-2	1.18E-2	9.11E-3
$h = 0.01$	3.35E-3	3.98E-3	2.34E-3	1.79E-3

Table 6: Relative errors with different h 's for fixed $N = 256$.

Appendix

First we recall definitions of smooth and analytic functions with values in Banach spaces.

Definition 18. Suppose F is a map from an open set $U \subset \mathbb{C}^N$ into a complex Banach space X . Then

- F is analytic at $z_0 \in U$ if there is an $R > 0$ and a series $\{f_n : n \in \mathbb{N}\} \subset X$ such that

$$F(z) = \sum_{n=0}^{\infty} \frac{(z - z_0)^n}{n!} f_n$$

converges uniformly for $z \in B(z_0, R) \cap U$ where $B(z_0, R)$ is the disk with center z_0 and radius R .

- F is smooth at $z_0 \in U$ if its Fréchet differentiable of any order exists.

With Definition 18, we can easily define spaces $C^\omega([a, b]; H^s(\Omega_0))$, $C^\infty([a, b]; H^s(\Omega_0))$ where the functions depend analytically or smoothly on the first variable. The space $C_{per}^\infty([a, b]; H^s(\Omega_0))$ is the subspace of $C^\infty([a, b]; H^s(\Omega_0))$ with an additional periodic condition on any order derivatives of the first variable.

We introduce the Floquet-Bloch transform. For a function $\varphi \in C_0^\infty(\Omega)$, define the transform

$$(\mathcal{J}\varphi)(\alpha, x) := (2\pi)^{-1/2} \sum_{j \in \mathbb{Z}} \varphi \left(x + \begin{pmatrix} j \\ 0 \end{pmatrix} \right) e^{-i\alpha(x_1 + j)}.$$

It is easily checked that with fixed $\alpha \in \mathbb{R}$, $(\mathcal{J}\varphi)(\alpha, \cdot)$ is 1-periodic in x_1 -direction. For fixed $x \in \Omega_0$, $e^{i\alpha x_1}(\mathcal{J}\varphi)(\alpha, x)$ is 2π -periodic in α . The properties of the transform is concluded in the following theorem. For details we refer to [21, 25].

Theorem 19. The Floquet-Bloch transform is extended to an isometry between $H^s(\Omega)$ and $L^2((-\pi, \pi); H_{per}^s(\Omega_0))$ for any $s \in \mathbb{R}$ and its inverse transform is:

$$(\mathcal{J}^{-1}\psi)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \psi(\alpha, x) e^{i\alpha x_1} d\alpha.$$

Moreover, the adjoint operator of \mathcal{J} with respect to the inner product of $L^2((-\pi, \pi); L^2(\Omega_0))$, denoted by \mathcal{J}^* , equals to \mathcal{J}^{-1} .

$\mathcal{J}\varphi$ depends analytically on α if and only if φ decays exponentially when $|x_1| \rightarrow \infty$.

Note here the subscript *per* is to indicate that the function is periodic for fixed first variable.

	Example 1 (CCI)	Example 1 (D)	Example 2 (CCI)	Example 2 (D)
$N = 16$	1.76E-1	1.17E-4	4.28E-2	8.02E-2
$N = 32$	3.87E-2	5.52E-5	5.82E-3	3.13E-2
$N = 64$	1.37E-3	4.06E-6	1.16E-4	3.71E-3
$N = 128$	1.57E-6	4.94E-8	4.30E-8	4.47E-5

Table 7: Relative errors with different N 's for fixed $h = 0.005$.

	$h = 0.04$	$h = 0.02$	$h = 0.01$	$h = 0.005$
Exampel 1	5.08E-2	1.76E-2	4.80E-3	1.25E-3
Example 2	1.25E-2	3.44E-3	8.68E-4	2.20E-4

Table 8: Relative error for Example 2, decomposition method.

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project-ID 258734477 SFB 1173. The author is grateful for Prof. Andreas Kirsch for valuable discussions and suggestions.

References

- [1] D. L. Colton and R. Kress. *Integral equation methods in scattering theory*. Krieger, repr. ed., with corr. edition, 1992.
- [2] T. Dohnal and B. Schweizer. A bloch wave numerical scheme for scattering problems in periodic wave-guides. *SIAM J. Numer. Anal.*, 56(3):1848–1870, 2018.
- [3] W. Dörfler, A. Lechleite, M. Plum, G. Schneider, and C. Wieners. *Photonic crystals: Mathematical analysis and numerical approximation*, 2011.
- [4] M. Ehrhardt, H. Han, and C. Zheng. Numerical simulation of waves in periodic structures. *Commun. Comput. Phys.*, 5:849–870, 2009.
- [5] M. Ehrhardt, J. Sun, and C. Zheng. Evaluation of scattering operators for semi-infinite periodic arrays. *Commun. Math. Sci.*, 7:347–364, 2009.
- [6] S. Fliss. *Analyse mathématique et numérique de problèmes de propagation des ondes dans des milieux périodiques infinis localement perturbés*. PhD thesis, Ecole Polytechnique, 2009.
- [7] S. Fliss. A dirichlet-to-neumann approach for the exact computation of guided modes in photonic crystal waveguides. *SIAM Journal on Scientific Computing*, 35(2):B438B461, 2013.
- [8] S. Fliss and P. Joly. Exact boundary conditions for time-harmonic wave propagation in locally perturbed periodic media. *Appl. Numer. Math.*, 59:2155–2178, 2009.
- [9] S. Fliss and P. Joly. Wave propagation in locally perturbed periodic media(case with absorption): Numerical aspects. *J. Comput. Phys.*, 231:1244–1271, 2012.
- [10] S. Fliss and P. Joly. Solutions of the time-harmonic wave equation in periodic waveguides: asymptotic behaviour and radiation condition. *Arch. Rational Mech. Anal.*, 2015.
- [11] G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities*. Cambridge Mathematical Library. Cambridge University Press, 2nd edition, 1988.
- [12] V. Hoang. The limiting absorption principle for a periodic semin-infinite waveguide. *SIAM J. Appl. Math.*, 71(3):791–810, 2011.

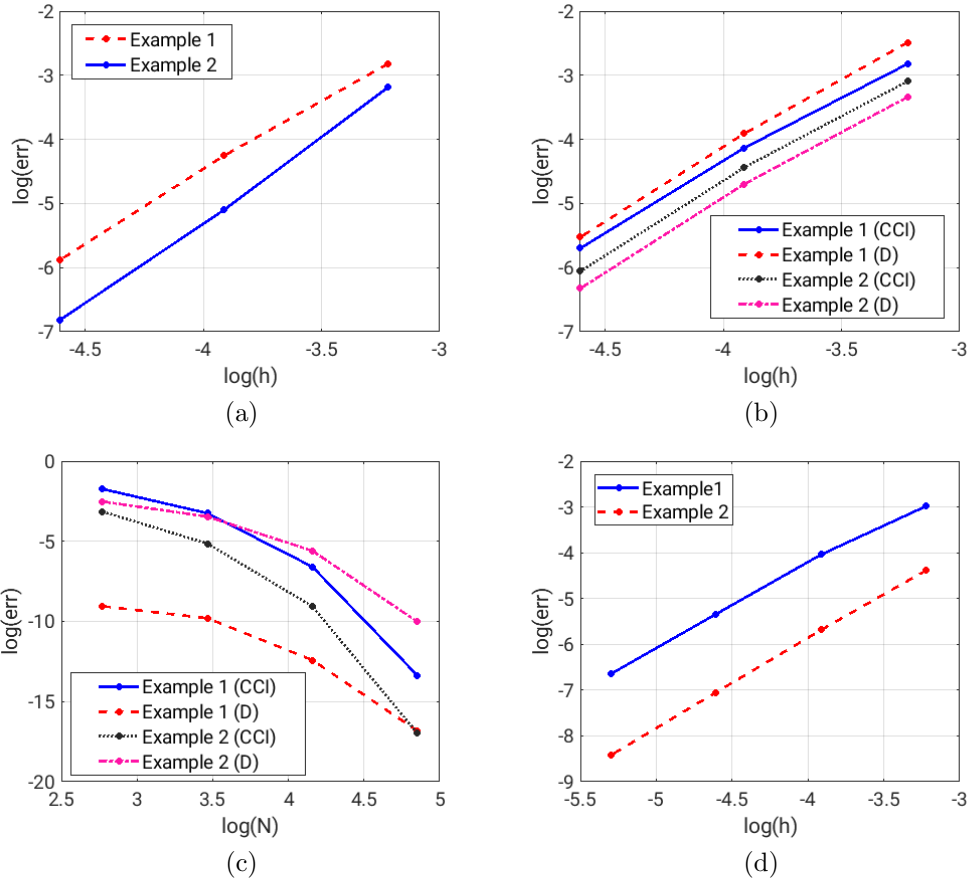


Figure 6: Convergence of exceptional values (a), relative errors depend on h (b) and N (c).

- [13] J.D. Joannopoulos, R.D. Meade, and N. J. Winn. *Photonic Crystal-Molding the Flow of Light*. Princeton University Press, 1995.
- [14] S.G. Johnson and J. D. Joannopoulos. *Photonic Crystal - The road from theory to practice*. Kluwer Academic Publishers, 2002.
- [15] P. Joly, J.-R. Li, and S. Fliss. Exact boundary conditions for periodic waveguides containing a local perturbation. *Commun. Comput. Phys.*, 1:945–973, 2006.
- [16] T. Kato. *Perturbation theory for linear operators*. Springer, repr. of the 1980 edition, 1995.
- [17] A. Kirsch. Scattering by a periodic tube in \mathbb{R}^3 : part i. the limiting absorption principle. *Inverse Problems*, 35(10):104004, 2019.
- [18] A. Kirsch. Scattering by a periodic tube in \mathbb{R}^3 : part ii. the radiation condition. *Inverse Problems*, 35(10):104005, 2019.
- [19] A. Kirsch and A. Lechleiter. A radiation condition arising from the limiting absorption principle for a closed full or halfwaveguide problem. *Math. Meth. Appl. Sci.*, 41(10):3955–3975, 2018.
- [20] R. Kress. *Numerical Analysis*. Springer, 1998.
- [21] P. Kuchment. *Floquet Theory for Partial Differential Equations*, volume 60 of *Operator Theory. Advances and Applications*. Birkhäuser, Basel, 1993.

- [22] P. Kuchment. The mathematics of photonic crystals. In *Mathematical modelling in optical science*, volume 22 of *Frontiers in Applied Mathematics*, page Chapter 7, Philadelphia, 2001. SIAM.
- [23] P. Kuchment and B. Ong. On guided waves in photonic crystal waveguides. In P. Kuchment, editor, *Waves in Periodic and Random Media*, volume 339 of *Contemporary Mathematics*, pages 105–116. AMS, 2003.
- [24] A. Lechleiter. The Floquet-Bloch transform and scattering from locally perturbed periodic surfaces. *J. Math. Anal. Appl.*, 446(1):605–627, 2017.
- [25] A. Lechleiter and D.-L. Nguyen. Scattering of Herglotz waves from periodic structures and mapping properties of the Bloch transform. *Proc. Roy. Soc. Edinburgh Sect. A*, 231:1283–1311, 2015.
- [26] A. Lechleiter and R. Zhang. A convergent numerical scheme for scattering of aperiodic waves from periodic surfaces based on the Floquet-Bloch transform. *SIAM J. Numer. Anal.*, 55(2):713–736, 2017.
- [27] A. Lechleiter and R. Zhang. A Floquet-Bloch transform based numerical method for scattering from locally perturbed periodic surfaces. *SIAM J. Sci. Comput.*, 39(5):B819–B839, 2017.
- [28] A. Lechleiter and R. Zhang. Non-periodic acoustic and electromagnetic scattering from periodic structures in 3d. *Comput. Math. Appl.*, 74(11):2723–2738, 2017.
- [29] S. A. Nazarov. Elliptic boundary value problems with periodic coefficients in a cylinder. *Math. USSR Izv.*, 18(1):89–99, 1982.
- [30] S. A. Nazarov. The mandelshtam energy radiation conditions and the umov-poynting vector in elastic waveguides. *Probl. Mat. Anal.*, 72:101–146, 2013.
- [31] S. A. Nazarov. Umov-mandelshtam radiation conditions in elastic periodic waveguides. *Sb. Math.*, 205(7):953–982, 2014.
- [32] S. A. Nazarov and B. A. Plamenevsky. On radiation conditions for selfadjoint elliptic problems. *Dokl. Akad. Nauk SSSR*, 311(3):532–536, 1990.
- [33] S. A. Nazarov and B. A. Plamenevsky. Radiation principles for selfadjoint elliptic problems. *Probl. Mat. Fiz.*, 13:192–244, 1991.
- [34] S. A. Nazarov and B. A. Plamenevsky. *Elliptic Problems in domains with piecewise smooth boundaries*, vol 13 of the *De Gruyter Expositions in Mathematics*. Walter de Gruyter & Co., Berlin, 1994.
- [35] K. Sadoka. *Optical Properties of Photonic Crystals*. Springer, 2001.
- [36] J. Sun and C. Zheng. Numerical scattering analysis of the plane waves by a metallic diffraction grating with local defects. *J. Opt. Soc. Am. A*, 26(1):156–162, 2009.
- [37] J.A.C. Weideman. Numerical integration of periodic functions: a few examples. *Am. Math. Mon.*, 109(1):21–35, 2002.
- [38] R. Zhang. A high order numerical method for scattering from locally perturbed periodic surfaces. *SIAM J. Sci. Comput.*, 40(4):A2286–A2314, 2018.
- [39] R. Zhang. Numerical method for scattering problems in periodic waveguides. <https://arxiv.org/pdf/1906.12283.pdf>, 2019.
- [40] R. Zhang. Spectrum decomposition of translation operators in periodic waveguide. *SIAM Journal on Applied Mathematics*, 81(1):233–257, 2021.