Bergische Universität Wuppertal

Fachbereich Mathematik und Naturwissenschaften

Lehrstuhl für Angewandte Mathematik
und Numerische Mathematik

Lehrstuhl für Optimierung und Approximation

Preprint BUW-AMNA-OPAP 09/03

Matthias Ehrhardt and Chunxiong Zheng

Fast Numerical Methods for Waves in Periodic Media

October 2009

http://www.math.uni-wuppertal.de

Fast Numerical Methods for Waves in Periodic Media

M. Ehrhardt* **and C. Zheng+**

* Bergische Universität Wuppertal, Fachbereich Mathematik und Naturwissenschaften, Lehrstuhl für Angewandte Mathematik und Numerische Mathematik, Gaußstrasse 20, 42119 Wuppertal, Germany.
+ Department of Mathematical Sciences, Tsinghua University, Beijing 100084, P.R. China.

Abstract: Periodic media problems widely exist in many modern application areas like semiconductor nanostructures (e.g. quantum dots and nanocrystals), semi-conductor superlattices, photonic crystals (PC) structures, meta materials or Bragg gratings of surface plasmon polariton (SPP) waveguides, etc. Often these application problems are modeled by partial differential equations with periodic coefficients and/or periodic geometries.
In order to numerically solve these periodic structure problems efficiently one usually confines the spatial domain to a bounded computational domain (i.e. in a neighborhood of the region of physical interest). Hereby, the usual strategy is to introduce so-called *artificial boundaries* and impose suitable boundary conditions. For wave-like equations, the ideal boundary conditions should not only lead to well-posed problems, but also mimic the perfect absorption of waves traveling out of the computational domain through the artificial boundaries.
In the first part of this chapter we present a novel analytical impedance expression for general second order ODE problems with periodic coefficients. This new expression for the kernel of the Dirichlet-to-Neumann mapping of the artificial boundary conditions is then used for computing the bound states of the Schrödinger operator with periodic potentials at infinity. Other potential applications are associated with the exact artificial boundary conditions for some time-dependent problems with periodic structures. As an example, a two-dimensional hyperbolic equation modeling the TM polarization of the electromagnetic field with a periodic dielectric permittivity is considered.
In the second part of this chapter we present a new numerical technique for solving periodic structure problems. This novel approach possesses several advantages. First, it allows for a fast evaluation of the Sommerfeld-to-Sommerfeld operator for periodic array problems. Secondly, this computational method can also be used for bi-periodic structure problems with local defects. In the sequel we consider several problems, such as the exterior elliptic problems with strong coercivity, the time-dependent Schrödinger equation and the Helmholtz equation with damping.
Finally, in the third part we consider periodic arrays that are structures consisting of geometrically identical subdomains, usually called periodic cells. We use the Helmholtz equation as a model equation and consider the definition and evaluation of the exact boundary mappings for general semi-infinite arrays that are periodic in one direction for any real wavenumber. The well-posedness of the Helmholtz equation is established via the *limiting absorption principle* (LABP).
An algorithm based on the doubling procedure of the second part of this chapter and an extrapolation method is proposed to construct the exact Sommerfeld-to-Sommerfeld boundary mapping. This new algorithm benefits from its robustness and the simplicity of implementation. But it also suffers from the high computational cost and the resonance wave numbers. To overcome these shortcomings, we propose another algorithm based on a conjecture about the asymptotic behaviour of limiting absorption principle solutions. The price we have to pay is the resolution of some generalized eigenvalue problem, but still the overall computational cost is significantly reduced. Numerical evidences show that this algorithm presents theoretically the same results as the first algorithm. Moreover, some quantitative comparisons between these two algorithms are given.

1. INTRODUCTION

Nowadays periodic media problems exist in many modern application areas like semiconductor nanostructures (e.g. quantum dots and nanocrystals), semi-conductor superlattices [11], [68], photonic crystals (PC) structures [10], [43], [53], meta materials [60] or Bragg gratings of surface plasmon polariton (SPP) waveguides [31], [61]. In many cases these problems are modeled by partial differential equations (PDEs) on unbounded domains with periodic coefficients and / or periodic geometries.

The most interesting property of these periodic media, especially in optical applications, is the capability to select the ranges of frequencies of the waves that are allowed to pass or blocked in the waveguide ('frequency filter'). Waves in (infinite) periodic media only exist if their frequency lies inside these allowed continuous bands separated by forbidden gaps. This fact corresponds mathematically to the gap structure of the differential operator having so-called *pass bands* and *stop bands*. Numerical simulations are necessary for the design, analysis and finally optimization of the waveguiding periodic structures. E.g. in a typical application the wanted frequencies of defect modes are the eigenvalues of a PDE eigenvalue problem posed on an *unbounded domain* [28].

In order to numerically solve these equations efficiently it is a standard practice to confine the spatial domain to a bounded computational region (usually in the neighborhood of the domain of physical interest). Hence it is necessary to introduce so-called *artificial boundaries* and impose adequate boundary conditions. Note that even in the case of a bounded but large domain, it is a common practice to reduce the original domain to a smaller one by introducing artificial boundaries, for example, see [44]. This technique is especially beneficial if these generated *exterior domains* consist of a huge number of periodicity cells. For wave-like equations, the ideal boundary conditions should not only lead to well-posed problems, but also mimic the perfect absorption of waves leaving the computational domain through the artificial boundaries. Moreover, these boundary conditions should allow for an easy implementation and a fast, efficient and accurate evaluation of the *Dirichlet-to-Sommerfeld* (DtS) mapping is essential. In the literature these boundary conditions are usually called *artificial* (or transparent, non-reflecting in the same spirit). The interested reader is referred to the review papers [6], [26], [29], [30], [65] on this fundamental research topic.

Artificial boundary conditions (ABCs) for the Schrödinger equation and related problems has been a hot research topic for many years, cf. [6] and the references therein. Since the first exact ABC for the Schrödinger equation was derived by Papadakis [48] 25 years ago in the context of underwater acoustics, many developments have been made on the designing and implementing of various ABCs, also for multi-dimensional and nonlinear problems. However, the question of exact ABCs for periodic structures still remained open, and it is a very up-to-date research topic, cf. the current papers [23], [25], [40], [56], [63], [64], [72], [73].

Let us note that recently Zheng [78] derived exact ABCs for the Schrödinger equation of the form

$$iu_t + u_{xx} = V(x)u, \quad x \in \mathbb{R}, \tag{1a}$$

$$u(x,0) = u_0(x), \quad x \in \mathbb{R}, \tag{1b}$$

$$u(x,t) \to 0, \quad x \to \pm\infty. \tag{1c}$$

Here, the initial function $u_0 \in L^2(\mathbb{R})$ is assumed to be compactly supported in an interval $[x_L, x_R]$, with $x_L < x_R$, and the real potential function $V \in L^\infty(\mathbb{R})$ is assumed to be sinusoidal on the interval $(-\infty, x_L]$ and $[x_R, +\infty)$. It is well-known that the system (1) has a unique solution $u \in C(\mathbb{R}^+, L^2(\mathbb{R}))$ for bounded potentials (cf. [50], e.g.):

Theorem 1. *Let $u_0 \in L^2(\mathbb{R})$ and $V \in L^\infty(\mathbb{R})$. Then the system* (1) *has a unique solution $u \in C(\mathbb{R}^+, L^2(\mathbb{R}))$. Moreover, it is a unitary evolution i.e. the "energy" is preserved:*

$$\|u(.,t)\|_{L^2(\mathbb{R})} = \|u_0\|_{L^2(\mathbb{R})}, \forall t \geq 0.$$

We remark that a recent paper [22] derives approximate ABCs for (1) with a more general class of possibly unbounded potentials.

In [78] Zheng considered the periodic potentials

$$V(x) = V_L + 2q_L \cos \frac{2\pi(x_L - x)}{S_L}, \ \forall x \in (-\infty, x_L],$$

$$V(x) = V_R + 2q_R \cos \frac{2\pi(x - x_R)}{S_R}, \ \forall x \in [x_R, +\infty),$$

where S_L and S_R are the periods, V_L and V_R are the average potentials, and the nonnegative numbers q_L and q_R relate to the amplitudes of sinusoidal part of the potential function V on $(-\infty, x_L]$ and $[x_R, +\infty)$, respectively.

Though *absorbing boundary conditions* (ABCs) for wave-like equations have been a hot research issue for many years and many developments have been made on their designing and implementing, the

question of exact ABCs for periodic structure problems is not fully settled yet. Some progresses can be found in the recent research articles [19], [20], [21], [22], [23], [25], [40], [56], [58], [63], [64], [72], [73], [74] and [78]. For a comprehensive review on the theory of waves in locally periodic media including a survey on physical applications we refer the interested reader to [27].

In the existing literature *frequency domain methods* (FDMs) are usually considered for wave problems with periodic structures [39]. These methods are able to exploit the special geometric structure and are based on an eigenmode expansion in every longitudinally uniform cell. Frequently, the FDMs are used in conjunction with the *perfectly matched layer* (PML) [12] technique for dealing with unbounded domains. Afterwards the *bidirectional beam propagation methods* (BiBPMs) [34] were introduced. Like the FDMs, they can utilize the periodic geometry but additionally they (and also the *eigenmode expansion methods* in [12] and [34]) are able to resolve the multiple reflections at the longitudinal interfaces.

The methods of Jacobsen [36] and Yuan & Lu [72] were developed to be more efficient than the eigenmode expansion methods, because it turns out that solving the eigenmodes in each segment is quite time consuming. Recently, a *DtN mapping method* [71] was developed by Yuan and Lu that is more accurate than the BiBPMs, since this approach works (mostly) without any approximation. In [73] the efficiency of this sequential DtN approach was further improved by a *recursive doubling process* for the DtN map.

In this chapter we study a numerical method for the *Helmholtz equation*

$$-\Delta u(\mathbf{x}) + (V - z)u(\mathbf{x}) = f(\mathbf{x}). \tag{2}$$

Here z denotes a complex parameter, and $V = V(\mathbf{x})$ is a sufficiently smooth real function bounded from both below and above. The domain of definition and the function V are assumed to be periodic at least on some part of the region.

The Helmholtz equation is one of the fundamental equations of mathematical physics and models time-harmonic wave propagation. In many cases, the Helmholtz equation (2) is posed on the unbounded domain $\mathbb{R}^2$ and solved as a boundary value problem with some radiation boundary conditions, for example, see [41].

In some special cases [36] it is possible to obtain analytic expressions of the solution, but in general, the Helmholtz equation (2) has to be solved numerically. However, If the number of periodic cells is large, then a direct discretization of the whole domain involves a huge number of unknowns which makes it costly and even impractical from an implementational point of view. In this chapter our goal is to find a smart resolution without naively solving the whole domain problem.

The most interesting property of periodic arrays, especially in optical applications in nano- and microtechnology, is the capability of selecting waves in a range of frequencies that are allowed to pass or blocked through the media. Waves in periodic arrays only exist when their frequency lies inside some allowed continuous bands separated by forbidden gaps. This fact corresponds mathematically to the dispersion diagram of suitable differential operator having so-called *pass bands* and *stop bands*. Since the governing wave equation is either of periodic variable coefficient, or defined on a domain consisting of periodic subregions, theoretical analysis is very limited, and numerical simulation is a fundamental tool for the design, analysis and finally optimization of the periodic arrays.

In many cases some defect cells are artificially introduced into a perfect periodic array for some additional interesting property. For example, if the defect cells are properly designed, some defect modes [59] can exist for certain frequencies in the band gaps. This phenomena has many important applications, e.g. in light emitting devices (LEDs) and photonic circuits [52].

The organization of this chapter is as follows. In Section 2, we present an elegant analytical expression of the impedance operator for problems with periodic coefficients. In Section 3 we use this result to compute bound states for the Schrödinger operator. In Section 4 we show how the results can be generalized to the time-dependent Schrödinger equation, a diffusion equation and a second order hyperbolic equation and present a concise numerical example.

In the sequel of the chapter we turn our considerations to more complicated periodic structures. We consider in Section 5 the Helmholtz equation (2) without the source term $f(\mathbf{x})$ on an array that is periodic in one direction and perform a cell analysis. Next, we explain how an improvement to the approach of Yuan and Lu [73] can be achieved by introducing *Sommerfeld-to-Sommerfeld (StS) mappings*. Moreover, we construct an efficient and robust method for numerically evaluating these StS operators. In Section 6 we present an application of the methods of Section 5 to waveguide prob-

lems discussing concisely the so-called pass and stop bands. We consider in Section 7 the transient Schrödinger equation on a semi-infinite array periodic in one direction and show how our fast evaluation method of Section 5 computes the exact StS mapping very efficiently. In Section 8 we discuss the numerical simulation of the time-dependent Schrödinger equation in two space dimensions with a bi-periodic potential function containing a defect. In Section 9 we return to the model problem of the Helmholtz equation now posed on a semi–infinite periodic array. Afterwards, we propose two different methods: the extrapolation method (Section 10) that is based on the limiting absorption principle (LABP) and the asymptotic method (Section 11) based on a conjecture about the asymptotic behavior of an LABP solution. Our proposed algorithm combines the doubling technique of Section 5 (now for evaluating the operator related to infinite arrays) and the limiting procedure (letting $\varepsilon \to 0$) with the extrapolation technique. The numerical tests in Section 12 supports the validity of our basic conjecture on how to identify the traveling Bloch waves which are compatible with the LABP, since from the numerical point of view the asymptotic method presents the same results as the extrapolation method does.

2. THE IMPEDANCE EXPRESSION

We consider the *general second order ODE*

$$-\frac{d}{dx}\left(\frac{1}{m(x)}\frac{dy}{dx}\right)+V(x)y=\rho(x)zy, \quad \forall x\geq 0, \tag{3}$$

where z denotes a complex parameter whose value space is to be determined. We assume that the functions $m(x)$, $V(x)$ and $\rho(x)$ are all *S-periodic* in $[0,+\infty)$ and *centrally symmetric* in each period, i.e.,

$$\begin{aligned} m(x)=m(S-x), \quad V(x)=V(S-x),\\ \rho(x)=\rho(S-x), \quad \text{a.e.} x\in[0,S]. \end{aligned} \tag{4}$$

The symmetry condition (4) simply implies that the *even* extensions of these functions to the whole real axis are still S-periodic. Moreover, we assume that the functions $m(x)$, $V(x)$ and $\rho(x)$ are sufficiently smooth and bounded, i.e. there exist several constants M_0, M_1, V_0 and ρ_0, such that

$$\begin{aligned} 0<M_0\leq m(x)\leq M_1<+\infty, \quad V(x)\geq V_0,\\ \rho(x)\geq\rho_0>0, \quad \forall x\in[0,S]. \end{aligned}$$

By introducing the new variable

$$w=\frac{1}{m(x)}\frac{dy}{dx},$$

the second order ODE (3) is transformed into a *first order ODE system*

$$\frac{d}{dx}\begin{pmatrix} w\\ y\end{pmatrix}=\begin{pmatrix} 0 & V(x)-\rho(x)z\\ m(x) & 0\end{pmatrix}\begin{pmatrix} w\\ y\end{pmatrix}, \tag{5}$$

for $x\geq 0$. The first part of this chapter deals with the L^2-solution of (3) in $[0,+\infty)$. To be more precise, we want to analyze

1. for which parameter z does the general ODE (3) possess a nontrivial L^2-solution $y(x)$?
2. and in this case, is it possible to formulate a closed form of the impedance $I:=y'(0)/y(0)$, i.e. the quotient of Neumann data over Dirichlet data evaluated at $x=0$?

For any two points x_1 and x_2, the ODE system (5) uniquely determines a linear transformation from the two-dimensional vector space associated with x_1, to the same space associated with x_2. We identify this transformation with the 2-by-2 matrix $T(x_1,x_2)$, which is an orthogonal change of basis matrix with periodicity properties. This matrix T satisfies the same form of equation as (5), namely:

$$\frac{d}{dx}T(x_1,x)=\begin{pmatrix} 0 & V(x)-\rho(x)z\\ m(x) & 0\end{pmatrix}T(x_1,x), \tag{6}$$

for all $x_1\geq 0$ and $x\geq 0$.

Lemma 2. *The* transformation matrix T *possess the following properties:*

$$T(x,x)=I_{2\times 2}, \tag{7a}$$

$$\det T(x_1,x_2)=\det T(x_1,x_1)=1, \tag{7b}$$

$$T(x_2,x_3)T(x_1,x_2)=T(x_1,x_3), \tag{7c}$$

$$T(x_1+S,x_2+S)=T(x_1,x_2). \tag{7d}$$

Proof. We prove a more general result. Suppose

$$T'=AT,$$

where T is an n-by-n matrix-valued function. Then

$$(\det T)'=\operatorname{trace}(A)\det T.$$

We write T into a column of row vectors $T=(t_1^\top,\cdots,t_n^\top)^\top$.Then

$$(\det T)'=\sum_{k=1}^{n}\det((t_1^\top,\cdots,t_{k-1}^\top,(t_k^\top)',t_{k+1}^\top,\cdots,t_n^\top)^\top).$$

Since

$$(t_k^\top)'=\sum_{l=1}^{n}a_{kl}t_l^\top,$$

we have

$$
\begin{aligned}
&(\det T)' \\
&= \sum_{k=1}^{n} \det((t_1^\top, \cdots, t_{k-1}^\top, \sum_{l=1}^{n} a_{kl} t_l^\top, t_{k+1}^\top, \cdots, t_n^\top)^\top) \\
&= \sum_{k=1}^{n} \sum_{l=1}^{n} \det((t_1^\top, \cdots, t_{k-1}^\top, a_{kl} t_l^\top, t_{k+1}^\top, \cdots, t_n^\top)^\top) \\
&= \sum_{k=1}^{n} \sum_{l=1}^{n} \delta_{lk} \det((t_1^\top, \cdots, t_{k-1}^\top, a_{kl} t_l^\top, t_{k+1}^\top, \cdots, t_n^\top)^\top) \\
&= \sum_{k=1}^{n} a_{kk} \det((t_1^\top, \cdots, t_{k-1}^\top, t_k^\top, t_{k+1}^\top, \cdots, t_n^\top)^\top) \\
&= \operatorname{trace}(A) \det T.
\end{aligned}
$$

Here, A is a 2-by-2 matrix with zero diagonal. According to the above result we have

$$\det T(x_1, x_2) = \det T(x_1, x_1) = 1. \quad \square$$

We proceed with a small illustrating example showing that there might not exist any nontrivial L^2 solutions of the second order ODE (3) for some z in the complex plane.

Example 1. We assume for simplicity that all coefficients $m(x)$, $\rho(x)$ and $V(x)$ are constant, hence the problem is periodic with any period. E.g. setting

$$V = 0,\ \rho = 1,\ m = 1,\ z = 1$$

leads to a constant system matrix A in (5) or (6)

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Then the matrix $T(x_1, x_2)$ is simply $\mathrm{Exp}((x_2 - x_1)A)$, where Exp denotes the matrix exponential. The two eigenvalues of $T(0, S)$ have modulus 1 and thus prevent any nontrivial L^2 solution.
In fact, as revealed later in this section (see Figs. (1)-(3) and the related discussion, if z lies in one of the so-called *pass bands*, then there exists no nontrivial L^2 solution. In this constant coefficient example there is only one pass band $(0, +\infty)$ and $z = 1$ lies exactly in this interval.

2.1 The Impedance Expression

The next step of the construction of the ABC is to consider the polar form of the eigenvalue σ with modulus lower than 1 and express the L^2-bounded solution in order to finally extract the impedance condition. According to (7a), the matrix $T(0, S)$ has two eigenvalues $\sigma (\neq 0)$ and $1/\sigma$ with $|\sigma| \leq 1$. Their associated eigenvectors are denoted by $(c_+, d_+)^\top$ and $(c_-, d_-)^\top$. If $|\sigma| < 1$, then $T(0,x)(c_\pm, d_\pm)^\top$ yields two linearly independent solutions of the ODE system (5). By setting $\sigma = e^{\mu S}$ with $\mathrm{Re}\,\mu < 0$ it is straightforward to verify that $e^{\mp \mu x} T(0,x)(c_\pm, d_\pm)^\top$ are periodic functions. Therefore, we conclude that

$$y_+ := T(0,x)(c_+, d_+)^\top = e^{\mu x} e^{-\mu x} T(0,x)(c_+, d_+)^\top$$

is L^2-bounded, while

$$y_- := T(0,x)(c_-, d_-)^\top = e^{-\mu x} e^{\mu x} T(0,x)(c_-, d_-)^\top$$

is not. For the L^2-bounded solution y_+, the impedance I is thus given as

$$I := \frac{y_+'(0)}{y_+(0)} = m(0)\, \frac{c_+}{d_+}. \tag{8}$$

We remark that σ and $(c_+, d_+)^\top$ depend on z, and hence the impedance I also depends on z. In the sequel we will refer to σ as the *Floquet's factor* [9, 42, 51]. It typically reflects how fast the L^2-bounded solution of the ODE (3) decays to zero when x tends to $+\infty$: the smaller its modulus, the faster. Also note that $\sigma(\bar{z}) = \overline{\sigma(z)}$ and $I(\bar{z}) = \overline{I(z)}$ holds.
For any fixed z, the impedance $I = I(z)$ in (8) can be computed numerically with arbitrary high accuracy. First we solve the ODE system (5) to get $T(0, S)$. Then we compute σ and its associated eigenvector (c_+, d_+). Finally we use (8) to determine the impedance (cf. the impedance plots in Figs. (5), (6) for some values of z).
In general, the matrix $T(0, S)$ cannot be represented with a simple analytical expression in terms of the functions $m(x)$, $V(x)$ and $\rho(x)$. However, it can be computed sufficiently accurately by integrating the ODE (6) numerically (setting $x_1 = 0$) in the interval $[0, S]$ with the initial data $T(0,0) = I_{2\times 2}$. Since this is a standard task, the detailed discussion is omitted here.

2.2 Numerical Tests A,B and C

In the sequel we present three numerical examples

$$
\begin{aligned}
&\textbf{Case A:} \quad m(x) = \rho(x) = 1, \quad V(x) = 2\cos(2x); \\
&\textbf{Case B:} \quad m(x) = \rho(x) = 1 + \cos(2x)/5, \\
&\qquad\qquad\ \ V(x) = \cos(2x); \\
&\textbf{Case C:} \quad m(x) = \rho(x) = 1 + \cos(2x)/5, \\
&\qquad\qquad\ \ V(x) = \sin(2x).
\end{aligned}
$$

that provide an illustration of what can be expected for the computation of the eigenvalues showing

some of its expected properties. This is an important issue since it shows where the solution to the periodic equation is bounded in L^2.
Figs. (1)-(3) show the modulus of σ, which denotes the eigenvalue of $T(0,S)$ with a smaller modulus. We observe that apart from some intervals in the real axis, for any z in the complex plane, σ has a modulus less than 1, thus the second order ODE (3) has a nontrivial L^2-solution. Furthermore, it turns out that the ending points of these intervals are exactly the eigenvalues of the following *characteristic problem*:
Find $\lambda \in \mathbb{R}$ and a nontrivial $y \in C^1_{\text{per}}[0,2S]$, such that

$$-\frac{d}{dx}\left(\frac{1}{m(x)}\frac{dy}{dx}\right)+V(x)y=\rho(x)\lambda y. \qquad (9)$$

We note that the symmetry condition (4) is not necessary for the above statements (In fact Case C does not satisfy (4)). We admit that the above statements have not been proven up to this time, but a vast number of other numerical evidences also support their validity.
If the coefficient functions $m(x)$, $V(x)$ and $\rho(x)$ satisfy the symmetry condition (4), then the characteristic problem (9) has a nice property: all the eigenvalues can be classified into two different groups

$$a_1 < a_2 < a_3 < \dots \quad \text{and} \quad b_1 < b_2 < b_3 < \dots,$$

where the eigenvalues a_r are associated with even eigenfunctions, and b_r with odd eigenfunctions. Besides, it holds that

$$a_1 < \min(a_2,b_1) \le \max(a_2,b_1) < \min(a_3,b_2) < \dots$$

For the Schrödinger equation (SE) with a periodic cosine potential, a special case of (3) with $m(x) = \rho(x) = 1$ and $V(x) = 2q\cos(2x)$, Zheng formulated in [78] a *conjecture* upon the impedance expression

$$I_{SE}(z) = -\sqrt[+]{-z+a_1}\prod_{r=1}^{+\infty}\frac{\sqrt[+]{-z+a_{r+1}}}{\sqrt[+]{-z+b_r}}, \qquad (10)$$

$\operatorname{Im} z > 0$, where $\sqrt[+]{\cdot}$ denotes the branch of the square root with positive real part and the branch cut is set along the negative real axis. While the validity (10) was checked numerically in [78] the analytical proof was done recently by Zhang and Zheng in [76] Since formally $I_{SE}(\bar{z}) = \overline{I_{SE}(z)}$ for any z with $\operatorname{Im} z \neq 0$, it is thus tempting to generalize the above conjecture to our general second order ODE (3), i.e.,

$$I(z) = -\sqrt{m(0)\rho(0)}\,\sqrt[+]{-z+a_1}\cdot \cdot\prod_{r=1}^{+\infty}\frac{\sqrt[+]{-z+a_{r+1}}}{\sqrt[+]{-z+b_r}}, \quad \operatorname{Im} z \neq 0. \qquad (11)$$

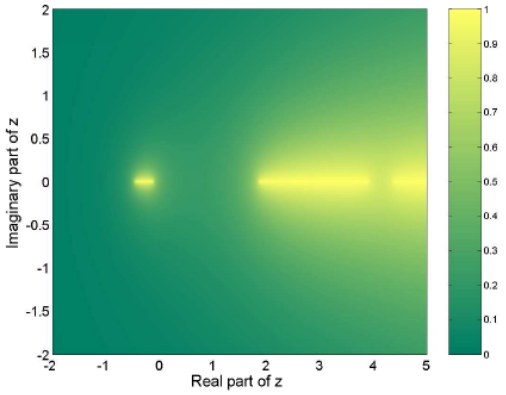


Fig. (1): **Case A:** Modulus of σ with respect to z.

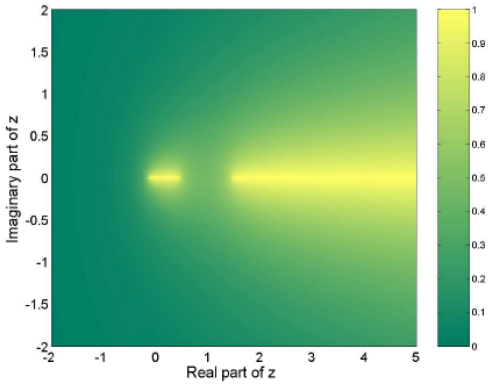


Fig. (2): **Case B:** Modulus of σ with respect to z.

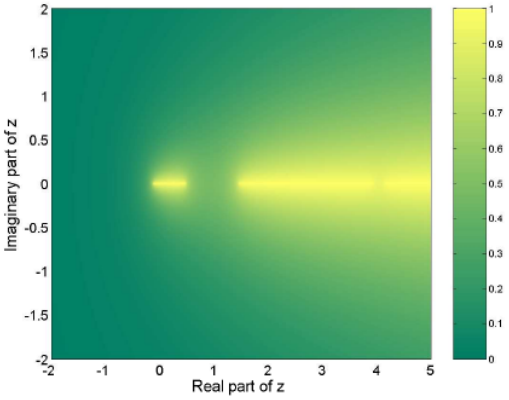


Fig. (3): **Case C:** Modulus of σ with respect to z.

Example 2. Let us briefly show how to obtain the constant coefficient case from the more general formula (11). The impedance for constant coefficients is given by

$$I(z) = -\sqrt{m\rho}\,\sqrt[+]{-z+\frac{V}{\rho}} = -\sqrt[+]{m(V-\rho z)}. \tag{12}$$

All the eigenvalues of (9) are

$$\lambda_n = \frac{(\frac{n\pi}{S})^2 + mV}{m\rho}.$$

The eigenspace of λ_0 is the set of constant functions. For $n > 0$, the eigenvalue λ_n is degenerate. Its eigenspace is two-dimensional, spanned by $\cos(\pi x/S)$ and $\sin(\pi x/S)$. Notice that cos is even and sin is odd. Thus we have

$$a_n = \lambda_{n-1},\ n \geq 1, \quad \text{and} \quad b_n = \lambda_n,\ n \geq 1.$$

Since $a_{r+1} = b_r$ for any $r \geq 1$, the equation (11) yields the correct impedance expression

$$I = -\sqrt{m\rho}\,\sqrt[+]{-z+a_1} = -\sqrt[+]{m(V-\rho z)}.$$

2.3 Numerical Tests D and E

Let us consider another two numerical tests:

Case D: $m(x) = \rho(x) = 1$,

$$V(x) = \sum_{n=-\infty}^{+\infty} e^{-16(x-\pi/2-n\pi)^2},$$

Case E: $m(x) = 1$,

$$V(x) = 0, \quad \rho(x) = 1+\cos(2x)/5.$$

Case D corresponds to the Schrödinger equation with a periodic Gaussian potential, cf. Fig. (4), and Case E could arise from a second order hyperbolic wave equation in a periodic medium.

Figs. (5) and (6) present the impedance function $I(z)$ when z is very close to the real axis. It can be clearly seen that the impedance turns out to be either real or purely imaginary. Those real intervals with purely imaginary impedance are exactly those values of z for which the ODE (3) has no nontrivial L^2-solution. Recall that this statement does not rely on the symmetry property of the coefficients (4). In the engineering literature these intervals are called *pass bands*, while their complementary intervals are called *stop bands*. This notion of 'pass' and 'stop' refers to allowing and preventing the existence of traveling wave solutions.

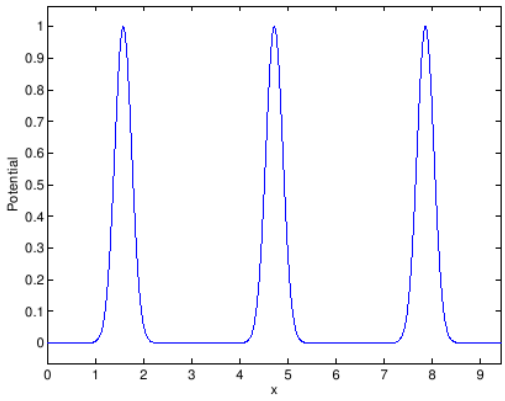


Fig. (4): Periodic Gaussian potential function $V(x) = \sum_{n=-\infty}^{+\infty} e^{-16(x-\pi/2-n\pi)^2}$.

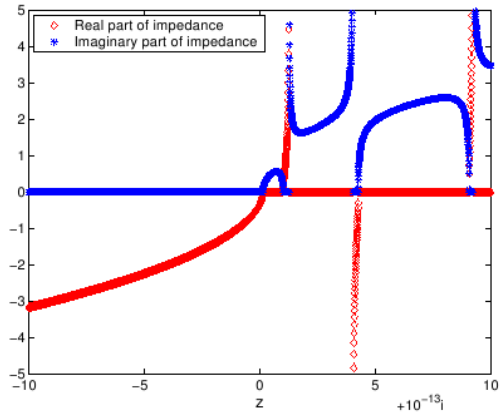


Fig. (5): **Case D:** Impedance $I(z)$ for the Schrödinger equation with a periodic Gaussian potential $V(x) = \sum_{n=-\infty}^{+\infty} e^{-16(x-\pi/2-n\pi)^2}$.

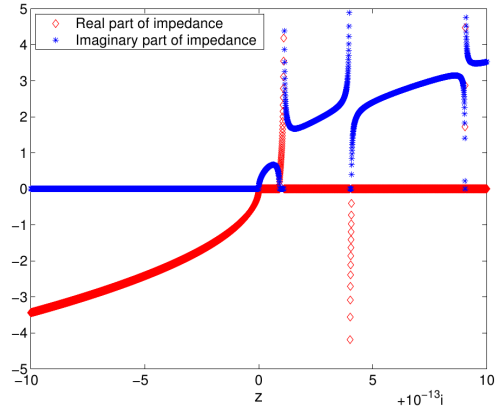


Fig. (6): **Case E:** Impedance plot for $m = 1$, $V = 0$ and $\rho = 1+\cos(2x)/5$.

Let us note that the impedance $I(z)$ becomes much more complicated as z approaches the real axis if one of the coefficient functions $m(x)$, $V(x)$ and $\rho(x)$ is not centrally symmetric, cf. (4).
Furthermore, we emphasize that the eigenvalues a_r and b_r can be computed with a high-accuracy solver for the characteristic problem (9). The first few eigenvalues are listed in Tables 1 and 2 with 6 digits. We observe that the relative difference between a_{r+1} and b_r decays quickly for increasing index r.

r	a_{r+1}	b_r	r	a_{r+1}	b_r
0	1.30811(-1)		7	4.91344(1)	4.91486(1)
1	1.00842(0)	1.26431(0)	8	6.41442(1)	6.41386(1)
2	4.25428(0)	4.03081(0)	9	8.11403(1)	8.11423(1)
3	9.06010(0)	9.22586(0)	10	1.00142(2)	1.00141(2)
4	1.61965(1)	1.60886(1)	11	1.21141(2)	1.21141(2)
5	2.51111(1)	2.51730(1)	12	1.44141(2)	1.44141(2)
6	3.61574(1)	3.61260(1)	13	1.69141(2)	1.69141(2)

Table 1: **Case D:** The first several eigenvalues of (9) with $m(x) = \rho(x) = 1$ and $V(x) = \sum_{n=-\infty}^{+\infty} e^{-16(x-\pi/2-n\pi)^2}$.

r	a_{r+1}	b_r	r	a_{r+1}	b_r
1	9.08164(-1)	1.10938	7	4.92536(1)	4.92537(1)
2	4.06748	3.98676	8	6.43296(1)	6.43296(1)
3	9.04010	9.06316	9	8.14157(1)	8.14157(1)
4	1.60896(1)	1.60838(1)	10	1.00512(2)	1.00512(2)
5	2.51315(1)	2.51328(1)	11	1.21618(2)	1.21618(2)
6	3.61880(1)	3.61877(1)	12	1.44735(2)	1.44735(2)

Table 2: **Case E:** The first few eigenvalues of (9), where $m(x) = 1$, $V(x) = 0$ and $\rho(x) = 1 + \cos(2x)/5$. Notice that $a_1 = 0$.

If the coefficient functions $m(x)$ and $\rho(x)$ are constant and $V(x) = 2q\cos(2x)$ with $q > 0$, then the general ODE (3) is reduced to the well-known *Mathieu's equation* [9, 51]. In this case, we obtain

$$a_1 < b_1 < a_2 < b_2 < a_3 < b_3 < \dots.$$

However, in general this property does not hold, and we can only expect the following

$$\begin{aligned} a_1 < \min(a_2,b_1) \le \max(a_2,b_1)) \\ < \min(a_3,b_2) \le \max(a_3,b_2) < \dots. \end{aligned}$$

Note that the stop bands are characterized as

$$\begin{aligned} (-\infty,a_1),\ (\min(a_2,b_1),\max(a_2,b_1)), \\ (\min(a_3,b_2),\max(a_3,b_2)),\ \dots \end{aligned}$$

and the pass bands are given by

$$\begin{aligned} (a_1,\min(a_2,b_1)),\ (\max(a_2,b_1),\min(a_3,b_2)), \\ (\max(a_3,b_2),\min(a_4,b_3)),\ \dots \end{aligned}$$

Now let us consider the expression (11) with the infinite product limited to R factors:

$$\begin{aligned} I_R(z) = -\sqrt{m(0)\rho(0)}\,\sqrt[+]{-z+a_1}\cdot \\ \cdot\prod_{r=1}^{R}\frac{\sqrt[+]{-z+a_{r+1}}}{\sqrt[+]{-z+b_r}}, \quad \operatorname{Im} z \neq 0. \end{aligned} \tag{13}$$

Figs. (7) and (8) show the maximum errors between the impedance $I(z)$ and $I_R(z)$ on 4001 equidistant points on three segments of the upper half complex plane. We detect that these errors become very small with increasing R. This observation has also been made for many other numerical tests.

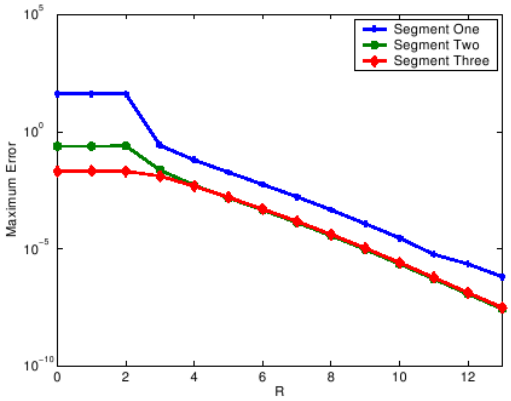


Fig. (7): **Case D:** Maximum error between the impedance $I(z)$ and $I_R(z)$. Segment One: $[-10,10]+10^{-13}i$. Segment Two: $[-10,10]+i$. Segment Three: $[-10,10]+10i$.

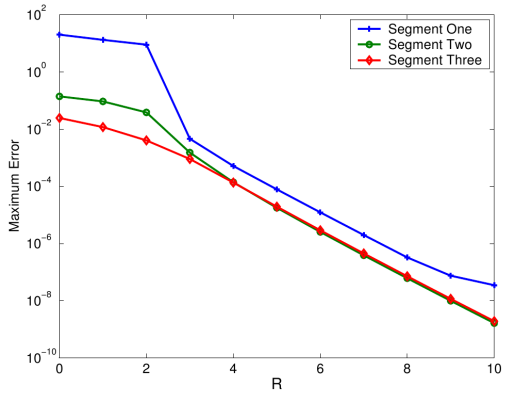


Fig. (8): **Case E:** Maximum error between the impedance $I(z)$ and $I_R(z)$. Segment One: $[-10,10]+10^{-13}i$. Segment Two: $[-10,10]+i$. Segment Three: $[-10,10]+10i$.

It is thus reasonable to conjecture that the limit of $I_R(z)$ as R tends to $+\infty$ is the impedance $I(z)$, i.e. the formula (11) states the correct impedance expression.
If $z = z_0$ is a real number, then the impedance expression (11) might not be well-defined. If z_0 lies in one of the stop bands, we already know that

$$\lim_{\varepsilon\to 0^+} \operatorname{Im} I(z_0+\varepsilon) = 0.$$

Due to the symmetry property of the impedance, i.e. $I(\bar{z}) = \overline{I(z)}$, we can define

$$I(z_0) = \lim_{\varepsilon\to 0^+} I(z_0 \pm \varepsilon).$$

Hence the impedance expression (11) still can be considered valid. If z_0 lies in one of the pass bands, the ODE (3) has no nontrivial bounded L^2-solution. In this case, we have to specify what kind of solution is really what we are seeking for. The impedance of this solution is thus the one-sided limit of $I(z_0+\varepsilon)$ as either $\varepsilon \to 0^+$ or $\varepsilon \to 0^-$. In most cases, this choice can be made naturally under physical considerations.
Let us finally remark that the impedance formulation was proven very recently by Zhang and Zheng [76].

3. BOUND STATES FOR THE SCHRÖDINGER OPERATOR

As a first application of the impedance expression (11), we consider the following *bound state problem for the Schrödinger operator*:
Find an energy $E \in \mathbb{R}$ and a nontrivial real function $u \in L^2(\mathbb{R})$, such that

$$-\frac{d^2u}{dx^2} + V(x)u = Eu, \quad x \in \mathbb{R}, \tag{14}$$

where

$$V(x) = \begin{cases} 2+2\cos(\pi x), & |x| > 1, \\ 0, & |x| < 1. \end{cases}$$

The potential function $V(x)$ is periodic in $\mathbb{R}\backslash(-1,1)$. In order to ensure that the solution u has a bounded L^2-norm, the energy E must be valued in the stop bands. The first few eigenvalues of the characteristic problem (9) with $m(x) = \rho(x) = 1$ and $V(x) = 2-2\cos(\pi x)$ (NOT $V(x) = 2+2\cos(\pi x)$) are listed in Table 3.
The first three stop bands are given by

$$(-\infty, 1.80087),\ (3.41926, 5.41414),\ (11.8359, 12.0349).$$

r	a_{r+1}	b_r	r	a_{r+1}	b_r
0	1.80087		3	2.42294(1)	2.42345(1)
1	3.41926	5.41414	4	4.14920(1)	4.14919(1)
2	1.20349(1)	1.18359(1)	5	6.36935(1)	6.36935(1)

Table 3: The first few eigenvalues of (9) with $m(x) = \rho(x) = 1$ and $V = 2-2\cos(\pi x)$.

If E is a bound state energy, then it must be an eigenvalue of the following *nonlinear characteristic problem*:
Find an energy $E \in \mathbb{R}$ and a nontrivial real function $u \in L^2(-1,1)$, such that

$$-\frac{d^2u}{dx^2} + V(x)u = Eu, \quad x \in (-1,1), \tag{15a}$$

$$-\frac{du}{dx}(-1) = I(E)u(-1), \tag{15b}$$

$$\frac{du}{dx}(1) = I(E)u(1). \tag{15c}$$

A direct discretization of the above problem (15) leads to a very complicated nonlinear algebraic equation with respect to E, and its solvability is not completely clear. Actually, the problem (15) is equivalent to the following *fixed point problem*.
For a given energy E we can solve the *linear characteristic problem*:
Find a function $\Phi(E) \in \mathbb{R}$ and a nontrivial real function $u \in L^2(-1,1)$, such that the following boundary value problem holds

$$-u_{xx} + V(x)u = \Phi(E)u, \quad x \in (-1,1), \tag{16a}$$

$$-\frac{du}{dx}(-1) = I(E)u(-1), \tag{16b}$$

$$\frac{du}{dx}(1) = I(E)u(1). \tag{16c}$$

The bound state energy thus satisfies $E = \Phi(E)$, i.e. E is a fixed point of the function $\Phi(E)$. Notice that $\Phi(E)$ is a multi-valued function and hence a series of bound states are expected.
Fig. (9) shows the first three branches of $\Phi(E)$ being restricted to $[-8,15]$. The time-harmonic Schrödinger equation is discretized by 50 eighth-order finite elements in $[-1,1]$. $I(E)$ is approximated by $I_{14}(E)$, which is equal to $I(E)$ within machine precision if $|E| < 20$. Three bound states exist in this energy range.
By performing the Newton-Steffenson iterations, the energies are found to be $E_0 = 0.642647$, $E_1 = 4.88651$ and $E_2 = 12.0164$. Our computations show that these values do not change within 6 digits by refining the finite element mesh.

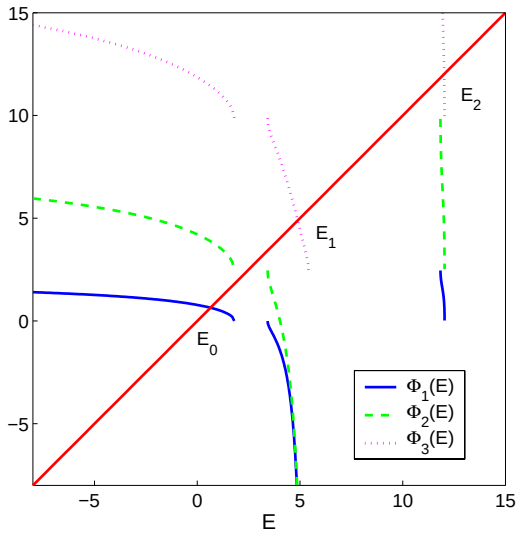


Fig. (9): The first three branches of $\Phi(E)$ being restricted to $[-8,15]$: $E_0 = 6.42647(-1)$. $E_1 = 4.88651$. $E_2 = 1.20164(1)$.

Next, the bound state wave functions (that are not normalized) are plotted in the Fig. (10). We observe in Fig. (10) that the ground state E_0 is well-localized, while the second excited bound state E_2 is greatly delocalized.

This demonstrates the advantage of the artificial boundary method and especially our ABCs (16b)–(16c), since a direct domain truncation method necessitates a very large computational domain to ensure the approximating accuracy of the wave function.

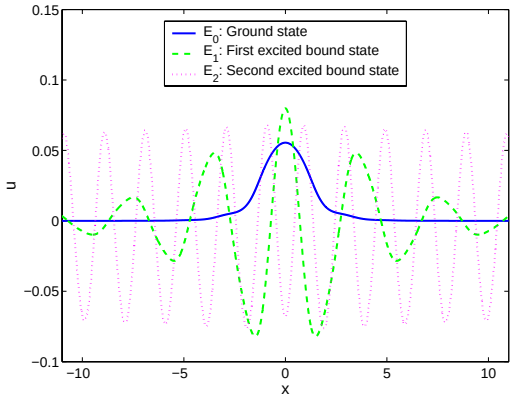


Fig. (10): The ground state E_0 and the first two excited bound states E_1, E_2.

4. EXACT ARTIFICIAL BOUNDARY CONDITIONS FOR TIME-DEPENDENT PROBLEMS

Based on the fundamental impedance expression (11), exact artificial boundary conditions can be derived for many time-dependent periodic structure problems, e.g., the Schrödinger equation (SE)

$$i\rho(x)\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}\left(\frac{1}{m(x)}\frac{\partial u}{\partial x}\right) = V(x)u,$$

the diffusion equation (DE)

$$\rho(x)\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(\frac{1}{m(x)}\frac{\partial u}{\partial x}\right) - L(x)u,$$

and the second order hyperbolic equation (HE)

$$\frac{\partial}{\partial x}\left(\frac{1}{m(x)}\frac{\partial u}{\partial x}\right) - L(x)u = \rho(x)\frac{\partial^2 u}{\partial t^2}.$$

Here, the coefficients $V(x)$, $\rho(x)$, $m(x)$ and $L(x)$ are supposed to be centrally symmetric periodic functions at infinity. Moreover, $\rho(x)$ and $m(x)$ are positive, and $L(x)$ is nonnegative. Now the free parameter involved during the derivation of the impedance operator for stationary problems plays the role of the Laplace variable s. The impedances for these three equations are given by

$$I_{\mathrm{SE}}(is) = -\sqrt{m(0)\rho(0)}\,\sqrt[+]{-is+a_1}\cdot \cdot\prod_{r=1}^{+\infty}\frac{\sqrt[+]{-is+a_{r+1}}}{\sqrt[+]{-is+b_r}}, \quad (17)$$

and

$$I_{\mathrm{DE}}(-s) = -\sqrt{m(0)\rho(0)}\,\sqrt[+]{s+a_1}\cdot \cdot\prod_{r=1}^{+\infty}\frac{\sqrt[+]{s+a_{r+1}}}{\sqrt[+]{s+b_r}}, \quad (18)$$

and

$$I_{\mathrm{HE}}(-s^2) = -\sqrt{m(0)\rho(0)}\,\sqrt[+]{s^2+a_1}\cdot \cdot\prod_{r=1}^{+\infty}\frac{\sqrt[+]{s^2+a_{r+1}}}{\sqrt[+]{s^2+b_r}}. \quad (19)$$

In equations (17)-(19) the variable s with $\mathrm{Re}\, s > 0$ denotes the free argument in the Laplace domain. Notice that due to our assumption, all coefficients a_r and b_r in (18) and (19) are nonnegative and thus the formulas (18), (19) are well-defined. The numerical solution to the Schrödinger equation in conjunction with the ABC (17) has been investigated in [78]. Similar techniques can be used for the diffusion equation with the ABC (18) with minor modifications.

4.1 A second order hyperbolic equation in 2D

We consider the propagation of electromagnetic waves in a waveguide with cavity, cf. the schematic map Fig. (11). For a TM polarized electromagnetic wave, the electric field E is governed by the equation

$$\frac{\partial^2 E}{\partial x^2}+\frac{\partial^2 E}{\partial z^2}-\frac{\varepsilon(x,z)}{c^2}\frac{\partial^2 E}{\partial t^2}=0. \qquad (20)$$

The relative dielectric permittivity ε, depending only on x after the artificial boundary, is supposed to be periodic. We assume that this waveguide is enclosed with a perfect conductor and hence we have a homogeneous Dirichlet boundary condition $E=0$ on the physical boundary.

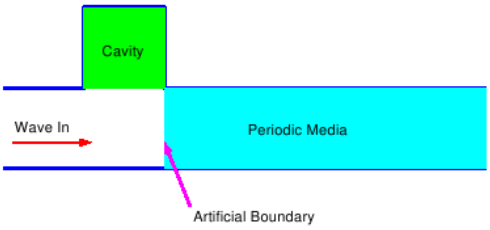


Fig. (11): Schematic of a waveguide with cavity.

On the semi-infinite slab region $[0,+\infty)\times[0,1]$, the characteristic decomposition can be applied with respect to the z variable. The eigenvalues are given by $n^2\pi^2$ and the eigenfunctions are $\sin(n\pi z)$, $n\geq 1$. An exact ABC in the frequency domain is thus set up as

$$\hat{E}_x^n(0,s)=-\frac{\sqrt{\varepsilon(0)}}{c}\sqrt[+]{s^2+a_1^n}\cdot\prod_{r=1}^{\infty}\frac{\sqrt[+]{s^2+a_{r+1}^n}}{\sqrt[+]{s^2+b_r^n}}\hat{E}^n(0,s),\quad n\geq 1. \qquad (21)$$

Here, $\hat{E}^n(x,s)$ denotes the n-th mode of $\hat{E}(x,z,s)$ in the z-direction defined as

$$\hat{E}^n(x,s)=2\int_0^1\hat{E}(x,z,s)\sin(n\pi z)dz,\quad x\geq 0,\ n\geq 1.$$

$\hat{E}(x,z,s)$ is determined by $\hat{E}^n(x,s)$ as

$$\hat{E}(x,z,s)=\sum_{n=1}^{+\infty}\hat{E}^n(x,s)\sin(n\pi z),\quad x\geq 0.$$

The constants a_r^n and b_r^n in (21) are the eigenvalues of the characteristic problem (9) with the coefficients $m(x)=1$, $V(x)=n^2\pi^2$ and $\rho(x)=\varepsilon(x)/c^2$.

By setting

$$\hat{w}_k^n(s)=\prod_{r=k}^{\infty}\frac{\sqrt[+]{s^2+a_{r+1}^n}}{\sqrt[+]{s^2+b_r^n}}\hat{E}^n(0,s),\quad k\geq 1,\ n\geq 1,$$

we get the recursion relation

$$\sqrt[+]{s^2+b_k^n}\,\hat{w}_k^n(s)=\sqrt[+]{s^2+a_{k+1}^n}\,\hat{w}_{k+1}^n(s),$$

$k\geq 1$, $n\geq 1$ and (21) reads

$$\hat{E}_x^n(0,s)=-\frac{\sqrt{\varepsilon(0)}}{c}\sqrt[+]{s^2+a_1^n}\,\hat{w}_1^n(s), \qquad (22)$$

$n\geq 1$. Returning to the physical domain yields

$$\frac{dw_k^n}{dt}=\frac{dw_{k+1}^n}{dt}+\frac{\sqrt{a_{k+1}^n}J_1(\sqrt{a_{k+1}^n}t)}{t}*w_{k+1}^n-\frac{\sqrt{b_k^n}J_1(\sqrt{b_k^n}t)}{t}*w_k^n,\ k\geq 1,\ n\geq 0,$$

and from (22) we obtain

$$\begin{aligned}\frac{\partial E^n}{\partial x}(0,t)&=-\frac{\sqrt{\varepsilon(0)}}{c}\Big(\frac{dw_1^n}{dt}+\frac{\sqrt{a_1^n}J_1(\sqrt{a_1^n}t)}{t}*w_1^n\Big)\\&=-\frac{\sqrt{\varepsilon(0)}}{c}\Big(\frac{\partial E^n}{\partial t}(0,t)\\&\quad+\sum_{k=0}^{+\infty}\frac{\sqrt{a_{k+1}^n}J_1(\sqrt{a_{k+1}^n}t)}{t}*w_{k+1}^n\\&\quad-\sum_{k=1}^{+\infty}\frac{\sqrt{b_k^n}J_1(\sqrt{b_k^n}t)}{t}*w_k^n\Big).\end{aligned} \qquad (23)$$

Here, $*$ denotes a convolution with respect to the time variable t and J_1 is the Bessel function of first order. In a real implementation the infinite summation in (23) has to be truncated by only keeping the first K_n terms:

$$\begin{aligned}\frac{\partial E^n}{\partial x}(0,t)&=-\frac{\sqrt{\varepsilon(0)}}{c}\Big(\frac{\partial E^n}{\partial t}(0,t)\\&\quad+\sum_{k=0}^{K_n}\frac{\sqrt{a_{k+1}^n}J_1(\sqrt{a_{k+1}^n}t)}{t}*w_{k+1}^n\\&\quad-\sum_{k=1}^{K_n}\frac{\sqrt{b_k^n}J_1(\sqrt{b_k^n}t)}{t}*w_k^n\Big),\end{aligned} \qquad (24)$$

and

$$w_{K_n+1}^n(t)=E^n(0,t).$$

If we want to resolve the n-th mode in the z-direction, we typically set $K_n\geq 0$. In order to ensure

the approximating accuracy of the ABC, K_n should be increased for larger values of n. Of course, if we are not interested in the n-th mode at all, we only need to set $K_n = -1$. In the next numerical example, we set $K_n = 10$ for any $n = 0, 1 \ldots, N$, and $K_n = -1$ for any $n = N+1, \ldots$, where N denotes the number of modes in the z-direction we want to resolve.

4.2 Numerical Example

We now study the wave field generated by a *periodic disturbance* at the left physical boundary

$$E(-2,z,t) = \sin(\pi z) \sum_{n=0}^{+\infty} e^{-160(t-(n+0.5))^2}, \; z \in (0,1).$$

The wave speed is set to 1, and the dielectric permittivity ε is set to be

$$\varepsilon(x,z) = \begin{cases} 1 & ,x < 0, \\ 1.2 - 0.2\cos(2\pi x) & ,x > 0. \end{cases}$$

We limit our computational time interval to $[0,6]$. Due to the finite wave propagation speed (at most 1), we can compute a *reference solution* E_{ref} in a large domain $(-2,4) \times (0,1) \cup (-1,0) \times (1,2)$ with small mesh sizes $\Delta x = \Delta z = 0.00125$ and $\Delta t = 0.000625$. The leap-frog central difference scheme is employed in all the computations. We use the standard *fast evaluation technique* proposed by Alpert, Greengard and Hagstrom [2] (cf. also [77]) for the convolution operations involved in the ABC (24). The poles and weights are taken from the web page of Hagstrom. The relative L^2-error is defined as

$$\frac{||E_{\text{ref}}(\cdot,\cdot,t) - E_{\text{num}}(\cdot,\cdot,t)||_{L^2}}{||E_{\text{ref}}(\cdot,\cdot,6)||_{L^2}},$$

where E_{ref} stands for the reference solution, while E_{num} denotes the numerical solution.

In Figs. (12) and (13) we compare the numerical solutions with the reference solutions at two different time steps $t = 3$ and $t = 3$. No difference can be observed with eyes.

In Fig. (14) we depict the errors when different number of modes in the z-direction are used. The accuracy of the numerical solutions is greatly improved for large number of modes.

The error evolution with respect to the time t is shown in Fig. (15). At the initial stage, the wave does not reach the artificial boundary, thus the ABC has no influence on the numerical solutions. The error arises completely from the interior discretization. After a critical time point (almost $t = 2.5$), the artificial boundary condition comes into effect.

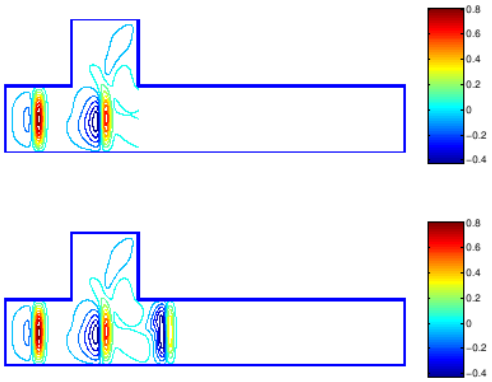


Fig. (12): Solutions at time $t = 3$. The number of modes is 10. The reference solution is obtained by taking $\Delta x = \Delta z = 0.00125$ and $\Delta t = 0.000625$.

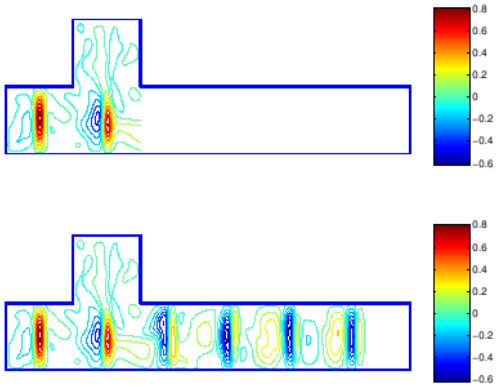


Fig. (13): Solutions at time $t = 6$. The number of modes is 10. The reference solution is obtained by taking $\Delta x = \Delta z = 0.00125$ and $\Delta t = 0.000625$.

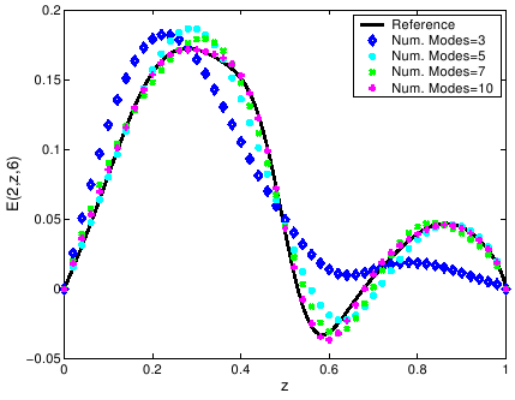


Fig. (14): Errors at time $t = 6$. $\Delta x = \Delta z = 0.02$. $\Delta t = 0.01$. The reference solution is obtained by taking $\Delta x = \Delta z = 0.00125$ and $\Delta t = 0.000625$. The line is $x = 0$.

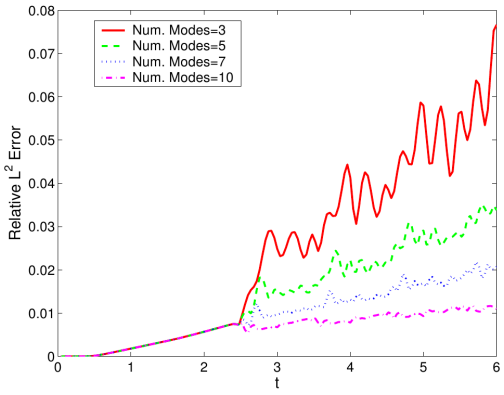


Fig. (15): Relative L^2 error. $\Delta x = \Delta z = 0.005$. $\Delta t = 0.0025$. The reference solution is obtained by taking $\Delta x = \Delta z = 0.00125$ and $\Delta t = 0.000625$.

We see that if enough number of modes are used, the error from the approximate boundary condition is nearly on the same level of interior discretization, which means the ABC is sufficiently accurate in this parameter regime.

Finally, we analyzed numerically in Fig. (16) the convergence rate of the relative L^2-errors at $t = 6$. Data-fitting reveals that the errors decay with an order of 1.851 in the parameter range $\Delta t \in [\frac{0.02}{7}, 0.01]$, when the number of modes in the z-direction is set to 10.

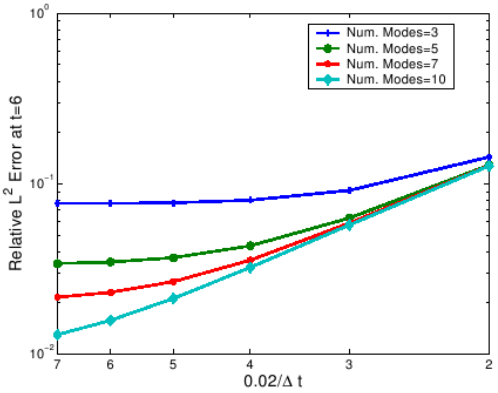


Fig. (16): Relative L^2 error. $\Delta x = \Delta z = 2\Delta t$. The reference solution is obtained by taking $\Delta x = \Delta z = 0.00125$ and $\Delta t = 0.000625$.

5. BOUNDARY MAPPINGS FOR PERIODIC ARRAYS

Let us consider the Helmholtz equation

$$-\Delta u(\mathbf{x}) + (V - z)u(\mathbf{x}) = f(\mathbf{x}). \qquad (25)$$

without source term, i.e., $V(\mathbf{x}) \equiv 0$ and $f(\mathbf{x}) \equiv 0$, on an array that is periodic in one direction, i.e. consisting of N identical cells as illustrated in Fig. (17).

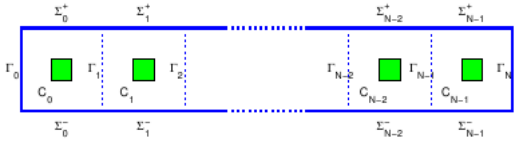


Fig. (17): Schematic view of a periodic array consisting of N cells.

We suppose that *appropriate* homogeneous linear boundary conditions are specified at the upper and lower, and the interior (if existing) boundaries, and these boundary conditions have the same periodicity consistent with that of the periodic structure. Here, "appropriate" means that these boundary conditions do not influence the well-posedness of the interior Helmholtz equation.
We define two *Sommerfeld mappings* of u as

$$\mathscr{G}_u^x = (\partial_x + \sqrt[+]{z})u, \qquad \mathscr{F}_u^x = (-\partial_x + \sqrt[+]{z})u.$$

It was proven in [20] that for given boundary data $\mathscr{F}_u^x$ on Γ_i and $\mathscr{G}_u^x$ on Γ_{i+k}, the Helmholtz equation (25), together with the boundary conditions on the upper and lower, and the interior (if existing) boundaries, is well-posed on the domain $\cup_{l=i}^{i+k-1} C_l$:

Lemma 3 (Lemma A.1. [20]). *The Helmholtz equation* (25) *is uniquely solvable in* $\cup_{l=i}^{i+k-1} C_l$ *for any* $f_i \in (H^{1/2}(\Gamma_i))'$, $g_k \in (H^{1/2}(\Gamma_k))'$ *and any* $z \in \mathbb{Z}\backslash\{0\}$

$$\begin{aligned}
&-\Delta u(\mathbf{x}) + zn^2 u(\mathbf{x}) = 0, \quad \mathbf{x} \in \Omega_{i,k} \equiv \cup_{l=i}^{i+k-1} C_l,\\
&\partial_y u(\mathbf{x}) = 0, \quad \mathbf{x} \in \cup_{i=0}^{N-1} \Sigma_i^{\pm},\\
&-\partial_x u(\mathbf{x}) + \sqrt[+]{z}u(\mathbf{x}) = f_i, \quad \mathbf{x} \in \Gamma_i,\\
&\partial_x u(\mathbf{x}) + \sqrt[+]{z}u(\mathbf{x}) = g_k, \quad \mathbf{x} \in \Gamma_k,
\end{aligned}$$

where $n \in C^1(\cup_{l\in\mathbb{Z}} C_l)$, C_l *and* $\Sigma_l^{\pm}$ *are defined as in Fig. (17).*

This implies that there exist four linear operators $\mathscr{A}_k$, $\mathscr{B}_k$, $\mathscr{C}_k$ and $\mathscr{D}_k$ satisfying

$$\begin{aligned}
\mathscr{G}_u^x|_{\Gamma_i} &= \mathscr{A}_k \mathscr{F}_u^x|_{\Gamma_i} + \mathscr{B}_k \mathscr{G}_u^x|_{\Gamma_{i+k}},\\
\mathscr{F}_u^x|_{\Gamma_{i+k}} &= \mathscr{C}_k \mathscr{F}_u^x|_{\Gamma_i} + \mathscr{D}_k \mathscr{G}_u^x|_{\Gamma_{i+k}}.
\end{aligned} \qquad (26)$$

Numerically, these operators can be derived by an appropriate spatial discretization of the domain $\cup_{l=0}^{k-1} C_l$. But if k is big, a large number of unknowns would get involved, which leads to a high computational effort. Our task is now to design an efficient and robust algorithm for evaluating these operators.

5.1 The recursive doubling method

Suppose for $k \in \{m,n\}$, the four linear operators $\mathscr{A}_k$, $\mathscr{B}_k$, $\mathscr{C}_k$ and $\mathscr{D}_k$ have already been obtained. From (26) we obtain

$$\begin{aligned}
\mathscr{G}_u^x|_{\Gamma_i} &= \mathscr{A}_m(\mathscr{C}_n\mathscr{F}_u^x|_{\Gamma_{i-n}} + \mathscr{D}_n\mathscr{G}_u^x|_{\Gamma_i}) + \mathscr{B}_m\mathscr{G}_u^x|_{\Gamma_{i+m}},\\
\mathscr{F}_u^x|_{\Gamma_i} &= \mathscr{C}_n\mathscr{F}_u^x|_{\Gamma_{i-n}} + \mathscr{D}_n(\mathscr{A}_m\mathscr{F}_u^x|_{\Gamma_i} + \mathscr{B}_m\mathscr{G}_u^x|_{\Gamma_{i+m}}).
\end{aligned}$$

It is easy to prove that $I-\mathscr{A}_m\mathscr{D}_n$ and $I-\mathscr{D}_n\mathscr{A}_m$ (I denotes the identity operator) are invertible and thus

$$\begin{aligned}
\mathscr{G}_u^x|_{\Gamma_i} &= (I-\mathscr{A}_m\mathscr{D}_n)^{-1}\mathscr{A}_m\mathscr{C}_n\mathscr{F}_u^x|_{\Gamma_{i-n}}\\
&\quad + (I-\mathscr{A}_m\mathscr{D}_n)^{-1}\mathscr{B}_m\mathscr{G}_u^x|_{\Gamma_{i+m}},\\
\mathscr{F}_u^x|_{\Gamma_i} &= (I-\mathscr{D}_n\mathscr{A}_m)^{-1}\mathscr{C}_n\mathscr{F}_u^x|_{\Gamma_{i-n}}\\
&\quad + (I-\mathscr{D}_n\mathscr{A}_m)^{-1}\mathscr{D}_n\mathscr{B}_m\mathscr{G}_u^x|_{\Gamma_{i+m}}.
\end{aligned} \tag{27}$$

Substituting the above expressions into (26) gives

$$\begin{aligned}
\mathscr{G}_u^x|_{\Gamma_{i-n}} &= [\mathscr{A}_n + \mathscr{B}_n(I-\mathscr{A}_m\mathscr{D}_n)^{-1}\mathscr{A}_m\mathscr{C}_n]\mathscr{F}_u^x|_{\Gamma_{i-n}}\\
&\quad + \mathscr{B}_n(I-\mathscr{A}_m\mathscr{D}_n)^{-1}\mathscr{B}_m\mathscr{G}_u^x|_{\Gamma_{i+m}},\\
\mathscr{F}_u^x|_{\Gamma_{i+m}} &= \mathscr{C}_m(I-\mathscr{D}_n\mathscr{A}_m)^{-1}\mathscr{C}_n\mathscr{F}_u^x|_{\Gamma_{i-n}} + [\mathscr{D}_m\\
&\quad + \mathscr{C}_m(I-\mathscr{D}_n\mathscr{A}_m)^{-1}\mathscr{D}_n\mathscr{B}_m]\mathscr{G}_u^x|_{\Gamma_{i+m}},
\end{aligned}$$

which imply the relations

$$\begin{aligned}
\mathscr{A}_{m+n} &= \mathscr{A}_n + \mathscr{B}_n(I-\mathscr{A}_m\mathscr{D}_n)^{-1}\mathscr{A}_m\mathscr{C}_n,\\
\mathscr{B}_{m+n} &= \mathscr{B}_n(I-\mathscr{A}_m\mathscr{D}_n)^{-1}\mathscr{B}_m,\\
\mathscr{C}_{m+n} &= \mathscr{C}_m(I-\mathscr{D}_n\mathscr{A}_m)^{-1}\mathscr{C}_n,\\
\mathscr{D}_{m+n} &= \mathscr{D}_m + \mathscr{C}_m(I-\mathscr{D}_n\mathscr{A}_m)^{-1}\mathscr{D}_n\mathscr{B}_m.
\end{aligned} \tag{28}$$

Hence, for any fixed cell number N, the operators $\mathscr{A}_N$, $\mathscr{B}_N$, $\mathscr{C}_N$, and $\mathscr{D}_N$ can be obtained by the following steps:

1. Derive $\mathscr{A}_1$, $\mathscr{B}_1$, $\mathscr{C}_1$, and $\mathscr{D}_1$ by the *cell analysis*. If $N=1$, it is done.

2. Write the number N into *binary form* $(j_L\cdots j_0)_2$, with $L=[\log_2 N]$ and $j_L=1$.

3. Use the doubling relations (28) L times by setting $m=n=2^{k-1}$ to get $\mathscr{A}_{2^k}$, $\mathscr{B}_{2^k}$, $\mathscr{C}_{2^k}$, and $\mathscr{D}_{2^k}$ for $k=1,\ldots,L$.

4. For $l=L-1,\ldots,0$, if $j_l\neq 0$, then use (28) by setting $m=(j_L\cdots j_{l+1}0\cdots 0)_2$ and $n=2^l$ to obtain $\mathscr{A}_{(j_L\cdots j_l0\cdots 0)_2}$, $\mathscr{B}_{(j_L\cdots j_l0\cdots 0)_2}$, $\mathscr{C}_{(j_L\cdots j_l0\cdots 0)_2}$ and $\mathscr{D}_{(j_L\cdots j_l0\cdots 0)_2}$.

This procedure uses (28) at most $2[\log_2 N]$ times.
Given the boundary data $\mathscr{F}_u^x|_{\Gamma_0}$ and $\mathscr{G}_u^x|_{\Gamma_N}$, in some cases it is necessary to obtain other data in a subdomain of $\cup_{l=0}^{N-1}C_l$, for example, $\mathscr{F}_u^y|_{\Sigma^-}$ and $\mathscr{G}_u^y|_{\Sigma^-}$ where $\Sigma^- = \cup_{i=0}^{N-1}\Sigma_i^-$. We need only to compute all $\mathscr{F}_u^x|_{\Gamma_i}$ and $\mathscr{G}_u^x|_{\Gamma_{i+1}}$ since for each i they completely determine the function u restricted to C_i. If N happens to be a power of 2, say $N=2^L$, this can be achieved efficiently with the following algorithm:
For $p=L,\ldots,1$ and $k=0,\ldots,2^{L-p}-1$, compute $\mathscr{G}_u^x|_{\Gamma_{k2^p+2^{p-1}}}$ and $\mathscr{F}_u^x|_{\Gamma_{k2^p+2^{p-1}}}$ using (27) by setting $i=k2^p+2^{p-1}$ and $n=m=2^{p-1}$.
For a general cell number N, we proceed in the following way:

1. Write N into binary form $(j_L\cdots j_0)_2$, with $L=[\log_2 N]$ and $j_L=1$;

2. For $l=0,\ldots,L$, if $j_l\neq 0$, compute $\mathscr{A}_k$, $\mathscr{B}_k$, $\mathscr{C}_k$, and $\mathscr{D}_k$ for $k=(j_l\cdots j_0)_2$ and $k=N-(j_l\cdots j_0)_2$, and use (27) by replacing i,n with $(j_l\cdots j_0)_2$ and m with $N-(j_l\cdots j_0)_2$ to derive $\mathscr{G}_u^x|_{\Gamma_{(j_l\cdots j_0)_2}}$ and $\mathscr{F}_u^x|_{\Gamma_{(j_l\cdots j_0)_2}}$. Then use the algorithm above for a power of 2 to derive $\mathscr{G}_u^x|_{\Gamma_i}$ and $\mathscr{F}_u^x|_{\Gamma_i}$ for any $i=(j_k\cdots j_0)_2+1,\cdots,(j_l\cdots j_0)_2-1$, where k is the largest number satisfying $k<l$ and $j_k\neq 0$.

For any $i=1,\ldots,N-1$, the above algorithm uses (28) at most $2[\log_2 N]$ times and (27) at most $[\log_2 N]$ times. After all $\mathscr{F}_u^x|_{\Gamma_i}$ and $\mathscr{G}_u^x|_{\Gamma_i}$ are derived, $\mathscr{F}_u^y|_{\Sigma_i^-}$ and $\mathscr{G}_u^y|_{\Sigma_i^-}$ are then obtained by the cell analysis. The final results can be written into the following form

$$\begin{aligned}
\mathscr{G}_u^y|_{\Sigma^-} &= (\mathscr{F}\to\mathscr{G})\mathscr{F}_u^x|_{\Gamma_0} + (\mathscr{G}\to\mathscr{G})\mathscr{G}_u^x|_{\Gamma_N},\\
\mathscr{F}_u^y|_{\Sigma^-} &= (\mathscr{F}\to\mathscr{F})\mathscr{F}_u^x|_{\Gamma_0} + (\mathscr{G}\to\mathscr{F})\mathscr{G}_u^x|_{\Gamma_N}.
\end{aligned}$$

Here $(\mathscr{F}\to\mathscr{G})$, $(\mathscr{G}\to\mathscr{G})$, $(\mathscr{F}\to\mathscr{F})$ and $(\mathscr{G}\to\mathscr{F})$ are four linear operators defined in suitable distributional spaces.

Remark. If the boundary condition on Γ_N is given as a *Sommerfeld-to-Sommerfeld* (StS) mapping

$$\mathscr{G}_u^x|_{\Gamma_N} = \mathscr{E}_N\mathscr{F}_u^x|_{\Gamma_N} + S_N, \tag{29}$$

where $\mathscr{E}_N$ is a linear operator and S_N is a function defined on Γ_N, we have

$$\begin{aligned}
\mathscr{G}_u^x|_{\Gamma_N} &= \mathscr{E}_N(I-\mathscr{D}_N\mathscr{E}_N)^{-1}\mathscr{C}_N\mathscr{F}_u^x|_{\Gamma_0}\\
&\quad + [I+\mathscr{E}_N(I-\mathscr{D}_N\mathscr{E}_N)^{-1}\mathscr{D}_N]S_N,
\end{aligned}$$

and

$$\mathscr{G}_u^x|_{\Gamma_0} = [\mathscr{A}_N + \mathscr{B}_N\mathscr{E}_N(I-\mathscr{D}_N\mathscr{E}_N)^{-1}\mathscr{C}_N]\mathscr{F}_u^x|_{\Gamma_0} + [\mathscr{B}_N + \mathscr{B}_N\mathscr{E}_N(I-\mathscr{D}_N\mathscr{E}_N)^{-1}\mathscr{D}_N]S_N. \quad (30)$$

The invertibility of $I-\mathscr{D}_N\mathscr{E}_N$ is obvious if the periodic array problem is well-posed with the StS boundary mapping (29) on Γ_N. This expression (30) yields an exact StS mapping at the leftmost boundary Γ_0. Furthermore, if the Dirichlet-to-Neumann (DtN) mapping is well-defined on Γ_0, it can be derived straightforwardly from (30).

Remark. Recently, Yuan and Lu [73] proposed an analogous technique for deriving the exact DtN mapping. In their cell analysis, instead of using Sommerfeld data on Γ_i and Γ_{i+1}, they used Dirichlet data to determine Neumann data. A problem will appear if $-z$ happens to be one of the eigenvalues of the operator $-\Delta$ on $\cup_{k=i}^{i+2^J-1}C_k$ for some J with homogeneous Dirichlet boundary conditions specified on Γ_i and Γ_{i+2^J}, since in this case, the Dirichlet-to-Neumann (DtN) mapping does not exist at all. One might argue that the probability for this to happen is very small, but if the total number of periodic cells is large, the eigenvalues of $-\Delta$ with Dirichlet boundary conditions are very dense in the pass bands. This implies that if there are some eigenvalues of $-\Delta$ very close but never equal to s, though the DtN mapping exists, it is very ill-conditioned.

6. APPLICATION TO WAVEGUIDES

Here we present a first application of the proposed technique to waveguide problems. Consider the Helmholtz equation in the waveguide shown in Fig. (18). The domain between Γ_1 and Γ_2 consists of four periodic cells. Each cell has a size of 1×2 with a hole of 0.5×1 in the center. The domain between Γ_3 and Γ_4 also contains four periodic cells. Each cell has a size of 1×1 with a hole of 0.5×0.5. These two periodic structures are joined with a junction region between Γ_2 and Γ_3. The domains left to Γ_1 and right to Γ_4 are homogeneous.

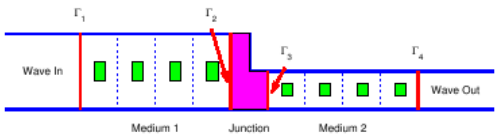


Fig. (18): Schematic of a model waveguide. Two waveguides with different periodic materials are joined with a junction zone between Γ_2 and Γ_3.

The governing equation is the Helmholtz equation (25) without source term $f(\mathbf{x})$ and $z(\mathbf{x})\equiv -k^2$, i.e.

$$\Delta u + k^2 u = 0, \quad (31)$$

where $k>0$ is the real wave number. Zero Dirichlet data is specified on the interior boundaries, and zero Neumann data on the top and bottom boundaries. A plane wave $u_0(x,y)=e^{-ikx}$ is traveling in the waveguide from the left side. It is well-known that the *disturbance part* $u-u_0$ satisfies the left-going boundary condition on Γ_1, i.e.,

$$\frac{\partial}{\partial x}(u-u_0) = \sqrt[+]{-\partial_y^2-k^2}\,(u-u_0), \quad (x,y)\in\Gamma_1,$$

or equivalently in the form of StS mapping,

$$\mathscr{F}_u^x = \frac{ik-\sqrt[+]{-\partial_y^2-k^2}}{ik+\sqrt[+]{-\partial_y^2-k^2}}\mathscr{G}_u^x + 2iku_0, \quad (x,y)\in\Gamma_1. \quad (32)$$

The wave function u satisfies the right-going boundary condition on Γ_4, i.e.,

$$\frac{\partial u}{\partial x} = -\sqrt[+]{-\partial_y^2-k^2}\,u, \quad (x,y)\in\Gamma_4,$$

or equivalently,

$$\mathscr{G}_u^x = \frac{ik-\sqrt[+]{-\partial_y^2-k^2}}{ik+\sqrt[+]{-\partial_y^2-k^2}}\mathscr{F}_u^x, \quad (x,y)\in\Gamma_4. \quad (33)$$

Now by using the technique in the last section, we could derive the StS mapping on Γ_2 and Γ_3. The wave function is then resolved by solving the Helmholtz equation *only* in the junction region between Γ_2 and Γ_3.

6.1 Band Structure Diagrams

To understand the typical wave behaviour in periodic waveguides we must consider the band structure diagrams of the characteristic equation $-\Delta u=\lambda u$ restricted to a single periodic cell. As assumed, the top and bottom boundary conditions are homogeneous Neumann, and the interior boundary condition is homogeneous Dirichlet. The boundary conditions at the left and right boundaries of the single cell are *pseudoperiodic*, namely,

$$u_{\text{right}} = e^{i\theta}u_{\text{left}}, \qquad \partial_x u_{\text{right}} = e^{i\theta}\partial_x u_{\text{left}},$$

where the parameter θ lies in the interval $[0,2\pi]$. For each value of θ, there exists a sequence of

real eigenvalues λ that are shown in the following *band structure diagrams*. These eigenvalues, also regarded as *discrete energies*, correspond to a series of Bloch waves which could travel through the waveguides without damping.
Figs. (19) and (20) show these band structure diagrams for the two periodic structures to the left and to the right.

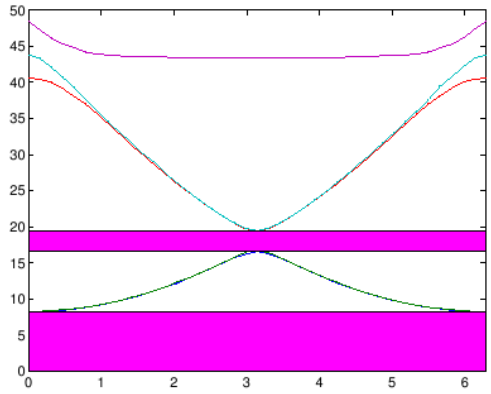


Fig. (19): Band structure with stop bands for the left periodic structure. The first two stop bands are the intervals $(-\infty, 8.27_{\pm 0.01})$ and $(16.69_{\pm 0.01}, 19.49_{\pm 0.01})$.

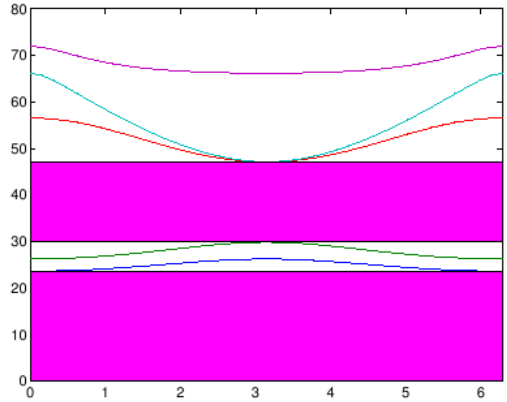


Fig. (20): Band structure with stop bands for the right periodic structure. The first two stop bands are the intervals $(-\infty, 23.61_{\pm 0.01})$ and $(29.85_{\pm 0.01}, 47.10 \pm 0.01)$.

The results are obtained by an eighth-order finite element discretization using the step sizes $\Delta x = \Delta y = 0.125$. For the left periodic structure between Γ_1 and Γ_2, the first two stop bands are $(-\infty, 8.27_{\pm 0.01})$ and $(16.69_{\pm 0.01}, 19.49_{\pm 0.01})$, while for the right periodic structure between Γ_3 and Γ_4, they are $(-\infty, 23.61_{\pm 0.01})$ and $(29.85_{\pm 0.01}, 47.10_{\pm 0.01})$. The first eigenvalue of the Dirichlet boundary value problem for the left periodic structure is $19.49_{\pm 0.01}$, while the first eigenvalue is $47.10_{\pm 0.01}$ for the right periodic structure.

We consider in the sequel *five cases*: $k = \sqrt{8}$, $k = \sqrt{19.49}$, $k = 6$, $k = \sqrt{47.10}$ and $k = 8$.

1. $k^2 = 8$ lies in stop bands of both two structures.
2. $k^2 = 19.49$ is the first eigenvalue of the Dirichlet boundary value problem for the left periodic structure.
3. $k^2 = 36$ lies in pass bands of both two structures.
4. $k^2 = 47.10$ is the first eigenvalue of the Dirichlet boundary value problem of the right periodic structure.
5. $k^2 = 64$ lies in pass bands of both two structures.

We point out the fact that the cases $k = \sqrt{19.49}$ and $k = \sqrt{47.10}$ cannot be solved with Yuan and Lu's method [73]. Figs. (21)-(25) show the real part of the wave function for the five chosen wave numbers. Again an eighth-order finite element code was used in the computation with the step sizes $\Delta x = \Delta y = 0.125$.

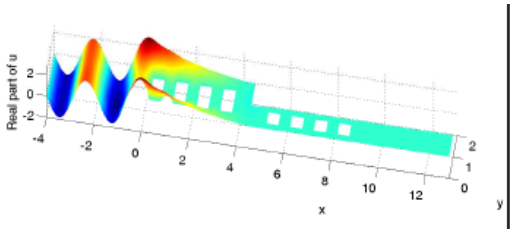


Fig. (21): Real part of the wave function for $k = \sqrt{8}$.

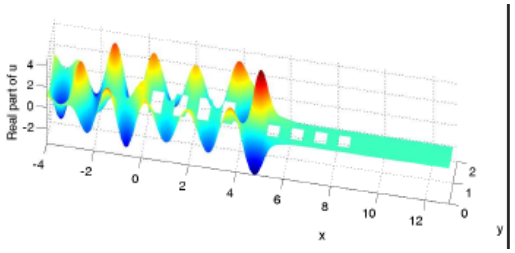


Fig. (22): Real part of the wave function for $k = \sqrt{19.49}$.

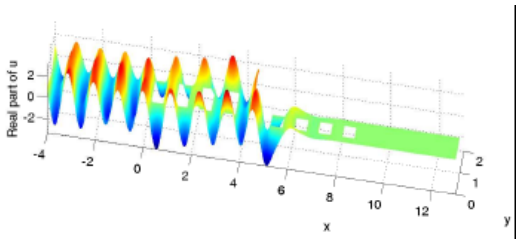


Fig. (23): Real part of the wave function for $k = 6$.

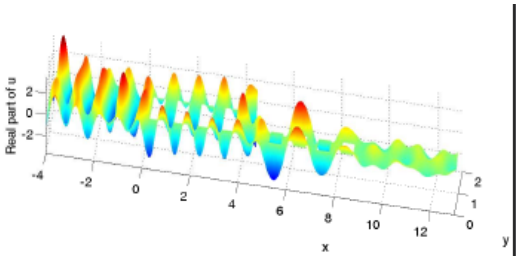


Fig. (24): Real part of the wave function for $k = \sqrt{47.10}$.

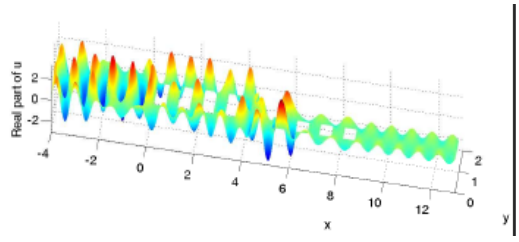


Fig. (25): Real part of the wave function for $k = 8$.

7. EXACT StS MAPPING FOR SEMI-INFINITE PERIODIC PROBLEMS

In many cases the exact StS (thus DtN or DtS) mapping is necessary to handle semi-infinite periodic array problems properly, see Fig. (26). Recently, Joly, Li and Fliss [40] presented a Newton-type method for the Helmholtz equation when z has a nonzero imaginary part. In this case any outgoing wave decays to zero exponentially fast at infinity. In the previous section, we have proposed a fast algorithm within $O(\log_2 N)$ operations for computing the exact StS mapping

$$g_0 = \mathscr{A}_N f_0 + \mathscr{B}_N g_N.$$

If the solution decays in one periodic cell with a factor of σ, by setting $N = \lceil -\frac{\ln \varepsilon}{\sigma} \rceil$, it is hopeful that $\mathscr{A}_N$ gives an approximation of the exact StS mapping on Γ_0 with an error of $O(\varepsilon)$. Here ε denotes the machine precision.

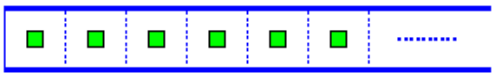


Fig. (26): Schematic view of a semi-infinite periodic array. Each cell has a size of 1×1, and a hole of 0.5×0.5 lies in the center.

It turns out that if $\operatorname{Im} z \neq 0$, or z is real but in the stop bands, the operator $\mathscr{A}_N$ converges with an exponential rate to the exact StS operator. In Fig. (27), we plot the relative errors of $\mathscr{A}_N$ w.r.t. $\mathscr{A}_{\text{ref}}$, which is obtained by setting $N = 1024$. In the computation we set $\Delta x = \Delta y = 0.125$ and use an eighth-order finite element method, thus in the discrete level A_N is expressed with a 65-by-65 matrix. Recall that $k^2 = 23,31$ are in stop bands, and $k^2 = 25,50$ in pass bands, cf. Fig. (20).

As a conclusion, using the doubling procedure of Section 5.1 at most $J = \lceil \log_2 \lceil -\frac{\ln \varepsilon}{\sigma} \rceil \rceil$ times gives the exact StS boundary mapping at the leftmost boundary up to machine precision. Our technique presents a very fast evaluation of the exact StS mapping.

If z lies in the stop bands, some traveling Floquet modes would appear, and the above argument ceases to hold. To obtain a well-posed PDE problem, we have to specify the outgoing waves and incoming waves. In a recent work of Joly et al. [40] a method is proposed to resolve this problem. However, we will not discuss this issue in this chapter.

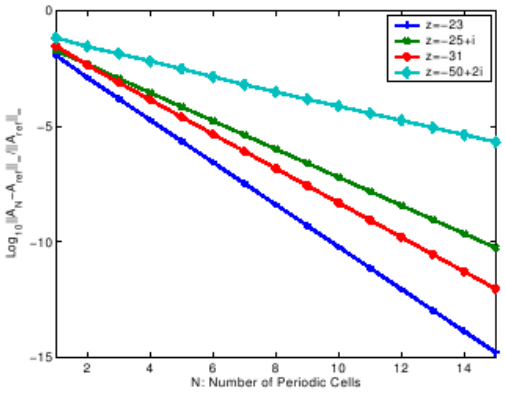


Fig. (27): Convergence of the StS mapping.

7.1 Application to the time-dependent Schrödinger equation

As an application, we consider the linear time-dependent Schrödinger equation

$$iu_t + u_{xx} = Vu, \quad x \in \mathbb{R}. \tag{34}$$

The initial data $u_0(x)$ is chosen as

$$u_0(x) = \exp(-x^2 + ik_0 x)$$

and the potential function V is set to

$$V(x) = \sum_{n\in\mathbb{Z}, n\neq 0} V_0 \exp\left(-(x-10n)^2\right).$$

In the Laplace domain the Schrödinger equation (34) is transformed into

$$-\hat{u}_{xx} + (V - is)\hat{u} = -iu_0, \quad x \in \mathbb{R}, \tag{35}$$

where Re $s > 0$ and $\hat{u}$ denotes the Laplace transformation of u defined by

$$\hat{u}(x,s) = \int_0^{+\infty} u(x,t)e^{-st}dt.$$

The function u_0 is well-supported in the interval $[-5,5]$. Outside of $[-5,5]$, the potential function V can be considered periodic with a period of 10. For any fixed s, the equation (35) can be solved in $[-5,5]$ with a high-order spatial discretization method. Here we use M eighth-order finite elements, which include $8M+1$ grid points. The StS boundary conditions at $x = \pm 5$ are derived by the method presented in the beginning of this section. The same number of grid points are used in the discrete periodic cell analysis.

The inverse Laplace transformation is evaluated numerically as

$$\begin{aligned} u(x,t) &= \frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} e^{st}\hat{u}(x,s)\,ds \\ &\approx \frac{1}{2\pi}\int_{-f_{\max}}^{f_{\max}} \chi(f)e^{(\gamma+if)t}\hat{u}(x,\gamma+if)\,df, \end{aligned} \tag{36}$$

and the integral is further approximated by the middle-point rule. Several parameters need to be tuned: the damping factor γ, the cutoff frequency f_{max}, and the number of quadrature points N_f. In principle, the bigger is γ, the smoother is the function $\hat{u}$, thus the number of quadrature points can be made smaller. But to guarantee stability γ cannot be too large. This is typically because there is an exponential factor $e^{\gamma t}$ involved in the integral. The *cutoff frequency* f_{max} depends on the regularity of the solution. The smoother is u, the smaller is f_{max}. We leave open the theoretical investigation on the optimal choice of these parameters in this chapter. For the considered model problem when $k_0 = 2$ and $M = 16$, we set

$$f_{max} = 200, \quad \gamma = 1, \quad N_f = 1024,$$

and the *filtering function* χ as

$$\chi(f) = \exp\left(-(1.2f/f_{max})^{20}\right).$$

If $V_0 = 0$, the exact solution is

$$u(x,t) = \sqrt{\frac{i}{-4t+i}}\exp\left(\frac{-ix^2 - k_0 x + k_0^2 t}{-4t+i}\right)..$$

The relative L^2-errors in the computational region $[-5,5]$ are listed in Table 4 at different time points. We observe that in this time regime the relative errors are very small. If $V_0 \neq 0$, the analytical exact solution is in general not available. In Fig. (28) we illustrate the solution at different time points for $V_0 = 10$. The dashed blue line shows the potential function scaled by $1/V_0$.

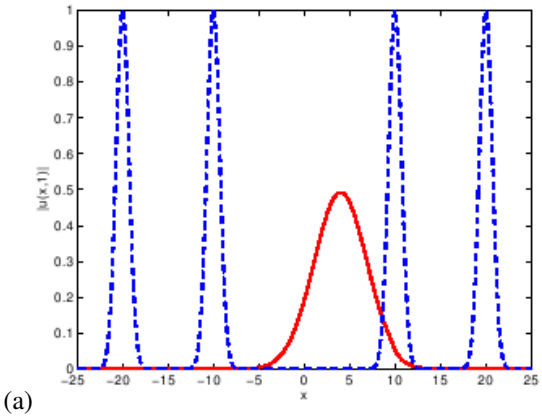
8. NUMERICAL SIMULATION OF THE 2D SCHRÖDINGER EQUATION

Here we consider the following two-dimensional time-dependent Schrödinger equation

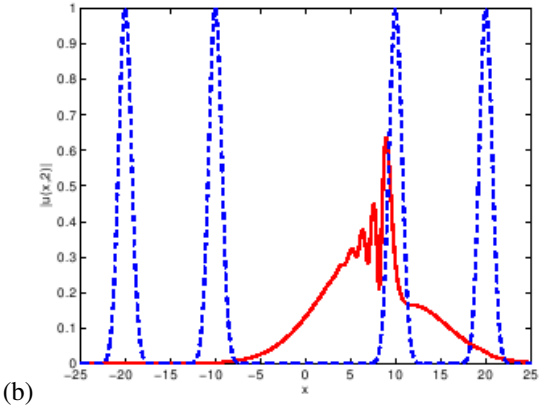
$$iu_t + u_{xx} + u_{yy} = Vu, \ \forall (x,y) \in \mathbb{R}^2, \ \forall t > 0, \tag{37}$$

$$u(x,y,0) = u_0(x,y), \ \forall (x,y) \in \mathbb{R}^2, \tag{38}$$

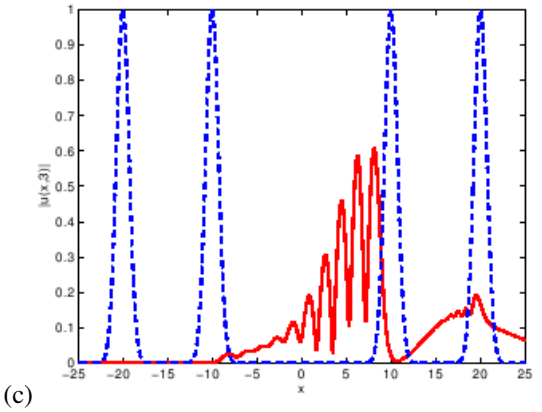
$$u(x,y,t) \to 0, \ r = \sqrt{x^2+y^2} \to +\infty, \ \forall t > 0. \tag{39}$$



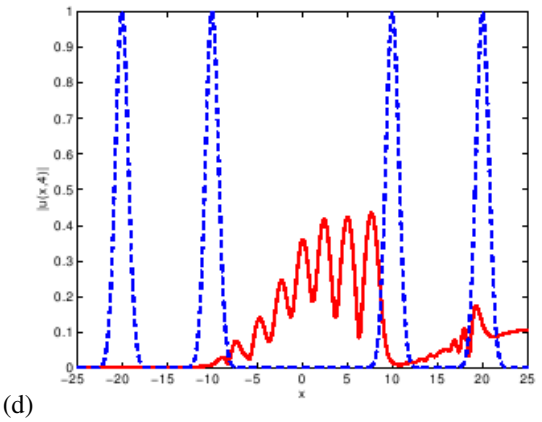
(a)



(b)



(c)



(d)

Fig. (28): Evolution of Gaussian packet in a periodic potential. (a) $t = 1$. (b) $t = 2$. (c) $t = 3$. (d) $t = 4$.

Time Point	Relative L^2-Error
1.0	1.83(-8)
1.5	2.12(-8)
2.0	2.66(-8)
2.5	3.14(-8)
3.0	3.56(-8)
3.5	3.92(-8)

Table 4: Relative L^2-errors in $[-5,5]$ at different time points for $V_0 = 0$.

The time evolution of the Gaussian wave packet is presented in Fig. (29). The potential function $V = V(x,y)$ is bi-periodic with a periodicity of 1×1 and a defect exists in the center of this periodic structure. The initial data u_0 is assumed locally supported, say in the defect cell.

The definition domain of the above problem is unbounded, and as a first step we could truncate the domain by introducing a rectangular artificial boundary and on it imposing the periodic boundary condition.
This treatment is justified if the time interval of simulation is finite and the number of cells enclosed by the artificial boundary is sufficiently large.

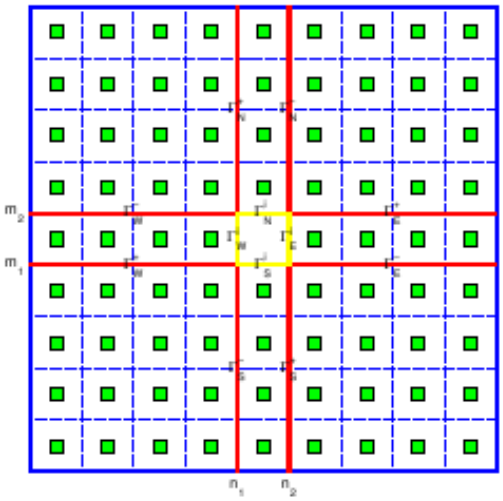


Fig. (29): A bi-periodic potential function with a defect in the center.

The next step is to find a suitable numerical scheme to resolve the wave field. Our basic idea is analogous to that in the last section for handling the one-dimensional Schrödinger equation with periodic potentials at infinity.

We first go to the frequency domain by solving the Helmholtz equation

$$-u_{xx}-u_{yy}+(V-is)u=-iu_0 \qquad (40)$$

with a series of complex parameters s, and then perform the inverse Laplace transformation with a frequency filter. Notice that in (40) we use the same notation u to represent its Laplace-transformed function. This is mainly for the brevity of notations used in the following of this section. Of course we do not indent to solve the equation (40) on the whole truncated domain, since a large number of unknowns would still get involved. Instead, we try to find in the following subsection an accurate boundary condition on the defect cell boundary $\Gamma_E^i \cup \Gamma_S^i \cup \Gamma_W^i \cup \Gamma_N^i$, and perform computation only on the defect cell.

8.1 The Boundary Condition on Defect Cell Boundary

Let us first consider the equation (40) on the geometry shown in Fig. (30). Suppose periodic boundary conditions are specified on Σ_0 and Σ_M. Set $\Gamma_W = \cup_{k=0}^{M-1}\Gamma_{W,k}$ and $\Gamma_E = \cup_{k=0}^{M-1}\Gamma_{E,k}$. In the y-direction, we have M periodic layers.

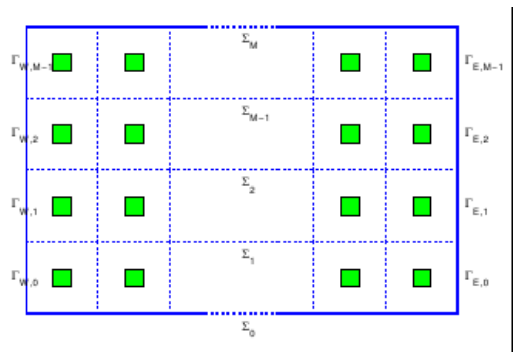


Fig. (30): Schematic view of a bi-periodic structure with periodic boundary conditions on Σ_0 and Σ_M.

We define the *discrete Fourier transformation* in the y-direction as

$$\hat{u}_k(x,y)=\sum_{m=0}^{M-1} u(x,y+mL)\,\omega^{km}, \quad \omega=e^{-2i\pi/M},$$

$k=0,1,\ldots,M-1$. The inverse transformation is given as

$$u(x,y+mL)=\frac{1}{M}\sum_{k=0}^{M-1}\hat{u}_k(x,y)\,\omega^{-km}.$$

It is straightforward to verify that

$$\hat{u}_k(x,y+L)=\omega^{-k}\hat{u}_k(x,y).$$

Thus the problem on the domain shown in Fig. (30) can be reduced to M periodic array problems with pseudo-periodic boundary conditions on Σ_0 and Σ_1. By the analysis in the first section, we get

$$\mathscr{G}^x_{\hat{u}_k}|_{\Gamma_{W,0}}=\hat{\mathscr{A}}_k\mathscr{F}^x_{\hat{u}_k}|_{\Gamma_{W,0}}+\hat{\mathscr{B}}_k\mathscr{G}^x_{\hat{u}_k}|_{\Gamma_{E,0}},$$
$$\mathscr{F}^x_{\hat{u}_k}|_{\Gamma_{E,0}}=\hat{\mathscr{C}}_k\mathscr{F}^x_{\hat{u}_k}|_{\Gamma_{W,0}}+\hat{\mathscr{D}}_k\mathscr{G}^x_{\hat{u}_k}|_{\Gamma_{E,0}},$$

and

$$\mathscr{G}^y_{\hat{u}_k}|_{\Sigma_0}=$$
$$(\widehat{\mathscr{F}\to\mathscr{G}})_k\mathscr{F}^x_{\hat{u}_k}|_{\Gamma_{W,0}}+(\widehat{\mathscr{G}\to\mathscr{G}})_k\mathscr{G}^x_{\hat{u}_k}|_{\Gamma_{E,0}},$$
$$\mathscr{F}^y_{\hat{u}_k}|_{\Sigma_0}=$$
$$(\widehat{\mathscr{F}\to\mathscr{F}})_k\mathscr{F}^x_{\hat{u}_k}|_{\Gamma_{W,0}}+(\widehat{\mathscr{G}\to\mathscr{F}})_k\mathscr{G}^x_{\hat{u}_k}|_{\Gamma_{E,0}}.$$

Then going back to the variable u of (40) yields

$$\mathscr{G}^x_u|_{\Gamma_{W,m}}=\mathscr{A}^m\mathscr{F}^x_u|_{\Gamma_W}+\mathscr{B}^m\mathscr{G}^x_u|_{\Gamma_E},$$
$$\mathscr{F}^x_u|_{\Gamma_{E,m}}=\mathscr{C}^m\mathscr{F}^x_u|_{\Gamma_W}+\mathscr{D}^m\mathscr{G}^x_u|_{\Gamma_E},$$

and

$$\mathscr{G}^y_u|_{\Sigma_m}=$$
$$(\mathscr{F}\to\mathscr{G})_m\mathscr{F}^x_u|_{\Gamma_W}+(\mathscr{G}\to\mathscr{G})_m\mathscr{G}^x_u|_{\Gamma_E},$$
$$\mathscr{F}^y_u|_{\Sigma_m}=$$
$$(\mathscr{F}\to\mathscr{F})_m\mathscr{F}^x_u|_{\Gamma_W}+(\mathscr{G}\to\mathscr{F})_m\mathscr{G}^x_u|_{\Gamma_E},$$

where

$$\mathscr{A}^m\mathscr{F}^x_u|_{\Gamma_W}=$$
$$\frac{1}{M}\sum_{n=0}^{M-1}\left[\sum_{k=0}^{M-1}\hat{\mathscr{A}}_k\omega^{k(n-m)}\right]\mathscr{F}^x_u|_{\Gamma_{W,n}},$$
$$\mathscr{B}^m\mathscr{G}^x_u|_{\Gamma_E}=$$
$$\frac{1}{M}\sum_{n=0}^{M-1}\left[\sum_{k=0}^{M-1}\hat{\mathscr{B}}_k\omega^{k(n-m)}\right]\mathscr{G}^x_u|_{\Gamma_{E,n}},$$
$$\mathscr{C}^m\mathscr{F}^x_u|_{\Gamma_W}=$$
$$\frac{1}{M}\sum_{n=0}^{M-1}\left[\sum_{k=0}^{M-1}\hat{\mathscr{C}}_k\omega^{k(n-m)}\right]\mathscr{F}^x_u|_{\Gamma_{W,n}},$$
$$\mathscr{D}^m\mathscr{G}^x_u|_{\Gamma_E}=$$
$$\frac{1}{M}\sum_{n=0}^{M-1}\left[\sum_{k=0}^{M-1}\hat{\mathscr{D}}_k\omega^{k(n-m)}\right]\mathscr{G}^x_u|_{\Gamma_{E,n}},$$

and

$$(\mathscr{F}\to\mathscr{G})_m\mathscr{F}^x_u|_{\Gamma_W}=$$
$$\frac{1}{M}\sum_{n=0}^{M-1}\left[\sum_{k=0}^{M-1}(\widehat{\mathscr{F}\to\mathscr{G}})_k\omega^{k(n-m)}\right]\mathscr{F}^x_u|_{\Gamma_{W,n}},$$

$$(\mathscr{G}\to\mathscr{G})_m\mathscr{G}_u^x|_{\Gamma_E}=\frac{1}{M}\sum_{n=0}^{M-1}\left[\sum_{k=0}^{M-1}\widehat{(\mathscr{G}\to\mathscr{G})}_k\omega^{k(n-m)}\right]\mathscr{G}_u^x|_{\Gamma_{E,n}},$$

$$(\mathscr{F}\to\mathscr{F})_m\mathscr{F}_u^x|_{\Gamma_W}=\frac{1}{M}\sum_{n=0}^{M-1}\left[\sum_{k=0}^{M-1}\widehat{(\mathscr{F}\to\mathscr{F})}_k\omega^{k(n-m)}\right]\mathscr{F}_u^x|_{\Gamma_{W,n}},$$

$$(\mathscr{G}\to\mathscr{F})_m\mathscr{G}_u^x|_{\Gamma_E}=\frac{1}{M}\sum_{n=0}^{M-1}\left[\sum_{k=0}^{M-1}\widehat{(\mathscr{G}\to\mathscr{F})}_k\omega^{k(n-m)}\right]\mathscr{G}_u^x|_{\Gamma_{E,n}}.$$

Note that the above operators can be evaluated efficiently by FFT.

Now come back to the equation (40) on the geometry shown in Fig. (29). Since periodic boundary conditions are specified on the boundary of the truncated domain, applying the above analysis we have

$$\begin{aligned}
\mathscr{G}_u^y|_{\Gamma_E^-\cup\Gamma_W^+}&=(\mathscr{F}\to\mathscr{G})_{m_1}^H\mathscr{F}_u^x|_{\Gamma_S^+\cup\Gamma_E^i\cup\Gamma_N^-}\\
&\quad+(\mathscr{G}\to\mathscr{G})_{m_1}^H\mathscr{G}_u^x|_{\Gamma_S^-\cup\Gamma_W^i\cup\Gamma_N^+},\\
\mathscr{F}_u^y|_{\Gamma_E^+\cup\Gamma_W^-}&=(\mathscr{F}\to\mathscr{F})_{m_2}^H\mathscr{F}_u^x|_{\Gamma_S^+\cup\Gamma_E^i\cup\Gamma_N^-}\\
&\quad+(\mathscr{G}\to\mathscr{F})_{m_2}^H\mathscr{G}_u^x|_{\Gamma_S^-\cup\Gamma_W^i\cup\Gamma_N^+},\\
\mathscr{G}_u^x|_{\Gamma_N^+\cup\Gamma_S^-}&=(\mathscr{F}\to\mathscr{G})_{n_1}^V\mathscr{F}_u^y|_{\Gamma_W^-\cup\Gamma_N^i\cup\Gamma_E^+}\\
&\quad+(\mathscr{G}\to\mathscr{G})_{n_1}^V\mathscr{G}_u^y|_{\Gamma_W^+\cup\Gamma_S^i\cup\Gamma_E^-},\\
\mathscr{F}_u^x|_{\Gamma_N^-\cup\Gamma_S^+}&=(\mathscr{F}\to\mathscr{F})_{n_2}^V\mathscr{F}_u^y|_{\Gamma_W^-\cup\Gamma_N^i\cup\Gamma_E^+}\\
&\quad+(\mathscr{G}\to\mathscr{G})_{n_2}^V\mathscr{G}_u^y|_{\Gamma_W^+\cup\Gamma_S^i\cup\Gamma_E^-},
\end{aligned}\tag{41}$$

and

$$\begin{aligned}
\mathscr{G}_u^x|_{\Gamma_E^i}&=\mathscr{A}_{m_1}^H\mathscr{F}_u^x|_{\Gamma_S^+\cup\Gamma_E^i\cup\Gamma_N^-}+\mathscr{B}_{m_1}^H\mathscr{G}_u^x|_{\Gamma_S^-\cup\Gamma_W^i\cup\Gamma_N^+},\\
\mathscr{F}_u^x|_{\Gamma_W^i}&=\mathscr{C}_{m_1}^H\mathscr{F}_u^x|_{\Gamma_S^+\cup\Gamma_E^i\cup\Gamma_N^-}+\mathscr{D}_{m_1}^H\mathscr{G}_u^x|_{\Gamma_S^-\cup\Gamma_W^i\cup\Gamma_N^+},\\
\mathscr{G}_u^y|_{\Gamma_N^i}&=\mathscr{A}_{n_1}^V\mathscr{F}_u^y|_{\Gamma_W^-\cup\Gamma_N^i\cup\Gamma_E^+}+\mathscr{B}_{n_1}^V\mathscr{G}_u^y|_{\Gamma_W^+\cup\Gamma_S^i\cup\Gamma_E^-},\\
\mathscr{F}_u^y|_{\Gamma_S^i}&=\mathscr{C}_{n_1}^V\mathscr{F}_u^y|_{\Gamma_W^-\cup\Gamma_N^i\cup\Gamma_E^+}+\mathscr{D}_{n_1}^V\mathscr{G}_u^y|_{\Gamma_W^+\cup\Gamma_S^i\cup\Gamma_E^-}.
\end{aligned}\tag{42}$$

Here we use the superscripts H and V to distinguish those operators in two different directions. Given $\mathscr{F}_u^x|_{\Gamma_E^i}$, $\mathscr{G}_u^x|_{\Gamma_W^i}$, $\mathscr{F}_u^y|_{\Gamma_N^i}$ and $\mathscr{G}_u^x|_{\Gamma_S^i}$, in principle $\mathscr{G}_u^y|_{\Gamma_E^-\cup\Gamma_W^+}$, $\mathscr{F}_u^y|_{\Gamma_E^+\cup\Gamma_W^-}$, $\mathscr{G}_u^x|_{\Gamma_N^+\cup\Gamma_S^-}$ and $\mathscr{F}_u^x|_{\Gamma_N^-\cup\Gamma_S^+}$ can be determined by the operator equations (41). Thus then (42) implicitly define an StS mapping from $\mathscr{F}_u^x|_{\Gamma_E^i}$, $\mathscr{G}_u^x|_{\Gamma_W^i}$, $\mathscr{F}_u^y|_{\Gamma_N^i}$ and $\mathscr{G}_u^x|_{\Gamma_S^i}$, to $\mathscr{G}_u^x|_{\Gamma_E^i}$, $\mathscr{F}_u^x|_{\Gamma_W^i}$, $\mathscr{G}_u^y|_{\Gamma_N^i}$ and $\mathscr{F}_u^x|_{\Gamma_S^i}$. A DtN mapping can be further derived on the boundary of the defect cell, and the computation can now be performed solely on the defect cell.

Unlike the periodic array problems which are periodic only in one direction, the derivation of StS mapping becomes much more complicated. On the discrete level we need to solve a linear system with unknowns $\mathscr{G}_u^y|_{\Gamma_E^-\cup\Gamma_W^+}$, $\mathscr{F}_u^y|_{\Gamma_E^+\cup\Gamma_W^-}$, $\mathscr{G}_u^x|_{\Gamma_N^+\cup\Gamma_S^-}$ and $\mathscr{F}_u^x|_{\Gamma_N^-\cup\Gamma_S^+}$. This operation is still much time-consuming. However, if the size of domain is enlarged, the number of unknowns is only increased *linearly* for two-dimensional problems.

8.2 A Numerical Example

The initial function is

$$u_0(x,y)=\exp(-100x^2-100y^2+20xi).$$

The potential function is

$$\begin{aligned}
V(x,y)=\sum_{m,n\in\mathbb{Z}}&V_0\exp(-100(x-m)^2-100(y-n)^2)\\
&-V_0\exp(-100x^2-100y^2).
\end{aligned}$$

The periodic cell is of size 1×1, and the origin is located in the center of the defect cell $[-0.5,0.5]\times[-0.5,0.5]$. The whole computational domain contains $9\times9=81$ periodic cells. We set the cut-off frequency as 40000, and use the middle-point quadrature rule to approximate the integral (36). The number of quadrature points is 1024. Each cell is discretized into $8\times8=64$ eighth-order finite elements. In Table 5 we list the relative L^2-errors at different time points when $V_0=0$. The reference solution is obtained by the spectral method with same grid points. We see that in this time regime, the errors are always less than 0.02 percent. In Figs. (31)-(33), we show several snapshots for the modulus of wave functions when $V_0=0$, and in Figs. (34)-(36) for the potential $V_0=4000$. Note that only 9 cells including the defect cell are shown in those figures.

Time Point ($\times0.0125$)	Relative L^2-Error
1.0	4.05(-5)
1.5	6.21(-5)
2.0	8.65(-5)
2.5	1.19(-4)
3.0	1.51(-4)

Table 5: Relative L^2-errors in $[-0.5,0.5]^2$ at different time points for $V_0=0$.

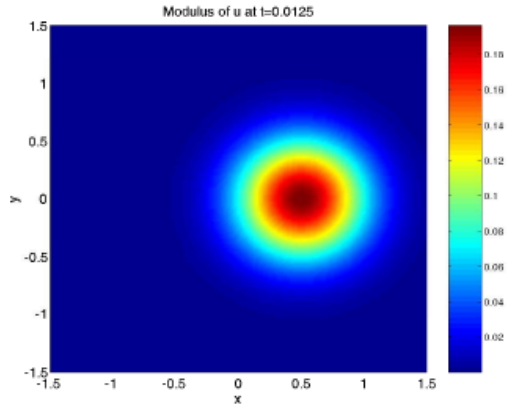


Fig. (31): $V_0 = 0$.

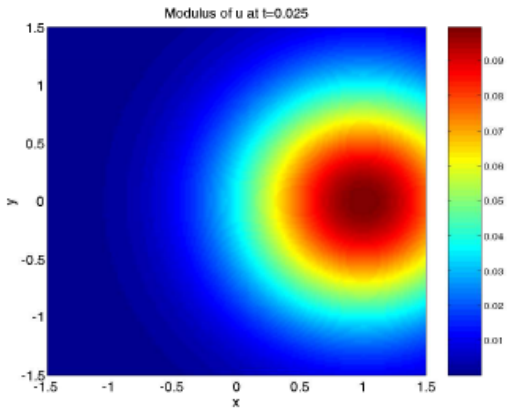


Fig. (32): $V_0 = 0$.

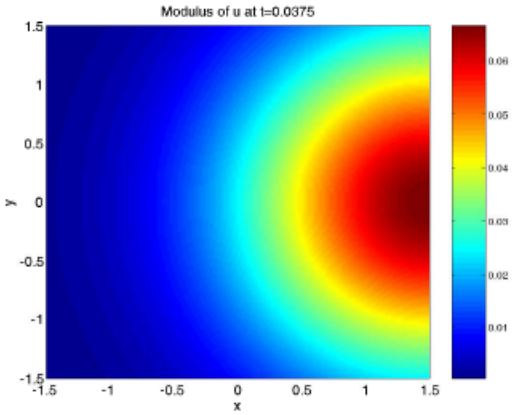


Fig. (33): $V_0 = 0$.

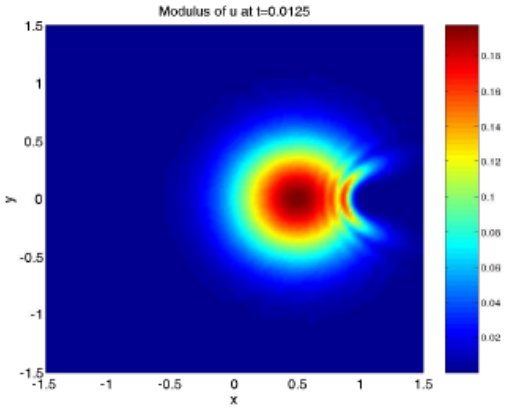


Fig. (34): $V_0 = 4000$.

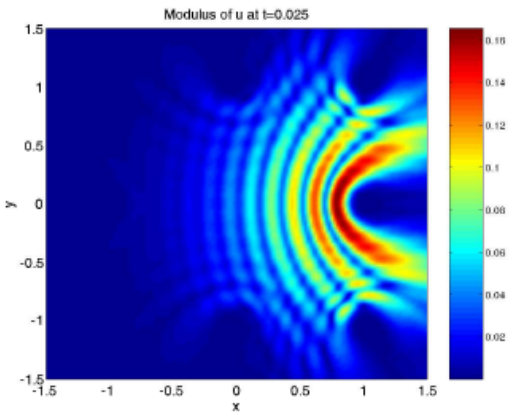


Fig. (35): $V_0 = 4000$.

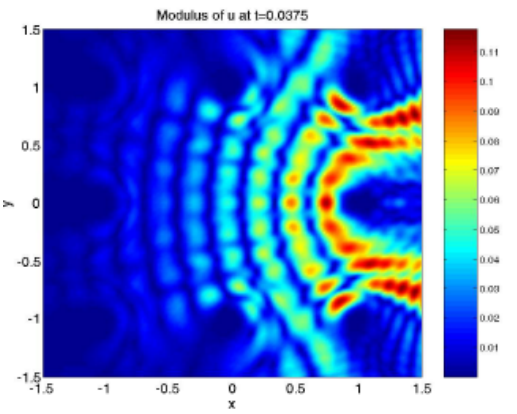


Fig. (36): $V_0 = 4000$.

9. A MODEL PROBLEM

We consider a closed waveguide consisting of an infinite number of identical cells, see Fig. (37). There C_j denotes the j-th periodic cell, and Γ_j the j-th cell boundary. The governing wave equation is the *Helmholtz equation*

$$\Delta u + k^2 n^2 u = 0, \quad (x,y) \in \Omega = \cup_{j=1}^{+\infty} \mathrm{C}_j, \qquad (43)$$

where k denotes the reference wave number, and $n = n(x,y)$ is the refraction index function. On each cell boundary Γ_j we define two *Sommerfeld data* associated with the function u as

$$f_j(u) = (\partial_x + ik)u|_{\Gamma_j}, \quad g_j(u) = (\partial_x - ik)u|_{\Gamma_j}, \quad (44)$$

where i denotes the imaginary unit. To clarify the physical meaning of these two data, let us first return to the one-dimensional constant coefficient Helmholtz equation

$$u_{xx} + k^2 u = 0.$$

Two linearly independent solutions are $e^{\pm ikx}$. As a common convention, e^{ikx} represents a wave traveling to the right, and e^{-ikx} to the left. An easy computation yields

$$(\partial_x + ik)e^{ikx} = 2ike^{ikx}, \quad (\partial_x - ik)e^{ikx} = 0,$$

and

$$(\partial_x + ik)e^{-ikx} = 0, \quad (\partial_x - ik)e^{-ikx} = -2ike^{-ikx}.$$

These expressions above imply that the operator $\partial_x + ik$ eliminates the left-going wave while the operator $\partial_x - ik$ eliminates the right-going wave. Thus, the functions f_j and g_j in (44) contain some information about the right-going and left-going waves respectively. They are further referred to as *incoming* or *outgoing* relying on the location of Γ_j with respect to (w.r.t.) the concerned part of the domain. For example, w.r.t. C_j, f_j is incoming and g_j is outgoing, but w.r.t. C_{j-1}, f_j is outgoing and g_j is incoming.

The boundary conditions on the top, bottom and interior (if existing) boundaries could be either Neumann or Dirichlet, or any combination, but they need to be consistent with the geometry periodicity. Moreover, these boundary conditions should guarantee the well-posedness of the Helmholtz equation (43) on the union of any finite number of periodic cells, say $\cup_{j=0}^{N-1} \mathrm{C}_j$, if the incoming Sommerfeld data are prescribed on its left and right boundaries, say Γ_0 and Γ_N.

We remark that these restrictions are in fact very mild thanks to the Holmgren uniqueness theorem [35, Section 5.3]. In the sequel, if not specified otherwise, we assume homogeneous Neumann boundary conditions at the top and bottom boundaries.

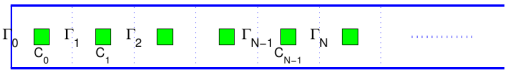


Fig. (37): Schematic of a semi-infinite periodic array. C_j denotes the j-th periodic cell. Γ_j is the left cell boundary of C_j and the right cell boundary of C_{j-1} (for $j \geq 1$).

9.1 The periodic Arrays

Three different periodic arrays (PA) will be considered in this chapter, and we will refer to them as PA-One, PA-Two and PA-Three. All of them consist of periodic cells with size of 1×1. More details are given below.

- **PA-One**. Homogeneous waveguide. $n = 1$.
- **PA-Two**. A hole of size 0.5×0.5 is located in the center of every periodic cell. Zero Dirichlet boundary condition is applied at the hole boundary. $n = 1$.
- **PA-Three**. Rectangular waveguide. $n(x,y) = 1 + 0.5\cos(2\pi x)\sin(2\pi y)$.

To explore the wave property in a periodic array, it is usually helpful to consider the dispersion diagram of the characteristic equation $-\Delta u = En^2 u$, restricted to a single periodic cell, say C_0. The boundary conditions at the left and right boundaries are *pseudoperiodic*, namely,

$$u|_{\Gamma_1} = e^{i\theta} u|_{\Gamma_0}, \qquad u_x|_{\Gamma_1} = e^{i\theta} u_x|_{\Gamma_0},$$

where the parameter θ is valued in $[0, 2\pi)$. For each θ, there exists a sequence of real eigenvalues E, usually called *energies*. All energies E w.r.t. θ then compose the dispersion diagram. The dispersion relation for PA-One, the homogeneous waveguide, can be obtained analytically as

$$E_{jm} = j^2\pi^2 + (\theta + 2\pi m)^2.$$

This multi-valued function is plotted in Fig. (38). For PA-Two and PA-Three, no analytical expressions of dispersion relation are available, and a spatial discretization method has to be employed.

We use the eighth-order FEM method with mesh sizes $\Delta x = \Delta y = 0.125$ for all the numerical tests reported in this chapter.

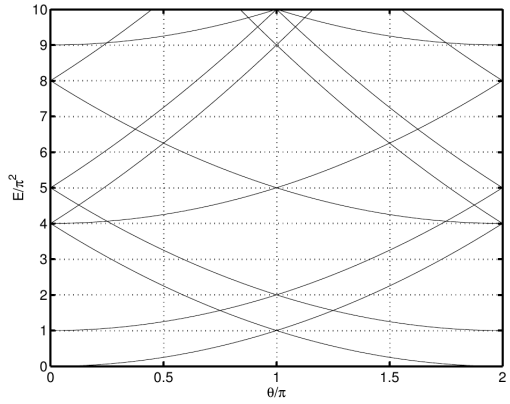


Fig. (38): Dispersion diagram of PA-One, an homogeneous waveguide.

The dispersion diagrams for PA-Two and PA-Three are shown in Figs. (39)-(40). A significant phenomena could be observed that unlike the homogeneous waveguide, there are some bands of energy values in the dispersion diagrams of PA-Two and PA-Three that could not be reached for any parameter θ.

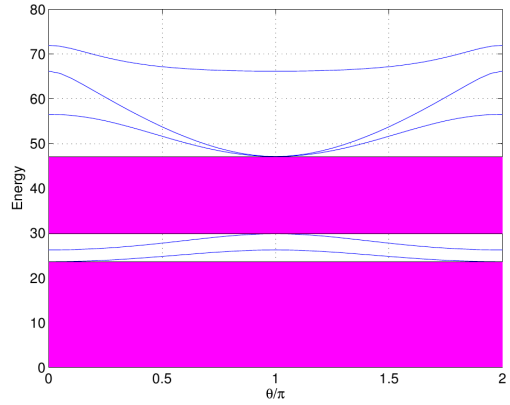


Fig. (39): Dispersion diagram of PA-Two. The first two stop bands are $(0, 23.61_{\pm 0.01})$ and $(29.85_{\pm 0.01}, 47.10_{\pm 0.01})$.

Physically, waves with energy (here k^2) in these bands could not propagate in the medium. Right in this context, they are usually referred to as *stop bands* in the literature. In fact, it is exactly this remarkable property which makes the periodic structures extremely useful, for example, they could be

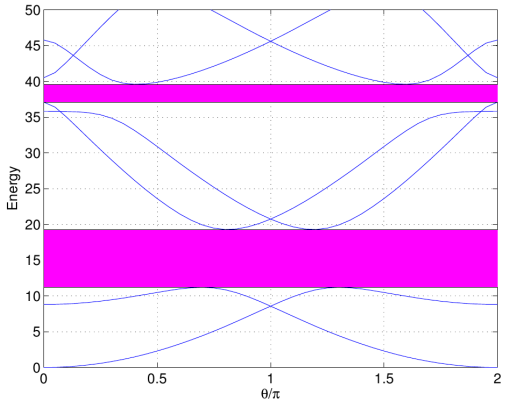


Fig. (40): Dispersion diagram of PA-Three.The first two stop bands are $(11.20_{\pm 0.01}, 19.29_{\pm 0.01})$ and $(37.08_{\pm 0.01}, 39.58_{\pm 0.01})$.

elaborately designed to act as some kind of frequency selecting modules in the microwave and optical engineering.

This work is aimed at developing an efficient method for deriving an exact boundary mapping of semi-infinite periodic arrays for *any real wavenumber* k.

10. THE LIMITING ABSORPTION PRINCIPLE

The first problem we are facing is how to guarantee the well-posedness of the Helmholtz equation (43), which naturally arises due to the absence of a radiation-like condition at infinity. Although the constant coefficient case with separable geometries is well solved, this problem is not trivial at all and largely remains open for the variable coefficient Helmholtz equation.

There are at least three methods of possibly deriving a unique solution of the Helmholtz equation in unbounded domains: asymptotic radiation condition, limiting absorption principle and limiting amplitude principle [67]. In this chapter we employ the *limiting absorption principle* (LABP). The LABP is said to hold at $k > 0$ if and only if for any $f_0(u) \in L^2(\Gamma_0)$ (take $f_0(u)$ as a unity), the solution $u^\varepsilon \in H^1(\Omega)$ of the following *damped Helmholtz equation*

$$\Delta u^\varepsilon + (k^2 + i\varepsilon) n^2 u^\varepsilon = 0 \qquad (45)$$

with the boundary condition

$$f_0(u^\varepsilon) = f_0(u),$$

converges to a unique solution $u \in H^1_{loc}(\Omega)$ of the Helmholtz equation (43), and the outgoing Sommerfeld datum $g_0(u^\varepsilon) = \mathscr{A}^\varepsilon_{\rm inf} f_0(u^\varepsilon)$ also converges to the unique function $g_0(u)$. This makes it possible to define a *Sommerfeld-to-Sommerfeld* (StS) mapping $\mathscr{A}_{\rm inf}$ as the limit of $\mathscr{A}^\varepsilon_{\rm inf}$, which maps $f_0(u)$ to $g_0(u)$, namely,

$$g_0(u) = \mathscr{A}_{\rm inf} f_0(u).$$

Let us start considering PA-One first. In this case the separation of variables method is available. We set

$$u^\varepsilon = \sum_{n=0}^{+\infty} u^{\varepsilon,n} \cos(n\pi y)$$

and

$$f_0(u) = \sum_{n=0}^{+\infty} f_0(u^n)\cos(n\pi y),$$
$$g_0(u^\varepsilon) = \sum_{n=0}^{+\infty} g_0(u^{\varepsilon,n})\cos(n\pi y).$$

Then (45) is transformed into a sequence of ODE problems:

$$u^{\varepsilon,n}_{xx} + (k^2 + i\varepsilon - n^2\pi^2)u^{\varepsilon,n}_{xx} = 0,$$
$$f_0(u^{\varepsilon,n}) = f_0(u^n), \quad \forall n = 0,1,\ldots.$$

The bounded solutions of the above problems are

$$u^{\varepsilon,n} = \frac{f_0(u^n)}{i\sqrt{k^2+i\varepsilon-n^2\pi^2}+ik} e^{i\sqrt{k^2+i\varepsilon-n^2\pi^2}x}.$$

Hence, we have

$$g_0(u^{\varepsilon,n}) = \frac{i\sqrt{k^2+i\varepsilon-n^2\pi^2}-ik}{i\sqrt{k^2+i\varepsilon-n^2\pi^2}+ik} f_0(u^n),$$

and

$$g_0(u^n) \overset{def}{=} \lim_{\varepsilon\to 0} g_0(u^{\varepsilon,n}) = \frac{i\sqrt{k^2-n^2\pi^2}-ik}{i\sqrt{k^2-n^2\pi^2}+ik} f_0(u^n). \tag{46}$$

Besides, it is straightforward to verify that

$$g_0(u^{\varepsilon,n}) = g_0(u^n) + \begin{cases} \frac{2\sqrt{i\varepsilon}f_0(u^n)}{k} + O(\varepsilon), & k = n\pi, \\ \frac{ik\varepsilon f_0(u^n)}{(\sqrt{k^2-n^2\pi^2}+k)^2\sqrt{k^2-n^2\pi^2}} + O(\varepsilon^2), & k \neq n\pi. \end{cases} \tag{47}$$

The expression (47) states that the convergence rate of $g_0(u^\varepsilon)$ to

$$g_0(u) = \sum_{n=0}^{+\infty} g_0(u^n)\cos(n\pi y)$$

is of first order with respect to ε if k is unequal to any $n\pi$ with $n \geq 0$. If k is equal to some $n_0\pi$, which implies the resonance of the n_0-th mode in the y-direction, the convergence rate would degenerate to half order. But the LABP holds independent of the wavenumber k.

Based on the above analysis, we conjecture that, under some mild restrictions on the geometry and the refraction index function, the LABP holds for every $k > 0$ for more general semi-infinite periodic arrays. Some numerical evidences will be reported in the end of this section.

The LABP itself suggests a method for deriving the exact StS mapping on the left boundary Γ_0: first compute the exact StS mapping of the problem (45) for a given ε, denoted by $\mathscr{A}^\varepsilon_{\rm inf}$, and then let ε tend to zero. In [20] the authors proposed a fast evaluation method for the exact StS mapping of the damped Helmholtz equation (45). The basic idea is as follows. For any $N > 0$, the damped Helmholtz equation (45) is well-posed on the domain $\cup_{j=0}^{N-1} \mathrm{C}_j$, with the incoming Sommerfeld data f^ε_0 and g^ε_N prescribed at the boundaries Γ_0 and Γ_N. Thus there are four linear scattering operators $\mathscr{A}^\varepsilon_N$, $\mathscr{B}^\varepsilon_N$, $\mathbb{C}^\varepsilon_N$ and $\mathscr{D}^\varepsilon_N$ satisfying

$$g^\varepsilon_0 = \mathscr{A}^\varepsilon_N f^\varepsilon_0 + \mathscr{B}^\varepsilon_N g^\varepsilon_N, \qquad f^\varepsilon_N = \mathbb{C}^\varepsilon_N f^\varepsilon_0 + \mathscr{D}^\varepsilon_N g^\varepsilon_N.$$

Since g^ε_N goes to zero exponentially fast as N tends to infinity, it is reasonable to expect that $\mathscr{A}^\varepsilon_N$ converges and the limit is just the exact StS mapping $\mathscr{A}^\varepsilon_{\rm inf}$. Note that the *fast doubling procedure* and the involved scattering operators are explained previously in Section

In Fig. (41) we plot the relative errors of the scattering operators $\mathscr{A}^\varepsilon_N$ compared to the reference operator $\mathscr{A}^\varepsilon_{\rm ref}$, which is obtained by using the doubling technique 20 times, i.e., $N = 2^{20}$. Since FEM is used, the scattering operators are approximated by matrices of rank 65×65. We could see that the doubling technique really leads to an efficient algorithm. Also notice that when k^2 lies in the stop bands, for example $k^2 = 23, 31$, $\mathscr{A}_N$ itself converges as N goes to infinity. This implies that when k^2 is in the stop bands, we could derive the StS mapping directly without considering the LABP.

Next we explain how to let ε tend to zero. In light of the expression (47), if the resonance does not occur, the exact StS mapping $\mathscr{A}_{\rm inf}$ is expected to bear an asymptotic expansion like

$$\mathscr{A}^\varepsilon_{\rm inf} = \mathscr{A}_{\rm inf} + \varepsilon\mathscr{A}^{(1)}_{\rm inf} + \varepsilon^2\mathscr{A}^{(2)}_{\rm inf} + \cdots. \tag{48}$$

Thus in most cases, the convergence rate of the LABP is of first order. This observation is supported

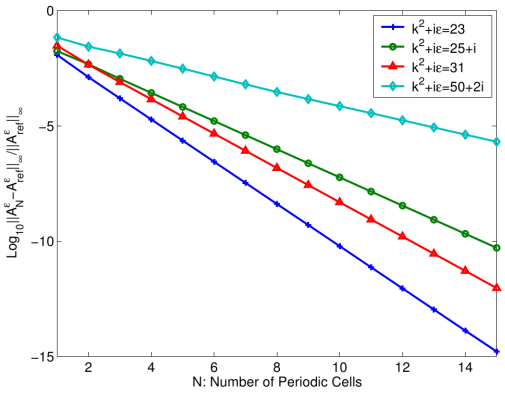


Fig. (41): Relative errors of $\mathscr{A}_N^{\varepsilon}$ to the reference StS mapping $\mathscr{A}_{\text{ref}}^{\varepsilon}$, which is obtained by setting $N = 2^{20}$.

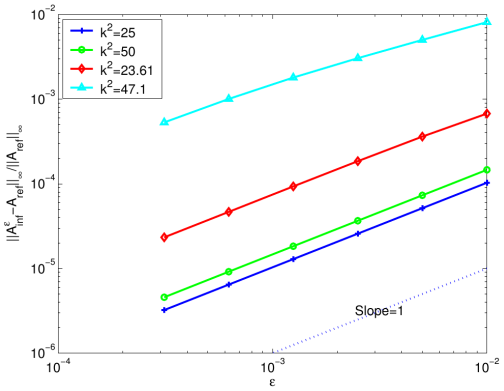


Fig. (42): The reference operator $\mathscr{A}_{\text{ref}}$ is obtained by setting $\varepsilon = 10^{-7}$.

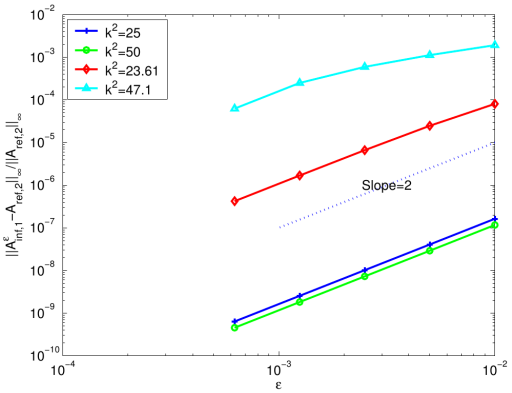


Fig. (43): The reference matrix $\mathscr{A}_{\text{ref},2}$ is obtained by using extrapolation technique twice with $\varepsilon_0 = 0.00125$, i.e., $\mathscr{A}_{\text{ref},2} = \mathscr{A}_{\text{inf}}^{\varepsilon_0}/3 - 2\mathscr{A}_{\text{inf}}^{\varepsilon_0/2} + 8\mathscr{A}_{\text{inf}}^{\varepsilon_0/4}/3$. $\mathscr{A}_{inf,1}^{\varepsilon} = -\mathscr{A}_{\text{inf}}^{\varepsilon} + 2\mathscr{A}_{\text{inf}}^{\varepsilon/2}$ is obtained by using extrapolation technique once.

by the numerical evidences shown in Fig. (42). Note that the convergence rate could be improved by standard extrapolation techniques. In Fig. (43) we show the errors of the StS operators extrapolated once to the reference operator, which is obtained by using extrapolations twice and setting a small damping parameter $\varepsilon_0 = 0.00125$. We could see that the accuracy is greatly improved, and second order rate can be clearly observed. We should also notice that if k is close to a resonance wave number, for example $k^2 = 23.61,\ 47.1$, the asymptotic convergence rate could only manifest for sufficiently small damping parameters.

11. ASYMPTOTIC BEHAVIOUR OF AN LABP SOLUTION

The last section showed that if k is not a resonance wave number, the extrapolation technique could yield very accurate solution. Obviously this algorithm needs to evaluate the scattering operators for a sequence of ε, and this turns out to be computationally quite expensive. Besides, though the chance of k being a resonance wave number is very rare, if k is close to a resonance wave number, the extrapolation method could not present very accurate result. In this section we will develop a new method by directly using the scattering operators for the undamped Helmholtz equation.

Recall from the last section that when k^2 lies in the stop bands, the exact StS mapping could be computed by the doubling technique without using the LABP. This is due to the fact that the solution lies in $L^2(\Omega)$, and thus it decays exponentially fast at infinity. If k^2 lies in the pass bands (complementary energy intervals of stop bands), in general an LABP solution cannot be expected to decay. Our basic idea is to separate those traveling (not-decaying) waves and evanescent (decaying) waves, and handle them by different means.

First let us introduce some notations. Suppose u and v are two solutions of the Helmholtz equation (43). Define the *co-related energy flux* of u and v as

$$\begin{aligned}\mathscr{E}(u,v) &= -2ik\left[(u_x,v)_{\Gamma_j} - (u,v_x)_{\Gamma_j}\right] \\ &= (f(u),f(v))_{\Gamma_j} - (g(u),g(v))_{\Gamma_j}.\end{aligned}$$

Besides, the energy flux of u is defined as $\mathscr{E}(u,u)$, which is also equal to

$$\mathscr{E}(u,u) = 4k\,\mathrm{Im}\int_{\Gamma_j} u_x\bar{u}\,dy.$$

We should remark that the co-related energy flux does not rely on the choice of Γ_j. Moreover, $\mathscr{E}(\cdot,\cdot)$ defines a sesquilinear form.

A nontrivial solution u of the Helmholtz equation (43) or (45) is regarded as a *Bloch wave* associated with the *Floquet multiplier* $\alpha \in \mathbb{C}$ if it satisfies the following two conditions

$$u|_{\Gamma_{i+1}} = \alpha u|_{\Gamma_i}, \quad u_x|_{\Gamma_{i+1}} = \alpha u_x|_{\Gamma_i}, \quad \forall\, i = 0, 1, \ldots.$$

We denote by $\mathbb{F}$ the set of all Floquet factors. A Bloch wave is referred to as evanescent, traveling, or anti-evanescent if the associated Floquet multiplier α satisfies $|\alpha| < 1$, $|\alpha| = 1$, or $|\alpha| > 1$. If $|\alpha| = 1$, we refer to α as a *unitary Floquet multiplier*. The set of unitary Floquet multipliers is denoted by $\mathbb{UF}$. Note that the Floquet factor cannot be zero due to the mentioned Holmgren uniqueness theorem. For any $\alpha \in \mathbb{F}$, all associated Bloch waves together with zero function form a linear space. This space, denoted by $\mathbb{E}_\alpha$, is called an (α-*periodic*) *eigenfunction space*. Here we list a couple of propositions about the Floquet theory from [42].

Proposition 4. *If $\alpha \in \mathbb{F}$, then $1/\alpha \in \mathbb{F}$ either.*

Proposition 5. *$\mathbb{UF}$ is a finite set. For any $\alpha \in \mathbb{UF}$, $N_\alpha = \dim \mathbb{E}_\alpha < +\infty$.*

Proposition 6. *Given two Floquet multipliers α_j and α_k, and two functions $\varphi_j \in \mathbb{E}_{\alpha_j}$ and $\varphi_k \in \mathbb{E}_{\alpha_k}$. If $\alpha_j \alpha_k^* \neq 1$, then $\mathscr{E}(\varphi_i, \varphi_j) = 0$.*

Proposition 7. *If u is an LABP solution, then the energy flux of u is nonnegative.*

Obviously, an LABP solution u cannot include the anti-evanescent Bloch waves, thus asymptotically, u is a combination of traveling Bloch waves. It is known that not every traveling Bloch wave is an LABP solution. We need to pick out those compatible with the LABP. To get some insight, let us consider the homogeneous waveguide problem.
Suppose $k = \pi$. Then the traveling Bloch wave space is given by

$$\mathrm{Span}\{e^{-i\pi x}, e^{i\pi x}, \cos(\pi y)\}.$$

If the x-period L is set as a non-integer positive number, then we get three unitary Floquet multipliers: $e^{-i\pi L}$ associated with $\mathrm{Span}\{e^{-i\pi x}\}$, $e^{i\pi L}$ with $\mathrm{Span}\{e^{i\pi x}\}$ and, 1 with $\mathrm{Span}\{\cos(\pi y)\}$. Since

$$\begin{aligned}\mathscr{E}(e^{-i\pi x}, e^{-i\pi x}) &= 4\pi\,\mathrm{Im} \int_0^1 (-i\pi e^{-i\pi x}) e^{i\pi x}\, dy \Big|_{x=0} \\ &= -4\pi^2,\end{aligned}$$

and an LABP solution has a nonnegative energy flux, $e^{-i\pi x}$ is thus not admissible. Comparatively, we have

$$\begin{aligned}\mathscr{E}(e^{i\pi x}, e^{i\pi x}) &= 4\pi\,\mathrm{Im} \int_0^1 (i\pi e^{i\pi x}) e^{-i\pi x}\, dy \Big|_{x=0} \\ &= 4\pi^2,\end{aligned}$$

and

$$\begin{aligned}\mathscr{E}(\cos(\pi y), \cos(\pi y)) &= 4\pi\,\mathrm{Im} \int_0^1 (0) e^{i\pi x} dy \Big|_{x=0} \\ &= 0.\end{aligned}$$

The problem appears when L is taken as an integer. For example, let us take $L = 1$. In this case there are two unitary Floquet multipliers 1 and -1, namely,

$$\begin{aligned}\alpha_1 = -1 &\longleftrightarrow \mathbb{E}_{\alpha_1} = \mathrm{Span}\{e^{-i\pi x},\ e^{i\pi x}\}, \\ \alpha_2 = \quad 1 &\longleftrightarrow \mathbb{E}_{\alpha_2} = \mathrm{Span}\{\cos(\pi y)\}.\end{aligned}$$

$\mathbb{E}_{\alpha_2}$ represents a resonance space, and two-dimensional space $\mathbb{E}_{\alpha_1}$ contains both the left-going and right-going traveling waves. The problem is how to classify these two kind of waves. One may say the energy principle could still work, since obviously the Bloch wave $e^{i\pi x}$ is outgoing, and $e^{-i\pi x}$ is incoming. But the question is that $\mathbb{E}_{\alpha_1}$ may have different basis representation, for example,

$$\begin{aligned}\mathbb{E}_{\alpha_1} &= \mathrm{Span}\{e^{-i\pi x} + 2e^{i\pi x},\ e^{-i\pi x} + 3e^{i\pi x}\} \\ &= \mathrm{Span}\{e^{i\pi x} + 2e^{-i\pi x},\ e^{i\pi x} + 3e^{-i\pi x}\}.\end{aligned}$$

For the first representation, both basis functions are right-going, and for the second, both are left-going. However, generally we could not distinguish an LABP outgoing traveling wave only through its energy flux.
The above problem becomes even more severe if we take $L = 2$. In this case there exists only one unitary Floquet multiplier

$$\alpha = 1 \longleftrightarrow \mathbb{E}_\alpha = \mathrm{Span}\{e^{-i\pi x},\ e^{i\pi x}, \cos(\pi y)\}.$$

It is not hard to find different basis representations for $\mathbb{E}_\alpha$, which have completely different signs of energy flux. As a conclusion, if α is a unitary Floquet multiplier and the associated eigenfunction space $\mathbb{E}_\alpha$ is multi-dimensional, we have to resort to other criterion to determine the LABP right-going Bloch waves.
Let us remark here that for a three-dimensional waveguide problem, the chance for $\mathbb{E}_\alpha$ being multi-dimensional is absolutely not rare, though it seems true for two-dimensional waveguide problems.

Suppose $\alpha \in \mathbb{UF}$, and $\{\varphi_j\}_{j=1}^{N_\alpha}$ constitute a set of basis functions of $\mathbb{E}_\alpha$, *orthonormal* w.r.t. the n^2-weighted inner product $(\cdot,\cdot)_{n^2}$ defined as

$$(\varphi_j, \varphi_k)_{n^2} = \int_{C_0} n^2 \varphi_j \bar{\varphi}_k \, dy.$$

We define the *energy flux matrix* $M = (m_{jk})$ as

$$m_{jk} = \mathscr{E}(\varphi_j, \varphi_k), \qquad \forall\, j,k = 1,2,\cdots,N_\alpha.$$

It is easy to verify that M is a Hermitian matrix, which implies the existence of a unitary matrix U, such that

$$U^\top M \bar{U} = \Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{N_\alpha}),$$

where λ_j are real eigenvalues of M ordered by

$$\begin{aligned}\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{m_1} > 0 = \lambda_{m_1+1} = \ldots \\ \cdots = \lambda_{m_2} = 0 > \lambda_{m_2+1} \geq \cdots \geq \lambda_{N_\alpha}.\end{aligned}$$

We introduce a new set of basis function $\{\psi_j\}_{j=1}^{N_\alpha}$ as

$$(\psi_1, \ldots, \psi_{N_\alpha}) = (\varphi_1, \ldots, \varphi_{N_\alpha})U,$$

which will be referred to as a *canonical set of basis functions* of $\mathbb{E}_\alpha$. Now we could separate $\mathbb{E}_\alpha$ into three parts, i.e.,

$$\mathbb{E}_\alpha = \mathbb{R}_\alpha \oplus \mathbb{S}_\alpha \oplus \mathbb{L}_\alpha,$$

with

$$\begin{aligned}\mathbb{R}_\alpha &= \mathrm{Span}\{\psi_1, \cdots, \psi_{m_1}\},\\ \mathbb{S}_\alpha &= \mathrm{Span}\{\psi_{m_1+1}, \cdots, \psi_{m_2}\},\\ \mathbb{L}_\alpha &= \mathrm{Span}\{\psi_{m_2+1}, \cdots, \psi_{N_\alpha}\}.\end{aligned}$$

Proposition 8. *For any $\alpha \in \mathbb{E}_\alpha$, $\{\lambda_j\}_{j=1}^{N_\alpha}$ are invariant quantities, and $\mathbb{R}$, $\mathbb{S}$ and $\mathbb{L}$ are invariant subspaces of $\mathbb{E}_\alpha$. Besides, for any $\varphi_1 \in \mathbb{R}_\alpha$, $\varphi_2 \in \mathbb{S}_\alpha$, $\varphi_3 \in \mathbb{L}_\alpha$, we have*

$$\mathscr{E}(\varphi_1, \varphi_1) > 0, \quad \mathscr{E}(\varphi_2, \varphi_2) = 0, \quad \mathscr{E}(\varphi_3, \varphi_3) < 0.$$

For the homogeneous waveguide problem, it is straightforward to verify that $\mathbb{R}_\alpha$ is the admissible LABP Bloch wave space with positive energy flux. $\mathbb{S}_\alpha$ is the resonance wave space, which is also admissible to the LABP. Note that if $\mathbb{S}_\alpha$ is excluded from the asymptotic solution space, the Helmholtz equation would loose solvability for some incoming Sommerfeld data f_0.

Based on these facts, for a general semi-infinite periodic array, we make the following conjecture.

Conjecture 1. Suppose $\alpha_1, \ldots, \alpha_M$ are all unitary Floquet multipliers, and $\varphi_1^{\alpha_j}, \ldots, \varphi_{N_{\alpha_j}}^{\alpha_j}$ constitute a set of orthonormal basis functions of $\mathbb{R}_{\alpha_j} \oplus \mathbb{S}_{\alpha_j}$. Then asymptotically, any LABP solution u lies in the space

$$\mathrm{Span}\{\varphi_k^{\alpha_j} | j = 1, \ldots, M, k = 1, \ldots, N_{\alpha_j}\}. \tag{49}$$

Although we have no proof of this conjecture yet, its validity is strongly supported by the numerical tests given in the next section. Let us remark here that according to Proposition 6, $\{\varphi_k^{\alpha_j}\}_{j=1,k=1}^{M,N_{\alpha_j}}$ in fact constitute a set of basis functions of the LABP right-going Bloch wave space.

12. EVALUATION OF THE EXACT StS MAPPING

Based on Conjecture 1, we know when N is large, asymptotically,

$$\begin{aligned}f_N(u) &\approx \sum_{j=1}^{M} \sum_{k=1}^{N_{\alpha_j}} t_k^j f_0(\varphi_k^{\alpha_j}),\\ g_N(u) &\approx \sum_{j=1}^{M} \sum_{k=1}^{N_{\alpha_j}} t_k^j g_0(\varphi_k^{\alpha_j}).\end{aligned}$$

Or in an abbreviated vector form,

$$f_N(u) \approx FT, \qquad g_N(u) \approx GT, \tag{50}$$

where

$$\begin{aligned}F = (F_1, \cdots, F_M), \quad G = (G_1, \cdots, G_M),\\ T = (T_1, \cdots, T_M)^\top\end{aligned}$$

with

$$\begin{aligned}F_j &= (f_0(\varphi_1^{\alpha_j}), \cdots, f_0(\varphi_{N_{\alpha_j}}^{\alpha_j})),\\ G_j &= (g_0(\varphi_1^{\alpha_j}), \cdots, g_0(\varphi_{N_{\alpha_j}}^{\alpha_j})),\\ T_j &= (t_1^{\alpha_j}, \cdots, t_{N_{\alpha_j}}^{\alpha_j}),\end{aligned} \tag{51}$$

Recall that

$$\begin{aligned}g_0(u) &= \mathscr{A}_N f_0(u) + \mathscr{B}_N g_N(u),\\ f_N(u) &= \mathbb{C}_N f_0(u) + \mathscr{D}_N g_N(u).\end{aligned}$$

Using (50) T could be derived by the least square method as

$$T \approx (F - \mathscr{D}_N G)^{-1} \mathbb{C}_N f_0(u). \tag{52}$$

Here, $^{-1}$ denotes the pseudo-inverse operator. We then have

$$g_0(u) = \mathscr{A}_N f_0(u) + \mathscr{B}_N g_N(u) \approx (\mathscr{A}_N + \mathscr{B}_N G(F - \mathscr{D}_N G)^{-1}\mathbb{C}_N) f_0(u),$$

which means that by putting

$$\tilde{\mathscr{A}_N} = \mathscr{A}_N + \mathscr{B}_N G(F - \mathscr{D}_N G)^{-1}\mathbb{C}_N,$$

the limit of $\tilde{\mathscr{A}_N}$ would give the exact StS mapping $\mathscr{A}_{\text{inf}}$ on the left boundary Γ_0.

The key step to implement the above algorithm is to derive a canonical set of basis functions for all unitary Floquet multipliers, i.e. we need to compute the functions F_j and G_j defined in (51). This objective can be achieved by the following steps:

1. Solve the generalized eigenvalue problem

$$\begin{pmatrix} -\mathscr{A}_1 & I \\ -\mathbb{C}_1 & 0 \end{pmatrix} \begin{pmatrix} f_0 \\ g_0 \end{pmatrix} = \alpha \begin{pmatrix} 0 & \mathscr{B}_1 \\ -I & \mathscr{D}_1 \end{pmatrix} \begin{pmatrix} f_0 \\ g_0 \end{pmatrix}$$

to obtain all (different) unitary Floquet multipliers $\{\alpha_j\}_{j=1}^M$ and their associated generalized eigenfunctions $(f_{0,k}^{\alpha_j}, g_{0,k}^{\alpha_j})$, $k = 1, \cdots, N_{\alpha_j}$.

2. If $\mathbb{E}_{\alpha_j}$ is one-dimensional, i.e. $N_{\alpha_j} = 1$, compute the energy flux of the eigenfunction φ_1^j associated with the Sommerfeld data $(f_{0,1}^{\alpha_j}, g_{0,1}^{\alpha_j})$ by

$$\mathscr{E}(\varphi_1^j, \varphi_1^j) = (f_{0,1}^{\alpha_j}, f_{0,1}^{\alpha_j})_{\Gamma_0} - (g_{0,1}^{\alpha_j}, g_{0,1}^{\alpha_j})_{\Gamma_0}.$$

If and only if $\mathscr{E}(\varphi_1^j, \varphi_1^j) \geq 0$, then φ_1^j is an admissible LABP traveling Bloch wave, i.e., $F_j = (f_{0,1}^{\alpha_j})$, $G_j = (g_{0,1}^{\alpha_j})$. Otherwise, $F_j = G_j = \emptyset$.

3. If $\mathbb{E}_{\alpha_j}$ is multi-dimensional, i.e., $N_{\alpha_j} > 1$, derive a set of orthonormal eigenfunctions $\{\varphi_k^{\alpha_j}\}_{k=1}^{N_{\alpha_j}}$ of the following problem

$$\Delta u + k^2 n^2 u = 0,$$
$$u|_{\Gamma_1} = \alpha_j u|_{\Gamma_0}, \quad u_x|_{\Gamma_1} = \alpha_j u_x|_{\Gamma_0}.$$

Compute the associated Sommerfeld data $\{f_0(\varphi_k^{\alpha_j})\}$ and $\{g_0(\varphi_k^{\alpha_j})\}$. Compute the energy matrix $M = (m_{kl})$ with

$$m_{kl} = (f_0(\varphi_k^{\alpha_j}), f_0(\varphi_l^{\alpha_j}))_{\Gamma_0} - (g_0(\varphi_k^{\alpha_j}), g_0(\varphi_l^{\alpha_j}))_{\Gamma_0},$$

for all $k, l = 1, \cdots, N_{\alpha_j}$. Find a unitary matrix $U = (u_{lk})$ to diagonalize M, such that

$$U^\top M \bar{U} = \Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{N_{\alpha_j}}),$$

where λ_j are real eigenvalues of M ordered by

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{m_1} > 0 = \lambda_{m_1+1} = \ldots \cdots = \lambda_{m_2} = 0 > \lambda_{m_2+1} \geq \cdots \geq \lambda_{N_{\alpha_j}}.$$

Set $F_j = (F_j^1, \cdots, F_j^{m_2})$ and $G_j = (G_j^1, \cdots, G_j^{m_2})$ with

$$F_j^k = \sum_{l=1}^{N_{\alpha_j}} f_0(\varphi_l^{\alpha_j}) u_{lk}, \qquad G_j^k = \sum_{l=1}^{N_{\alpha_j}} g_0(\varphi_l^{\alpha_j}) u_{lk},$$

for all $k = 1, \cdots, m_2$.

4. Finally, set $F = (F_1, \cdots, F_N)$ and $G = (G_1, \cdots, G_N)$.

In the following we will report our numerical tests. For simplicity, we refer to the StS mapping derived with the LABP as LABP-StS, and the StS mapping based on the asymptotic expansion of the traveling Bloch waves as ASYM-StS. First we consider the PA-One. In this case the analytical StS mapping is available. For the n-th mode in the y-direction, the exact StS mapping is given as in (46). The computed StS mapping, no matter which method is employed, is diagonalizable. In Table 6 we list the errors of ASYM-StS. We see generally the asymptotic method presents very accurate results except on the resonance wave number. For example, if $k = \pi$, the first y-mode is resonant.

	$k=\pi$	$k=\frac{5\pi}{4}$	$k=\sqrt{2}\pi$	$k=\sqrt{3}\pi$
$n=0$	1.50(-9)	4.60(-9)	7.02(-12)	5.91(-13)
$n=1$	7.58(-6)	1.78(-9)	1.07(-9)	9.44(-13)
$n=2$	2.13(-12)	3.52(-12)	1.31(-11)	3.23(-12)
$n=3$	5.44(-13)	8.74(-13)	2.80(-12)	5.40(-13)
$n=4$	2.28(-13)	3.24(-13)	1.00(-12)	2.10(-13)

Table 6: Errors of Direct computation.

In Table 7 we list the errors of the LABP-StS. They are derived with two times of extrapolation. We see that except at the resonance wave numbers, this method presents the results at least of the same quality of those derived by the asymptotic method. But when resonance occurs, the extrapolation technique is only of little use. In order to obtain high accuracy, one has to make the damping parameter very small, but this probably implies a numerical stability problem.

For the other two periodic structures PA-Two and PA-Three, no analytical expression is available on the exact StS mapping. We compare the numerical solutions by two different methods. From Table 8-9, we could conclude in principle these two methods

	$k=\pi$	$k=\frac{5\pi}{4}$	$k=\sqrt{2}\pi$	$k=\sqrt{3}\pi$
$n=0$	5.03(-9)	3.53(-12)	6.91(-12)	1.07(-12)
$n=1$	5.68(-3)	7.30(-12)	1.52(-8)	1.49(-12)
$n=2$	2.26(-12)	3.43(-12)	1.26(-11)	3.40(-12)
$n=3$	7.51(-13)	8.22(-13)	2.99(-12)	5.82(-13)
$n=4$	2.39(-13)	2.74(-13)	1.06(-12)	2.37(-13)

Table 7: $\varepsilon = 0.00125$. Extrapolation.

bring the same results. When k is away from the resonance wave number, these two methods present the results of same quality. But their difference becomes big when k approaches the resonance wave number. Considering the results for the homogeneous waveguide problem, we thus believe at the resonance wave numbers, the asymptotic method presents better solution.

	Relative error
$k^2=25$	1.31(-12)
$k^2=50$	3.26(-12)
$k^2=23.61$	3.89(-8)
$k^2=47.1$	6.76(-5)

Table 8: $\varepsilon = 0.00125$. Comparison. PA-Two

	Relative error
$k^2=5$	9.58(-13)
$k^2=25$	9.26(-13)
$k^2=11.20$	7.16(-9)
$k^2=19.29$	6.23(-10)

Table 9: $\varepsilon = 0.00125$. Comparison. PA-Three.

CONCLUSIONS

In this chapter we have generalized a recent result of Zheng [78] and derived an exact Dirichlet-to-Neumann artificial boundary condition for general problems with periodic structures at infinity. We considered in detail the bound state problem for the Schrödinger operator and a second order hyperbolic equation in two space dimensions. The proof of this new kernel expression for the artificial boundary condition was presented recently by Zhang and Zheng [76].

Secondly, we introduced a fast evaluation method of the Sommerfeld-to-Sommerfeld (StS) mapping for periodic structure problems. Our proposed strategy is an improvement of the recently developed recursive doubling process by Yuan and Lu for the evaluation of Dirichlet-to-Neumann maps. We presented numerical results for the Helmholtz equation and the time-dependent Schrödinger equation in one and two space dimensions with periodic structures including cases where the method of Yuan and Lu fails.

In the last part of this chapter we considered the Helmholtz equation in the semi-infinite periodic array in this paper. Since no radiation-like boundary condition is specified at infinity, the Helmholtz equation is in general not well-posed. To solve this problem we employed the limiting absorption principle. We have proposed a new algorithm which combines the doubling procedure of the second part and the extrapolation technique to obtain high-accuracy approximation to the exact StS mappings. Considering the computational complexity, we present another method which uses the asymptotic behavior of a limiting absorption principle solution. Though we could not prove, the validity of this method is strongly supported by our numerical evidences.

We believe that these results can be generalized to the derivation of fully discrete artificial boundary conditions in the spirit of [15] for periodic potential problems. These boundary conditions are directly derived for the numerical scheme. Another very challenging task would be the extension of the present work to multi-dimensional problems with periodic structures. Furthermore, our ideas can be extended to more complicated wave-like equations, such as Maxwell's equations and elastic wave equations. Besides, many relevant theoretical problems are left open in this chapter.

REFERENCES

[1] T. Abboud, Electromagnetic waves in periodic media, in: R. Kleinman, T. Angell, D. Colton, F. Santosa and I. Stakgold (Eds.), Proceedings of the Second International Conference on Mathematical and Numerical Aspects of Wave Propagation, Newark, DE, 1993, SIAM, Philadelphia, 1993: 1-9.

[2] B. Alpert, L. Greengard and T. Hagstrom, Nonreflcting boundary conditions for the time-dependent wave equation, J Comput Phys 2002; 180: 270-296.

[3] X. Antoine and C. Besse, Unconditionally stable discretization schemes of non-reflecting boundary conditions for the one-dimensional Schrödinger equation, J Comput Phys 2003; 181: 157-175.

[4] X. Antoine, C. Besse and V. Mouysset, Numerical Schemes for the simulation of the two-dimensional Schrödinger equation using non-reflecting boundary conditions, Math. Comp 2004; 73: 1779-1799.

[5] X. Antoine, C. Besse and S. Descombes, Artificial boundary conditions for one-dimensional cubic nonlinear Schrödinger equations, SIAM J. Numer. Anal. 2006; 43: 2272-2293.

[6] X. Antoine, A. Arnold, C. Besse, M. Ehrhardt and A. Schädle, A Review of Transparent and Artificial Boundary Conditions Techniques for Linear and Nonlinear Schrödinger Equations, Commun Comput Phys 2008; 4: 729-796. (open-access article)

[7] X. Antoine, C. Besse and J. Szeftel, Towards accurate artificial boundary conditions for nonlinear PDEs through examples, CUBO A Mathematical Journal 2009; 11: 29-48.

[8] X. Antoine, C. Besse and P. Klein, Absorbing Boundary Conditions for the One-Dimensional Schrödinger Equation with an Exterior Repulsive Potential, J Comput Phys 2009; 228: 312-335.

[9] F.M. Arscott, Periodic differential equations, Pergamon Press, Oxford, 1964.

[10] M. Barth and O. Benson Manipulation of dielectric particles using photonic crystal cavities, Appl Phys Lett 2006; 89: 253114.

[11] G. Bastard, Wave mechanics applied to semiconductor heterostructures, les éditions de physique, Les Ulis Cedex, France, 1988.

[12] P. Bienstman and R. Baets, Optical modelling of photonic crystals and VCSELs using eigenmode expansion and perfectly matched layers, Opt Quant Electron 2001; 33: 327-341.

[13] A.M.B. Braga and G. Hermann, Floquet waves in anisotropic periodically layered composites, J Acoust Soc Am 1992; 91: 1211-1227.

[14] L. Brillouin, Wave propagation in periodic structures, Dover Publications, New york, 1953.

[15] M. Ehrhardt, Discrete Artificial Boundary Conditions, Dissertation, TU Berlin, 2001.

[16] M. Ehrhardt and A. Arnold, Discrete transparent boundary conditions for the Schrödinger equation, Riv Mat Univ Parma 2001; 6: 57-108.

[17] M. Ehrhardt and R.E. Mickens, Solutions to the discrete Airy equation: Application to parabolic equation calculations, J Comput Appl Math 2004; 172: 183-206.

[18] M. Ehrhardt, Discrete transparent boundary conditions for Schrödinger-type equations for non-compactly supported initial data, Appl Numer Math 2008; 58: 660-673.

[19] M. Ehrhardt and C. Zheng, Exact artificial boundary conditions for problems with periodic structures, J Comput Phys 2008; 227: 6877-6894.

[20] M. Ehrhardt, H. Han and C. Zheng, Numerical simulation of waves in periodic structures, Commun Comput Phys 2009; 5: 849-870.

[21] M. Ehrhardt, J. Sun and C. Zheng, Evaluation of scattering operators for semi-infinite periodic arrays, Commun Math Sci 2009; 7: 347-364.

[22] M. Ehrhardt and C. Zheng, Implementing exact absorbing boundary condition for the linear one-dimensional Schrödinger problem with variable potential by Titchmarsh-Weyl theory, Preprint No. 1426, WIAS Berlin, July 2009,

[23] S. Fliss and P. Joly, Exact boundary conditions for time-harmonic wave propagation in locally perturbed periodic media, Appl Numer Math 2009; 59(9): 2155-2178

[24] W.M.C. Foulkes, L. Mitas, R. J. Needs and G. Rajagopal, Quantum Monte Carlo simulations of solids, Rev Mod Phys 2001; 73: 33-83.

[25] C. Fox, V. Oleinik and B. Pavlov, A Dirichlet-to-Neumann map approach to resonance gaps and bands of periodic networks, In: N. Chernov, Y. Karpeshina, I. W. Knowles, R. T. Lewis and R. Weikard (Eds.), Recent advances in differential equations and mathematical physics volume 412 of Contemp. Math., Amer. Math. Soc., Providence, RI, 2006: 151-169.

[26] D. Givoli, Non-reflecting boundary conditions, J Comput Phys 1991; 94: 1-29.

[27] D.J. Griffiths and C.A. Steinke, Waves in locally periodic media, Amer J Phys 2001; 69: 137-154.

[28] S.P. Guo and S. Albin, Numerical techniques for excitation and analysis of defect modes in photonic crystals, Opt Express 2003; 11: 1080-1089.

[29] T. Hagstrom, Radiation boundary conditions for the numerical simulation of waves, Acta Numerica 1999; 8: 47-106.

[30] H. Han, The artificial boundary method-numerical method of partial differential equations on unbounded domains, in Frontiers and Propests of Contemporary Applied Mathematics, edited by T. Li and P. Zheng, Higher Education Press, World Scientific, 2005: 33-58.

[31] Z. Han, E. Forsberg and S. He, Surface plasmon Bragg gratings formend in metal-insulator-metal waveguides, IEEE Photonics Techn Lett 2007; 19: 91-93.

[32] I. Harari, I. Patlashenko and D. Givoli, Dirichlet-to-Neumann maps for unbounded wave guides, J Comput Phys 1998; 143: 200-223.

[33] S.F. Helfert and R. Pregla, Efficient analysis of periodic structures, J Lightwave Technology 1998; 16:1694-1702.

[34] P.L. Ho and Y.Y. Lu, A bidirectional beam propagation method for periodic waveguides, IEEE Photonics Techn Lett 2002; 14: 325-327.

[35] L. Hörmander, Linear Partial Differential Operators, 4th Printing, Grundlehren der mathematischen Wissenschaften 116, Springer-Verlag, 1969.

[36] J. Jacobsen, Analytical, numerical, and experimental investigation of guided waves on a periodically strip-loaded dielectric slab, IEEE Trans Antennas and Propagation 1970; 18: 379-388.

[37] Z. Hu and Y.Y. Lu, Efficient analysis of photonic crystal devices by Dirichlet-to-Neumann maps , Optics Express 2008; 16: 17383-17399.

[38] Z. Hu and Y.Y. Lu, Efficient numerical method for analyzing coupling structures of photonic crystal waveguides, to appear in: IEEE Photonics Technology Letters.

[39] S.G. Johnson and J.D. Joannopoulos, Photonic crystals : the road from theory to practice, Kluwer Academic Publishers, 2002.

[40] P. Joly, J.-R. Li and S. Fliss, Exact Boundary Conditions for Periodic Waveguides Containing a Local Perturbation, Commun Comput Phys 2006; 1: 945-973.

[41] J.B. Keller and D. Givoli, Exact non-reflecting boundary conditions, J Comput Phys 1989; 82: 172-192.

[42] P. Kuchment, Floquet theory for partial differential equations, volume 60 of Operator Theory: Advances and Applications Birkhäuser Verlag, Basel, 1993.

[43] P. Kuchment, The mathematics of photonic crystals, Chapter 7 in: Mathematical modeling in optical science volume 22 of Frontiers in applied mathematics SIAM, Philadelphia, 2001.

[44] S. Li and Y.Y. Lu, Computing photonic crystal defect modes by Dirichlet-to-Neumann maps, Optics Express 2007; 15: 14454-14466.

[45] B.E. Little and H.A. Haus, A variational coupled-mode theory for periodic waveguides, IEEE J Quantum Elect 1995; 31: 2258-2264.

[46] Y.Y. Lu, Computing Dirichlet-to-Neumann maps for numerical simulation of photonic crystal structures, Proc Nat Inst Math Sci 2008; 3: 65-70.

[47] A. Ludwig and Y. Leviatan, Analysis of bandgap characteristics of two-dimensional periodic structures by using the source-model technique, J Opt Soc Am A 2003; 20: 1553-1562.

[48] J.S. Papadakis, Impedance formulation of the bottom boundary condition for the parabolic equation model in underwater acoustics, NORDA Parabolic Equation Workshop, NORDA Tech Note 143, 1982.

[49] C. Potel, Ph. Gatignol and J. F. De Belleval, Energetic criterion for the radiation of floquet waves in infinite anisotropic periodically multilayered media, Acustica-Acta Acustica 2001; 87: 340-351.

[50] M. Reed and B. Simon, Methods of modern mathematical physics II: Fourier analysis, self-adjointness, Academic Press, San Diego 1975.

[51] J.A. Richards, Analysis of Periodically Time-Varying Systems, Springer-Verlag, 1983.

[52] V.F. Rodríguez-Esquerre, M. Koshiba and H.E. Hernández-Figueroa, Finite-element analysis of photonic crystal cavities: Time and frequency domains, J Lightw Technol 2005; 23: 1514-1521.

[53] K. Sakoda, Optical Properties of Photonic Crystals Springer-Verlag, Berlin, 2001.

[54] F. Schmidt, Solution of Interior-Exterior Helmholtz-Type Problems Based on the Pole Condition Concept: Theory and Algorithms, Habilitation thesis, Free University Berlin, 2002.

[55] J.R. Schulenberger and C.H. Wilcox, The limiting absorption principle and spectral theory for steady-state wave propagation in inhomogeneous anisotropic media, Arch Ration Mech Anal 1971; 41: 46-65.

[56] I.A. Semenikhin, B.S. Pavlov and V.I. Ryzhii, Plasma waves in two-dimensional electron channels: propagation and trapped modes, Preprint No. NI07028-AGA of the Isaac Newton Institute for Mathematical Sciences, 2007.

[57] S. Shipman and D. Volkov, Guided modes in periodic slabs: existence and nonexistence, SIAM J Appl Math 2007; 67: 687-713.

[58] D. Sjöberg, Analysis of large finite periodic structures using infinite periodicity methods, Technical Report TEAT-7143, Lund Institute of Technology, Sweden, 2006.

[59] D.R. Smith, R. Dalichaouch, N. Kroll, S. Schultz, S.L. McCall and P.M. Platzman, Photonic band structure and defects in one and two dimensions, J Opt Soc Am B 1993; 10: 314-321.

[60] D.R. Smith, J.B. Pendry and M.C.K. Wiltshire, Metamaterials and Negative Refractive Index, Science 2004; 305: 788-792.

[61] T. Søndergard, S.I. Bozhevolnyi and A. Boltasseva, Theoretical analysis of ridge gratings for long-range surface plasmon polaritons, Phys Rev B 2006; 73: 045320.

[62] S. Soussi, Convergence of the supercell method for defect modes calculations in photonic crystals, SIAM J Numer Anal 2005; 43: 1175-1201.

[63] J. Tausch and J. Butler, Floquet Multipliers of periodic Waveguides via Dirichlet-to-Neumann Maps, J Comput Phys 2000; 159: 90-102.
[64] J. Tausch and J. Butler, Efficient Analysis of Periodic Dielectric Waveguides using Dirichlet-to-Neumann Maps, J Opt Soc Amer A 2002; 19: 1120-1128.
[65] S.V. Tsynkov, Numerical solution of problems on unbounded domains. A review, Appl Numer Math 1998; 27: 465-532.
[66] M.N. Vouvakis, Z. Kezhong and J.-F. Lee, Finite-element analysis of infinite periodic structures with nonmatching triangulations, IEEE Trans Magnetics 2006; 42: 691-694.
[67] B.R. Vainberg, Principles of radiation, limit absorption and limit amplitude in the general theory of partial differential equations, Russ Math Surv 1966; 21: 115-193.
[68] A. Wacker, Semiconductor Superlattices: A model system for nonlinear transport, Phys Rep 2002; 357: 1-111.
[69] Y. Wu and Y.Y. Lu, Dirichlet-to-Neumann map method for analyzing crossed arrays of circular cylinders, to appear in: Journal of the Optical Society of America B.
[70] A. Yariv and P. Yeh, Optical Waves in Crystals – Propagation and Control of Laser Radiation, Wiley Series in Pure and Applied Optics, Wiley, 2002.
[71] L. Yuan and Y.Y. Lu, An efficient bidirectional propagation method based on Dirichlet-to-Neumann maps, IEEE Photonics Techn Lett 2006; 18: 1967-1969.
[72] L. Yuan and Y.Y. Lu, Dirichlet-to-Neumann map method for second harmonic generation in piecewise uniform waveguides, J Opt Soc of Am B 2007; 24: 2287-2293.
[73] L. Yuan and Y.Y. Lu, A Recursive Doubling Dirichlet-to-Neumann Map Method for Periodic Waveguides, J Lightwave Technology 2007; 25: 3649-3656.
[74] J. Yuan, Y.Y. Lu and X. Antoine, Modeling photonic crystals by boundary integral equations and Dirichlet-to-Neumann maps, J Comput Phys 2008; 227: 4617-3629.
[75] L. Yuan and Y.Y. Lu, An efficient numerical method for optical waveguides with holes, J Lightwave Technology 2009; 27: 2557-2562.
[76] M. Zhang and C. Zheng, Closed form impedance expression for periodic Schrödinger operators with symmetric coefficient functions, to be submitted to: Frontiers of Mathematics in China, 2009.
[77] C. Zheng, Approximation, stability and fast evaluation of an exact artificial boundary condition for the one-dimensional heat equation, J Comput Math 2007; 25: 730-745.
[78] C. Zheng, An exact boundary condition for the Schrödinger equation with sinusoidal potentials at infinity, Commun Comput Phys 2007; 3: 641-658.