A High Order Numerical Method for Scattering from Locally Perturbed Periodic Surfaces

Ruming Zhang*

Abstract

In this paper, we will introduce a high order numerical method to solve the scattering problems with non-periodic incident fields and (locally perturbed) periodic surfaces. For the problems we are considering, the classical methods to treat quasi-periodic scattering problems no longer work, while a Bloch transform based numerical method was proposed in [LZ17b]. This numerical method, on one hand, is able to solve this kind of problems convergently; on the other hand, it takes up a lot of time and memory during the computation. The motivation of this paper is to improve this numerical method, from the regularity results of the Bloch transform of the total field, which have been studied in [Zha17]. As the set of the singularities of the total field is discrete in $\mathbb{R}$, and finite in one periodic cell, we are able to improve the numerical method by designing a proper integration contour with special conditions at the singularities. With a good choice of the transformation, we can prove that the new numerical method could possess a super algebraic convergence rate. This new method improves the efficiency significantly. At the end of this paper, several numerical results will be provided to show the fast convergence of the new method. The method also provides a possibility to solve more complicated problems efficiently, e.g., three dimensional problems, or electromagnetic scattering problems.

1 Introduction

In this paper, we will propose an efficient numerical method of solving scattering problems from locally perturbed periodic surfaces with non-periodic incident fields. For the special case that a quasi-periodic incident field scattered by a smooth enough periodic surface, a classical way is to reduce the problem into one periodic cell, see [Kir93, Sch03]. Then the reduced problem could be solved by numerical methods, such as the integral equation method (see [MACK00]), and the finite element method (see [Bao05, HNPX11]). However, when the incident field or the surface is no longer (quasi-)periodic, the classical method fails, new numerical methods are required for these more difficult problems. In [LZ17b], a Floquet-Bloch based numerical method was applied to solve this kind of problems. In this paper, the method is improved by reducing the computational complexity significantly.

The Floquet-Bloch based methods were used in [Cox12, HNi01, LZ17a, LZ17c, LZ17b] for the numerical solutions in (locally perturbed) periodic scattering problems with non-periodic incident fields. For theoretical results, see [LN15] and [Lec17]. There is also another way to solve the locally perturbed periodic problems, by studying the operators defined on the transparent edge of a periodic cell. In [JLF06, FJ09, FJ15], a method that approximates the

*Center for Industrial Mathematics, University of Bremen; rzhang@uni-bremen.de
Dirichlet-to-Neumann map by solving an operator equation is applied, while in [EHZ09,ESZ09], the problems were solved by approximating the Scattered-to-Scattered map by a so-called doubling recursive procedure.

The method in this paper follows the results in [LZ17b,Zha17]. To solve the scattering problems with a locally perturbed surface, firstly we have to transform the problem into one defined on a periodic domain. With the Bloch transform, the problem is then transformed into an equivalent coupled system of quasi-periodic problems, which is a three dimensional problem defined in a bounded domain. The error comes from two parts in the algorithm – the finite element method and the numerical inverse Bloch transform. As the error brought by the finite element method could be estimated by classical methods, in this paper, we will improve the numerical scheme by the study of the numerical inverse Bloch transform. To study this, a detailed behaviour of the Bloch transformed solutions is required, which has been shown in [Zha17].

Denote by \( J_\Omega \) the Bloch transform defined on the periodic domain \( \Omega \), then the total field \( u \) is transformed into \( w(\alpha,x) = (J_\Omega u)(\alpha,x) \). From [Zha17], the singularities in \( \alpha \) locate in a discrete set in \( \mathbb{R} \), with a square root like singularity in the neighbourhoods. As the inverse Bloch transform has the representation of the integral on a bounded interval, in this paper, we will pick up a new curve and let the inverse Bloch transform be defined on the curve such that the newly defined integrand satisfies some smoothness and periodic conditions. This technique to treat integrations with singularities is quite classical, and it has been described in detail and applied to the numerical solution of the scattering problems from bounded obstacles with corners in Chap 3.5, [CK13] (pp. 83-85). From the well known result of the convergent rule for periodic functions, see [Atk89], we can prove that with the new integral curve, the numerical method will converge much faster, depending on the choice of the new integration curve.

The high order numerical method is available when the incident fields satisfy certain conditions. As far as we know, both the point sources and the Herglotz wave functions satisfy the conditions, thus the scattering problems with these incident fields could be solved numerically by the high order method introduced in this paper. Numerical examples for these cases will be presented at the end of this paper. However, the plane waves do not satisfy the conditions, so the numerical method fails in this case.

The rest of the paper is organized as follows. In Section 2, we will review some results for Bloch transform and the locally perturbed periodic problems. In Section 3, we recall the singularity results for the Bloch transform of the total field. In Section 4, we will introduce the inverse Bloch transform defined on a new contour. In Section 5, we will study the finite element method with the new inverse Bloch transform. In Section 6, some numerical examples will be shown.

2 Bloch transform and the locally perturbed periodic problems

2.1 Introduction to the Scattering problems

Suppose \( \zeta \) is a Liptschiz periodic function with period \( \Lambda \), and the periodic surface is defined by

\[
\Gamma := \{(x_1, \zeta(x_1)) : x_1 \in \mathbb{R}\}.
\]
Let $\zeta_p$ be a function such that $\text{supp}(\zeta_p - \zeta)$ is compact in $\mathbb{R}$. For simplicity, assume that $\text{supp}(\zeta_p - \zeta) \subset (-\frac{\Lambda}{2}, \frac{\Lambda}{2})$, then define the locally perturbed surface by

$$\Gamma_p := \{(x_1, \zeta_p(x_1)) : x_1 \in \mathbb{R}\}.$$ 

Define $\Lambda^* = \frac{2\pi}{\Lambda}$, and

$$W_{\Lambda} = \left( -\frac{\Lambda}{2}, \frac{\Lambda}{2} \right), \quad W_{\Lambda^*} = \left( -\frac{\Lambda^*}{2}, \frac{\Lambda^*}{2} \right) = \left( -\frac{\pi}{\Lambda}, \frac{\pi}{\Lambda} \right).$$

We can define the space above the surfaces $\Gamma$ and $\Gamma_p$ by

$$\Omega := \{(x_1, x_2) : x_1 \in \mathbb{R}, x_2 > \zeta(x_1)\}$$
$$\Omega_p := \{(x_1, x_2) : x_1 \in \mathbb{R}, x_2 > \zeta_p(x_1)\}.$$ 

For simplicity, we assume that $\Gamma$ and $\Gamma_p$ are both above the straight line $\mathbb{R} \times \{0\}$, and $H$ be a positive number such that $H > \max\{\sup\{\zeta\}, \sup\{\zeta_p\}\}$. Define the line $\Gamma_H = \mathbb{R} \times \{H\}$ and the spaces with finite height

$$\Omega^H := \{(x_1, x_2) : x_1 \in \mathbb{R}, \zeta(x_1) < x_2 < H\}$$
$$\Omega^H_p := \{(x_1, x_2) : x_1 \in \mathbb{R}, \zeta_p(x_1) < x_2 < H\}.$$ 

We can also define the curves and domains in one periodic cell:

$$\Gamma^\Lambda = \Gamma \cap W_\Lambda, \quad \Gamma_H^\Lambda = \Gamma_H \cap W_\Lambda;$$
$$\Omega^\Lambda = \Omega \cap W_\Lambda \times \mathbb{R}, \quad \Omega_H^\Lambda = \Omega^H \cap W_\Lambda \times \mathbb{R}.$$ 

![Diagram of locally perturbed periodic structure.](image)

Figure 1: Locally perturbed periodic structure.

Given an incident field $u^i$, the scattering problem is to find a total field $u$ such that it satisfies the Helmholtz equation in the space $\Omega_p$

$$\Delta u + k^2 u = 0 \quad \text{in} \ \Omega_p \quad (1)$$

3
and the scattered field \( u^s : = u - u^i \) propagates upward, i.e.,

\[
u^s(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ix_1 \xi + i\sqrt{k^2 - |\xi|^2}(x_2 - H)} \hat{u}^s(\xi, H) \, d\xi \quad x_2 > H,
\]

where \( \hat{u}^s(\xi, H) \) is the Fourier transform of \( u^s(\cdot, H) \). The upward propagating condition is equivalent to

\[
\frac{\partial u}{\partial x_2} = T^+ u|_{\Gamma_H} + \left( \frac{\partial u^i}{\partial x_2} - T^+ [u^i|_{\Gamma_H}] \right) \quad \text{on } \Gamma_H,
\]

where \( T^+ \) is the Dirichlet-to-Neumann map (see [CM05]) defined by

\[
T^+ \varphi(x_1) = \frac{i}{2\pi} \int_{\mathbb{R}} \sqrt{k^2 - |\xi|^2} e^{ix_1 \xi + i\sqrt{k^2 - |\xi|^2}(x_2 - H)} \tilde{\varphi}(\xi) \, d\xi
\]

that is bounded from \( H^{1/2}_r(\Gamma_H) \) to \( H^{-1/2}_r(\Gamma_H) \) for all \(|r| < 1\). Define the total field \( u = u^i + u^s \), then we can formulate the variational problem, i.e., to find \( u \in \tilde{H}^1_{r}(\Omega^H_p) \) such that for any \( \varphi \in H^1(\Omega^H_p) \) with a compact support such that

\[
\int_{\Omega^H_p} \left[ \nabla u \cdot \nabla \varphi - k^2 u \varphi \right] \, dx - \int_{\Gamma_H} T^+ (u|_{\Gamma_H}) \varphi \, ds = \int_{\Gamma_H} \left[ \frac{\partial u^i}{\partial x_2} - T^+ [u^i|_{\Gamma_H}] \right] \varphi \, ds,
\]

where \( H^s_r(\Omega^H_p) \) is a weighted space defined by

\[
H^s_r(\Omega^H_p) := \left\{ \varphi \in \mathcal{D}'(\Omega^H_p) : (1 + |x_1|^2)^{r/2} \varphi(x) \in H^s(\Omega^H_p) \right\},
\]

and \( \tilde{H}^s_r(\Omega^H_p) \) is a subspace defined by

\[
\tilde{H}^s_r(\Omega^H_p) = \left\{ \varphi \in H^s_r(\Omega^H_p) : \varphi|_{\Gamma_p} = 0 \right\}.
\]

In [CE10], the unique solvability of the variational form was proved.

**Theorem 1.** For any \(|r| < 1\), given any \( u^i \in H^1(\Omega^H_p) \), there is a unique solution \( u \in \tilde{H}^1_r(\Omega^H_p) \) of the variational problem

2.2 Bloch transform and well-posedness of the scattering problems

Firstly, we have to introduce the Bloch transform defined on the periodic space \( \Omega^H \). For any function \( \varphi \in C_0^\infty(\Omega^H) \), define

\[
(J_{\Omega^H} \varphi)(\alpha, x) = \sum_{j \in \mathbb{Z}} \varphi(x_1 + \Lambda j, x_2)e^{-i\alpha \Lambda j}, \quad \alpha \in W_\Lambda^*; \ x \in \Omega^H.
\]

**Remark 2.** In this paper, we denote the one dimensional Bloch transform defined on any periodic domain in \( \mathbb{R}^2 \) by \( J_\Omega \).

For any fixed \( \alpha \in W_\Lambda^* \), the function \( (J_{\Omega^H} \varphi) \) is \( \alpha \)-quasi-periodic with period \( \Lambda \), i.e.,

\[
(J_{\Omega^H} \varphi) \left( \alpha, \left( x_1 + \Lambda j \atop x_2 \right) \right) = e^{i\alpha \Lambda} (J_{\Omega^H} \varphi).
\]
For the positive integer \( m \), let the space \( H^m(W_{\Lambda^r}; H^s(\Omega_H^\lambda)) \) be defined with the norm
\[
\| \varphi \|_{H^m(W_{\Lambda^r}; H^s(\Omega_H^\lambda))} = \left[ \sum_{\ell=0}^m \int_{W_{\Lambda^r}} \| \partial_\alpha^\ell \varphi(x,\cdot) \|_{H^s(\Omega_H^\lambda)}^2 \right]^{1/2},
\]
then extend the definition to any real number \( r \) by interpolation and duality arguments. Define the subspace \( H^r_0(W_{\Lambda^r}; H^s(\Omega_H^\lambda)) \) with a further condition that \( \varphi \in H^r(W_{\Lambda^r}; H^s(\Omega_H^\lambda)) \) is \( \Lambda^r \)-periodic in \( \alpha \) and \( \varphi(\alpha,\cdot) \in H^s(\Omega_H^\lambda) \). The mapping property of the Bloch transform was proved in [Lec17].

**Theorem 3.** The Bloch transform extends to an isomorphism between \( H^r_s(\Omega) \) and \( H^0_0(W_{\Lambda^r}; H^s(\Omega_H^\lambda)) \) for any \( s, r \in \mathbb{R} \). When \( s = r = 0 \), \( J_{\Omega} \) is an isometry with the inverse operator
\[
(J_{\Omega})^{-1} \left( \begin{array}{c} x_1 + \Lambda^j \vspace{1mm} \\ x_2 \end{array} \right) = \left[ \frac{\Lambda}{2\pi} \right]^{1/2} \int_{W_{\Lambda^r}} w(\alpha,x)e^{i\Lambda^j \alpha} \, d\alpha, \quad x \in \Omega_H^\lambda, \; j \in \mathbb{Z}. \tag{6}
\]
Moreover, the \( L^2 \)-adjoint operator \( J_{\Omega}^* \) equals to its inverse.

As the Bloch transform is defined on periodic domains, we have to transform the problem defined on a locally perturbed domain into one defined on a periodic domain. Following [Lec17, LZ17], define a diffeomorphism \( \Phi_p \) from \( \Omega_H^\lambda \) into \( \Omega_H^\lambda \) by
\[
\Phi_p(x_1,x_2) = \left( x_1, x_2 + \frac{(x_2 - H)^3}{(\zeta(x_1) - H)^3}(\zeta_p(x_1) - \zeta(x_1)) \right),
\]
which has a support contained in \( \Omega_H^\lambda \). Then \( u_T := u \circ \Phi_p \) satisfies the following variational problem
\[
\int_{\Omega_H^\lambda} \left[ A_p \nabla u_T \cdot \nabla \varphi_T - k^2 c_p u_T \varphi_T \right] \, dx - \int_{\Gamma_H^\lambda} T^+(u_T|\Gamma_H^\lambda) \varphi_T \, ds = \int_{\Gamma_H^\lambda} \left[ \frac{\partial u_T}{\partial x_2} - T^+(u_T|\Gamma_H^\lambda) \right] \varphi_T \, ds \tag{7}
\]
for all \( \varphi_T \in H^1(\Omega_H^\lambda) \) with compact support, where
\[
A_p(x) = \left| \det \nabla \Phi_p(x) \right| [([\nabla \Phi_p(x)]^{-1}(\nabla \Phi_p(x))^{-1})^T] \in L^\infty(\Omega_H^\lambda, \mathbb{R}^{2 \times 2}),
\]
\[
c_p(x) = \left| \det \nabla \Phi_p(x) \right| \in L^\infty(\Omega_H^\lambda).
\]
Let \( w = J_{\Omega} u_T \) then \( w \in L^2(W_{\Lambda^r}; \tilde{H}^1_0(\Omega_H^\lambda)) \), and satisfies the following variational problem for any \( \varphi \in L^2(W_{\Lambda^r}; \tilde{H}^1_0(\Omega_H^\lambda)) \)
\[
\int_{W_{\Lambda^r}} a_\alpha(w(\cdot,\cdot),\varphi(\cdot,\cdot)) \, d\alpha + b(J_{\Omega}^{-1} w, J_{\Omega}^{-1} \varphi) = \int_{W_{\Lambda^r}} \int_{\Gamma_H^\lambda} f(\alpha,\cdot) \varphi(\alpha,\cdot) \, ds \, d\alpha, \quad (8)
\]
where
\[ a_\alpha(w(\alpha, \cdot), \varphi(\alpha, \cdot)) = \int_{\Omega_H^\Lambda} \left[ \nabla w(\alpha, \cdot) \cdot \nabla \varphi(\alpha, \cdot) - k^2 w(\alpha, \cdot) \varphi(\alpha, \cdot) \right] \, dx \]
\[ - \int_{\Gamma_H} T_\alpha^+(w(\alpha, \cdot)|\Gamma_H) \varphi(\alpha, \cdot) \]
\[ b(\xi, \psi) = \left[ \frac{\Lambda}{2\pi} \right]^{1/2} \int_{\Omega_H} (A_p - I_2) \nabla \xi \cdot \nabla \psi \, dx - k^2 \left[ \frac{\Lambda}{2\pi} \right]^{1/2} \int_{\Omega_H} (c_p - 1) \xi \psi \, dx, \]
\[ f(\alpha, \cdot) = \frac{\partial J_{\Omega} u^i(\alpha, \cdot)}{\partial x_2} - T_\alpha^+ \left[ (J_{\Omega} u^i)(\alpha, \cdot)|\Gamma_H^\Lambda \right], \]
with \( T_\alpha^+ \) defined by
\[ T_\alpha^+(\psi) = i \sum_{j \in Z} \sqrt{k^2 - |\Lambda^* j - \alpha|^2} \hat{\psi}(j) e^{i(\Lambda^* j - \alpha)x_1} \qquad \psi = \sum_{j \in Z} \hat{\psi}(j) e^{i(\Lambda^* j - \alpha)x_1}. \]  

**Remark 4.** The system is decoupled when the surface is purely periodic. In this case, \( \Phi_p \) is the identity operator, thus \( A_p = I_2 \) and \( c_p = 1 \). From the definition of the term \( b \), it disappears in this case, thus the system has the form of
\[ \int_{W_\alpha^*} a_\alpha(w(\alpha, \cdot), \varphi(\alpha, \cdot)) \, d\alpha = \int_{W_\alpha^*} \int_{\Gamma_H} f(\alpha, \cdot) \varphi(\alpha, \cdot) \, ds \, d\alpha. \]

At the end of this subsection, we will list some of the useful results in [LZ17b].

**Theorem 5.** Assume \( u^i \in H^2_r(\Omega_p^H) \) for some \( r \in [0,1) \), then \( u_T \in \tilde{H}^2(\Omega^H) \) satisfies [5] if and only if \( w = J_{\Omega} u_T \in H_0^1(W_\Lambda^*; \tilde{H}^1(\Omega_p^H)) \) satisfies [3]. If \( \Gamma_p \) is Lipschitz, then [3] is uniquely solvable in \( H_0^1(W_\Lambda^*; \tilde{H}^1(\Omega_p^H)) \) for all \( u^i \in H^1(\Omega_p^H), r \in [0,1). \)

**Theorem 6.** If \( u^i \in H^2_r(\Omega_p^H) \) for \( r \in [0,1) \) and \( \zeta, \zeta_p \in C^{2,1}(\mathbb{R}) \) then \( w(\alpha, \cdot) \in \tilde{H}^2(\Omega^\Lambda_H) \) and \( u_T = J_{\Omega}^{-1} w \in \tilde{H}^2(\Omega^H) \).

**Theorem 7.** If \( \Gamma_p \) is Lipschitz continuous and \( u^i \in H^1(\Omega_p^H) \) for \( r \in (1/2,1) \), then \( w \in L^2(W_\Lambda^*; \tilde{H}^1(\Omega_p^H)) \) equivalently satisfies for all \( \alpha \in W_\Lambda^* \) and \( \varphi_\alpha \in \tilde{H}^1(\Omega_p^H) \) that
\[ a_\alpha(w(\alpha, \cdot), \varphi_\alpha) + b(J_{\Omega}^{-1} w, \varphi_\alpha) = \int_{\Gamma_H} f(\alpha, \cdot) \varphi_\alpha \, ds, \]  

**2.3 Numerical method**

Assume that \( M_h \) is a family of regular and quasi-uniform meshes (see [BS94]) of \( \Omega_H^\Lambda \), such that the width of each mesh is not larger than \( h \). Suppose \( \{ \varphi^{(1)}_{\alpha_M} \}^M_{i=1} \) is the piecewise linear nodal function that equals to one at the \( l \)-th nodal point \( x^{(l)} = \left( x^{(l)}_1, x^{(l)}_2 \right) \) while equals to zero at other nodal points. Then the space spanned by \( \{ \varphi^{(1)}_{\alpha_M} \}^M_{i=1} \) is a subset of \( H^1_0(\Omega_H^\Lambda) \).

For the domain \( W_\Lambda^* \), define the uniformly located grid points
\[ \alpha^{(j)}_N = -\frac{\pi}{\Lambda} + \frac{2\pi j}{N \Lambda}, \quad j = 1, 2, \ldots, N, \]
Theorem 8. Assume that \( u^i \in H^2_r(\Omega^H_\alpha) \) for \( r \geq 1/2 \), \( \zeta, \zeta_p \in C^{2,1}(\mathbb{R}) \), then is uniquely solvable in \( X_{N,h} \). The error between the numerical solution \( w_{N,h} \) and exact one \( w \) is bounded by

\[
\|w_{N,h} - w\|_{L^2(W_{\alpha,h};H^0_\alpha(\Omega^H_\alpha))} \leq Ch^{1-l}(N^{-r} + h)\|f\|_{H^0_\alpha(H^{1/2}(\Gamma_h))}.
\]

Remark 9. For the special case that the surface is non-perturbed, the same error estimate of the numerical result is also obtained. Moreover, when \( u^i \) satisfies that \( J_\Omega u^i \in W^{1,p}_p(W_{\alpha,h};H^1_\alpha(\Omega^H_\alpha)) \) for some \( p \in [1,2) \), then

\[
\|w_{N,h} - w\|_{L^2(W_{\alpha,h};H^0_\alpha(\Omega^H_\alpha))} \leq C(h^{2-l} + N^{-1})\|f\|_{L^2(W_{\alpha,h};H^{1/2}(\Gamma_h))}.
\]

The space \( W^{1,p}_p(W_{\alpha,h};H^1_\alpha(\Omega^H_\alpha)) \) is defined by

\[
\|w\|^p_{W^{1,p}_p(W_{\alpha,h};H^1_\alpha(\Omega^H_\alpha))} = \int_{W_{\alpha,h}} \left[ \|w(\alpha, \cdot)\|^p_{H^1_\alpha(\Omega^H_\alpha)} + \left\| \frac{\partial}{\partial \alpha} w(\alpha, \cdot) \right\|^p_{H^1_\alpha(\Omega^H_\alpha)} \right] d\alpha.
\]

For details see [LZ17c].

We will introduce the numerical method in [LZ17c] briefly in the rest of this section. Let \( w^s(\alpha, x) = w(\alpha, x) - (J_\Omega u^i)(\alpha, x) \) and \( w_0(\alpha, x) = e^{i\alpha x}; w^s(\alpha, x) \) and \( w_0(\alpha, x) \) in \( W^2(\Omega^H_\alpha;H^0_\alpha(\Omega^H_\alpha)) \).

Recall that for the choice of \( X_{N,h} \), piecewise constant basis functions are chosen for \( \alpha \) in \( W_j \), we can define the discrete space as follows. Still denote by \( \{\varphi^{(j)}_M\}_{j=1}^M \) the piecewise linear nodal functions in \( \Omega^H_\alpha \), and \( \{\psi^{(l)}_N\}_{l=1}^N \) the piecewise constant basis function in \( W_j \). Suppose there is a positive integer \( M' < M \) such that \( \varphi_\ell|_{\Gamma_\alpha} = 0 \) for any \( \ell = 1, \ldots, M' \) and \( \varphi_\ell|_{\Gamma_\alpha} \neq 0 \) for any \( \ell = M' + 1, \ldots, M \). Thus span \( \{\varphi^{(j)}_M\}_{j=1}^{M'} \subset H^1_\alpha(\Omega^H_\alpha) \).

Define the discrete space by

\[
Y^{(j)}_{N,h} = \left\{ v^{(j)}_{N,h} = \sum_{\ell=1}^{M'} e^{(j)}_{N,h} \psi^{(j)}_N(\alpha) \varphi^{(l)}_M(x) : v^{(j)}_{N,h} \in \mathbb{C} \right\} \subset L^2(\Omega^H_\alpha;H^1_\alpha(\Omega^H_\alpha))
\]
for $j = 1, \ldots, N$ and also the subspace of functions that vanish on $\Gamma^A$

$$
\tilde{Y}^{(j)}_{N,h} = \left\{ v^{(j)}_{N,h} = \sum_{\ell=1}^{M'} v^{(j)}_{N,h}(\ell) \phi_M^{(j)}(x) : v^{(j)}_{N,h} \in \mathbb{C} \right\} \subset L^2(W_\Lambda^*; \tilde{H}_0^0(\Omega_H^\ell)).
$$

Then setting

$$
Y^0_{N,h} = Y^{(1)}_{N,h} \oplus \cdots \oplus Y^{(N)}_{N,h}, \quad \tilde{Y}^0_{N,h} = \tilde{Y}^{(1)}_{N,h} \oplus \cdots \oplus \tilde{Y}^{(N)}_{N,h},
$$

then for any $w_0 \in X_{N,h}$, there is a unique vector $(w^{(j)}_0)_{j=1}^N \in Y^0_{N,h}$ such that $w_0 = \sum_{j=1}^N \exp(-\alpha^{(j)}_N x_1) w^{(j)}_0$. Then the inverse Bloch transform $J^{-1}_\Omega^N$ applied to the function $w_0$ equals to the numerical inverse Bloch transform $J^{-1}_{\Omega,N}$ applied to \( \{ e^{-\alpha^{(j)}_N x_1} w^{(j)}_0(\alpha^{(j)}_N, x) \}_{j=1}^N \)

\[
J^{-1}_\Omega w_0 = J^{-1}_{\Omega,N} \left( \{ e^{-\alpha^{(j)}_N x_1} w^{(j)}_0(\alpha^{(j)}_N, x) \}_{j=1}^N \right) = \frac{\Lambda^*}{N} \sum_{j=1}^N e^{-\alpha^{(j)}_N x_1} w^{(j)}_0(\alpha^{(j)}_N, x).
\]

Suppose $w_0 = \sum_{j=1}^N \exp(-\alpha^{(j)}_N x_1) w^{(j)}_0$ where $w^{(j)}_0 = \sum_{m=1}^M w^{(j,m)}_{N,h} \psi^{(j)}_M \tilde{\gamma}^{(m)}_j$, omitting the details in [LZ17b], the final linear system is set up as follows.

\[
\begin{pmatrix}
A_1 & 0 & \cdots & 0 & C_1 \\
0 & A_2 & \cdots & 0 & C_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & A_N & C_N & W_N \\
B_1 & B_2 & \cdots & B_N & I_M
\end{pmatrix}
= \begin{pmatrix}
W_1 \\
W_2 \\
\vdots \\
W_N \\
U
\end{pmatrix}
= \begin{pmatrix}
F_1 \\
F_2 \\
\vdots \\
F_N \\
0
\end{pmatrix}, \tag{12}
\]

where $A_j, B_j, C_j, I_m$ are $M \times M$ matrices and $F_j$ are $M \times 1$ vertices, $W_j = (w^{(j,1)}_{N,h}, \ldots, w^{(j,m)}_{N,h})^\top$ and $U = (u^{(1)}_{N,h}, \ldots, u^{(M)}_{N,h})^\top$, where

$$u^{(N)}_{N,h} = \frac{\Lambda^*}{N} \sum_{j=1}^N e^{-\alpha^{(j)}_N x_1} w^{(j)}_{N,h}.$$

The matrix linear system (12) is a $M(N+1) \times M(N+1)$-sparse matrix with $3N+1$ non-zero $M \times M$ blocks. In [LZ17b], a GMRES iteration method with an incomplete LU-decomposition based pre-conditioner was applied to the solution of the linear system. The convergent rate with respect to $N$ is $N^{-r}$ where $1/2 \leq r < 1$, which means that if the mesh size $h$ is small enough, a small error requires a relatively large $N$. However, if $N$ gets larger, there are several disadvantages to solve the linear system (12).

- Setting up the matrix takes up a lot of time and memory.
- The time of solving the system is long.

To solve the scattering problems more efficiently, we would like to look for a new method that converges faster in $N$. 

8
3 Regularity results for the scattering problems

A natural way to look for a new numerical method that converges faster as \( N \) tends to \( \infty \) is to go deeper into the regularity of the field \( w(\alpha, \cdot) \) with respect to \( \alpha \in \mathbb{W} \). Theorem a in [Kir93] gave a result for non-perturbed surfaces with incident plane waves. It could be a good inspiration for the problems in this paper.

**Theorem 10** ([Kir93], Theorem a). Suppose the incident field \( u^i(\alpha, x) = e^{i\alpha x_1 - \sqrt{k^2 - \alpha^2} x_2} \) for \( \alpha \in (-k, k) \), and it is scattered by a smooth enough sound soft periodic surface.

a) The total field, denoted by \( u(\alpha, x) \), depends continuously on \( \alpha \in (-k, k) \) and analytically on \( \alpha \) in the set
\[
\{ \alpha \in (-k, k) : (n\Lambda^* - \alpha)^2 \neq k^2 \text{ for every } n \in \mathbb{Z} \}.
\]

b) Let \( \alpha_0 \in (-k, k) \) such that \( (n_0\Lambda^* - \alpha_0)^2 = k_0^2 \) for some \( n_0 \in \mathbb{Z} \). Then there is a neighbourhood \( U \) of \( \alpha_0 \) and quasi-periodic functions \( v(\alpha, \cdot), w(\alpha, \cdot) \in \tilde{H}^1(\Omega H^\Lambda) \) that depend analytically on \( \alpha \in U \) and satisfies \( u(\alpha, x) = v(\alpha, x) + \sqrt{k^2 - |n_0\Lambda^* - \alpha|^2} w(\alpha, x) \).

From the proof of Theorem a in [Kir93], the singularity of the total fields with respect to incident plane waves comes from the DtN map. The singularity occurs at the points where at least one term of the DtN map vanishes (so called Wood anomalies, see [May84]). However, for the scattering problems in this paper, the singularities may also come from the slow decay of the incident fields. In this section, we will only focus on the singularity comes from the DtN map. We will firstly consider the case that the incident field with compact support, and then extend the results to more generalized and commonly used incident fields. We will give a brief description of these singularity results, for details see [Zha17].

### 3.1 Notations and function spaces

Define the set of Wood anomalies by
\[
S := \{ \alpha \in \mathbb{R} : \exists n \in \mathbb{Z}, \text{ such that } |\Lambda^* n - \alpha| = k \},
\]
which is a countable set in \( \mathbb{R} \). Let \( \{ \alpha_j \}_{j \in \mathbb{Z}} \) be an ascending sequence of all the points in \( S \). To define the continuous or analytical "dependence" rigorously, we have to define the following function spaces first.

Let \( \mathcal{I} \subset \mathbb{R} \) be an interval (\( \mathcal{I} = \mathbb{R} \) included), \( W \subset \mathbb{R}^2 \) be a bounded domain, and define the space of functions defined in the \( \mathcal{I} \times W \) that depend analytically on the first variable by
\[
C^\omega(\mathcal{I}; S(W)) := \left\{ f \in \mathcal{D}'(\mathcal{I} \times W) : \forall \alpha_0 \in \mathcal{I}, \exists \delta > 0, \text{ s.t., } \forall \alpha \in (\alpha_0 - \delta, \alpha_0 + \delta) \cap \mathcal{I}, \exists C > 0, f_n \in S(W), \text{ s.t., } f(\alpha, x) = \sum_{n=0}^{\infty} (\alpha - \alpha_0)^n f_n(x), \|f_n\|_{S(W)} \leq C^n \right\},
\]
where \( S(W) \) is a Sobolev space defined on the domain \( W \). In this paper, \( W = \Omega_H^\Lambda \), the function space \( S(W) \) could be either \( H^n(\Omega_H^\Lambda) \) or \( \tilde{H}^n(\Omega_H^\Lambda) \) for \( n = 0, 1, 2 \). Define the space of functions
that depend \( C^n \)-continuously on the first variable

\[
C^n(\mathcal{I}; S(W)) := \left\{ f \in \mathcal{D}'(\mathcal{I} \times W) : \forall \alpha \in \mathcal{I}, j = 0, \ldots, n, \frac{\partial^j f(\alpha, \cdot)}{\partial \alpha^j} \in S(W), \right. \\
\left. \text{moreover, } \left\| \frac{\partial^j f(\alpha, \cdot)}{\partial \alpha^j} \right\|_{S(W)} \text{ is uniformly bounded for } \alpha \in \mathcal{I} \right\},
\]

thus we can define the space \( C^\infty(\mathcal{I}; S(W)) \) in the same way

\[
C^\infty(\mathcal{I}; S(W)) := \left\{ f \in \mathcal{D}'(\mathcal{I} \times W) : \forall \alpha \in \mathcal{I}, j = 0, \ldots, \infty, \frac{\partial^j f(\alpha, \cdot)}{\partial \alpha^j} \in S(W), \right. \\
\left. \text{moreover, } \left\| \frac{\partial^j f(\alpha, \cdot)}{\partial \alpha^j} \right\|_{S(W)} \text{ is uniformly bounded for } \alpha \in \mathcal{I} \right\},
\]

We can also define the subspace of \( C^n(\mathcal{I}; S(W)) \) for finite interval \( \mathcal{I} = [A_0, A_1] \) by

\[
C^n_p(\mathcal{I}; S(W)) := \left\{ f \in C^n(\mathcal{I}; S(W)) : \forall j = 0, \ldots, n, \right. \\
\left. \lim_{t \to A^+_j} \left\| \frac{\partial^j f(t, \cdot)}{\partial x^j} \right\|_{S(W)} = \lim_{t \to A^-_j} \left\| \frac{\partial^j f(t, \cdot)}{\partial x^j} \right\|_{S(W)} = 0 \right\}
\]

Similarly, we can also define \( C^\infty_p(\mathcal{I}; S(W)) \). Define the subspace \( C^\omega(\mathcal{I}; S_\alpha(W)) \) by

\[
C^\omega(\mathcal{I}; S_\alpha(W)) := \{ f(\alpha, \cdot) \in C^\omega(\mathcal{I}; S(W)) : f \text{ is } \alpha \text{-quasi-periodic for any } \alpha \in \mathcal{I} \},
\]

and also define \( C^n(\mathcal{I}; S_\alpha(W)) \) and \( C^\infty(\mathcal{I}; S_\alpha(W)) \) in the same way.

### 3.2 Regularity properties for Bloch transformed solutions

Firstly, assume that the incident field \( w^i \) is compactly supported in \( \Omega^H_p \), thus \( w^i|_{\Gamma_H} \in H^1_0(\Gamma_H) \) for any \( r \in \mathbb{R} \), especially, for any \( r > 1/2 \). Thus \( f(\alpha, \cdot) \) in \([10]\) belongs to the space \( H^r_0(W_{\Lambda^*}; H^{-1/2}_a(\Gamma_H^1)) \cap C^\omega(W_{\Lambda^*}; H^{-1/2}_a(\Gamma_H^1)). \) From Theorem 7, the Bloch transformed scattered field \( w \in H^r_0(W_{\Lambda^*}; \tilde{H}^1_0(\Omega^H_p)) \). Moreover, for any fixed \( \alpha \in W_{\Lambda^*} \) and \( \varphi_\alpha \in \tilde{H}^1_0(\Omega^H_p) \),

\[
a_\alpha(w(\alpha, \cdot), \varphi_\alpha) + b(w_T, \varphi_\alpha) = \int_{\Gamma_H^1} f(\alpha, \cdot)\varphi_\alpha \, ds. \tag{13}
\]

Set \( w_\alpha(x) := \exp(i\alpha x_1)w(\alpha, x) \), then \( w_\alpha \) is a periodic function for any fixed \( \alpha \in W_{\Lambda^*} \). Define \( \varphi_\alpha := \exp(-i\alpha x_1)\varphi(x) \) for any periodic function \( \varphi \). Then replace \( w \) and \( \varphi \) by \( w_\alpha \) and \( \varphi_\alpha \) for each \( \alpha \), it satisfies the following variational equation

\[
\int_{\Omega^H_p} \left[ \nabla_\alpha w_\alpha \cdot \nabla_\varphi - k^2 w_\alpha \varphi \right] - \int_{\Gamma_H^1} \tilde{T}_\alpha^+ \left[ w_\alpha|_{\Gamma_H^1} \right] \varphi \, ds = \int_{\Gamma_H^1} g(\alpha, \cdot)\varphi \, ds - \tilde{b}_\alpha(w_T, \varphi), \tag{14}
\]

where \( \nabla_\alpha = \nabla - i\alpha(1, 0)^T \), \( \tilde{T}_\alpha^+ \) is defined by

\[
\tilde{T}_\alpha^+(\psi) = i \sum_{j \in \mathbb{Z}} \sqrt{k^2 - |j|^2} \tilde{\psi}(j)e^{i\lambda^* x_1}, \quad \psi = \sum_{j \in \mathbb{Z}} \tilde{\psi}(j)e^{i\lambda^* x_1},
\]
the sesquilinear form \( \tilde{b}_\alpha(\cdot, \cdot) \) is defined by
\[
\tilde{b}_\alpha(\xi, \psi) = \left[ \frac{\Lambda}{2\pi} \right]^{1/2} \int_{\Omega^\Lambda} e^{i\alpha x^1} \left[ (A_p - I) \nabla \xi \cdot \nabla^\alpha \psi - k^2(c_p - 1)\xi \psi \right] \, dx,
\]
and
\[
g(\alpha, x) = e^{i\alpha x^1} f(\alpha, x).
\]

As \( f \in C^\omega(W_{\Lambda^*}; H_0^{-1/2}(\Gamma^\Lambda_H)) \), the function \( g(\alpha, \cdot) \in C^\omega(W_{\Lambda^*}; H_0^{-1/2}(\Omega^\Lambda_H)) \). So the right hand side of equation (14) depends analytically on \( \alpha \), for any \( g \in C^\omega(W_{\Lambda^*}; H_0^{-1/2}(\Omega^\Lambda_H)) \) and \( \varphi \in H_0^1(\Omega^\Lambda_H) \). From Riesz representation theorem, there is an operator \( F_\alpha \) that maps \( C^\omega(W_{\Lambda^*}; H_0^{-1/2}(\Omega^\Lambda_H)) \) to the space \( H_0^1(\Omega^\Lambda_H) \), such that
\[
\int_{\Gamma^\Lambda_H} g(\alpha, \cdot) \overline{\varphi} \, ds - \tilde{b}_\alpha(u_T, \varphi) = \int_{\Omega^\Lambda_H} F_\alpha(g) \overline{\varphi} \, dx. \tag{15}
\]
Moreover, the operator \( F_\alpha \) depends analytically on \( \alpha \), see Chapter 7, Section 1 in [Kat95]. Following the proof of Theorem a in [Kir93], we can obtain the following property of \( w_\alpha \) from the singularity with respect to \( \alpha \) in the DtN map \( \tilde{T}_b^+ \).

**Theorem 11.** Suppose the incident field \( u^i \in H^1(\Omega_H) \) has a compact support. Then \( w_\alpha(x) \) that solves (31) depends continuously in \( \alpha \in W_{\Lambda^*} \) and analytically on \( \Omega_{\Lambda^*} \setminus S \). If \( \alpha_0 \in S \), i.e., there is an \( n_0 \in \mathbb{Z} \) such that \( |\Lambda^*\alpha - \alpha_0| = k \). Then there is a neighbourhood \( U \) of \( \alpha_0 \) and two functions \( v_1^0, v_2^0 \in H_0^1(\Omega^\Lambda_H) \) that depend analytically on \( \alpha \), such that \( w_\alpha = v_1^0 + \sqrt{k^2 - |v_0\Lambda^* - \alpha|^2} v_2^0 \).

The singularity result could have a simplified form.

**Corollary 12.** If \( \alpha_0 \in S \), then there is a neighbourhood \( U \) of \( \alpha_0 \) and two functions \( w_1, w_2 \in C^\omega(U; H_0^1(\Omega^\Lambda_H)) \) such that \( w(\alpha, \cdot) = w_1(\alpha, \cdot) + \sqrt{\alpha - \alpha_0} w_2(\alpha, \cdot) \).

**Proof.** For any \( \alpha_0 \in S \), there is either 1) an \( n_0 \in \mathbb{Z} \) such that \( |\Lambda^*n - \alpha_0| = k \) or 2) two \( n_1, n_2 \in \mathbb{Z} \) such that \( |\Lambda^*n_1 - \alpha_0| = |\Lambda^*n_2 - \alpha_0| = k \).

Suppose \( |\Lambda^*n_j - \alpha_0| = k \), then \( \Lambda^*n_j - k = \alpha_0 \) or \( \Lambda^*n_j + k = \alpha_0 \). If \( \Lambda^*n_j - k = \alpha_0 \),
\[
\sqrt{k^2 - |\Lambda^*n_j - \alpha|^2} = \sqrt{\alpha - \alpha_0} \sqrt{2k + \alpha_0 - \alpha},
\]
where the second term is analytic in a small neighbourhood of \( \alpha_0 \). Thus the total field \( w_\alpha \) could be written into the form \( w_\alpha = w_1^0 + \sqrt{\alpha - \alpha_0} w_2^0 \), where \( w_1^0 = v_1^0 \) and \( w_2^0 = \sqrt{2k + \alpha_0 - \alpha} v_2^0 \). Let \( w_1(\alpha, \cdot) = e^{-i\alpha x^1} w_1^0 \) and \( w_2(\alpha, \cdot) = e^{-i\alpha x^1} w_2^0 \), then \( w(\alpha, \cdot) = w_1(\alpha, \cdot) + \sqrt{\alpha - \alpha_0} w_2(\alpha, \cdot) \). The case that \( \Lambda^*n_j + k = \alpha_0 \) is similar. The proof is finished.

We have just investigated the decomposition of the Bloch transformed field \( w(\alpha, \cdot) \), when \( u^i \) is compactly supported. However, the result for this special case could be extended to more generalized cases. Firstly, define the following two conditions for a function \( u \in C^0(\mathcal{I}; S(W)) \).

1. For any subinterval \( \mathcal{I}_0 \subset \mathcal{I} \setminus S, u \in C^\omega(\mathcal{I}_0; S(W)) \).
2. For any \( \alpha_j \in \mathcal{I} \cap \mathcal{S} \), there is a small enough \( \delta > 0 \) with two pairs \( v_1, w_1 \in C^\omega(\{\alpha_j - \delta, \alpha_j\}; S(W')) \) and \( v_2, w_2 \in C^\omega([\alpha_j, \alpha_j + \delta); S(W')) \) such that
\[
\begin{align*}
  u &= v_1 + \sqrt{\alpha - \alpha_j} w_1 \text{ for } \alpha \in (\alpha_j - \delta, \alpha_j], \\
  u &= v_2 + \sqrt{\alpha - \alpha_j} w_2 \text{ for } \alpha \in [\alpha_j, \alpha_j + \delta).
\end{align*}
\]
Then the function spaces are defined as follows.
\[
A_c^\omega(\mathcal{I}; S(W); \mathcal{S}) := \{ u \in C^0(\mathcal{I}; S(W)) : u \text{ satisfies conditions 1 and 2} \},
\]
and denote its subspace of all \( \alpha \)-quasi-periodic functions by \( A_c^\omega(\mathcal{I}; S_\alpha(\mathcal{W}); \mathcal{S}) \).

**Assumption 13.** The incident field \( u^i \in H^r_\omega(\Omega^H) \) for \( r > 1/2 \) satisfies that its Bloch transform \( \mathcal{F}_\Omega u^i \in A_c^\omega(W_{\Lambda^*}; H^r_\omega(\Omega^H); \mathcal{S}) \).

**Remark 14.** From \([\text{Zha17}]\), the two types of incident fields, i.e., the half space Green’s function and the Herglotz wave functions, satisfy Assumption 13. Set \( \Lambda = 2\pi \) for simplicity, then \( \Lambda^* = 1 \).

The first step is to investigate the function \( e^{\sqrt{k^2 - |n - \alpha|^2} x_2} \). From the Taylor’s expansion,
\[
e^{\sqrt{k^2 - |n - \alpha|^2} x_2} = \sum_{\ell \in \mathbb{Z}} \frac{(i\sqrt{k^2 - |n - \alpha|^2} x_2)^\ell}{\ell!}
\]
For any \( \alpha \in W_{\Lambda^*} \\setminus \mathcal{S} \), \( e^{\sqrt{k^2 - |n - \alpha|^2} x_2} \) is an analytic function with respect to \( \alpha \). For any \( \alpha_0 \in \mathcal{S} \), there is an integer \( n \) such that either \( k = n - \alpha_0 \) or \( -k = n - \alpha_0 \). Take \( k = n - \alpha_0 \) for example,
\[
e^{\sqrt{k^2 - |n - \alpha|^2} x_2} = \sum_{\ell \in \mathbb{Z}} \frac{(i\sqrt{k^2 - |n - \alpha|^2} x_2)^{2\ell}}{(2\ell)!} + \sum_{\ell \in \mathbb{Z}} \frac{(i\sqrt{k^2 - |n - \alpha|^2} x_2)^{2\ell+1}}{(2\ell+1)!}
\]
As the two series above converges absolutely, they are both analytic functions with respect to \( \alpha \). In a small neighbourhood of \( \alpha_0 \), \( \sqrt{2k + \alpha_0 - \alpha} \) is also analytic, thus there are two functions \( \psi_1(\alpha, x_1) \) and \( \psi_2(\alpha, x_1) \) that depend analytically in a small neighbourhood of \( \alpha_0 \) such that
\[
e^{\sqrt{k^2 - |n - \alpha|^2} x_2} = \psi_1(\alpha, x_1) + \sqrt{\alpha - \alpha_0} \psi_2(\alpha, x_1).
\]

1) The half space Green’s function, defined by
\[
G(x, y) = \frac{1}{4} \left[ H_0^{(1)}(k|x - y|) - H_0^{(1)}(k|x - y'|) \right]
\]
with the point \( y = (y_1, y_2)^\top \) be above the (locally perturbed) periodic surface and \( y' = (y_1, -y_2)^\top \).
It is easy to check that the Bloch transform has the following representation
\[
(\mathcal{F}_\Omega G(., y))(\alpha, x) = \frac{1}{2\pi} \sum_{j \in \mathbb{Z}} e^{i(j-\alpha)(x_1 - y_1)} e^{i\sqrt{k^2 - |j - \alpha|^2} y_2} \text{sinc} \left( \sqrt{k^2 - |j - \alpha|^2} y_2 \right) x_2,
\]
where sinc is an analytic function defined by \( \sin t/t \) for \( t \neq 0 \) with sinc(0) = 1. As the Taylor’s series sinc = \( \sum_{j=0}^\infty \frac{(-t)^n}{(2n+1)!} \), the function sinc(\( \sqrt{k^2 - |j - \alpha|^2} y_2 \)) depends analytically on \( \alpha \in \mathbb{R} \).
Thus the only factor that may cause singularity is $e^{\sqrt{k^2-|j\alpha|^2}}$ for finite number of $j$’s. From (16), the function $J_0 G(\cdot, y)$ also have the similar form of decomposition as (16).

2) The upward/downward propagating Herglotz wave function defined by

$$(H^\pm g)(x) = \int S_+ e^{i k x \cdot d} d s (d) = \int_{-\pi/2}^{\pi/2} e^{i k (\cos t x_1 \pm \sin t x_1)} \varphi(t) \ dt.$$ 

where $S_+$ ($S_-$) is the upper (lower) unit circle. $(H^\pm g)$ satisfies Assumption (13) when $\varphi(t) = h(\cos(t)) \cos(t)$ where $h$ is an analytic function defined in $[0, 1]$ with $h(0) = 0$.

Take $H^- g$ for example. From [LN13], the Bloch transform has the following representation

$$(J_\Omega H^- g)(\alpha, x) = \sum_{|j\alpha| < k} e^{i(j\alpha)x_1 - i \sqrt{k^2 - |j\alpha|^2} x_2} \varphi(\arcsin(j\alpha/k)) / \sqrt{k^2 - |j\alpha|^2}$$

$$= \frac{1}{k} \sum_{|j\alpha| < k} e^{i(j\alpha)x_1 - i \sqrt{k^2 - |j\alpha|^2} x_2} h(\sqrt{1 - |j\alpha|^2/k^2})$$

This is a finite series, we only need to investigate it by the decompositions of the exponential terms and the property of the analytic function $h$. The calculation process is omitted.

**Remark 15.** The plane waves do not satisfy Assumption 13, for the plane waves belong to $H^2(\Omega^1_H)$ for $r = -1/2$.

With the incident fields satisfying Assumption 13, the following results hold, see [Zha17].

**Theorem 16.** Suppose $u^i$ satisfies Assumption 13, then $w \in H^0_0 \left(W_{\Lambda^*}; \tilde{H}^2(\Omega^1_H)\right)$ satisfies (31) belongs to the space $A^w_\epsilon \left(W_{\Lambda^*}; \tilde{H}^2(\Omega^1_H); \mathcal{S}\right)$.

**Proof.** When $\alpha_0 \in W_{\Lambda^* \setminus \mathcal{S}}$, as $J_\Omega u^i \in A^\omega_\epsilon (W_{\Lambda^*}; H^2_0(\Omega^1_H))$, it depends analytically on $\alpha$ in a neighborhood $U \subset W_{\Lambda^* \setminus \mathcal{S}}$. From Corollary 12 $w(\alpha, \cdot)$ also depends analytically on $\alpha \in U$.

When $\alpha_0 \in W_{\Lambda^* \cap \mathcal{S}}$. Consider the case that $\alpha \in (\alpha_0 - \delta, \alpha_0]$. Suppose there are $v_1(\alpha, \cdot), v_2(\alpha, \cdot) \in C^\omega((\alpha_0 - \delta, \alpha_0]; H^2(\Omega^1_H))$ such that $J_\Omega u^i = v_1(\alpha, \cdot) + \sqrt{\alpha - \alpha_0} v_2(\alpha, \cdot)$, then the Bloch transformed fields $w_1(\alpha, \cdot), w_2(\alpha, \cdot)$ with respect to $v_1(\alpha, \cdot), v_2(\alpha, \cdot)$ have the following decompositions

$$w_1(\alpha, \cdot) = w_1^i(\alpha, \cdot) + \sqrt{\alpha - \alpha_0} w_2^i(\alpha, \cdot); \quad w_2^i(\alpha, \cdot) = w_2^i(\alpha, \cdot) + \sqrt{\alpha - \alpha_0} w_2^i(\alpha, \cdot),$$

where $w_j^i(\alpha, \cdot) \in C^\omega((\alpha_0 - \delta, \alpha_0]; H^2(\Omega^1_H))$ for $j = 1, 2$. Then the Bloch transformed field $w$ with respect to the field $J_\Omega u^i$ could be written into

$$w(\alpha, \cdot) = w_1(\alpha, \cdot) + \sqrt{\alpha - \alpha_0} w_2(\alpha, \cdot)$$

$$= \left[w_1^i(\alpha, \cdot) + (\alpha - \alpha_0) w_2^i(\alpha, \cdot)\right] + \sqrt{\alpha - \alpha_0} \left[w_1^2(\alpha, \cdot) + w_2(\alpha, \cdot)\right].$$

Similar result could also be obtained for $\alpha \in [\alpha_0, \alpha_0 + \delta]$.

Thus $w(\alpha, \cdot) \in A^\omega_\epsilon \left(W_{\Lambda^*}; \tilde{H}^2(\Omega^1_H); \mathcal{S}\right)$, the proof is finished. \hfill \Box

The decomposition of the Bloch transformed total field could also be written by smooth functions.

**Corollary 17.** Suppose $u^i$ satisfies Assumption 13, then the Bloch transform $w(\alpha, \cdot) = (J_\Omega u^i)(\alpha, \cdot)$ depends continuously in $\alpha \in W_{\Lambda^*}$. For each $j \in \mathbb{Z}$, for $\alpha \in [\alpha_j, \alpha_{j+1}]$, there are three functions $w_\ell(\alpha, \cdot) \in C^\infty([\alpha_j, \alpha_{j+1}]; \tilde{H}^2(\Omega^1_H))$ where $\ell = 0, 1, 2$ such that

$$w(\alpha, \cdot) = w_0^i(\alpha, \cdot) + \sqrt{\alpha - \alpha_j} w_1^i(\alpha, \cdot) + \sqrt{\alpha - \alpha_{j+1}} w_2^i(\alpha, \cdot). \quad (17)$$
4 The inverse Bloch transform

In [LZ17a, LZ17c, LZ17b], the accuracy of the numerical inverse Bloch transform plays an important role in the convergence of the numerical schemes. In this section, we will discuss a new numerical method for the inverse Bloch transform $J_{Ω}^{-1}w$. For references we refer to [CK13], Chap 3.5.

From the previous section, the Bloch transformed solution $w(\alpha, \cdot)$ only has square-root like singularities at the points in $S$, i.e., any $\alpha \in W_{Λ^*}$ such that $|Λ^*n - \alpha| = k$ for some $n \in \mathbb{N}$. Thus there is an $n \in \mathbb{Z}$, such that $\alpha = Λ^*n - k$ or $\alpha = Λ^*n + k$. Thus we can define the real number $k$ by

$$k := \min \{|Λ^*n - k| : n \in \mathbb{Z}\},$$

then $k \leq Λ^*/2$ and $S$ could also be represented by $k$.

$$S = \{Λ^*n \pm k : n \in \mathbb{Z}\}.$$

There are two different kinds of representations of points in $S$ by $k$.

- Case 1, $k = \frac{mΛ^*}{2}$ where $m = 0, 1$, the set of singularities $S = \{k + nΛ^* : n \in \mathbb{Z}\}$.
- Case 2, $k \neq \frac{mΛ^*}{2}$ for $m = 0, 1$, $S = \{k + mΛ^*, -k + nΛ^* : m, n \in \mathbb{Z}\}$.

For simplicity, we redefine one periodic cell $W_{Λ^*}$ by $W_{Λ^*} = (-k, Λ^* - k]$. Define the singular points in $W_{Λ^*}$ as follows.

- For Case 1, let $a_0 = -k$, $a_1 = Λ^* - k$.
- For Case 2, let $a_0 = -k$, $a_1 = k$, $a_2 = Λ^* - k$.

The redefinition is proper as the field $w(\alpha, \cdot)$ is $Λ^*$-periodic in $\alpha$ for any fixed $x$.

$$\int_{Λ^*/2}^{−Λ^*/2} w(α, \cdot) dα = \int_{−k}^{Λ^*−k} w(α, \cdot) dα = J_{Ω}^{-1}w.$$

4.1 Change of variables

To design a new algorithm with a greater convergence rate, we intend to use a "better" contour to take place of the straight line in the integral of the inverse Bloch transform. The strategy adopted in this paper is illustrated and utilized in Chap 3.5, [CK13]. In this section, we will recall the method and apply it to the numerical approximation of the inverse Bloch transform.

Define a monotonic function $g \in C^\infty[a_j, a_{j+1}]$ such that

$$g(a_j) = a_j, \ g(a_{j+1}) = a_{j+1}; \ g' > 0 \text{ in } (a_j, a_{j+1}).$$

Moreover, assume that the following conditions are also satisfied.

**Assumption 18.** Assume that there is a $δ > 0$, such that

$$g(t) - a_j = C_jh^2(t - a_j), \ t \in [a_j, a_j + δ);$$

$$a_{j+1} - g(t) = C_jh^2(a_{j+1} - t), \ t \in (a_{j+1} - δ, a_{j+1}],$$

where the function $h \in C^\infty[0, δ)$ satisfies either of the following two conditions.
1. There is a positive integer \( m \in \mathbb{N} \) such that \( h(t) = O(t^{m+1}) \) as \( t \to 0^+ \).

2. \( h(t) = o(t^n) \) for any \( n \in \mathbb{N} \) as \( t \to 0^+ \).

The contours in one periodic cell \( W_{\Lambda^*}(\Lambda = 2\pi) \) are shown in Figure 2. For Case 1, we take \( k = 1 \) as an example, and the contour is plotted in (a); for Case 2, \( k = 1.2 \) is taken as an example, see (b).

![Figure 2](image)

Figure 2: (a)-(b): The two cases of the locations of singularities and contours.

With the transform of \( \alpha = g(t) \), the inverse Bloch transform has the representation

\[
(\mathcal{J}^{-1}w)(x) = \int_{W_{\Lambda^*}} w(\alpha, \cdot) d\alpha = \int_{W_{\Lambda^*}} w(g(t), \cdot)g'(t)dt. \tag{18}
\]

For Case 1, \( W_{\Lambda^*} = (a_0, a_1] \),

\[
(\mathcal{J}^{-1}w)(x) = \int_{a_0}^{a_1} w(g(t), \cdot)g'(t)dt;
\]

for Case 2, \( W_{\Lambda^*} = (a_0, a_2] \),

\[
(\mathcal{J}^{-1}w)(x) = \int_{a_0}^{a_1} w(g(t), \cdot)g'(t)dt + \int_{a_1}^{a_2} w(g(t), \cdot)g'(t)dt.
\]

Let the interval \([A_0, A_1]\) be either \([a_0, a_1]\) (Case 1 & 2) or \([a_1, a_2]\) (Case 2). Define

\[
v(t, \cdot) := w(g(t), \cdot)g'(t), \quad t \in [A_0, A_1], \tag{19}
\]

then the inverse Bloch transform

\[
(\mathcal{J}^{-1}_{\Omega^*}w)(x) = \int_{W_{\Lambda^*}} v(t, x)dt, \quad x \in \Omega^*_H. \tag{20}
\]

We will study the Bloch transform when \( w \) satisfies certain conditions.

**Assumption 19.** \( w \in C^0([A_0, A_1]; S_{\alpha}(W)) \) has a form of

\[
w = w_2 + \sqrt{\alpha - A_0}w_0 + \sqrt{\alpha - A_1}w_1,
\]

where \( w_0, w_1, w_2 \in C^\infty([A_0, A_1]; S(W)) \).
Theorem 20. If there is an \( m \in \mathbb{N} \) such that \( h(t) = O(t^{m+1}) \) as \( t \to 0^+ \), for any \( w \) satisfying Assumption 19 \( v \) defined by 19 belongs to the function space \( C_{p}^{2m}([A_0,A_1];S_{g(t)}(W)) \cap C^{\infty}([A_0,A_1];S_{g(t)}(W)) \).

Proof. First, consider the case that \( S(W) = L^{2}(\Omega_H) \). Consider the function with \( \alpha \) near the point \( A_0 \). From Assumption 19, for any \( 0 < \delta < A_1 - A_0 \), in the interval \([A_0,A_0 + \delta]\), \( w(\alpha,x) \) has the form of

\[
w(\alpha,x) = \tilde{w}_0(\alpha,x) + \sqrt{\alpha - A_0} w_0(\alpha,x),
\]

where \( \tilde{w}_0(\alpha,x) = w_2(\alpha,x) + \sqrt{\alpha - A_1} w_1(\alpha,x) \) belongs to \( C^{\infty}([A_0,A_0 + \delta];S(W)) \). Then for any \( t \in [A_0,A_0 + \delta] \),

\[
v(t,x) = \tilde{w}_0(g(t),x)g'(t) + w_0(g(t),x)\sqrt{g(t) - A_0} g'(t)
\]

\[
= 2C_0 \tilde{w}_0(A_0 + h^2(t - A_0),x)h(t - A_0)h'(t - A_0)
\]

\[
+ 2C_0 w_0(A_0 + h^2(t - A_0),x)h^2(t - A_0)h'(t - A_0).
\]

From the representation of \( \tilde{w}_0 \) and \( v \),

\[
\|v(t,\cdot)\|_{S(W)} \leq C(g) \left[ \|w_0\|_{S(W)} + \|w_1\|_{S(W)} + \|w_2\|_{S(W)} \right].
\]

Take the first term for example. As \( h \in C^{\infty}[0,\delta] \) and \( g \in C^{\infty}[A_0,A_1] \), the function \( v \in C^{\infty}([A_0,A_1];S_{g(t)}(W)) \). As \( h(t) = O(t^{m+1}) \), \( h(t - A_0)h'(t - A_0) = O((t - A_0)^{2m+1}) \), then

\[
\lim_{t \to A_0^+} [h(t - A_0)h'(t - A_0)]^{(\ell)} (t) = 0, \quad \text{for any } \ell = 0,1, \ldots, 2m.
\]

From the \( \ell \)-th derivative w.r.t. \( t \)

\[
\frac{\partial^\ell}{\partial t^\ell} \left[ \tilde{w}_0(A_0 + h^2(t - A_0),x)h(t - A_0)h'(t - A_0) \right]
\]

\[
= \sum_{j=0}^{\ell} \binom{\ell}{j} \frac{\partial^j}{\partial t^j} \left[ \tilde{w}_0(A_0 + h^2(t - A_0),x) \right] \frac{\partial^{\ell-j}}{\partial t^{\ell-j}} \left[ h(t - A_0)h'(t - A_0) \right],
\]

in each term, the first element \( \frac{\partial^j}{\partial t^j} \left[ \tilde{w}_0(A_0 + h^2(t - A_0),x) \right] \) is a finite sum of the terms if form of \( \frac{\partial^j}{\partial t^j} \left[ \tilde{w}_0(A_0 + h^2(t - A_0),x) \right] \Pi_{m=0}^{N} [h^{(m)}(t - A_0)]^{n(m)} \), where \( N \) is some positive integer and \( n(m) \geq 0 \) in \( \mathbb{N} \). Thus the norm \( \left\| \frac{\partial^\ell}{\partial t^\ell} \left[ \tilde{w}_0(A_0 + h^2(t - A_0),x) \right] \right\|_{S(W)} \) is uniformly bounded for \( t \to A_0^+ \). From \( \lim_{t \to A_0^+} [h(t - A_0)h'(t - A_0)]^{(\ell)} (t) = 0 \),

\[
\left\| \frac{\partial^\ell}{\partial t^\ell} v(t,x) \right\|_{S(W)} \to 0 \quad \text{as } t \to A_0^+.
\]

Similarly, we can also prove that \( \left\| \frac{\partial^\ell}{\partial t^\ell} v(t,x) \right\|_{S(W)} \to 0 \) as \( t \to A_0^- \), \( \ell = 0,1, \ldots, 2m \), thus \( v \in C_{p}^{2m}([A_0,A_1];S_{g(t)}(W)) \).

The case that \( S(W) = H^n(\Omega_H) \) for any \( n \in \mathbb{N} \) could be proved similarly, thus we omit the proof here.

A direct corollary shows a similar result for the case that \( h(t) = o(t^n) \) as \( t \to 0^+, \forall n \).
Corollary 21. If \( h(t) = o(t^n) \) as \( t \to 0^+ \) for any positive integer \( n \), for any \( w \) satisfies Assumption 19, \( v \in C_p^\infty([A_0, A_1]; S_{g(t)}(W)) \).

From the smoothness of the function \( v(t, \cdot) \), the bounds of the derivatives of \( v \) depend on the functions \( w \) and \( g \), and the order of the derivative.

Theorem 22. Suppose \( w \) satisfies Assumption 19, \( v \) is defined by (19). If \( h(t) = O(t^{m+1}) \) for some \( m \in \mathbb{N} \), then \( v \) belongs to the functions space \( C_p^{2m}([A_0, A_1]; S_{g(t)}(W)) \cap C^\infty([A_0, A_1]; S_{g(t)}(W)) \) with

\[
\left\| \frac{\partial^j}{\partial t^j} v(t, \cdot) \right\|_{S(W)} \leq C(j, g) \|w\|_{S(W)}, \quad j = 0, 1, \ldots, 2m;
\]

if \( h(t) = o(t^n) \) for any \( n \in \mathbb{N} \), then \( v \in C_p^\infty([A_0, A_1]; S_{g(t)}(W)) \) with

\[
\left\| \frac{\partial^j}{\partial t^j} v(t, \cdot) \right\|_{S(W)} \leq C(j, g) \|w\|_{S(W)}, \quad j = 0, 1, \ldots, \infty,
\]

where \( C(j, g) \) is a constant that depends on \( j \) and \( g \), where

\[
\|w\|_{S(W)} := \|w_0\|_{S(W)} + \|w_1\|_{S(W)} + \|w_2\|_{S(W)}.
\]

4.2 Numerical method for the inverse Bloch transform

In this subsection, we will analyse the new numerical method of the inverse Bloch transform. Consider the definite integral

\[
\int_{A_0}^{A_1} v(t, \cdot) \, dt = \int_{A_0}^{A_1} w(\alpha, \cdot) \, d\alpha,
\]

where \( w \) satisfies Assumption 19 and \( v \) is defined by (19). We will approximate the function \( v \) by trigonometrical interpolation (for references see [Atk89]), and then study the convergence of the numerical integration based on the approximation. Let \([A_0, A_1]\) be divided uniformly into \( N \) subintervals, where \( N \) is assumed to be even in this paper. Let the uniformly located grid points

\[
t_j = A_0 + \frac{A_1 - A_0}{N} j, \quad j = 1, \ldots, N,
\]

and define the basic functions by

\[
\varphi_N^{(m)}(t) = \frac{1}{N} \sum_{t=-N/2+1}^{N/2} \exp \left( i(t - t_m) \frac{2\pi}{A_1 - A_0} \right), \quad j = 1, 2, \ldots, N,
\]

then \( \varphi_N^{(m)}(t) = \delta_{\ell,m} \), where

\[
\delta_{\ell,m} = \begin{cases} 
1, & \ell = m; \\
0, & \ell \neq m.
\end{cases}
\]

Lemma 23. The basic functions \( \left\{ \varphi_N^{(m)} \right\}_{m=1}^N \) are orthogonal.
Lemma 25. Suppose \( (S_1, \mu_1) \) and \((S_2, \mu_2)\) are two measure spaces and \( F: S_1 \times S_2 \rightarrow \mathbb{R} \) is measurable. Then the following inequality holds for any \( p \geq 1 \)

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

**Proof.** For any \( j, j' \in \{1, \ldots, N\} \), the inner product shows

\[
\int_{A_0}^{A_1} \phi_N^{(j)}(t) \overline{\phi_N^{(j')}}(t) \, dt = \frac{1}{N^2} \int_{A_0}^{A_1} \sum_{\ell=-N/2+1}^{N/2} \sum_{\ell'=-N/2+1}^{N/2} \exp \left( i\ell (t - t_j) \frac{2\pi}{A_1 - A_0} \right) \exp \left( -i\ell' (t - t_{j'}) \frac{2\pi}{A_1 - A_0} \right) \, dt,
\]

\[
= \frac{A_1 - A_0}{N^2} \sum_{\ell=-N/2+1}^{N/2} \exp \left( i\ell (t_{j'} - t_j) \frac{2\pi}{A_1 - A_0} \right) = \frac{A_1 - A_0}{N^2} \sum_{\ell=-N/2+1}^{N/2} \exp \left( i\ell (j' - j) \frac{2\pi}{N} \right)
\]

then the integral equals to 0 if and only if \( j \neq j' \). Thus

\[
\int_{A_0}^{A_1} \phi_N^{(j)}(t) \overline{\phi_N^{(j')}}(t) \, dt = \frac{A_1 - A_0}{N} \delta_{j,j'}.
\]

The basic functions \( \{ \phi_N^{(m)} \}_{m=1}^N \) are orthogonal. The proof is finished. \( \square \)

Let \( v_N \) be the interpolation of \( v \) in the space span \( \{ \phi_N^{(m)} \}_{m=1}^N \), with the representation

\[
v_N(t, \cdot) = \sum_{\ell=1}^N \phi_N^{(\ell)}(t) v(t_\ell, \cdot),
\]

then \( v_N(t_\ell, \cdot) = v(t_\ell, \cdot) \in S(W) \) for \( \ell = 1, 2, \ldots, N \). Then the definite integration

\[
\int_{A_0}^{A_1} v_N(t, \cdot) \, dt = \sum_{\ell=1}^N v(t_\ell, \cdot) \int_{A_0}^{A_1} \phi_N^{(\ell)}(t) \, dt = \frac{A_1 - A_0}{N} \sum_{j=1}^N v(t_j, \cdot),
\]

which coincides with the formula of the trapezoidal rule.

We will study the error estimate of the trapezoidal rule of the integration of \( v \). Let’s begin with the Fourier series of the function \( v(t, \cdot) \). A classical Minkowski integral inequality is needed, see Theorem 202, [HLPSS].

**Lemma 24.** Suppose \( (S_1, \mu_1) \) and \( (S_2, \mu_2) \) are two measure spaces and \( F: S_1 \times S_2 \rightarrow \mathbb{R} \) is measurable. Then the following inequality holds for any \( p \geq 1 \)

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

**Lemma 25.** Suppose \( v \) belongs to \( C_2^{2m}([A_0, A_1]; S_g(t)(W)) \cap C_\infty([A_0, A_1]; S_g(t)(W)) \). If \( h = O(t^{m+1}) \) as \( t \to 0^+ \) for some positive integer \( m \), then \( v \) has the Fourier series with respect to \( t \) has the form

\[
v(t, x) = \sum_{j \in \mathbb{N}} \hat{v}_j(x) e^{ijt}_{A_1^{A_0}},
\]

where \( \hat{v}_j \in S(W) \) and its norm for \( j \in \mathbb{N}^+ \)

\[
\|\hat{v}_j\|_{S(W)} \leq \left| \frac{A_1 - A_0}{2\pi j} \right|^{2m+1} \|v\|_{C_2^{2m+1}([A_0, A_1]; S(W))}.
\]
Proof. First, consider the case that $S(W) = L^2(\Omega^1_H)$. As $v(t, \cdot) \in C_p^2([A_0, A_1]; S_g(t) (W)) \cap C^\infty([A_0, A_1]; S_g(t) (W))$, for any $j = 0, \ldots, 2m$, $\frac{\partial}{\partial t} v(t, x) \in L^2([A_0, A_1]; S(W))$. So the Fourier series of $v(t, x)$ with respect to $t$ exists. From Theorem 20 the function $v \in C_p^m ([A_0, A_1]; S(W))$. By integration by part, for $j \neq 0$

$$\hat{v}_j(x) = \frac{1}{A_1 - A_0} \int_{A_0}^{A_1} v(t, x) e^{-ijt \frac{2\pi}{A_1 - A_0}} dt$$

$$= \frac{1}{A_1 - A_0} \left( \frac{A_1 - A_0}{2\pi j} \right)^{m+1} \frac{d^{m+1}}{dt^{m+1}} \int_{A_0}^{A_1} v(t, x) e^{-ijt \frac{2\pi}{A_1 - A_0}} dt,$$

then from Lemma 24

$$\left\| \frac{d^{m+1}}{dt^{m+1}} v(t, x) e^{-ijt \frac{2\pi}{A_1 - A_0}} dt \right\|_{S(W)}$$

$$= \left( \int_{\Omega_H^1} \left| \frac{d^{m+1}}{dt^{m+1}} v(t, x) e^{-ijt \frac{2\pi}{A_1 - A_0}} dt \right|^2 dx \right)^{1/2}$$

$$\leq \int_{A_0}^{A_1} \left( \int_{\Omega_H^1} \left| \frac{d^{m+1}}{dt^{m+1}} v(t, x) e^{-ijt \frac{2\pi}{A_1 - A_0}} \right|^2 dx \right)^{1/2} dt$$

$$\leq \int_{A_0}^{A_1} \left| \frac{d^{m+1}}{dt^{m+1}} v(t, \cdot) \right|_{S(W)} dt \leq (A_1 - A_0) \| v \|_{C_p^{m+1}([A_0, A_1]; S(W))},$$

then

$$\left\| \hat{v}_j(x) \right\|_{S(W)} \leq \left| \frac{A_1 - A_0}{2\pi j} \right|^{m+1} \| v \|_{C_p^{m+1}([A_0, A_1]; S(W))}.$$

The case that $S(W) = H^n(\Omega^1_H)$ for any $n \in \mathbb{N}$ could be proved similarly, thus we omit the proof here. The proof is finished.

A direct corollary of the lemma shows the case that $h(t) = o(t^n)$ for any $n \in \mathbb{N}$.

Corollary 26. If $h = o(t^n)$ as $t \to 0^+$ for any positive integer $n$, then $v$ has the Fourier series with respect to $t$ has the form

$$v(t, x) = \sum_{j \in \mathbb{N}} \hat{v}_j(x) e^{ijt \frac{2\pi}{A_1 - A_0}}, \quad (24)$$

where $\hat{v}_j \in S(W)$ and its norm

$$\| \hat{v}_j \|_{S(W)} = o(j^{-2n}). \quad (25)$$

Then we can get the error between the exact function $v$ and its Fourier interpolation $v_N$ defined in (24) for $\alpha \in W^\alpha_{A^*}$, when $N$ is large enough.

Theorem 27. If $v \in C_p^m ([A_0, A_1]; S_g(t) (W)) \cap C^\infty([A_0, A_1]; S_g(t) (W))$ for some $n \in \mathbb{N}$, then the difference between $v(t, \cdot)$ and its approximation $v_N(t, \cdot)$ is bounded by

$$\| v - v_N \|_{L^2([A_0, A_1]; S(W))} \leq C N^{-2n-1/2} \| v \|_{C_p^m([A_0, A_1]; S(W))}. \quad (26)$$

19
Proof. First, let \( S(W) = L^2(\Omega_H^A) \). From the definition of \( v_N(t, \cdot) \), i.e., the projection of \( v \) into the subspace span \( \{ c_n(x) \exp \left[ \ln \frac{2n}{A_1-A_0} \right], n = -N/2 + 1, \ldots, N/2, c_n \in S(W) \} \),

\[
\| v - v_N \|_{L^2([A_0, A_1]; S(W))}^2 = \int_{A_0}^{A_1} \int_{\Omega_H^A} (v - v_N)^2 \, dx \, dt
\]

\[
= \int_{A_0}^{A_1} \int_{\Omega_H^A} \left( \sum_{j \in \mathbb{Z} \cap [-N/2+1,N/2]} \hat{v}_j e^{i(\frac{2\pi}{A_1-A_0}) j t} \right)^2 \, dx \, dt
\]

\[
\leq (A_1 - A_0) \sum_{j \in \mathbb{Z} \cap [-N/2+1,N/2]} \| \hat{v}_j \|_{S(W)}^2.
\]

With the result of Lemma 25,

\[
\| v - v_N \|_{L^2([A_0, A_1]; S(W))}^2 \leq (A_1 - A_0) \sum_{j \in \mathbb{Z} \cap [-N/2+1,N/2]} \left| \frac{A_1 - A_0}{2\pi j} \right|^{4n+2} \| v \|_{C^{2n+1}([A_0, A_1]; S(W))}^2
\]

\[
\leq C N^{-4n-1} \| v \|_{C^{2n}([A_0, A_1]; S(W))}^2.
\]

The case that \( S(W) = H^n(\Omega_H^A) \) for any \( n \in \mathbb{N} \) could be proved similarly, thus we omit the proof here. The proof is finished.

\[\square\]

The error estimation of the trapezoidal rule will be shown in the following theorem.

**Theorem 28.** If \( v \in C^{2n}_p([A_0, A_1]; S_{g(t)}(W)) \cap C^{\infty}([A_0, A_1]; S_{g(t)}(W)) \) for some \( n \in \mathbb{N} \), then the error of the numerical integration of \( S \) is bounded by

\[
\left\| \int_{A_0}^{A_1} (v - v_N)(t, \cdot) \, dt \right\|_{S(W)} \leq C N^{-2n-1/2} \| v \|_{C^{2n+1}([A_0, A_1]; S(W))}.
\]

**Proof.** First, let \( S(W) = L^2(\Omega_H^A) \). With the result of Theorem 27 and Theorem 24

\[
\left\| \int_{A_0}^{A_1} (v - v_N)(t, \cdot) \, dt \right\|_{S(W)}^2 = \int_{\Omega_H^A} \left( \int_{A_0}^{A_1} (v - v_N)(t, \cdot) \, dt \right)^2 \, dx
\]

\[
\leq \int_{A_0}^{A_1} \| v - v_N \|_{L^2(\Omega_H^A)}^2 \, dt
\]

\[
\leq \| v - v_N \|_{L^2([A_0, A_1]; S(W))}^2.
\]

Then

\[
\left\| \int_{A_0}^{A_1} (v - v_N)(t, \cdot) \, dt \right\|_{S(W)} \leq \| v - v_N \|_{L^2([A_0, A_1]; S(W))}
\]

\[
\leq C N^{-2n-1/2} \| v \|_{C^{2n+1}([A_0, A_1]; S(W))}.
\]

The case that \( S(W) = H^n(\Omega_H^A) \) for any \( n \in \mathbb{N} \) could be proved similarly, thus we omit the proof here. The proof is finished.

\[\square\]

The result could be applied to the error estimation of the numerical inverse Bloch transform [18].

20
Theorem 29. Given a positive integer $N$, define the nodal points and inverse Bloch transform for the two conditions.

1. If $W_{A^*} = [a_0, a_1]$, then let $t_j = a_0 + j(a_1 - a_0)/N$, then the inverse Bloch transform

$$ \langle J_{\Omega,N}^{-1}w \rangle(x) = \frac{a_1 - a_0}{N} \sum_{j=1}^{N} v(g(t_j), x)g'(t_j) $$

(28)

2. If $W_{A^*} = [a_0, a_2]$, then let $t_j = a_0 + j(a_1 - a_0)/N$, $t_{j+N} = a_1 + j(a_2 - a_1)/N$, then the inverse Bloch transform

$$ \langle J_{\Omega,N}^{-1}w \rangle(x) = \frac{a_1 - a_0}{N} \sum_{j=1}^{N} v(g(t_j), x)g'(t_j) + \frac{a_2 - a_1}{N} \sum_{j=N+1}^{2N} v(g(t_j), x)g'(t_j). $$

(29)

If $h(t) = O(h^{m+1})$ for some $m \in \mathbb{N}$, the error of the numerical inverse Bloch transform is bounded by

$$ \| \langle J_{\Omega,N}^{-1}w - J_{\Omega}^{-1}w \rangle_{S(W)} \|_{C^2} \leq CN^{-2m-1/2} \| v \|_{C^{2m+1}([A_0, A_1]; S(W))}, \quad \ell = 1, 2; $$

if $h(t) = o(h^{n+1})$ for any $n \in \mathbb{N}$, the error of the numerical inverse Bloch transform is bounded by

$$ \| \langle J_{\Omega,N}^{-1}w - J_{\Omega}^{-1}w \rangle_{S(W)} \|_{C^2} \leq CN^{-2n-1/2} \| v \|_{C^{2n+1}([A_0, A_1]; S(W))}, \quad \ell = 1, 2, $$

for any $n \in \mathbb{N}$.

5 Finite element method for the scattering problems

In this section, we assume that the incident field $u'$ satisfies Assumption 30. With the transform that $\alpha = g(t)$ in $W_{A^*}$, the transformed inverse Bloch transform

$$ \langle \tilde{J}_{\Omega}^{-1}v \rangle(x_1 + \Lambda_j, x_2) = \left[ \frac{\Lambda}{2\pi} \right]^{1/2} \int_{W_{A^*}} v(t, x)e^{-i\theta(t)\Lambda_j} dt \quad x \in \Omega_H^\Lambda. $$

Replace $\alpha$ by $g(t)$ in (30), then the function $v(t, x)$ satisfies the following equation for any $\varphi \in L^2(W_{A^*}; \lambda_{g(t)}(\Omega_H^\Lambda))$

$$ \int_{W_{A^*}} a_{g(t)}(v(t, \cdot), \varphi(t, \cdot)) dt + b(\tilde{J}_{\Omega}^{-1}v, \tilde{J}_{\Omega}^{-1}[\varphi(t, \cdot)g'(t)]) = \int_{W_{A^*}} \int_{\Gamma_H^\Lambda} F(t, \cdot) \langle \varphi(t, \cdot) \rangle dt \quad (30) $$

where $F(t, \cdot) = f(g(t), \cdot)g'(t)$.

Define the space $L^2(W_{A^*}; \lambda_{g(t)}(\Omega_H^\Lambda))$ by

$$ L^2(W_{A^*}; \lambda_{g(t)}(\Omega_H^\Lambda)) := \{ \varphi(t, \cdot) = \psi(g(t), \cdot)g'(t); \psi(\alpha, \cdot) \in L^2(W_{A^*}; \lambda_{g(t)}(\Omega_H^\Lambda)) \}. $$

Then define the norm in $L^2(W_{A^*}; \lambda_{g(t)}(\Omega_H^\Lambda))$ by

$$ \| \varphi \|_{L^2(W_{A^*}; \lambda_{g(t)}(\Omega_H^\Lambda))} := \| \psi \|_{L^2(W_{A^*}; \lambda_{g(t)}(\Omega_H^\Lambda))} = \int_{W_{A^*}} \| \psi(g(t), \cdot) \|_{H_{g(t)}(\Omega_H^\Lambda)}^2 dt. $$

21
We can also define the function space $H^r_0(W_{\Lambda^*}; H^s_{g(t)}(\Omega^A_H); g)$ in the same way. By replacing $\varphi(t, \cdot) g'(t)$ with $\psi(t, \cdot) \in L^2(W_{\Lambda^*}; H^s_{g(t)}(\Omega^A_H); g)$, the variational problem (30) is equivalent to

$$\int_{W_{\Lambda^*}} a_{g(t)}(v(t, \cdot), \psi(t, \cdot)) \left[ g'(t) \right]^{-1} \, dt + b \left( \tilde{J}_{\Omega}^{-1} v, \tilde{J}_{\Omega}^{-1} \psi \right) = \int_{W_{\Lambda^*}} F(t, \cdot) \psi(t, \cdot) \left[ g'(t) \right]^{-1} \, dt.$$  

(31)

The transformed inverse Bloch transform $\tilde{J}_{\Omega}^{-1}$ is also defined in $H^r_0(W_{\Lambda^*}; H^s_{g(t)}(\Omega^A_H); g)$, we will show that the operator is bounded. Take $v \in L^2(W_{\Lambda^*}; L^2(\Omega^A_H); g)$,

$$\| \tilde{J}_{\Omega}^{-1} v \|^2_{L^2(\Omega_H^A)} = \| J_{\Omega}^{-1} w \|^2_{L^2(\Omega^A_H)} \leq \| w \|^2_{L^2(W_{\Lambda^*}; L^2(\Omega^A_H); g)} = \| v \|^2_{L^2(W_{\Lambda^*}; L^2(\Omega^A_H); g)}.$$  

We can also prove that $\tilde{J}_{\Omega}^{-1}$ is a bounded operator from $H^r_0(W_{\Lambda^*}; H^s_{g(t)}(\Omega^A_H); g)$ to $H^s(\Omega_H^A)$.

From the procedure in the arguments above, the following equivalence result is obtained by simply change of variables.

**Lemma 30.** Suppose the incident field $u^i \in H^r_r(\Omega^H_p)$ for some $r \in ]0, 1]$ and the functions $\zeta$, $\zeta_p$ are Lipschitz continuous functions. Then $w = J_{\Omega^H} u_T \in H^r_0(W_{\Lambda^*}; \tilde{H}^A_{\Omega^H}(\Omega^A_H))$ satisfies variational problem (3) if and only if $w \in H^r_0(W_{\Lambda^*}; \tilde{H}^A_{\Omega^H}(\Omega^A_H); g)$ is the solution to (31).

From the equivalence between (31) and (3), the following result is a simple corollary of Theorem 5 and Theorem 6.

**Theorem 31.** The variational problem (31) is uniquely solvable in $H^r_0(W_{\Lambda^*}; \tilde{H}^A_{g(t)}(\Omega^A_H); g)$. Further more, if $u^i$ satisfies Assumption 13, $\zeta, \zeta_p \in C^{2,1}(\mathbb{R})$, when $h(t) = O(t^{m+1})$, the unique solution $w$ to (31) belongs to the space $C^2_p(W_{\Lambda^*}; \tilde{H}^A_{g(t)}(\Omega^A_H))$.

**Proof.** We only need to prove the unique solvability. Suppose $v \in H^r_0(W_{\Lambda^*}; H^s_{g(t)}(\Omega^A_H); g)$ is the solution to (31) for $F = 0$, from the definition of the space $H^r_0(W_{\Lambda^*}; H^s_{g(t)}(\Omega^A_H); g)$, there is a $w \in H^r_0(W_{\Lambda^*}; \tilde{H}^A_{\Omega^H}(\Omega^A_H))$ such that $v(t, \cdot) = w(g(t), \cdot) g'(t)$. From Lemma 30 $w(\alpha, \cdot)$ is the unique solution to (3) with $f = 0$. From Theorem 5 the variational problem (3) has at most one solution, so $w = 0$, then $v = 0$. Then the injectivity is proved. From the Bounded Inverse Theorem, the problem is uniquely solvable.

When $u^i$ satisfies Assumption 13 and $\zeta, \zeta_p \in C^{2,1}(\mathbb{R})$, the solution $w$ to (3) belongs to the space $H^r_0(W_{\Lambda^*}; \tilde{H}^A_{\Omega^H}(\Omega^A_H))$ and belongs to $A_c^c \left( W_{\Lambda^*}; \tilde{H}^A_{\Omega^H}(\Omega^A_H); S \right)$, then from Corollary 17 $w$ satisfies Assumption 19. From Theorem 22 when $h = O(t^{m+1})$, $v$ belongs to the space $C^2_p([a_0, a_1]; \tilde{H}^A_{g(t)}(\Omega^A_H))$ for any $a_0, a_1 \in S$ such that $(a_0, a_1) \cap S$. From the definition of the space $C^2_p(W_{\Lambda^*}; \tilde{H}^A_{g(t)}(\Omega^A_H))$ and the new definition of $W_{\Lambda^*}$, $v \in C^2_p(W_{\Lambda^*}; \tilde{H}^A_{g(t)}(\Omega^A_H)) \subset H^r_0(W_{\Lambda^*}; \tilde{H}^A_{g(t)}(\Omega^A_H)).$

We will build up the finite element space, take $W_{\Lambda^*} = [a_0, a_1]$ for example. Let the uniformly located grid points be $\{ t_j \}_{j=1}^N$ and Fourier basic functions $\varphi_{N}^{(j)}$ be defined as in (21)-(22). For notations defined in $\Omega^A_H$, recall the triangular mesh $\mathcal{M}_H$, the piecewise linear basis functions $\left\{ \psi_M^{(j)} \right\}_{j=1}^M$ and the finite dimensional subspace $V_0 = \text{span} \left\{ \psi_M^{(j)} \right\}_{j=1}^M \subset H^r_0(\Omega^A_H)$. Define the finite element space $\tilde{X}_{N,h}$ by

$$\tilde{X}_{N,h} = \left\{ v_{N,h}(t, x) = \sum_{j=1}^{N} \sum_{l=1}^{M} \psi^{(j)}_{M} \exp \left( -ig(t_j)x_1 \right) \varphi_{N}^{(j)}(t) \psi_{M}^{(l)}(x) \right\}.$$
Theorem 32. If \( v \in C^2_p([a_0, a_1]; H^2_{g(t)}(\Omega^n_H)) \), then the error between the approximation of \( v \) in \( X_{N,h} \), i.e., \( v_N \), and \( v \), is bounded by

\[
\min_{v_{N,h} \in X_{N,h}} \| v - v_{N,h} \|_{L^2([a_0, a_1]; H^1_{g(t)}(\Omega^n_H))} \leq C \left( N^{-2n-1/2} + h \right) \| v \|_{C^{2n+1}([a_0, a_1]; H^2_{g(t)}(\Omega^n_H))},
\]

for large enough \( N \) and small enough \( h > 0 \), where \( C \) depends on \( n \) and \( g \).

Proof. From Theorem 27, \( v_N(t, x) = \sum_{j=1}^N \varphi_N^{(j)}(t) v(t_j, x) \) is an approximation of \( v \) with the error

\[
\| v - v_N \|_{L^2(W_{N,h}; H^1_{g(t)}(\Omega_H^n))} \leq C N^{-2n-1/2} \| v \|_{C^{2n+1}([a_0, a_1]; H^2_{g(t)}(\Omega_H^n))}. \tag{32}
\]

For each \( t_j \), the function \( \exp(ig(t_j)x_1) v(t_j, x) \) is approximated by functions in the subspace \( V_0 = \text{span} \{ \psi_M^{(\ell)}, \ell = 1, \ldots, M \} \), i.e., there is a series of coefficients \( v_{N,h}^{(j, \ell)} \in \mathbb{C} \) such that

\[
v(t_j, x) = \sum_{\ell=1}^M v_{N,h}^{(j, \ell)} \exp(-ig(t_j)x_1) \psi_M^{(\ell)}(x)
\]

for any \( x \) in the grid points of \( M_h \), with the error estimate

\[
\left\| \left( v(t_j, \cdot) - \sum_{\ell=1}^M v_{N,h}^{(j, \ell)} \exp(-ig(t_j)x_1) \psi_M^{(\ell)}(\cdot) \right) \right\|_{H^1_{g(t)}(\Omega_H^n)} \leq C h \| v(t_j, \cdot) \|_{H^2_{g(t)}(\Omega_H^n)}.
\]

For \( v_{N,h} = \sum_{j=1}^N \sum_{\ell=1}^M v_{N,h}^{(j, \ell)} \exp(-ig(t_j)x_1) \varphi_N^{(j)}(t) \psi_M^{(\ell)}(x) \), and let \( \psi_{h}^{(j)}(x) := v(t_j, x) - \sum_{\ell=1}^M v_{N,h}^{(j, \ell)} \exp(-ig(t_j)x_1) \psi_M^{(\ell)}(x) \), then

\[
\left\| \psi_{h}^{(j)} \right\|_{H^1_{g(t)}(\Omega_H^n)} \leq C h \| v \|_{H^2_{g(t)}([a_0, a_1]; H^1_{g(t)}(\Omega_H^n))}. \tag{33}
\]

With the result of Lemma 23,

\[
\frac{1}{N} \sum_{j=1}^N \left\| \psi_{h}^{(j)} \right\|_{L^2(\Omega_H^n)}^2 = \int_{W_{N,h}} \int_{\Omega_H^n} \sum_{j=1}^N \varphi_N^{(j)}(t) \psi_{h}^{(j)}(x) \, dx \, dt
\]

\[
= \int_{W_{N,h}} \int_{\Omega_H^n} \left| \sum_{j=1}^N \varphi_N^{(j)}(t) \psi_{h}^{(j)}(x) \right|^2 \, dx \, dt
\]

\[
\leq \int_{W_{N,h}} \int_{\Omega_H^n} \sum_{j=1}^N \varphi_N^{(j)}(t) \varphi_N^{(j)}(t) \psi_{h}^{(j)}(x) \psi_{h}^{(j)}(x) \, dx \, dt
\]

\[
= \frac{a_1 - a_0}{N} \sum_{j=1}^N \left\| \psi_{h}^{(j)} \right\|_{L^2(\Omega_H^n)}^2.
\]

Similarly, we can have the \( H^1 \)-estimate

\[
\| v_N - v_{N,h} \|_{L^2([a_0, a_1]; H^1_{g(t)}(\Omega_H^n))} \leq \frac{a_1 - a_0}{N} \sum_{j=1}^N \left\| \psi_{h}^{(j)} \right\|_{H^1_{g(t)}(\Omega_H^n)}^2 \leq C^2 \left[ a_0 - a_1 \right] h^2 \| v \|_{L^2([a_0, a_1]; H^2_{g(t)}(\Omega_H^n))}^2 \tag{33}
\]

23
With the inequalities (32) and (33),
\[
\|v - v_{N,h}\|_{L^2(W_{\Lambda^*}; H^1(\Omega_T^1))} \leq \|v - v_N\|_{L^2(W_{\Lambda^*}; H^1(\Omega_T^1))} + \|v_N - v_{N,h}\|_{L^2(W_{\Lambda^*}; H^1(\Omega_T^1))} \\
\leq CN^{-2n-1/2}\|v\|_{C^{2n+1}([A_0,A_1]; H^1(\Omega_T^1))} + C_{h}\|v\|_{C^{2n}([A_0,A_1]; H^2(\Omega_T^1))} \\
\leq CN^{-2n-1/2} + h\|v\|_{C^{2n+1}([A_0,A_1]; H^2(\Omega_T^1))},
\]
where \(C\) depends on \(n\) and \(g\).

With this result, we can estimate the error of the approximation of the Bloch transformed total field \(w\) by functions in \(\tilde{X}_{N,h}\), when \(W_{\Lambda^*} = [a_0,a_1]\).

**Theorem 33.** Suppose \(u^i\) satisfies Assumption \(\ref{assumption:13}\); \(g\) satisfies Assumption \(\ref{assumption:18}\); \(\zeta, \zeta_p \in C^{2,1}(\mathbb{R})\). \(u\) is the total field with the incident field \(u^i\) and surface \(\Gamma_p\), and \(w_T := u \circ \Phi_p\), and define \(w\) is the Bloch transform of \(w_T\), i.e., \(w := J_\Omega w_T\). Let \(v\) be defined by (19), then \(v\) satisfies the variational equation (31), for any \(\varphi \in L^2(W_{\Lambda^*}; H^1(\Omega_T^1))\). Moreover, if \(h = O(t^{m+1})\) for some \(m \in \mathbb{N}\) as \(t \to 0^+\), the approximation of \(v\) in \(\tilde{X}_{N,h}\) is bounded by
\[
\min_{v_{N,h} \in \tilde{X}_{N,h}} \|v - v_{N,h}\|_{L^2(W_{\Lambda^*}; H^1(\Omega_T^1))} \leq C \left( N^{-2m-1/2} + h \right) \|v\|_{C^{2m+1}([A_0,A_1]; H^2(\Omega_T^1))};
\]
if \(h = o(t^{n+1})\) for any \(n \in \mathbb{N}\) as \(t \to 0^+\), the approximation of \(v\) in \(\tilde{X}_{N,h}\) is bounded by
\[
\min_{v_{N,h} \in \tilde{X}_{N,h}} \|v - v_{N,h}\|_{L^2(W_{\Lambda^*}; H^1(\Omega_T^1))} \leq C \left( N^{-2n-1/2} + h \right) \|v\|_{C^{2n+1}([A_0,A_1]; H^2(\Omega_T^1))},
\]
for any \(n \in \mathbb{N}\).

**Proof.** As \(u^i\) satisfies Assumption \(\ref{assumption:13}\) and \(\zeta, \zeta_p \in C^{2,1}(\mathbb{R})\), from Corollary \(\ref{corollary:17}\) for the interval \(W_{\Lambda^*} = [a_0,a_1]\), there are three functions \(w_0, w_1, w_2 \in C^\infty(W_{\Lambda^*}; H^2(\Omega_T^1))\) such that
\[
w = w_2 + \sqrt{\alpha - A_0}w_0 + \sqrt{\alpha - A_1}w_1.
\]
So \(w\) also satisfies the conditions in Assumption \(\ref{assumption:19}\). From the definition
\[
v(t,\cdot) = w(g(t),\cdot)g'(t).
\]
From Theorem \(\ref{theorem:22}\) there are two different cases.

1) If \(h(t) = O(t^{m+1})\) as \(t \to 0^+\) for some \(m \in \mathbb{N}\), \(v \in C^{2m}([A_0,A_1]; S(W)) \cap C^{2m+1}([A_0,A_1]; S(W))\), then from Theorem \(\ref{theorem:22}\) the approximation of \(v\) in \(\tilde{X}_{N,h}\) is bounded by
\[
\min_{v_{N,h} \in \tilde{X}_{N,h}} \|v - v_{N,h}\|_{W_{\Lambda^*}; H^1(\Omega_T^1)} \leq C(N^{-2m-1/2} + h)\|v\|_{C^{2m+1}([A_0,A_1]; H^2(\Omega_T^1))};
\]

2) If \(h(t) = o(t^{n+1})\) as \(t \to 0^+\) for any \(n \in \mathbb{N}\), \(v \in C^\infty([A_0,A_1]; S(W))\), then from Theorem \(\ref{theorem:32}\) the approximation of \(v\) in \(\tilde{X}_{N,h}\) is bounded by
\[
\min_{v_{N,h} \in \tilde{X}_{N,h}} \|v - v_{N,h}\|_{L^2(W_{\Lambda^*}; H^1(\Omega_T^1))} \leq C(N^{-2n-1/2} + h)\|v\|_{C^{2n+1}([A_0,A_1]; H^2(\Omega_T^1))}
\]
for any \(n \in \mathbb{N}\).
Similar settings and results could be obtained for \( W_{\Lambda^*} = [a_0, a_2] \), with nodal points \( t_j \) are chosen in the same way as Theorem 28. Following the proof of Theorem 9 in [LZ17b], the error estimate of the finite element method is obtained.

**Theorem 34.** Assume that \( u^i \in H^2(\Omega_H) \) for \( r \geq 1/2 \) satisfies Assumption 14, \( \zeta, \zeta_p \in C^{2,1}(\mathbb{R}) \), then the linear system (31) is uniquely solvable in \( \tilde{X}_{N,h} \) for large enough \( N \) and small enough \( h > 0 \).

1) If \( h(t) = O(t^{m+1}) \) for some positive integer \( m \), then the solution \( w_{N,h} \in \tilde{X}_{N,h} \) satisfies the error estimate

\[
\| v - v_{N,h} \|_{L^2(W_{\Lambda^*};H^\ell(\Omega_H^\Lambda))} \leq C h^{1-\ell} \left( N^{-2m-1/2} + h \right) \| v \|_{C^{2m+1}([A_0,A_1];H^2(\Omega_H^\Lambda))}, \quad \ell = 0, 1, \quad (34)
\]

where \( C \) depends on \( m, u, g \).

2) If \( h(t) = o(t^{n+1}) \) for any positive integer \( n \), then the solution \( w_{N,h} \in \tilde{X}_{N,h} \) satisfies the error estimate

\[
\| v - v_{N,h} \|_{L^2(W_{\Lambda^*};H^\ell(\Omega_H^\Lambda))} \leq C h^{1-\ell} \left( N^{-2n-1/2} + h \right) \| v \|_{C^{2n+1}([A_0,A_1];H^2(\Omega_H^\Lambda))}, \quad \ell = 0, 1, \quad (35)
\]

for any \( n \in \mathbb{N} \), where \( C \) depends on \( n, u, g \).

### 6 Numerical Results

In this section, we will show some numerical results with the finite element method introduced in this paper. The periodic function is defined by

\[
\zeta(t) = 1.9 + \frac{\sin(t)}{3} - \frac{\cos(2t)}{4}
\]

with the local perturbation defined by

\[
\zeta_p(t) - \zeta(t) = \begin{cases} 
\exp \left( \frac{1}{t^2-1} \right) \sin [\pi(t + 1)], & \text{where } t \in [-1, 1]; \\
0, & \text{otherwise}.
\end{cases}
\]

In all of the numerical examples, we fix \( H = 4 \), and the period \( \Lambda = 2\pi \), thus \( \Lambda^* = 1 \).

![Figure 3: (a): Periodic surface \( \Gamma \); (b): locally perturbed periodic surface \( \Gamma_p \).](image-url)

We will show the results for two groups. For each group, we consider two different choices of the integration contour, i.e., the function \( g \).
Choice 1. $g(t)$ is defined in $[A_0, A_1]$ by

$$
g_1(t) = A_0 + \frac{A_1 - A_0}{\int_{A_0}^{A_1} (s - A_0)^3(s - A_1)^3 \, ds} \left[ \int_{A_0}^{t} (s - A_0)^3(s - A_1)^3 \, ds \right],
$$

thus $h(t) = O(t^2)$ as $t \to 0^+$.

Choice 2. $g(t)$ is defined in $[A_0, A_1]$ by

$$
g_2(t) = A_0 + \frac{A_1 - A_0}{\int_{A_0}^{A_1} \exp\left(\frac{1}{(s-A_0)(s-A_1)}\right) \, ds} \left[ \int_{A_0}^{t} \exp\left(\frac{1}{(s-A_0)(s-A_1)}\right) \, ds \right],
$$

thus $h(t) = o(t^n)$ as $t \to 0^+$ for any $n \in \mathbb{N}$.

Each group contains four different examples. In the four examples, two of them are with non-perturbed surface $\Gamma$ and another two are with locally perturbed surface $\Gamma_p$. The incident field is defined by a half-space Green’s function, that is defined as

$$u^i_G(x, y) = \frac{i}{4} \left[ H_0^{(1)}(k|x - y|) - H_0^{(1)}(k|x - y'|) \right],$$

where $y = (y_1, y_2)^\top$ and $y' = (y_1, -y_2)^\top$, or the upward propagating Herglotz wave function

$$u^i_+(x) = \int_{-\pi/2}^{\pi/2} \exp[ik(\sin \theta x_1 + \cos \theta x_2)] \cos^2(\theta) \, d\theta,$$

or the downward propagating Herglotz wave function

$$u^i_-(x) = \int_{-\pi/2}^{\pi/2} \exp[ik(\sin \theta x_1 - \cos \theta x_2)] \cos^2(\theta) \, d\theta.$$

Group 1. In this group, we show four upward propagating incident fields scattered by both perturbed and non-perturbed surface. As the upward propagating incident fields satisfy the Helmholtz equation above the surface and also satisfy the radiation condition, from the boundary condition that $u^s = -u^i$ and the unique solvability of the direct scattering problem, $u^s = -u^i$ above the (perturbed) surface. The parameter $h$ is fixed to be 0.01. The $L^2$-relative errors, defined by $\|u^s_{N,h} - u^s\|_{L^2(\Gamma^A_H)} / \|u^s\|_{L^2(\Gamma^A_H)}$, for $N = 4, 8, 16, 32, 64$ in Table 1 for $g = g_1$ and in Table 2 for $g = g_2$. 

![Figure 4: (a): function $g_1$; (b): function $g_2$.](image-url)
1. The surface is perturbed. The incident field is the half-space Green’s function \( u_G^i(\cdot, y) \) with \( y = (0.5, 0.4)^\top, k = 1 \).

2. The surface is not perturbed. The incident field is the half-space Green’s function \( u_G^i(\cdot, y) \) with \( y = (-1, 0.4)^\top, k = 1 \).

3. The surface is perturbed. The incident field is the upward propagating Herglotz wave function \( u_+^i, k = \sqrt{2} \).

4. The surface is not perturbed. The incident field is the upward propagating Herglotz wave function \( u_+^i, k = 1.5 \).

\[ \begin{array}{|c|cccc|}
\hline
& E_g 1 & E_g 2 & E_g 3 & E_g 4 \\
\hline
N = 4 & 7.95\text{E}−02 & 9.14\text{E}−02 & 4.76\text{E}−02 & 4.76\text{E}−02 \\
N = 8 & 8.28\text{E}−04 & 8.43\text{E}−04 & 5.24\text{E}−04 & 5.27\text{E}−04 \\
N = 16 & 5.78\text{E}−05 & 3.25\text{E}−05 & 5.26\text{E}−05 & 3.20\text{E}−05 \\
N = 32 & 2.23\text{E}−05 & 1.50\text{E}−05 & 3.75\text{E}−05 & 2.38\text{E}−05 \\
N = 64 & 2.12\text{E}−05 & 1.30\text{E}−05 & 3.09\text{E}−05 & 2.38\text{E}−05 \\
\hline
\end{array} \]

Table 1: Group 1, \( g = g_1 \).

\[ \begin{array}{|c|cccc|}
\hline
& E_g 1 & E_g 2 & E_g 3 & E_g 4 \\
\hline
N = 4 & 1.99\text{E}−01 & 2.22\text{E}−01 & 1.15\text{E}−01 & 1.15\text{E}−01 \\
N = 8 & 5.03\text{E}−04 & 6.27\text{E}−04 & 3.42\text{E}−04 & 3.43\text{E}−04 \\
N = 16 & 2.35\text{E}−05 & 2.05\text{E}−05 & 3.75\text{E}−05 & 2.49\text{E}−05 \\
N = 32 & 2.17\text{E}−05 & 1.49\text{E}−05 & 3.73\text{E}−05 & 2.38\text{E}−05 \\
N = 64 & 2.13\text{E}−05 & 1.49\text{E}−05 & 3.70\text{E}−05 & 2.38\text{E}−05 \\
\hline
\end{array} \]

Table 2: Group 1, \( g = g_2 \).

**Group 2.** In this group, we show four downward propagating incident fields scattered by both perturbed and non-perturbed surface. As the exact value of the scattered field could not be written out, we set the solution with \( N = 256 \) be the "exact value", and show the relative errors for \( N = 4, 8, 16, 32, 64 \) in Table 3 for \( g = g_1 \) and in Table 4 for \( g = g_2 \). Figure 5 plots in logarithmic scale of the relative \( L^2 \)-errors for the examples, see (a) for \( g = g_1 \) and (b) for \( g = g_2 \). The mesh size \( h \) is fixed to be 0.02.

1. The surface is perturbed. The incident field is the half-space Green’s function \( u_G^i(\cdot, y) \) with \( y = (0.5, 3)^\top, k = 1 \).

2. The surface is non-perturbed. The incident field is the half-space Green’s function \( u_G^i(\cdot, y) \) with \( y = (-1, 3)^\top, k = 1.5 \).

3. The surface is perturbed. The incident field is the downward propagating Herglotz wave function \( u_-^i, k = \sqrt{2} \).

4. The surface is non-perturbed. The incident field is the downward propagating Herglotz wave function \( u_-^i, k = 2.01 \).
we are confident to say that the new method is very efficient to solve the (locally-perturbed) tables are good illustrations for the result in Theorem 34. With these results shown above, it shows the super algebraic convergence rate of the numerical method. The results in the two thus the error decays faster than expected. In (b), the error decreases much more faster, thus

$$L \text{ by our new method in this paper at } N = 320. This means during the computational procession, we can save a lot of time and memory space in setting up the matrix and solving the linear system}^{12}.$$ 

Table 5: Numerical results in previous papers.

|       | $Eg 1$ | $Eg 2$ | $Eg 3$ | $Eg 4$ |
|-------|--------|--------|--------|--------|
| $N = 4$ | 2.19E−01 | 6.12E−01 | 1.23E−02 | 2.49E−02 |
| $N = 8$ | 1.71E−03 | 3.84E−03 | 4.98E−04 | 5.62E−04 |
| $N = 16$ | 9.24E−05 | 1.44E−04 | 3.17E−05 | 3.50E−05 |
| $N = 32$ | 5.64E−06 | 8.72E−06 | 1.99E−06 | 2.19E−06 |
| $N = 64$ | 3.49E−07 | 5.38E−07 | 1.24E−07 | 1.36E−07 |

Table 3: Group 2, $g = g_1$

|       | $Eg 1$ | $Eg 2$ | $Eg 3$ | $Eg 4$ |
|-------|--------|--------|--------|--------|
| $N = 4$ | 4.76E−01 | 1.05E+00 | 3.24E−02 | 6.20E−02 |
| $N = 8$ | 4.28E−03 | 2.34E−02 | 9.30E−05 | 2.39E−04 |
| $N = 16$ | 2.30E−05 | 3.76E−05 | 7.15E−06 | 8.05E−06 |
| $N = 32$ | 2.61E−08 | 4.04E−08 | 9.20E−09 | 1.01E−08 |
| $N = 64$ | 2.30E−12 | 2.90E−12 | 6.72E−13 | 6.60E−13 |

Table 4: Group 2, $g = g_2$

First let’s go to the examples in Choice 1. In Table 1 and 2, the relative error decays fast at first, and then stays at the level of $10^{-5}$ as $N$ gets larger ($N \geq 16$). As we have shown, the $L^2$-relative error is bounded by $C(N^{-2n-1.2} + h)h$, when $h$ is small enough, the error is brought by $N$ is comparatively larger, then when $N$ is relatively small, the error decays fast when $N$ increases. However, when $h$ is not small enough, the dominant part of error is brought by $h$, then the error will stay at a certain level when $N$ increases. Note that the first two examples Group 1 were shown in [LZ17a] and [LZ17b], as we list in Table 5.

Table 5: Numerical results in previous papers.

|       | $Eg 1$ | $Eg 2$ |
|-------|--------|--------|
| $N = 20$ | 7.76E−03 | 1.58E−02 |
| $N = 40$ | 2.83E−03 | 5.60E−03 |
| $N = 80$ | 1.04E−03 | 1.98E−03 |
| $N = 160$ | 4.04E−04 | 7.01E−04 |
| $N = 320$ | 1.93E−04 | 2.49E−04 |

Compared to the first and second columns in Table 1 and Table 2, even the result produced by our new method in this paper at $N = 16$ is much better than that produced by the old method at $N = 320$. This means during the computational procession, we can save a lot of time and memory space in setting up the matrix and solving the linear system.\[12\]

Now let’s go to the examples in Choice 2. In Table 3 and 4, the error decays as $N$ gets larger, and the logarithmic scale of the $L^2$-errors have been plotted in Figure 5. Figure 5 shows the convergences of the new method for both $g = g_1$ and $g_2$. In (a), the error decreases at the rate that is about $O(N^{-4})$. For $m = 1$, the expected error is $O(N^{-2m-1/2}) = O(N^{-2-1/2})$, thus the error decays faster than expected. In (b), the errors decrease much more faster, thus it shows the super algebraic convergence rate of the numerical method. The results in the two tables are good illustrations for the result in Theorem 34. With these results shown above, we are confident to say that the new method is very efficient to solve the (locally-perturbed)
Dedication.

This paper is dedicated to Prof. Dr. Armin Lechleiter, who passed away in January 2018 at a very young age. He was a talented mathematician, a helpful colleague, and a nice friend. His mathematical theory and numerical analysis for the scattering problems from locally perturbed periodic surfaces, which have been cited several times, are essential for the work in this paper. The author is very grateful for his scientific contribution to this area and his advise during their collaboration. He will always be missed.

Acknowledgements.

The author was supported by the University of Bremen and the European Union FP7 COFUND under grant agreement no 600411.

References

[Atk89] K. E. Atkinson. An Introduction to Numerical Analysis. John Wiley & Sons, Inc., 2nd edition, 1989.

[Bao95] G. Bao. Finite element approximation of time harmonic waves in periodic structures. SIAM Journal on Numerical Analysis, 32(4):1155–1169, 1995.

[BS94] S. C. Brenner and L. R. Scott. The Mathematical Theory of Finite Element Methods. Springer, New York, 1994.

[CE10] S. N. Chandler-Wilde and J. Elschner. Variational approach in weighted Sobolev spaces to scattering by unbounded rough surfaces. SIAM. J. Math. Anal., 42:2554–2580, 2010.
[CK13] D. L. Colton and R. Kress. *Inverse acoustic and electromagnetic scattering theory*. Springer, 3. edition, 2013.

[CM05] S. N. Chandler-Wilde and P. Monk. Existence, uniqueness, and variational methods for scattering by unbounded rough surfaces. *SIAM. J. Math. Anal.*, 37:598–618, 2005.

[Coo12] J. Coatléven. Helmholtz equation in periodic media with a line defect. *J. Comp. Phys.*, 231:1675–1704, 2012.

[EHZ09] M. Ehrhardt, H. Han, and C. Zheng. Numerical simulation of waves in periodic structures. *Commun. Comput. Phys.*, 5:849–870, 2009.

[ESZ09] M. Ehrhardt, J. Sun, and C. Zheng. Evaluation of scattering operators for semi-infinite periodic arrays. *Commun. Math. Sci.*, 7:347–364, 2009.

[FJ15] S. Fliss and P. Joly. Solutions of the time-harmonic wave equation in periodic waveguides: asymptotic behaviour and radiation condition. *Arch. Rational Mech. Anal.*, 2015.

[Fli09] S. Fliss. *Analyse mathématique et numérique de problèmes de propagation des ondes dans des milieux périodiques localement perturbés*. PhD thesis, Ecole Polytechnique, 2009.

[HLP88] G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities*. Cambridge Mathematical Library. Cambridge University Press, 2nd edition, 1988.

[HN16] H. Haddar and T. P. Nguyen. A volume integral method for solving scattering problems from locally perturbed infinite periodic layers. *Appl. Anal.*, 96(1):130–158, 2016.

[HNPX11] G. C. Hsiao, N. Nigam, J. E. Pasciak, and L. Xu. Error analysis of the DtN-FEM for the scattering problem in acoustics via Fourier analysis. *J. Comp. Appl. Math.*, 235(17):4949–4965, 2011.

[JLF06] P. Joly, J.-R. Li, and S. Fliss. Exact boundary conditions for periodic waveguides containing a local perturbation. *Commun. Comput. Phys.*, 1:945–973, 2006.

[Kat95] T. Kato. *Perturbation theory for linear operators*. Springer, repr. of the 1980 edition, 1995.

[Kir93] A. Kirsch. Diffraction by periodic structures. In L. Pävarinta and E. Somersalo, editors, *Proc. Lapland Conf. on Inverse Problems*, pages 87–102. Springer, 1993.

[Lec17] A. Lechleiter. The Floquet-Bloch transform and scattering from locally perturbed periodic surfaces. *J. Math. Anal. Appl.*, 446(1):605–627, 2017.

[LN15] A. Lechleiter and D.-L. Nguyen. Scattering of Herglotz waves from periodic structures and mapping properties of the Bloch transform. *Proc. Roy. Soc. Edinburgh Sect. A*, 231:1283–1311, 2015.
[LZ17a] A. Lechleiter and R. Zhang. A convergent numerical scheme for scattering of aperiodic waves from periodic surfaces based on the Floquet-Bloch transform. *SIAM J. Numer. Anal.*, 55(2):713–736, 2017.

[LZ17b] A. Lechleiter and R. Zhang. A Floquet-Bloch transform based numerical method for scattering from locally perturbed periodic surfaces. *SIAM J. Sci. Comput.*, 39(5):B819–B839, 2017.

[LZ17c] A. Lechleiter and R. Zhang. Non-periodic acoustic and electromagnetic scattering from periodic structures in 3d. *Comput. Math. Appl.*, 74(11):2723–2738, 2017.

[MACK00] A. Meier, T. Arens, S. N. Chandler-Wilde, and A. Kirsch. A Nyström method for a class of integral equations on the real line with applications to scattering by diffraction gratings and rough surfaces. *J. Int. Equ. Appl.*, 12:281–321, 2000.

[May84] D. Maystre. Rigorous vector theories of diffraction gratings. *Progress in Optics*, 21:1–67, 1984.

[Sch03] G. Schmidt. On the diffraction by biperiodic anisotropic structures. *Appl. Anal.*, 82:75–92, 2003.

[Zha17] R. Zhang. The study of the bloch transform of the fields scattered by locally perturbed periodic surfaces. *Preprint*, https://arxiv.org/pdf/1708.07560.pdf, 2017.