



**HAL**  
open science

# Direct and inverse solvers for scattering problems from locally perturbed infinite periodic layers

Thi Phong Nguyen

► **To cite this version:**

Thi Phong Nguyen. Direct and inverse solvers for scattering problems from locally perturbed infinite periodic layers. Analysis of PDEs [math.AP]. Université Paris Saclay (COMUE), 2017. English. NNT : 2017SACLX004 . tel-01475424v2

**HAL Id: tel-01475424**

**<https://inria.hal.science/tel-01475424v2>**

Submitted on 16 Apr 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NNT : 2017SACLX004

THESE DE DOCTORAT  
DE L'UNIVERSITE PARIS-SACLAY

préparée à

L'ÉCOLE POLYTECHNIQUE

ÉCOLE DOCTORALE N°574

École Doctorale de Mathématiques Hadamard (EDMH)

Spécialité de doctorat : Mathématiques

par

Mme. Thi Phong Nguyen

Direct and inverse solvers for scattering problems  
from locally perturbed infinite periodic layers

Thèse présentée et soutenue à Palaiseau, le 11 Janvier 2017.

Composition du jury :

M.	TILO ARENS	Examineur
M.	LAURENT BOURGEOIS	Examineur
Mme.	FIORALBA CAKONI	Examinatrice
M.	HOUSSEM HADDAR	Directeur de Thèse
M.	PATRICK JOLY	Président du jury
M.	ARMIN LECHLEITER	Examineur
M.	KARIM RAMDANI	Rapporteur
M.	JIGUANG SUN	Rapporteur



# Contents

<b>1</b>	<b>Introduction</b>	<b>11</b>
<b>2</b>	<b>Scattering problems for unbounded inhomogeneous layers</b>	<b>17</b>
2.1	Introduction . . . . .	17
2.2	Setting of the problem . . . . .	19
2.2.1	Construction of the DtN operators . . . . .	19
2.2.2	Variational formulation of the problem . . . . .	21
2.3	A Rellich identity for the case with absorption . . . . .	22
2.4	Scattering problem from periodic layers . . . . .	30
2.4.1	The DtN operator for $\xi$ -quasi-periodic problems . . . . .	31
2.4.2	Well-posedness of the problem . . . . .	33
2.5	The case of infinite layers with local perturbations . . . . .	34
2.6	On the case of the scattering problem for half space . . . . .	35
2.6.1	Variational formulation of the problem . . . . .	36
2.6.2	The Rellich identity for the case without absorption . . . . .	37
<b>3</b>	<b>A volume integral method for solving the scattering problem</b>	<b>41</b>
3.1	Introduction . . . . .	41
3.2	Setting of the problem . . . . .	43
3.2.1	Introduction of the problem and notation . . . . .	43
3.2.2	Formulation of the problem using the Floquet-Bloch transform . . . . .	44
3.3	Volume integral formulation of the $\xi$ -quasi-periodic problem . . . . .	46
3.3.1	Setting of the volume integral equation . . . . .	46
3.3.2	Periodization of the integral equation . . . . .	47
3.3.3	Spectral approximation of problem (3.19) . . . . .	49
3.4	Discretization of the problem and convergence analysis . . . . .	54
3.4.1	Discretization and convergence in the Floquet-Bloch variable . . . . .	55
3.4.2	Discretization in the spatial variable . . . . .	59
3.5	Numerical Algorithm and Experiments . . . . .	65
3.5.1	Numerical examples for real wave numbers . . . . .	69
<b>4</b>	<b>Inverse scattering problem from locally perturbed periodic media</b>	<b>75</b>
4.1	Introduction . . . . .	76
4.2	Setting of the direct scattering problem . . . . .	77
4.3	Setting of the inverse problem . . . . .	80
4.3.1	Definition of the sampling operator . . . . .	80
4.3.2	Some useful properties for sampling methods . . . . .	81
4.4	Application to Sampling methods . . . . .	86
4.4.1	The Linear Sampling Method (LSM) . . . . .	86
4.4.2	The Factorization Method . . . . .	87
4.4.3	The Generalized Linear Sampling Method (GLSM) . . . . .	88
4.4.4	Reconstruction of the periodic domain $D_p$ from $N^\pm$ . . . . .	89
4.4.5	Validating Numerical Experiments . . . . .	91
4.5	On the Use of Differential measurements . . . . .	95

4.5.1	Theory of the Differential LSM . . . . .	95
4.5.2	A numerical example . . . . .	97
4.6	Sampling methods for a single Floquet-Bloch mode . . . . .	98
4.6.1	Near field operator for a fixed Floquet-Bloch mode . . . . .	98
4.6.2	Some properties of the operators $\mathcal{H}_q^\pm$ , $N_q^\pm$ and $G_q^\pm$ . . . . .	99
4.6.3	A new differential imaging functional . . . . .	106
4.6.4	Numerical validating examples . . . . .	108
<b>5</b>	<b>The case of TM-mode</b>	<b>113</b>
5.1	Introduction . . . . .	113
5.2	The spectral volumetric method for solving the direct scattering problem . . . . .	114
5.2.1	Volume integral formulation of the $\xi$ -quasi-periodic problem . . . . .	115
5.2.2	Discretization of the locally perturbed periodic problem . . . . .	117
5.3	The sampling methods for the TM-mode . . . . .	123
5.3.1	Setting of the inverse problem and definition of the sampling operator . . . . .	125
5.3.2	Some key properties of the introduced operators . . . . .	127
5.3.3	Sampling methods for a single Floquet-Bloch mode . . . . .	132
<b>6</b>	<b>Surface potential formulation of a scattering problem</b>	<b>141</b>
6.1	Introduction . . . . .	141
6.2	Setting of the problem . . . . .	142
6.3	Reformulation of the problem via surface potentials . . . . .	143
6.3.1	Surface potentials and some classical results . . . . .	144
6.3.2	The system of surface integral equations . . . . .	145
6.4	The Fredholm property of the surface integral operator $Z(k_i, k_e)$ . . . . .	147
6.4.1	The existence and uniqueness of solutions to the scattering problem . . . . .	150
<b>7</b>	<b>Electromagnetic scattering with sign-changing coefficients</b>	<b>153</b>
7.1	Introduction . . . . .	153
7.2	Periodic Electromagnetic Scattering . . . . .	154
7.3	$T$ -Coercivity Framework . . . . .	156
7.4	Fredholm Alternative . . . . .	157
<b>8</b>	<b>Conclusion and Perspectives</b>	<b>161</b>
<b>A</b>	<b>Abstract theoretical foundations of the sampling methods</b>	<b>163</b>
A.1	Main theorem for the $F_{\sharp}$ method . . . . .	163
A.2	Theoretical foundations of GLSM . . . . .	164
	<b>Bibliography</b>	<b>167</b>

# Résumé en Français

Nous nous intéressons principalement dans cette thèse à l'analyse des problèmes de diffraction directs et inverses pour des couches périodiques infinies localement perturbées. Notre travail a été motivé (et financé) par le projet ANR METAMATH dédié à la simulation des métamatériaux et des matériaux avec des indices négatifs. Un métamatériau est généralement composé de petits résonateurs périodiquement espacés comme indiqué sur la Fig. 1.1. Comme autre structure périodique typique que l'on rencontre dans les technologies récentes, nous pouvons également citer les nanograss (voir Fig. 1.2). Dans ces structures, les nanotubes sont cultivés chimiquement sur un substrat et peuvent être disposés de manière périodique (éventuellement avec des échelles de périodicité multiples).

Les développements mathématiques et numériques qui sont entrepris visent tout d'abord à résoudre le problème du contrôle non destructif de ces structures périodiques. Dans ce problème, une onde (électromagnétique) incidente est utilisée pour éclairer la structure et la réponse de la structure est mesurée à une certaine distance de celle-ci. Il s'agit alors d'utiliser ces mesures pour identifier la présence de défauts et si possible reconstruire l'emplacement et la forme de ces défauts. Nous considérerons le problème à une fréquence fixe signifiant que l'onde est harmonique dans le temps avec une pulsation connue. Pour l'essentiel de notre travail, seul un problème scalaire simplifié est considéré (modélisant par exemple certaines polarisations spécifiques de l'onde électromagnétique ou une configuration spéciale d'ondes acoustiques [48]).

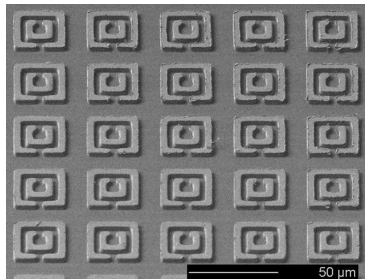


Figure 1: An example of a metamaterial micro-structure

Avant d'aborder le problème inverse, nous allons d'abord analyser le problème direct de diffraction puis concevoir une méthode numérique pour simuler la diffraction des ondes par des couches périodiques non bornées contenant un défaut. La difficulté de l'analyse du problème direct réside dans le fait que le milieu périodique est infini et par conséquent, certaines ondes guidées peuvent contribuer à la solution du problème. De plus, la perturbation locale du milieu périodique ne permet pas de réduire le problème de diffusion en un problème plus simple posé sur une cellule de périodicité. C'est pourquoi l'analyse du problème localement perturbé apparaît aussi difficile que l'analyse des couches non périodiques infinies.

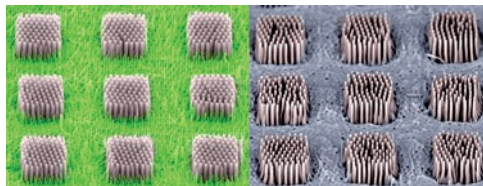


Figure 2: An example of a nanograss structure

Il existe une grande littérature consacrée à l'analyse des ondes guidées (dans les guides d'ondes ouverts) [62, 20, 104] ou la diffraction à partir de structures non bornées [33, 36, 37, 38, 39, 35, 8, 5, 24]. Il est par exemple bien connu que si la couche périodique est homogène avec un indice de réfraction constant supérieur à un, alors certains modes guidés existent [104, 25]. Il est également prouvé dans [86, 57, 38] que, par exemple, si l'indice de réfraction admet une propriété de monotonie stricte dans la direction orthogonale à la direction de périodicité, alors il n'existe aucun mode guidé. Du point de vue du problème inverse, nous sommes plus intéressés à analyser ce dernier cas que le premier. Cela est principalement dû au fait que pour les applications que nous avons en vu, la mesure des ondes guidées ne serait pas possible. Nous revisitons dans le chapitre 2 l'analyse du problème de diffraction pour les couches non bornées et les conditions qui assurent l'absence d'ondes guidées. L'outil principal est l'identité de Rellich qui est largement utilisée dans l'analyse de ce type de problèmes [37, 38]. L'objectif de notre analyse est d'affaiblir l'hypothèse de la monotonie requise dans [86] qui nous empêcherait de traiter la diffraction d'une couche périodique dans l'espace libre, ce qui correspond à la configuration que nous aimerions aborder pour le problème inverse. De plus, cette condition de monotonie n'est pas compatible avec certaines des hypothèses géométriques (techniques) que nous allons faire dans l'étude du problème inverse.

L'étape suivante consiste à concevoir une méthode numérique pour résoudre le problème de diffraction que nous considérons dans le chapitre 3. Cette étape est nécessaire par exemple pour générer les données pour notre problème inverse. Nous étendons une méthode qui a été introduite dans le cadre de la thèse de D.L. Nguyen [91, 84] pour la simulation du problème de diffraction par des couches périodiques pour gérer la présence d'une perturbation locale. On peut transformer, via la transformée de Floquet-Bloch, le problème localement perturbé en problèmes couplés quasi-périodiques. Le facteur de quasi-périodicité est lié au paramètre de Floquet-Bloch. Notre méthode est ensuite basée sur la discrétisation du paramètre Floquet-Bloch et d'appliquer ensuite la même stratégie de discrétisation que celle proposée dans [91] au problèmes semi-discrètes. Cette stratégie est basée sur la reformulation du problème de diffraction en utilisant la représentation de la solution par un potentiel volumique (l'équation intégrale de Lippmann-Schwinger), puis en discrétisant le problème en utilisant une base spectrale de Fourier. Afin de choisir efficacement la base discrète (permettant l'utilisation de la transformée de Fourier rapide pour évaluer les produits matrice-vecteur), une périodisation du noyau du potentiel volumique doit être incorporée. Pour un aperçu de ce type de techniques, nous nous référons à [99]. Nous obtenons un ensemble d'équations discrètes couplées qui est résolu itérativement en utilisant une méthode point de fixe de type Jacobi. La procédure en deux étapes que nous suivons est largement inspirée du travail de [43] où l'on étudie l'utilisation d'une méthode d'éléments finis pour la discrétisation en espace. L'analyse de convergence de cette méthodologie repose sur l'observation que si pas uniforme est utilisé pour discrétiser la variable de Floquet-Bloch, le problème semi-discrète obtenu est équivalent (à une petite modification près du terme source) à tronquer le domaine spatial dans la direction de la périodicité avec une longueur proportionnelle au nombre de points de discrétisation. On comprend alors pourquoi notre méthode serait convergente seulement s'il n'y a pas de modes guidés. Pour notre analyse nous nous limitons au cas d'un nombre d'onde avec une partie imaginaire non nulle (qui modéliserait une certaine absorption dans le milieu). L'obtention de résultats de convergence plus faibles pour le cas non absorbant est envisageable comme dans [87] si certaines propriétés de régularité et de décroissance sur la solution sont supposées.

La conception d'une méthode capable de gérer les modes guidés est une perspective intéressante et une possibilité serait de construire des opérateurs Dirichlet-à-Neumann appropriées comme dans [68, 54, 55, 16]. Nous citons également la méthode proposée dans [23, 15] qui gèrent

des guides d'ondes ouverts comme une approche numérique alternative ou la méthode dans [4] basée sur des équations intégrales de surface. Nous pensons aussi que l'on pourrait étendre notre méthode qui s'appuie sur la transformée de Floquet-Bloch si l'on peut adapter la règle de quadrature aux paramètres de Floquet-Bloch où apparaissent certains modes guidés.

Comme indiqué précédemment, notre première motivation est de traiter le problème inverse où l'on veut imager un défaut dans la couche périodique à partir des mesures des ondes diffractées à une fréquence donnée. Nous nous intéressons à étendre ce qu'on appelle les méthodes d'échantillonnage [46, 29, 76, 30] pour gérer ce problème inverse. Ce type de méthodes a été introduit en 1996 dans [45] et a suscité depuis un grand intérêt dans la communauté problème inverse. Les principaux avantages de ce type de méthodes est qu'ils ne nécessitent pas un solveur du problème direct. De plus, ils fournissent, pour certains d'entre eux, une preuve d'unicité pour la reconstruction de la géométrie sans une connaissance a priori des propriétés physique du milieu. En comparaison aux méthodes de type Born [48], les méthodes d'échantillonnage ne reposent pas sur une linéarisation du problème inverse. Les méthodes d'échantillonnage reposent cependant sur des données dites multistatistiques: c'est-à-dire des données associées à un grand nombre de sources et à un grand nombre de récepteurs. Ils fournissent également seulement une information géométrique sur la cible.

Le principe d'une méthode d'échantillonnage est de construire, à partir de l'opérateur de mesure et de fonctions test soigneusement choisies, une fonction indicatrice qui renseigne sur si un point se trouve ou non à l'intérieur de la géométrie recherchée. Les différences entre les méthodes d'échantillonnage se trouvent principalement dans la manière de construire cette fonction indicatrice. Nous étudierons la méthode d'échantillonnage linéaire (LSM) [45, 30], la méthode de factorisation (FM) [70, 72, 73, 74, 76, 80] et la méthode l'échantillonnage linéaire généralisée [14, 13, 12, 30]. Ces méthodes sont liées les unes aux autres et chacune a ses avantages et ses inconvénients. La LSM est probablement la méthode d'échantillonnage qui a la formulation la plus simple et la gamme d'applications la plus large (divers modèles d'ondes et des configurations plutôt souples pour les sources et les récepteurs). La fonction indicatrice LSM a également une interprétation en termes de solutions au problème dit de la transmission intérieure (ITP), qui joue un rôle important dans la méthode que nous allons développer. Cependant, cette méthode n'a pas une justification mathématique rigoureuse. La FM est cependant plus robuste en terme de justification mathématique car on a une caractérisation exacte de la géométrie par la fonction indicatrice construite. L'analyse de la méthode de factorisation repose cependant sur une factorisation spéciale de l'opérateur de mesure qui correspond en général à des mesures à ouverture complète. Dans la méthode de factorisation, on perd cependant le lien avec ITP. La GLSM est une méthode qui se situe entre les deux méthodes. elle donne une caractérisation exacte de la géométrie recherchée en utilisant un critère asymptotique explicitement construit. En outre, elle fournit un lien direct avec des solutions au problème de transmission intérieure. L'inconvénient principal est probablement l'absence d'une méthode qui ajuste le facteur de pénalisation au bruit présent dans l'opérateur de mesure.

Il existe plusieurs travaux qui ont étendu les méthodes d'échantillonnage aux problèmes périodiques [82, 91, 10, 7, 98, 53, 105, 66]. Cependant, nous pensons que le cas du milieu périodique infini localement perturbé (dans l'espace libre) n'a pas été abordé dans la littérature. En fait, nous ne traiterons pas exactement le problème périodique localement perturbé, mais plutôt considérons la version discrétisée de ce problème par rapport à la variable de Floquet-Bloch. Comme cela a déjà été mentionné, ceci correspond à la troncature du milieu dans le sens de la périodicité avec des conditions de périodicité de période égale à un multiple de la périodicité du milieu de référence. Cette restriction est principalement due à des aspects techniques qui sont plus faciles à traiter au niveau discret. Pour ce faire, nous considérons un milieu à deux échelles de péri-



odicité dont l'une est multiple de l'autre. Notre première contribution est de montrer comment les méthodes d'échantillonnage peuvent être utilisées pour résoudre le problème d'imagerie pour la plus grande longueur de périodicité. Ce problème est très similaire aux problèmes de guide d'onde comme dans [28, 27, 82, 101]. Nous expliquons ensuite comment, dans ce cadre spécial, on peut utiliser ces méthodes pour reconstituer le fond périodique sans connaître le défaut.

Notre contribution la plus originale est sur la façon dont on peut utiliser les méthodes d'échantillonnage pour reconstruire directement le défaut. C'est le sujet du chapitre 4. La méthode que nous proposons est inspirée de la LSM différentielle introduite dans [13, 12]. Pour cette dernière méthode, on s'appuie sur des mesures différentielles (mesures pour la configuration avec défauts et mesures pour la configuration sans défaut) pour construire une fonction indicatrice qui caractérise la géométrie du défaut. La construction de cette fonction indicatrice repose sur le formalisme GLSM et le lien avec les solutions à l'ITP susmentionné. Pour notre problème nous exploiterons l'information que le milieu de référence est périodique avec une longueur de périodicité connue pour développer une fonction indicatrice qui n'utilise pas de mesures différentielles. Nous construisons à partir de l'opérateur de mesure un opérateur de mesure associé à un seul mode de Floquet-Bloch. Cet opérateur jouerait le rôle de l'opérateur de référence dans le cas des mesures différentielles. L'analyse des méthodes d'échantillonnage associées à ces opérateurs à un seul mode de Floquet-Bloch pour le problème périodique perturbé constitue notre principale contribution théorique. Par exemple, nous montrons comment ces opérateurs reconstruiraient le milieu de référence périodique plus des copies périodiques du défaut. L'analyse de la méthode GLSM pour ces opérateurs nécessite l'analyse d'une nouvelle forme du problème ITP. L'étude de l'unicité des solutions à ce problème (prouvant l'injectivité des opérateurs en dehors des valeurs propres de transmission) n'est faite que pour le cas de variables de Floquet-Bloch discrètes, et c'est là que nous avons besoin de l'hypothèse technique (troncature) mentionnée précédemment. La comparaison des solutions ITP nous permet de construire une fonction indicatrice capable d'identifier le défaut. Ceci n'est prouvé que dans le cas où il n'y a pas d'intersection entre le défaut et le domaine périodique de référence. En effet, la méthode que nous proposons peut être appliquée dans le cas général (c'est-à-dire en dehors des hypothèses techniques que nous avons mentionnées). Les résultats numériques illustrent comment la fonction indicatrice est capable d'identifier le défaut, même si les méthodes d'échantillonnage ne parviennent pas à saisir la structure (complexe) du fond périodique. Ceci constitue le principal avantage de notre fonctionnelle d'imagerie différentielle.

Les méthodes numériques que nous avons développées pour les problèmes directs et inverses sont présentées d'abord dans le cas simple où la modification des propriétés du milieu est modélisée par des changements dans l'indice de réfraction (de vant le terme d'ordre inférieur de l'edp). Les méthodes sont cependant générales et peuvent être appliquées à d'autres modèles plus compliqués. C'est ce que nous essayons de démontrer au chapitre 5 en donnant un aperçu de l'extension au cas scalaire où les changements dans le milieu affectent aussi l'opérateur principal de l'équation de Helmholtz.

La dernière partie de notre travail est faite de quelques "contributions mineures" à l'analyse de certains problèmes de diffraction liés aux métamatériaux modélisés comme des matériaux avec des indices négatifs. La première contribution, présentée au chapitre 6, est l'analyse du problème de diffraction pour le cas où le contraste = -1. Il est bien connu que ce cas est le plus difficile à analyser puisque, par exemple, le problème n'est pas de type Fredholm dans l'espace  $H^1$  [18, 17, 22, 52]. Nous nous référons à [41, 97, 108, 93] pour une vue d'ensemble sur l'analyse mathématique et numérique de certains problèmes de diffraction avec des coefficients qui changent de signe. En utilisant une méthode de potentiels surfaciques, nous montrons que le problème est en fait bien posé dans le sens  $L^2$  pour le cas bidimensionnel. Nous nous

appuyons sur les propriétés des potentiels de surface pour des densités qui sont des traces de solutions  $L^2$  et qui ont été analysées dans [50, 49] dans le but d'établir la propriété Fredholm de ITP. Notre résultat complète les résultats dans [94] où le problème tridimensionnel est analysé. Nous nous référons également à [93] où des résultats complémentaires sont dérivés à l'aide de techniques variationnelles. Notre deuxième contribution, présentée au chapitre 7, est sur le caractère Fredholm de l'équation intégrale de volume qui est utilisée dans notre solveur numérique lorsque l'indice du milieu est négatif. Cette analyse est faite pour les équations de Maxwell et s'appuie sur l'approche T-coercivité et les résultats de [41].

## Résumés chapitre par chapitre

Le manuscrit est divisé en six chapitres et nous avons choisi de rendre chaque chapitre autonome dans le sens où il peut être lu indépendamment des autres chapitres. Une exception est faite pour le chapitre 5 puisqu'il ne correspond qu'à des généralisations de nos résultats dans les chapitres 2 - 4 à un modèle scalaire différent.

**Chapitre 2: Scattering problems for unbounded inhomogeneous layers** Il existe une abondante littérature sur les problèmes de diffraction par des structures non bornées [6, 9, 37, 86, 34, 107, 106]. Une première question est de spécifier la condition de rayonnement imposée à l'infini et donc l'espace de solutions. La condition de rayonnement peut également être substituée par l'utilisation d'espaces à poids adhoc [35, 9]. Nous adoptons ici la forme la plus simple de conditions de rayonnement basée sur l'application de la transformée de Fourier dans la direction longitudinale à la couche [86, 37]. Cela permet d'introduire un opérateur de Dirichlet-à-Neumann pour tronquer le domaine dans la direction orthogonale à la couche, puis d'utiliser une technique variationnelle pour étudier le problème de diffraction. Dans le cas des couches non bornées, le domaine variationnel reste non borné dans les directions transversales et donc l'utilisation de l'alternative de Fredholm n'est pas possible (comme dans le cas des problèmes de diffraction des domaines bornés). Une technique couramment employée pour remédier à cette difficulté consiste à dériver des estimations a priori sur la solution en utilisant des multiplicateurs spécifiques (technique de Rellich). On vérifie alors une condition inf-sup pour la formulation variationnelle pour conclure sur le caractère bien posé du problème. Ceci est par exemple l'approche appliquée dans [86] pour étudier le problème de diffraction par des couches diélectriques. Après avoir rappelé quelques résultats de cette dernière référence, nous les compléterons en étudiant un cas spécifique avec absorption. La raison principale de considérer ce cas est de permettre à l'indice de réfraction d'être le même des deux côtés de la couche (ce ne peut être le cas dans [86]). Nous obtenons une identité de Rellich qui ne demande pas une monotonie stricte de l'indice de réfraction en supposant une partie imaginaire de l'indice de réfraction définie positive sur une couche d'épaisseur positive (voir Fig. 2.2).

Dans le cas de couches périodiques, le signe de la partie imaginaire n'a pas besoin de tenir sur toute la couche d'épaisseur positive. Ceci est dû au fait que le domaine variationnel devient borné (en utilisant la condition de périodicité). Le caractère bien posé du problème peut alors être déduit à l'utilisation d'une alternative Fredholm dans les cas de sources quasi périodiques (ou ondes incidentes planes). Lorsque l'onde incidente n'est plus quasi-périodique, l'analyse ne peut plus être réduite à l'utilisation d'une alternative de Fredholm (puisque les estimations a priori sur la solution ne sont pas explicites en terme du paramètre de quasi-périodicité). On peut se fier dans ce cas aux résultats pour des couches non bornées qui ont besoin d'hypothèses plus fortes sur le coefficient d'absorption ou sur la monotonie stricte de la partie réelle.

Le cas des couches périodiques localement perturbées peut être considéré comme une perturbation compacte du problème de diffraction par des couches non bornées. Une alternative de Fredholm est utilisée pour déduire des résultats sur le caractère bien posé de ce problème.

Un autre cas où l'hypothèse sur la monotonie de la partie réelle de l'indice de réfraction peut être assouplie est le cas de problèmes de diffraction dans le demi-espace où une condition de limite de Dirichlet est imposée sur la frontière du demi-espace. Bien que des problèmes très similaires aient déjà été étudiés dans [86, 38], nous fournirons ici une preuve de ce cas pour la commodité du lecteur. Inclure ce cas dans notre présentation n'est pas essentiel, mais peut être utile pour le lecteur afin d'avoir une idée sur la façon dont les choses se simplifient dans ce cas. Remarquons également que, de la même manière, les résultats des chapitres 3-5 peuvent être facilement étendus au cas de demi-espace avec Dirichlet (ou conditions aux limites de Neumann, si l'on suppose que la problème direct est bien posé) avec des adaptations mineures et évidentes.

### Chapitre 3: A Volume integral method for solving scattering problems from locally perturbed infinite periodic layers

Nous sommes concernés dans ce chapitre par la conception d'une méthode numérique pour calculer la solution du problème de diffraction par des couches périodiques infinies localement perturbées. Notre travail peut être considéré comme une extension de la méthode numérique dans [103, 85] où la solution pour les milieux périodiques est calculée en exploitant la quasi-périodicité de l'onde incidente. Le principe de notre méthode est similaire à celui utilisé dans [42, 44] et repose sur l'utilisation de la transformée de Floquet-Bloch dans les directions de périodicité pour transformer le problème en un système de problèmes quasi périodiques couplés. Nos méthodologies diffèrent alors dans la façon dont nous formulons le problème et le discrétisons par rapport à la variable spatiale. Dans [42, 44] on utilise une méthode d'éléments finis utilisant un opérateur Dirichlet à Neumann pour borner le domaine de calcul comme introduit dans [68, 54]. Nous nous appuyons dans le présent travail sur une formulation intégrale volumique du problème. La discrétisation efficace de l'équation intégrale volumique grâce à l'utilisation d'une méthode spectrale et de la technique FFT pour évaluer le produit matrice-vecteur a été introduite dans [103] et étudiée dans un certain nombre de travaux [103, 84, 63, 64]. Nous nous basons ici principalement sur l'algorithme numérique mis en oeuvre dans [85] pour les problèmes quasi-périodiques, mais nous considérerons les modes TE. Le mode TM peut être traité de la même manière en utilisant les adaptations proposées dans [85]. Comme dans [42], la difficulté principale dans l'analyse de convergence est d'établir la convergence de la discrétisation du problème par rapport à la variable de Floquet-Bloch. L'utilisation d'une méthode des trapèzes pour évaluer la transformée de Floquet-Bloch induit un problème semi-discret qui est équivalent, à un changement du terme source près, au problème périodique avec une période  $ML$  où  $M$  est le nombre de points de discrétisation et  $L$  est la longueur de périodicité. La convergence (par rapport au paramètre  $M$ ) repose alors sur la décroissance de la solution dans les directions de périodicité. Cette dernière n'est en effet pas garantie en général et nécessite habituellement certaines hypothèses spéciales sur les propriétés du milieu, telle que la monotonie stricte dans la direction orthogonale aux directions de périodicité [86, 25, 38, 58, 83]. Nous nous limiterons ici à un problème simplifié où l'absorption (petite) dans le milieu est supposée en considérant des nombres d'ondes complexes avec des parties imaginaires positives. Cette hypothèse a également été faite dans [42]. En effet, cela simplifie grandement l'analyse puisque l'opérateur de Helmholtz devient coercitif. Cette hypothèse nous permet de nous concentrer davantage sur la spécificité de la discrétisation spectrale et le couplage de la méthode intégrale de volume avec la discrétisation dans la variable de Floquet-Bloch. En effet, cette étape nous permettra, dans un travail futur, de traiter le cas sans absorption et de nous concentrer sur la technicité liée à l'établissement des identités de type Rellich qui peuvent être exploitées dans

l'analyse de convergence. Notre hypothèse permet aussi d'évacuer les problèmes qui se produiraient aux anomalies du Wood. Remarquons cependant que l'incorporation d'une condition de rayonnement pour des problèmes sans absorption (comme dans [83]) n'entraînerait pas de difficultés supplémentaires majeures (pour la mise en oeuvre numérique de la méthode) puisque cette condition est codée dans l'équation intégrale volumique du problème.

L'analyse de convergence de l'approximation spectrale repose sur l'hypothèse que le système discret obtenu coïncide avec celui de l'équation intégrale volumique associée au problème périodique de période  $ML$ . Cette propriété est en effet surprenante et on ne s'attendrait pas à ce qu'elle tienne en général. En fait, elle est spécifique au choix d'une règle trapézoïdale pour évaluer la transformée de Floquet-Bloch. Elle a été également observée pour la discrétisation par éléments finis dans [42]. L'optimalité de ce choix de discrétisation n'est pas évidente et peut par exemple ne pas être appropriée au voisinage de l'anomalie de Wood (dans le cas sans absorption).

Notre système discret peut alors être considéré comme une réorganisation spéciale de la discrétisation du problème  $ML$ - périodique (avec un terme source légèrement différent). Un des principaux avantages de ce réarrangement est que l'on obtient une croissance linéaire du coût de calcul par rapport au nombre de points de discrétisation dans la variable de Floquet-Bloch. Il s'agit d'un gain d'un facteur logarithmique par rapport à l'application de la méthode dans [103] au problème  $ML$  périodique. En effet, ceci serait intéressant pour de grands problèmes (tridimensionnels) ou lorsque l'absorption est faible (nécessitant l'utilisation d'un grand nombre de points de discrétisation dans la variable de Floquet-Bloch). Nous discuterons cette question dans un travail futur. Dans le présent, nous nous contenterons de résultats de validation préliminaires en 2D du problème.

**Chapitre 4: Inverse scattering problem from locally perturbed periodic media** Nous étudions dans ce chapitre le problème inverse où l'on s'intéresse à reconstruire le support de la perturbation du milieu périodique à partir de mesures d'ondes diffractées. Nous nous intéressons à la conception d'une méthode d'échantillonnage qui reconstruirait le support des inhomogénéités sans reconstruire l'indice de réfraction. Le développement des méthodes d'échantillonnage a suscité un grand intérêt ces dernières années et de nombreuses méthodes ont été introduites dans la littérature pour traiter une large variété de problèmes de diffraction inverse [46, 29, 76, 30]. Le cas des milieux périodiques a été traité dans [82, 91, 10, 7, 98, 53, 105, 66]. Pour le problème inverse dans un guide d'onde périodique localement perturbé, on se réfère à [101, 27] et à leurs références.

A notre connaissance, les méthodes d'échantillonnage pour des couches périodiques infinies perturbées localement n'ont pas été traitées dans la littérature. Même si ce problème est celui qui motive notre étude, nous considérerons ici un problème légèrement différent qui sera appelé le problème  $ML$ - périodique: il correspond à une couche périodique infinie localement perturbée avec la période  $L$  qui a été réduit à un domaine de taille  $ML$  (avec  $M$  un paramètre suffisamment grand) avec des conditions aux limites périodiques. Cette limitation provient de des raisons techniques, car notre analyse pour la fonctionnelle d'imagerie différentielle nouvellement introduite repose fortement sur la transformée de Floquet-Bloch discrète. D'après le chapitre 3, le problème  $ML$ - périodique peut être considéré comme le problème discrétisé par rapport à la variable Floquet-Bloch en utilisant  $M$  points de discrétisation dans la règle trapézoïdale (voir Section 5.2.2 du Chapitre 3). Nous nous référons au chapitre 2 pour l'étude du caractère bien posé du problème  $ML$ - périodique.

Comme méthode d'inversion, nous utiliserons la version généralisée de la LSM (GLSM) récemment introduite (voir [12, 14, 30]) et considérons également la méthode Factorization (voir [70, 71, 76]). Nous considérons le cas où les données correspondent aux séquences de Rayleigh

du champ diffracté associées aux ondes incidentes qui sont des modes propagatifs et évanescents. Pour des problèmes inverses semblables dans les guides d'onde, nous nous référons à [28, 27, 82, 101]. Nous allons prouver dans une première partie comment la GLSM et aussi la méthode de factorisation peuvent être appliquées à notre problème.

La principale contribution de notre travail est la conception d'une nouvelle méthode d'échantillonnage qui permet d'imager la géométrie du défaut sans reconstruire milieu  $L$ -périodique. Cette méthode est dans l'esprit de la LSM différentielle qui a été introduite dans [13] pour l'imagerie de défauts dans des milieux complexes utilisant des mesures différentielles. Cependant, dans notre cas, nous introduirons une méthode qui n'exige pas l'opérateur de mesure pour le milieu de référence. Nous exploitons la  $L$ -périodicité du milieu sans défaut et la transformation de Floquet-Bloch pour concevoir un critère différentiel entre différentes périodes. Ce critère est basé sur l'étude des méthodes d'échantillonnage où un seul mode de Floquet-Bloch est utilisé. Cette étude constitue le principal ingrédient théorique de notre méthode. L'opérateur d'échantillonnage pour un seul mode de Floquet-Bloch joue en quelque sorte le rôle de l'opérateur de mesure pour milieu de référence. En effet, le principal intérêt pour cette nouvelle méthode d'échantillonnage est qu'elle est capable d'identifier le défaut même si les méthodes classiques d'échantillonnage échouent dans l'obtention d'une reconstruction précise du milieu de référence (complexe).

Les performances des méthodes d'échantillonnage sont testées par rapport aux données synthétiques générées par le solveur développé au chapitre 3.

**Chapitre 5: Extensions to TM-mode like equations** Lorsque le milieu est invariant dans une direction d'espace et lorsque la direction de l'onde plane incidente est choisie perpendiculairement à l'axe d'invariance, les équations de Maxwell sont découplées en deux équations de Helmholtz scalaires. Ils sont connus sous le nom de mode électrique transverse (TE) et mode magnétique transverse (TM). Dans le chapitre 3, nous avons considéré l'équation de Helmholtz pour le mode TM lorsque la perméabilité magnétique est constante et étudié le problème inverse associé dans le chapitre 4. Nous complétons ici l'étude des problèmes de diffraction à partir de couches périodiques infiniment perturbées localement en considérant le mode TM où la perméabilité magnétique peut varier, ce qui a la forme

$$\operatorname{div}(\mu^{-1}\nabla u) + k^2 n u = f \quad \text{in } \mathbb{R}^d, \quad (1)$$

où  $f \in L^2(\mathbb{R}^d)$  et  $\mu \in L^\infty(\mathbb{R}^d)$  désigne la perméabilité magnétique, telle que  $\mu^{-1} \in L^\infty(\mathbb{R}^d)$ . Cette équation peut également modéliser le mode TE où la permittivité électrique n'est pas constante. Bien que l'équation soit significative, dans le cas de problèmes de diffraction électromagnétique, seulement pour  $d = 2$ , nous gardons pour la généralité  $d = 2$  ou  $3$ . Le cas  $d = 3$  peut également modéliser les problèmes de diffraction acoustique pour les milieux à masse volumique non homogène.

Bien que ce ne soit pas exactement physiquement vrai, afin de distinguer le problème étudié dans le chapitre étudié ici, nous allons nous référer au cas  $\mu = 1$  comme le mode TE et le cas  $\mu \neq 1$  comme mode TM. L'étude des problèmes directs et inverses pour le mode TM est quasiment similaire au mode TE, mais elle est techniquement plus compliquée car la perméabilité magnétique affecte la partie principale de l'opérateur de Helmholtz. Nous décrirons dans ce chapitre les principaux résultats et incluons les principales preuves indiquant les principales différences (techniques). Le contenu de ce chapitre s'appuie donc fortement sur le contenu du chapitre 3 pour la méthode numérique pour résoudre le problème de diffraction et le contenu du chapitre 4 pour la partie inverse.

**Chapitre 6: Surface potential formulation of a scattering problem with sign changing coefficients** Le présent chapitre est consacré à l’analyse du problème de diffraction pour des milieux à indice négatif. Il contient un travail qui a été développé au début de ma thèse et qui n’est pas directement lié avec le sujet principal sur la diffraction par des milieux périodiques localement perturbés. Il est toutefois lié au thème principal du projet ANR de Metamath qui a financé mon travail de doctorat et a été motivé par l’analyse du problème de la transmission intérieure dans [50].

Nous considérons précisément le problème de diffraction pour le cas où le contraste entre deux milieux  $= -1$ . Il est bien connu que ce cas est le cas le plus difficile à analyser puisque, par exemple, le problème n’est plus du type Fredholm dans le sens  $H^1$  classique du problème [18, 17, 22, 52]. Nous nous référons à [41, 97, 108, 93] pour une vue d’ensemble sur l’analyse mathématique et numérique de certains problèmes de diffraction avec des coefficients de changeant de signe. En utilisant une méthode de potentiels surfaciques, nous montrons que le problème est en fait bien posé dans le sens  $L^2$  dans le cas bidimensionnel. Nous nous appuyons sur les propriétés des potentiels de surface pour les densités qui sont des traces de solutions  $L^2$  et qui ont été analysées dans [50, 49] dans le but d’établir la propriété Fredholm du problème de transmission intérieure. Notre résultat complète les résultats dans [94] où le problème tridimensionnel est analysé. On se réfère également à [93] où des résultats complémentaires sont obtenus à l’aide de techniques variationnelles.

Puisque ce n’est pas le coeur de mon travail de thèse, nous avons gardé ce chapitre aussi court que possible avec seulement les principaux points clés des preuves.

**Chapitre 7: Electromagnetic scattering by periodic structures with sign-changing coefficients** La théorie de la diffraction électromagnétique dans des structures périodiques est bien connue comme étant un sujet d’un grand intérêt dans des applications, par exemple pour la construction et l’optimisation de filtres optiques, de lentilles et de diviseurs de faisceaux en optique (voir [96]). Il existe une grande partie de la littérature mathématique appliquée sur les problèmes de diffraction directe et inverse liés aux matériaux diélectriques périodiques, voir par exemple [5, 91] et leurs références. Récemment, des métamatériaux périodiques ayant des applications en photonique et en optique ont reçu un intérêt de recherche intense pour la communauté de l’ingénierie et de la physique appliquée (voir par exemple [2, 95]). Cependant, à notre connaissance, il n’y a eu que peu de travaux sur les métamatériaux périodiques dans la littérature mathématique appliquée. Le papier [84] étudie la propriété de Fredholm d’un problème de diffraction pour des métamatériaux périodiques en deux dimensions pour le problème scalaire.

Nous considérons dans ce travail un problème de diffraction électromagnétique pour les réseaux bipériodiques constitués de matériaux diélectriques et de métamatériaux. Par bipériodique, nous voulons dire que la grille est périodique dans la direction, par exemple,  $x_1$  et  $x_2$ , alors qu’elle est bornée dans la direction  $x_3$ . Ceci est modélisé par un problème de diffraction périodique pour les équations de Maxwell 3D avec des coefficients changeant de signe. La transmission des problèmes de Maxwell avec des coefficients de changeant de signe dans des domaines bornés a été étudiée récemment dans [19] dans le cadre de la  $T$ -coercivité. Dans ce chapitre, nous généralisons un tel cadre pour les problèmes de diffraction périodique en utilisant une approche d’équation intégrale volumique. D’après les travaux [75, 91] on sait que les problèmes de diffraction pour les équations de Maxwell peuvent être formulés comme une équation intégrale hypersingulière de type Lipmann-Schwinger. Nous cherchons à prouver une estimation de type Garding pour l’équation intégrale avec des coefficients changeant de signe. L’idée est d’étudier ces estimations pour l’équation intégrale dans une cellule unitaire tronquée. Ceci permet l’utilisation de la formulation variationnelle des opérateurs intégraux, via l’incorporation de la  $T$ -coercivité.

Il s'avère que l'on a besoin d'hypothèses similaires à celles de [19] lors de l'application du cadre  $T$ -coercivité.

L'approche de l'équation intégrale dans ce travail présente certains avantages en ce sens que la propriété de Fredholm obtenue est valide aux fréquences de Rayleigh qui est typiquement exclue dans l'approche variationnelle. Elle évite en outre la complication technique que l'on pourrait avoir lors du traitement des termes de frontières, avec des opérateurs de Calderon, perturbées par l'opérateur abstrait  $T$  dans le cadre de la  $T$ -coercivité.

# Introduction

---

We are mainly interested in this thesis by the analysis of scattering and inverse scattering problems for locally perturbed periodic layers. This problem has connexions with many real life applications like photonics structures, optical fibers, gratings, etc. More specifically, our work has been motivated (and funded) by the ANR METAMATH where a special focus is put on simulating metamaterials and materials with negative indices. A metamaterial is generally composed of periodically spaced small resonators as indicated in Fig. 1.1. As other typical periodic structure that occur in recent technologies we can also quote the so-called nanoglass (see Fig. 1.2). In these structures nanotubes are chemically grown on a substrate and can be arranged in a periodic way (possibly with multiple periodicity scales).

The mathematical and numerical developments that are undertaken aim first at addressing the problem of non destructive testing of these periodic structures. In this problem, some incident (electromagnetic) wave is used to illuminate the structure and the response of the structure is measured at some distance from the structure. One is interested in identifying the presence of defects and if possibly find the location and the shape of these defects. We shall consider the problem at a fixed frequency meaning that the wave is harmonic in time with a known pulsation. For most of the thesis, only a simplified scalar problem is considered (modeling for instance some specific polarizations of the electromagnetic wave or some special configuration of acoustic waves [48]).

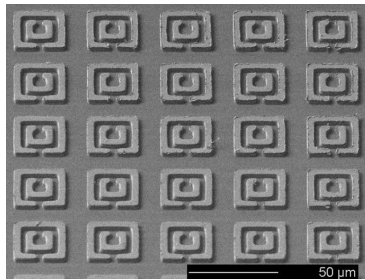


Figure 1.1: An example of a metamaterial micro-structure

Before addressing the inverse problem, we shall first analyze the forward scattering problem and design a numerical method that can simulate the scattering of waves from unbounded periodic layers that would contain a defect. The analysis of the direct problem is not straightforward since the periodic background is unbounded and therefore, for instance, some guided waves may contribute to the solution of the problem. The local perturbation of the periodic media also destroy the periodicity of the problem and therefore does not allow to reduce the scattering problem into a simpler problem posed on a periodicity cell. This is why the analysis for the locally perturbed problem appears as difficult as the analysis for infinite non periodic layers.



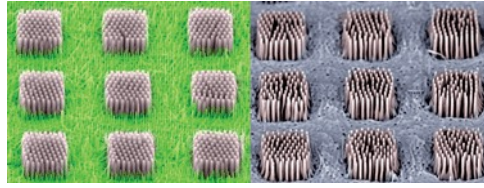


Figure 1.2: An example of a nanoglass structure

There exists a large literature dedicated to the analysis of either the guided waves (in open waveguides) [62, 20, 104] or the scattering from unbounded structures [33, 36, 37, 38, 39, 35, 8, 5, 24]. It is for instance well known that if the periodic layer is homogeneous with a constant refractive index greater than one, then some guided modes do exist [104, 25]. It is also proved in [86, 57, 38] for instance that if the index of refraction obeys some strict monotonicity property in the direction orthogonal to the periodicity direction then no guided modes exist. From the inverse problem perspective, we are more interested in analyzing the latter case than the former one. It is mainly due to the fact that for the applications we have in mind, measuring the guided waves would not be possible. We shall revisit in Chapter 2 the analysis of the scattering problem for unbounded layers and the conditions that would ensure the non existence of guided waves. The main tool is the so-called Rellich identity that is widely used in the analysis of this type of problem [37, 38]. The motivation behind our analysis is to weaken the monotonicity assumption in [86] that would prevent us to treat the scattering from a periodic layer embedded in the free space, which correspond with the configuration we would like to tackle for the inverse problem. Also this monotonicity condition is not compatible with some of the geometrical (technical) assumptions that we shall make in the study of the inverse problem.

The next step is the design of a numerical method to solve the direct scattering problem which we consider in Chapter 3. This step is needed for instance for generating the data for our inverse problem. We shall extend a method that has been introduced in the framework of the PhD thesis of D.L. Nguyen [91, 84] for the simulation of the scattering problem from periodic layers to handle the presence of a local perturbation. One can transform, using the Floquet-Bloch transform, the locally perturbed problem into coupled quasi-periodic problems. The quasi-periodicity factor is related to the Floquet-Bloch parameter. Our method is then based on discretizing first the Floquet-Bloch parameter and then apply to the semi-discrete problems the same discretization strategy as the one proposed in [91]. The latter is based of reformulating the scattering problem using volume potential representation of the solution (the Lippmann-Schwinger integral equation), then discretizing the problem using a spectral Fourier basis. In order to efficiently choose the discrete basis (making possible the use Fast Fourier Transform to evaluate matrix-vector products), some periodization of the volume potential kernel has to be incorporated. For an overview of this type of technique we refer to [99]. We obtain a set of coupled discrete equations that is solved iteratively using a Jacobi-type iterative fix point iterations. The two-steps procedure that we are following is largely inspired by the work in [43] where the use of a finite element method is studied for the space discretization. The convergence analysis for this methodology relies on observing that, if a uniform grid point is used to discretize the Floquet-Bloch variable, then the obtained semi-discrete problem is equivalent (up to a small change in the source term) to truncate the layer in the periodicity direction with a length proportional to the number of discretization points. One then understands why our method would be convergent only if there are no guided modes. For our analysis we restrict ourselves to the case of a wave number with non vanishing imaginary part (that would model some absorption in the media). Obtaining some weaker convergence results for the non absorbing case can be obtained as in [87]

if some regularity and decay properties on the solution is assumed.

The design of a method that can handle the propagative modes is an interesting perspective and one possibility would be to construct suitable Dirichlet-to-Neumann maps as in [68, 54, 55, 16]. We also quote the method proposed in [23, 15] that handle open waveguides as an alternative numerical approach or the method in [4] based on surface integral equations. We also think that one could extend our method that rely on the Floquet-Bloch transform if one can adapt the quadrature rule at the Floquet-Bloch parameters where some guided modes would appear. This is a major perspective of our work on the numerical part for the forward scattering problem.

As indicated earlier, our first motivation is to adress the inverse problem where one would like to image a defect in the periodic layer from measurements of the scattered waves at a given frequency. We are interested in extending so-called sampling methods [46, 29, 76, 30] to handle this inverse problem. This type of methods has been introduced in 1996 [45] and have gained since that time a large interest in the inverse problem community. The main advantages of this type of methods is that they do not rely on a solver for the forward problem. This make them computationally fast. Moreover, they provide, for some of them, a uniqueness proof for the reconstruction of the geometry without strong a priori knowledge on the material properties. Also, and compared to Born type methods [48], the sampling methods do not rely on a linearization of the inverse problem. The sampling methods rely however on so-called multistatic data: i.e. data associated with a large number of sources and a large number of receivers. They also provide only a geometrical information on the target.

The principle of a sampling method is to construct, from the measurement operator and carefully chosen test functions, an indicator function that would tell whether a point lies in the interior of the sought geometry or not. The differences between sampling methods is indeed in the way of constructing this indicator function. We shall mainly consider the case of the Linear Sampling Method (LSM) [45, 30], the Factorization Method (FM) [70, 72, 73, 74, 76, 80] and the Generalized Linear Sampling Method (GLSM) [14, 13, 12, 30]. These methods are related to each others and each one has its pros and cons. The LSM is probably the sampling method that has the most simple formulation and wider range of applications (various wave models and rather flexible configurations for the sources and the receivers). The LSM indicator function has also a direct interpretation in terms of solutions to the so-called interior transmission problem, which plays an important role in the method that we shall develop. However, this method has not a rigorous mathematical justification as one employs an approximation argument that is not explicitly constructed. The FM is however more robust in term of mathematical background since one has an exact characterization of the geometry by the constructed indicator function. The analysis of the factorization method relies however on a special factorization of the measurement operator that correspond in general to full aperture measurements. In the factorization method, one however loses the link with ITP. The GLSM is a method that comes in between the two methods. It gives an exact characterization of the sought geometry using an explicitly constructed asymptotic criterion. Moreover, it provides a direct link with solutions to the interior transmission problem. The principal drawback is probably the absence of a method that adjust the penalty factor to the noise present in the measurement operator. For some formulations that are relevant to limited aperture cases, the method is also numerically involving [12].

There has been several works that extended the sampling methods to periodic problems [82, 91, 10, 7, 98, 53, 105, 66]. However, we think that the case of locally perturbed periodic medium embedded in the whole space has not been adressed in the literature. In fact, we shall not treat exactly the full locally perturbed periodic problem but rather consider the discretized version of this problem with respect to the Floquet-Bloch variable. As already mentioned, this corresponds to truncating the medium in the direction of periodicity with periodicity conditions

and a period which is multiple of the background periodicity (roughly speaking, the periodicity length times the number of discretization points). This restriction is mainly due to technicalities that are easier to handle at the discrete level and that will be pointed out later. Doing so, we are then considering a medium that has two scales of periodicity. One very large scale compared to the wavelength and the other one is comparable to the wavelength. Our first contribution is to show how samplings methods can be used to solve the imaging problem for the larger periodicity length. This problem is very similar to waveguide problems as in [28, 27, 82, 101]. We then explain how in this special, setting one can use these methods to reconstruct the periodic background without knowing the defect.

Our most original contribution is on how one can use the sampling methods to directly reconstruct the defect. This is the subject of Chapter 4. The method that we propose is inspired by the Differential LSM introduced in [13, 12]. In the latter, one relies on differential measurements (measurements for the configuration with defects and measurements for the configuration without defects) to construct an indicator function that characterises the defect geometry. The construction of this indicator function relies on the GLSM formalism and the link with solutions to the over-mentioned ITP. For our problem we shall exploit the information that the background is periodic with known periodicity length to develop an indicator function that does not use differential measurements. We construct from the measurement operator a so-called single Floquet-Bloch mode measurement operator. This operator would play the role of the background operator in the case of differential measurements. The analysis of sampling methods associated with these single Floquet-Bloch modes operators for the perturbed periodic problem constitutes our main theoretical contribution. For instance, we show how these operators would reconstruct the background domain plus periodic copies of the defect. Analyzing the GLSM method for these operators requires the analysis of a new form of the ITP problem. The study of uniqueness of solutions to this problem (proving injectivity of the operators outside transmission eigenvalues) is done only for the discrete setting of the Floquet-Bloch variables, and this is where we need the technical (truncation) assumption previously mentioned. Comparison of ITP solutions allows us to construct an indicator function that is capable of identifying the defect. This is proved only in the case where there is no intersection between the defect and the background periodic domain. Indeed the method that we propose can be applied in the general case (i.e. outside the technical assumptions that we mentioned). The numerical results illustrate how the indicator function is capable of identifying the defect even though the sampling methods fail to capture the (complex) structure of the periodic background. This constitutes the main advantage of our differential imaging functional.

The numerical methods that we developed for the forward and inverse problems are presented first in the simple case where the change in the medium properties are modeled by changes in the index of refraction in front of the lower order term in the Helmholtz equation. The methods are however general and can be applied to other more complicated models. This is what we try to demonstrate in Chapter 5 by giving a sketch of the extensions to the scalar case where the changes in the medium affects also the main operator of the Helmholtz equation.

The last part of our work is made of some “minor contributions” to the analysis of some scattering problems related to metamaterials modeled as materials with negative indices. The first contribution, presented in Chapter 6, is the analysis of the scattering problem for the case where the contrast = -1. It is well known that this case is the most difficult case to analyze since for instance the problem is no longer of Fredholm type in the classical  $H^1$  setting of the problem [18, 17, 22, 52]. We refer to [41, 97, 108, 93] for an overview on the mathematical and numerical analysis of some scattering problems with sign changing coefficients. Employing a surface potential method, we show that the problem is in fact well posed in the  $L^2$  sense for the

---

two dimensional setting. We rely on the surface potential mapping properties for densities that are traces of  $L^2$  solutions and that were analyzed in [50, 49] for the purpose of establishing the Fredholm property of ITP. Our result complement the results in [94] where the three dimensional problem is analyzed. We also refer to [93] where complementary results are derived using variational techniques. Our second contribution, presented in Chapter 7, is on the Fredholm properties of the volume integral equation that is used in our numerical solver when the medium index is negative. This analysis is done for Maxwell's equations and rely on the T-coercivity approach and results from [41].

## Outline of the manuscript

The manuscript is divided into six chapters and we choose to make each chapter self-contained in the sense that it can be read independently from the other chapters. An exception is made for Chapter 5 since it only corresponds to generalizations of our results in Chapters 2-4 to a different scalar model.

**Chapter 2:** This chapter is dedicated to establishing non existence of guided modes of the scattering problem from unbounded dielectric layers using the Rellich type identities. We show in particular that if absorption is present in the layer, then one can avoid the strict monotonicity assumption for the real part of the refractive index. This allows in particular considering the problem where the layer is embedded in the free space. We also sketch the simpler results for periodic problems and the cases where the layer lies on a substrate with Dirichlet boundary conditions.

**Chapter 3:** We introduce and analyze a numerical method to solve the scattering problem from locally perturbed infinite periodic layers. Employing a Floquet-Bloch transform in the periodicity directions, we decompose the scattering problem into coupled quasi-periodic problems. The quasi-periodic problems are reformulated using the Lipmann-Schwinger integral equations. We first analyse the spectral method for discretizing these volume integral equations and prove convergence properties independent from the quasi-periodicity parameter. We then present the two-steps discretization procedure for the coupled problem and prove convergence for case with complex wave numbers with positive imaginary parts. We end the chapter with validating numerical results, discussing in particular the convergence rates-versus regularity of the refractive index. Most of the material in this chapter is extracted from the article [60].

**Chapter 4:** This chapter is dedicated to the design and analysis of sampling methods to recover the shape of a perturbation from measurements of scattered waves at a fixed frequency. We first introduce the model problem that corresponds with the semi-discretized version of the continuous model with respect to the Floquet-Bloch variable. We then present the inverse problem setting where (propagative and evanescent) plane waves are used to illuminate the structure and measurements of the scattered wave at a parallel plane to the periodicity directions are performed. We introduce the near field operator and analyze two possible factorizations of this operator. We then establish sampling methods (LSM, FM and GLSM) to identify the defect and the periodic background geometry from this operator measurement. We also show how one can recover the geometry of the background independently from the defect. Some validating results are given in 2D and simple configurations. We introduce and analyze in the last part the single Floquet-Bloch

mode measurement operators and show how one can exploit them to build an indicator function of the defect. Numerical validating results are provided for simple and complex backgrounds.

**Chapter 5:** We sketch in this chapter how one can extend the theory in Chapters 2 to 4 to handle the cases where the contrast in the medium is present in the main operator of the Helmholtz equation.

**Chapter 6:** This chapter is dedicated to the analysis of the well posedness of the scattering problem for negative materials with contrast  $= -1$ . After introducing the problem and recalling some known results on surface potentials, we reformulate the problem as a surface integral equation problem. We then recall the potential properties for the traces of  $L^2$  solutions and exploit them to prove the Fredholm properties of the problem in dimension 2.

**Chapter 7:** This chapter is also concerned with the scattering problem for negative materials. It corresponds to a joint work with D.L. Nguyen that was finalized at the beginning of my PhD thesis on the Fredholm properties of volume integral formulation of the scattering problem for Maxwell's equations using the T-Coercivity approach. This work has been published and we reproduce here the text of the article [92].

# Scattering problems for unbounded inhomogeneous layers

---

## Abstract

The goal of this chapter is two-folds. It first introduces the scattering problem that motivates the main developments in this thesis, namely the scattering problem for perturbed infinite periodic layers. It second complements the results of the literature on non existence of guided modes using Rellich type identities. We show in particular that if absorption is present in the layer, then one can avoid the strict monotonicity assumption for the real part of the refractive index. This allows in particular considering the problem where the layer is embedded in the free space. We also sketch the simpler results for periodic problems and the cases where the layer lies on a substrate with Dirichlet boundary conditions.

## Contents

---

<b>2.1</b>	<b>Introduction</b>	<b>17</b>
<b>2.2</b>	<b>Setting of the problem</b>	<b>19</b>
2.2.1	Construction of the DtN operators	19
2.2.2	Variational formulation of the problem	21
<b>2.3</b>	<b>A Rellich identity for the case with absorption</b>	<b>22</b>
<b>2.4</b>	<b>Scattering problem from periodic layers</b>	<b>30</b>
2.4.1	The DtN operator for $\xi$ -quasi-periodic problems	31
2.4.2	Well-posedness of the problem	33
<b>2.5</b>	<b>The case of infinite layers with local perturbations</b>	<b>34</b>
<b>2.6</b>	<b>On the case of the scattering problem for half space</b>	<b>35</b>
2.6.1	Variational formulation of the problem	36
2.6.2	The Rellich identity for the case without absorption	37

---

## 2.1 Introduction

We examine in this chapter the well-posedness of the scattering problem from unbounded dielectric layers in the harmonic regime. The case of locally perturbed periodic layers, which is central in this thesis, can be seen as a particular case of this problem. The analysis for the latter can also rely on results for periodic layers but we were not able to obtain more general results than those for "general" unbounded layers.

There exists an abundant literature on scattering problems from unbounded structures [6, 9, 37, 86, 34, 107, 106]. A first issue is to specify the radiation condition imposed at infinity and

accordingly the solution space. The radiation condition can also be substituted with the use of adhoc weighted spaces [35, 9]. We adopt here the simplest form of radiation conditions based on the application of the Fourier transform in the longitudinal direction of the layer [86, 37]. This allows the introduction of a Dirichlet-to-Neumann mapping to bound the domain in the direction “orthogonal” to the layer and then to use a variational technique to study the scattering problem. In the case on unbounded layers, the variational domain is still unbounded in transverse directions and therefore the use of Fredholm alternative is not possible (as in the case of scattering problems from bounded domains). A by now widely used technique to get around this difficulty, is to derive a priori estimates on the solution using specific multipliers (and yielding the so-called Rellich identities). One then verifies an inf-sup condition for the variational formulation to conclude on the well-posedness of the problem. This is for instance the approach applied in [86] to study the scattering problem from dielectric layers. After recalling some results from the latter reference, we shall complement them by studying a specific case with absorption. The main reason of considering this case is to allow the index of refraction to be the same on both sides of the layer (this cannot be the case in [86]). We obtain a Rellich identity independently from strict monotonicity of the refractive index by assuming a definite positive imaginary part of the refractive index on a layer of positive thickness (see Fig 2.2).

In the case of periodic layers, the sign for the imaginary part does not need to hold on a the whole layer of positive thickness. This is due to the fact that the variational domain becomes bounded (using the periodicity condition). The well posedness of the problem can then be deduced from the use of a Fredholm alternative in the cases of quasi-periodic sources (or incident waves). When the incident wave is no longer quasi-periodic, the analysis can no longer be recast to the use of a Fredholm alternative (since the obtained bound on the solution is not explicit in terms of the quasi-periodicity parameter). One can rely for this case on the results for unbounded layers that need stronger assumptions of the absorption coefficient or the strict monotonicity of the real part.

The case of locally perturbed periodic layers can be seen as a compact perturbation of the scattering problem from unbounded layers. A Fredholm alternative is used to infer results on the well-posedness for this problem.

Another case where the assumption on the monotonicity of the real part of the refractive index can be relaxed is the case of half-space scattering problems where a Dirichlet boundary condition is imposed on the half-space boundary. Although very similar problems are already studied in number of papers [86, 38], we shall provide here a proof of this case for the reader’s convenience. Including this case in our presentation is not essential but can be useful for the reader to get a hint on how things would simplify in this case. Let us also point out that similarly to this chapter, the results of Chapters 3-5 can be easily extended to the case of half space with Dirichlet (or Neumann boundary conditions, if one assumes that the direct problem is well posed) with minor and obvious adaptations.

The outline of the chapter is as follows. We first present the setting of the scattering problem for unbounded layers in Section (2.2). We introduce the Rayleigh radiation condition and establish the variational formulation of the problem. We then recall some results from the literature on the well-posedness of the problem under some monotonicity conditions on the real part of the refractive index. We present in Section 2.3 the main result of this chapter on the well-posedness of the problem for the case where absorption is present in some layer of positive thickness. We establish the Rellich identity that allows the derivation of well-posedness results. We indicate in Section 2.4 how the assumptions can be simplified for the case of periodic layers and how one can simply address the case of locally perturbed layers. Section 2.5 is dedicated to the study of problems that can be seen as “compact perturbations” of the problem studied in the previous

sections. We end up this chapter with Section 2.6 where the case of the scattering problem in half space with Dirichlet boundary conditions is studied.

## 2.2 Setting of the problem

We consider the Helmholtz equation given by

$$\Delta u + k^2 n u = g \quad \text{in } \mathbb{R}^d \quad (2.1)$$

with a source term  $g \in L^2(\mathbb{R}^d)$  and where the total field  $u$  satisfies some radiation condition (that will be specified later) with respect to the direction  $x_d$  and where  $n$  denotes the index of refraction. We consider the case where the *wavenumber*  $k$  is positive and real valued. In the following,  $D$  denotes a domain such that  $n = 1$  outside  $D$  and we shall assume that  $D \subset \Omega^h$  for some  $h > 0$  where  $\Omega^h := \mathbb{R}^{d-1} \times ]-h, h[$  (See Fig. 2.1). In a first step, we shall rephrase equation

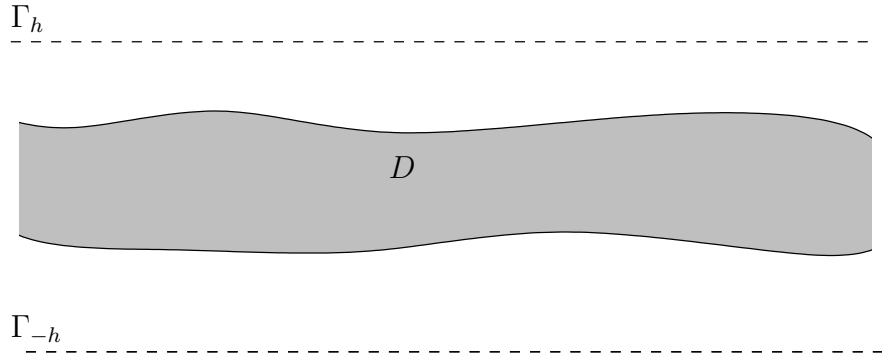


Figure 2.1: Sketch of the problem configuration and notations

(2.1) into an equation on  $\Omega^h$  by using some Dirichlet-to-Neumann (DtN) maps on

$$\Gamma_{\pm h} := \mathbb{R}^{d-1} \times \{\pm h\}$$

that are constructed by expressing the outgoing solutions on  $\Omega_{\pm}^h := \{x = (\bar{x}, x_d) \in \mathbb{R}^d \mid \pm x_d > h\}$  in terms of the solution trace on  $\Gamma_{\pm h}$ . The construction of DtN maps relies on the application of a Fourier transform in the directions orthogonal to  $x_d$  as explained in the following section.

### 2.2.1 Construction of the DtN operators

Our construction is classical and follows the one in [37, 86]. For  $u \in L^2(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ , the Fourier transform of  $u$  with respect to first  $d - 1$  variables is given by

$$\mathcal{F}u(\xi, x_d) := (2\pi)^{-(d-1)/2} \int_{\mathbb{R}^{d-1}} e^{-i\bar{x}\cdot\xi} u(\bar{x}, x_d) d\bar{x}, \quad \xi \in \mathbb{R}^{d-1}.$$

Applying the Fourier transform to (2.1) in  $\Omega_{\pm}^h$  we obtain an ordinary differential equation

$$\frac{\partial^2}{\partial x_d^2} \mathcal{F}u + (k^2 n - \xi^2) \mathcal{F}u = \mathcal{F}g$$

Let us choose a harmonic time dependence of the physical solution in the form  $u(x)e^{-i\omega t}$  where  $\omega$  is the time pulsation. Then, in order to ensure that  $u$  corresponds with outgoing solutions, we



have the following representation in the domains  $\Omega_h^\pm$

$$\mathcal{F}(u)(\xi, x_d) = \mathcal{F}(u|_{\Gamma_{\pm h}}) e^{i(x_d \mp h) \sqrt{k^2 - \xi^2}}, \quad \forall \pm x_d \geq h$$

where  $\sqrt{k^2 - \xi^2}$  is chosen to have a non negative imaginary part. Therefore

$$u(x) = (2\pi)^{-(d-1)/2} \int_{\mathbb{R}} e^{i((x_d \mp h) \sqrt{k^2 - \xi^2} + \bar{x} \cdot \xi)} \mathcal{F}(u|_{\Gamma_{\pm h}}) d\xi, \quad \forall x \in \Omega_{\pm}^h. \quad (2.2)$$

Taking the normal derivative of  $u$  on the boundary  $\Gamma_{\pm h}$  we then have

$$\left. \frac{\partial u}{\partial x_d} \right|_{\Gamma_{\pm h}} = i(2\pi)^{-(d-1)/2} \int_{\mathbb{R}^{d-1}} \sqrt{k^2 - \xi^2} e^{i\bar{x} \cdot \xi} \mathcal{F}(u|_{\Gamma_{\pm h}})(\xi) d\xi.$$

Identifying  $\Gamma_{\pm h}$  with  $\mathbb{R}^{d-1}$ , let  $H^s(\Gamma_{\pm h})$  denotes the completion of  $C_0^\infty(\Gamma_{\pm h})$  with respect to the norm  $\|\cdot\|_{H^s(\Gamma_{\pm h})}$  defined by

$$\|\varphi\|_{H^s(\Gamma_{\pm h})}^2 := \int_{\mathbb{R}^{d-1}} |k^2 + \xi^2|^s |\mathcal{F}\varphi(\xi)|^2 d\xi.$$

The DtN operator  $T^\pm : H^{1/2}(\Gamma_{\pm h}) \rightarrow H^{-1/2}(\Gamma_{\pm h})$  is given by

$$T^\pm \phi(\bar{x}, x_d) = i(2\pi)^{-(d-1)/2} \int_{\mathbb{R}^{d-1}} \sqrt{k^2 - \xi^2} e^{i\bar{x} \cdot \xi} \mathcal{F}(\phi|_{\Gamma_{\pm h}})(\xi) d\xi.$$

Let us denote by  $\langle \cdot, \cdot \rangle$  the sesquilinear duality product between  $H^{-1/2}(\mathbb{R}^{d-1})$  and  $H^{1/2}(\mathbb{R}^{d-1})$ .

**Lemma 2.2.1.** *The operators  $T^\pm : H^{1/2}(\Gamma_{\pm h}) \rightarrow H^{-1/2}(\Gamma_{\pm h})$  are bounded with  $\|T^\pm\| \leq 1$  and*

$$\operatorname{Re} \langle T^\pm u, u \rangle \leq 0 \quad \text{and} \quad \operatorname{Im} \langle T^\pm u, u \rangle \geq 0, \quad \forall u \in H^{1/2}(\Gamma_{\pm h}). \quad (2.3)$$

*Proof.* From Plancherel's identity and the definition of the operator  $T^\pm$  we have that

$$\begin{aligned} \|T^\pm u\|_{H^{-1/2}(\Gamma_{\pm h})} &= \left( \int_{\mathbb{R}^{d-1}} |k^2 + \xi^2|^{-1/2} |\mathcal{F}(T^\pm u)|^2 d\xi \right)^{1/2} \\ &= \left( \int_{\mathbb{R}^{d-1}} |k^2 + \xi^2|^{-1/2} |k^2 - \xi^2| |\mathcal{F}(u|_{\Gamma_{\pm h}})|^2 d\xi \right)^{1/2} \leq \|u\|_{H^{1/2}(\Gamma_{\pm h})}, \end{aligned}$$

which implies  $\|T^\pm\| \leq 1$ . We now prove (2.3). Using again the Plancherel identity we have that

$$\begin{aligned} \langle T^\pm u, u \rangle &= i(2\pi)^{-(d-1)} \int_{\Gamma_{\pm h}} \sqrt{k^2 - \xi^2} |\mathcal{F}(u|_{\Gamma_{\pm h}})(\xi)|^2 d\xi \\ &= i(2\pi)^{-(d-1)} \int_{k^2 \geq \xi^2} \sqrt{k^2 - \xi^2} |\mathcal{F}(u|_{\Gamma_{\pm h}})(\xi)|^2 d\xi \\ &\quad - (2\pi)^{-(d-1)} \int_{k^2 \leq \xi^2} \sqrt{|k^2 - \xi^2|} |\mathcal{F}(u|_{\Gamma_{\pm h}})(\xi)|^2 d\xi \end{aligned}$$

which proves that  $\operatorname{Re} \langle T^\pm u, u \rangle \leq 0$  and  $\operatorname{Im} \langle T^\pm u, u \rangle \geq 0$ .  $\square$

### 2.2.2 Variational formulation of the problem

The solution of problem (2.1) in  $\mathbb{R}^d$  subjected to the radiation condition with respect to the  $x_d$  direction is sought in  $H^1(\Omega^R)$  for all  $R > 0$ . More precisely, the problem can be formulated in the layer  $\Omega^h$  as: Find  $u \in H^1(\Omega^h)$  such that

$$\begin{aligned} \Delta u + k^2 n u &= g & \text{in } \Omega^h \\ \frac{\partial u}{\partial x_d} \Big|_{\Gamma_{\pm h}} &= T^{\pm} u & \text{on } \Gamma_{\pm h} \end{aligned} \quad (2.4)$$

where the source term  $g$  is now assumed to have support in  $\Omega^h$ . Multiplying with a test function  $\bar{v} \in H^1(\Omega^h)$  then integrating by part the first equation of (2.4) we obtain the equivalent variational equation: Find  $u \in H^1(\Omega^h)$  such that

$$\int_{\Omega^h} (\nabla u \cdot \nabla \bar{v} - k^2 n u \bar{v}) dx - \langle T^+ u, \bar{v} \rangle - \langle T^- u, \bar{v} \rangle = - \int_{\Omega^h} g \bar{v} dx, \text{ for all } v \in H^1(\Omega^h). \quad (2.5)$$

Let us introduce the sesquilinear form  $a(\cdot, \cdot) : H^1(\Omega^h) \times H^1(\Omega^h) \rightarrow \mathbb{C}$  defined by

$$a(u, v) = \int_{\Omega^h} (\nabla u \cdot \nabla \bar{v} - k^2 n u \bar{v}) dx - \langle T^+ u, \bar{v} \rangle - \langle T^- u, \bar{v} \rangle.$$

Then problem (2.5) is rewritten in the form

$$a(u, v) = - \int_{\Omega^h} g \bar{v} dx \quad \forall v \in H^1(\Omega^h). \quad (2.6)$$

Since the operators  $T^{\pm}$  are bounded then the sesquilinear form  $a$  is also bounded, i.e, there is a positive constant  $C$  independent from  $u, v$  such that

$$|a(u, v)| \leq C \|u\|_{H^1(\Omega^h)} \|v\|_{H^1(\Omega^h)}, \quad \forall u, v \in H^1(\Omega^h).$$

Hence, there exists a solution of problem (2.6) if the sesquilinear form  $a$  satisfies an inf-sup condition as the following (see in [67])

$$\inf_{u \in H^1(\Omega^h)} \sup_{v \in H^1(\Omega^h)} \frac{|a(u, v)|}{\|u\|_{H^1(\Omega^h)} \|v\|_{H^1(\Omega^h)}} \geq \gamma > 0. \quad (2.7)$$

and

$$\sup_{u \in H^1(\Omega^h)} \frac{|a(u, v)|}{\|u\|_{H^1(\Omega^h)}} > 0, \quad \forall 0 \neq v \in H^1(\Omega^h). \quad (2.8)$$

Since  $a(u, v) = a(\bar{v}, \bar{u})$  then (2.8) can be obtained from (2.7) and therefore the latter is sufficient. In fact, from (2.7) we deduce that

$$\inf_{u \in H^1(\Omega^h)} \sup_{v \in H^1(\Omega^h)} \frac{|a(\bar{v}, \bar{u})|}{\|u\|_{H^1(\Omega^h)} \|v\|_{H^1(\Omega^h)}} \geq \gamma > 0 \quad (2.9)$$

Changing the variables  $\bar{v}$  by  $u$  and  $\bar{u}$  by  $v$  we then have

$$\sup_{u \in H^1(\Omega^h)} \frac{|a(u, v)|}{\|u\|_{H^1(\Omega^h)}} \geq \gamma \|v\|_{H^1(\Omega^h)} > 0. \quad (2.10)$$

Following [67] we then have the following lemma

**Lemma 2.2.2.** *Let  $a : H^1(\Omega^h) \times H^1(\Omega^h) \rightarrow \mathbb{C}$  be a sesquilinear continuous form and  $F \in H^1(\Omega^h)'$ . If for all  $u \in H^1(\Omega^h)$  and verifying*

$$a(u, v) = F(v), \quad \forall v \in H^1(\Omega^h) \quad (2.11)$$

we have the estimate

$$\|u\|_{H^1(\Omega^h)} \leq C \|F\|_{H^1(\Omega^h)'} \quad (2.12)$$

for some constant  $C$  independent from  $u$  and  $F$ , then the sesquilinear form  $a$  verifies the inf-sup condition (2.7).

*Proof.* From (2.12) we have that

$$\frac{1}{C} =: \gamma \leq \frac{\|F\|_{H^1(\Omega^h)'}}{\|u\|_{H^1(\Omega^h)}} = \sup_{v \in H^1(\Omega^h)} \frac{|F(v)|}{\|u\|_{H^1(\Omega^h)} \|v\|_{H^1(\Omega^h)}} = \sup_{v \in H^1(\Omega^h)} \frac{|a(u, v)|}{\|u\|_{H^1(\Omega^h)} \|v\|_{H^1(\Omega^h)}} \quad (2.13)$$

Since (2.12) holds true for all  $F \in H^1(\Omega^h)'$  then (2.13) holds true for all  $u \in H^1(\Omega^h)$ , which implies (2.7).  $\square$

As indicated above, problem (2.6) has been already studied in [86]. In their study, a strict monotonicity of the real part of the index of refraction is assumed. We summarize their result in the following theorem. We assume that  $n$  equals two (real) constants  $n^\pm$  in the two domains  $\Omega_\pm^h$  respectively and such that  $n^\pm \geq c_0 > 0$ ,  $n^+ \geq \operatorname{Re} n$  and  $\operatorname{Im} n \geq 0$ . Moreover consider a surface  $\Gamma \subset \Omega^h$  on which  $n$  may have a discontinuity with  $\Gamma := \{x \in \mathbb{R}^d, x_d = f(\bar{x})\}$  for  $f \in C^{0,1}(\overline{\Omega^+}, \mathbb{R}) \cap C^{0,1}(\overline{\Omega^-}, \mathbb{R})$  with  $\Omega^\pm := \{x \in \Omega, x_d \gtrless f(\bar{x})\}$ . We further assume that

$$\frac{\partial \operatorname{Re} n}{\partial x_d} \geq 0 \text{ in } \Omega^h \quad \text{and} \quad [\operatorname{Re} n]_\Gamma := \operatorname{Re} n|_\Gamma^+ - \operatorname{Re} n|_\Gamma^- \geq 0 \quad (2.14)$$

where the  $\pm$  sign corresponds with limiting values on  $\Gamma$  of  $n$  in  $\Omega^\pm$ . Then one has the following theorem [86].

**Theorem 2.2.3.** *We assume in addition to the hypothesis above that either  $[n]_\Gamma \geq c_0 > 0$  or  $\partial n / \partial x_d \geq c_0 > 0$  in a layer  $\Omega^\delta$  of thickness  $2\delta > 0$ . Then there exists a unique solution  $u \in H^1(\Omega^h)$  of the variational problem (2.6) and the solution depends continuously on the data  $g \in L^2(\Omega^h)$  as*

$$\|u\|_{H^1(\Omega^h)} \leq C(k_0)(1 + k^3) \|g\|_{L^2(\Omega^h)} \quad (2.15)$$

with  $C(k_0)$  independent of  $k \geq k_0 > 0$ .

This result excludes the case  $n^+ = n^-$  that we are considering in this thesis. It is known that in this case some propagative modes (along the transverse directions of the layer) may exist [104, 25]. In the following we shall prove that if there is absorption in  $\Omega^\delta$  then the problem is still well posed (without strict monotonicity of the real part). The case where the absorption is not in a layer of uniform thickness is still an open question for us.

## 2.3 A Rellich identity for the case with absorption

We shall study in this section the well-posedness of the problem for the case where some absorption is present in the media. More specifically, let us denote by  $\tilde{D}$  the domain such that  $\operatorname{Im} n = 0$  outside  $\tilde{D}$ . Then we shall make the following assumption (the assumption is illustrated in Fig. 2.2)

**Assumption 2.3.1.** *The index of refraction  $n \in L^\infty(\mathbb{R}^d)$  with non negative imaginary part. Moreover, the domain  $\tilde{D}$  contains a thin infinite layer  $\Omega^\delta := \mathbb{R}^{d-1} \times ]-\delta, \delta[$ ,  $\delta > 0$  and the refractive index  $n$  has positive definite imaginary part in  $\tilde{D}$ , i.e.  $\exists c_0 > 0$  such that  $\text{Im } n \geq c_0$  in  $\tilde{D}$ .*

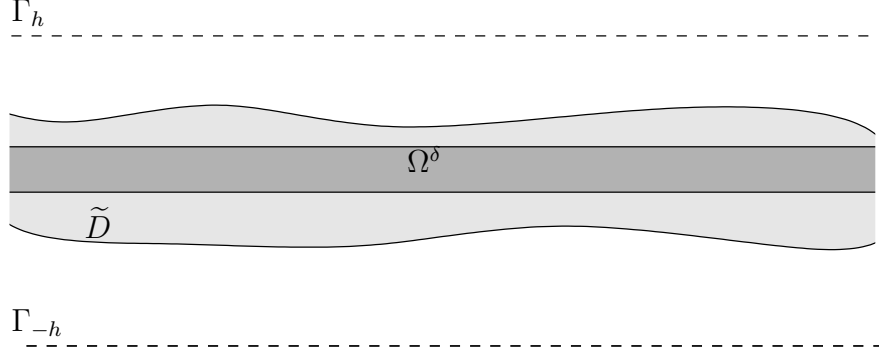


Figure 2.2: Illustration of Assumption 2.3.1

In order to prove the main theorem in this section, we need some preliminary technical results that are given in the form of four lemmas.

**Lemma 2.3.2.** *Assume that  $\tilde{D}$  is as in Assumption 2.3.1. Then the following (Poincaré's like) inequality holds*

$$\|u\|_{L^2(\Omega^h)}^2 \leq C \left( \|u\|_{L^2(\tilde{D})}^2 + \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}^2 \right), \quad \forall u \in H^1(\Omega^h). \quad (2.16)$$

where  $C$  is a positive constant independent from  $u$ .

*Proof.* For  $x = (\bar{x}, x_d) \in \Omega^h$ , let  $x_s := (\bar{x}, s) \in \tilde{D}$  with  $s \in ]-\delta, \delta[$ . Then we have

$$u(\bar{x}, x_d) = u(\bar{x}, s) + \int_s^{x_d} \frac{\partial u}{\partial x_d}(\bar{x}, t) dt. \quad (2.17)$$

Therefore,

$$\begin{aligned} |u(\bar{x}, x_d)|^2 &\leq 2 \left( |u(\bar{x}, s)|^2 + \left| \int_s^{x_d} \frac{\partial u}{\partial x_d}(\bar{x}, t) dt \right|^2 \right) \\ &\leq 2|u(\bar{x}, s)|^2 + 4h \int_{-h}^h \left| \frac{\partial u}{\partial x_d}(\bar{x}, t) \right|^2 dt \end{aligned} \quad (2.18)$$

Integrating inequality (2.18) with respect to  $x_d$  over  $[-h, h]$  we get

$$\int_{-h}^h |u(\bar{x}, x_d)|^2 dx_d \leq 4h|u(\bar{x}, s)|^2 + 8h^2 \int_{-h}^h \left| \frac{\partial u}{\partial x_d}(\bar{x}, t) \right|^2 dt \quad (2.19)$$

Integrating inequality (2.19) over  $\tilde{D}$  we finally obtain

$$\delta \|u\|_{L^2(\Omega^h)}^2 \leq 4h \|u\|_{L^2(\tilde{D})}^2 + 8h^2 \delta \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}^2 \quad (2.20)$$

which implies the result of the lemma with  $C := \max\{4h/\delta, 8h^2\}$ .  $\square$

**Lemma 2.3.3.** *Assume that Assumption 2.3.1 holds and let  $u \in H^1(\Omega^h)$  satisfying (2.6). Then the following a priori estimate holds*

$$\|u\|_{L^2(\tilde{D})}^2 \leq \frac{1}{c_0 k^2} \|g\|_{L^2(\Omega^h)} \|u\|_{L^2(\Omega^h)}. \quad (2.21)$$

*Proof.* Considering the imaginary part of equation (2.6) with  $v$  is replaced by  $u$ , we get

$$k^2 \int_{\Omega^h} \operatorname{Im} n |u|^2 dx + \operatorname{Im} \langle T^+(u), u \rangle + \operatorname{Im} \langle T^-(u), u \rangle = \operatorname{Im} \int_{\Omega^h} g \bar{u} dx. \quad (2.22)$$

This implies, using Lemma 2.2.1 and Assumption 2.3.1, that

$$c_0 k^2 \|u\|_{L^2(\tilde{D})}^2 \leq k^2 \int_{\Omega^h} \operatorname{Im} n |u|^2 dx \leq \|g\|_{L^2(\Omega^h)} \|u\|_{L^2(\Omega^h)}. \quad (2.23)$$

which proves the lemma.  $\square$

From Lemma 2.3.2 and Lemma 2.3.3 we obtain the following bound for the  $L^2$  norm of the solution in terms of derivatives with respect to  $x_d$  only. This type of estimate is the main motivation behind the derivation of the Rellich estimate later.

**Corollary 2.3.4.** *Assume that Assumption 2.3.1 holds and let  $u \in H^1(\Omega^h)$  satisfying (2.6). Then there exists a constant  $C$  independent from  $u$  and  $g$  that*

$$\|u\|_{L^2(\Omega^h)} \leq C \left( \|g\|_{L^2(\Omega^h)} + \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)} \right). \quad (2.24)$$

*Proof.* From Lemma 2.3.2 and Lemma 2.3.3, there is  $\tilde{C} > 0$  such that

$$\|u\|_{L^2(\Omega^h)}^2 \leq \tilde{C} \left( \frac{1}{c_0 k^2} \|g\|_{L^2(\Omega^h)} \|u\|_{L^2(\Omega^h)} + \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}^2 \right). \quad (2.25)$$

Therefore,

$$\left( \|u\|_{L^2(\Omega^h)}^2 - \frac{\tilde{C}}{2c_0 k^2} \|g\|_{L^2(\Omega^h)} \right)^2 \leq \frac{\tilde{C}^2}{c_0^2 k^4} \|g\|_{L^2(\Omega^h)} + \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}^2 \quad (2.26)$$

which implies

$$\|u\|_{L^2(\Omega^h)} \leq \frac{3\tilde{C}}{2c_0 k^2} \|g\|_{L^2(\Omega^h)} + \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}. \quad (2.27)$$

and proves the lemma with  $C := \max\{3\tilde{C}/2c_0 k^2, 1\}$ .  $\square$

We now remark that for a source term  $g \in L^2(\mathbb{R}^d)$ , a solution  $u \in H^1(\Omega^h)$  of problem (2.1) belongs to  $H^2(\mathcal{O})$  for all compact sets  $\mathcal{O} \subset \mathbb{R}^d$  (using classical elliptic regularities). Follows [88] we also have:

**Lemma 2.3.5.** *Assume that Assumption 2.3.1 holds. Then any solution  $u \in H^1(\Omega^h)$  of (2.4) belongs to  $H^2(\Omega^h)$ .*

*Proof.* Let  $\tilde{h} > h$  and  $\delta > 0$ . For all  $j \in \mathbb{Z}$ ,  $j > 0$  we consider  $\tilde{C}_j := [-j - \delta, j + \delta] \times [-\tilde{h}, \tilde{h}]$  and  $C_j := [-j, j] \times [-h, h]$ . Choose a cut-off function  $\chi \in C_c^\infty(\tilde{C}_j)$  such that  $\chi = 1$  in  $C_j$ . Assume  $u$  verifies (2.4) then  $\chi u$  verifies

$$\Delta(\chi u) + k^2 n(\chi u) = g_1, \quad \text{in } \tilde{C}_j \quad (2.28)$$

with

$$\|g_1\|_{L^2(\tilde{C}_j)} \leq \|\chi g\|_{L^2(\tilde{C}_j)} + C\|u\|_{H^1(\tilde{C}_j)}. \quad (2.29)$$

Squaring and integrating the first equation of (2.28) with moving the term  $k^2 n(\chi u)$  to right hand side, we then have:

$$\int_{\tilde{C}_j} |\Delta(\chi u)|^2 dx = \int_{\tilde{C}_j} |k^2 n(\chi u) + g_1|^2 dx \quad (2.30)$$

We perform integration by part the left hand side and obtain

$$\int_{\tilde{C}_j} |\Delta(\chi u)|^2 dx = \int_{\tilde{C}_j} \sum_{i,j=1}^d \partial_{ii}^2(\chi u) \partial_{jj}^2(\chi u) dx = \int_{\tilde{C}_j} \sum_{i,j=1}^d \partial_{ij}^2(\chi u) \partial_{ij}^2(\chi u) dx. \quad (2.31)$$

Therefore, from two previous equations we get

$$\begin{aligned} \|(\chi u)\|_{H^2(\tilde{C}_j)}^2 &= \|(\chi u)\|_{H^1(\tilde{C}_j)}^2 + \int_{\tilde{C}_j} \sum_{i,j=1}^d \partial_{ij}^2(\chi u) \partial_{ij}^2(\chi u) dx \\ &= \|(\chi u)\|_{H^1(\tilde{C}_j)}^2 + \int_{\tilde{C}_j} |k^2 n(\chi u) + g_1|^2 dx \\ &\leq \|(\chi u)\|_{H^1(\tilde{C}_j)}^2 + 2k^2 \|n\|_{L^\infty(\tilde{C}_j)} \|(\chi u)\|_{L^2(\tilde{C}_j)}^2 + 2\|g_1\|_{L^2(\tilde{C}_j)}^2. \end{aligned} \quad (2.32)$$

Combine with (2.29) and with  $\chi = 1$  in  $C_j$  we arrive that

$$\|u\|_{H^2(C_j)} \leq \theta \left( \|u\|_{H^1(\tilde{C}_j)} + \|g\|_{L^2(\tilde{C}_j)} \right), \quad (2.33)$$

where  $\theta$  independent from  $C_j$  and  $\tilde{C}_j$ . Therefore, let  $j$  tend to  $+\infty$  we obtain

$$\|u\|_{H^2(\Omega^h)} \leq \theta \left( \|u\|_{H^1(\Omega^{\tilde{h}})} + \|g\|_{L^2(\Omega^{\tilde{h}})} \right) \quad (2.34)$$

Follows [37, 86] we have that

$$\|u\|_{H^1(\mathbb{R} \times [h, \tilde{h}])} \leq \tilde{\theta} \|u\|_{H^1(\Omega^h)}. \quad (2.35)$$

and

$$\|u\|_{H^1(\mathbb{R} \times [-\tilde{h}, -h])} \leq \tilde{\theta} \|u\|_{H^1(\Omega^h)}. \quad (2.36)$$

This gives the existence of a positive constant  $C$  such that

$$\|u\|_{H^2(\Omega^h)} \leq C \left( \|u\|_{H^1(\Omega^h)} + \|g\|_{L^2(\Omega^h)} \right), \quad (2.37)$$

which proves the lemma.  $\square$

We now prove the main identity that will allow us to obtain a priori estimate on the  $H^1$  norm of the solution.

**Lemma 2.3.6.** *Assume that Assumption 2.3.1 holds and that*

$$\frac{\partial \operatorname{Re} n}{\partial x_d} \in L^\infty(\mathbb{R}^d).$$

Then the following identity holds for  $u \in H^1(\Omega^h)$  satisfying (2.6):

$$\begin{aligned} & \int_{\Omega^h} \left( 2 \left| \frac{\partial u}{\partial x_d} \right|^2 dx + k^2(x_d + h) \frac{\partial \operatorname{Re} n}{\partial x_d} |u|^2 + 2k^2(x_d + h) \operatorname{Im} n \operatorname{Im} \left( u \frac{\partial}{\partial x_d} \bar{u} \right) \right) dx \\ & + 2h \int_{\Gamma_h} \left( |\nabla u|^2 - 2 \left| \frac{\partial u}{\partial x_d} \right|^2 - k^2 |u|^2 \right) ds = \operatorname{Re} \langle T^+ u, u \rangle + \operatorname{Re} \langle T^- u, u \rangle \\ & - \operatorname{Re} \int_{\Omega^h} g \bar{u} dx - 2 \operatorname{Re} \int_{\Omega^h} (x_d + h) g \frac{\partial \bar{u}}{\partial x_d} dx \end{aligned}$$

*Proof.* From Lemma 2.3.5, a solution  $u$  of equation (2.4) belongs to  $H^2(\Omega^h)$ . Therefore, the partial derivative  $\partial u / \partial x_d$  belongs to  $H^1(\Omega^h)$ . Multiplying the first equation of (2.4) with  $(x_d + h) \partial \bar{u} / \partial x_d$  and then integrating over  $\Omega^h$  gives

$$\int_{\Omega^h} (x_d + h) \frac{\partial \bar{u}}{\partial x_d} (\Delta u + k^2 n u) dx = \int_{\Omega^h} (x_d + h) \frac{\partial \bar{u}}{\partial x_d} g dx. \quad (2.38)$$

We first observe that

$$\begin{aligned} I & := \int_{\Omega^h} (x_d + h) \frac{\partial \bar{u}}{\partial x_d} (\Delta u + k^2 n u) \\ & = \int_{\partial \Omega^h} (x_d + h) \frac{\partial u}{\partial \nu} \frac{\partial \bar{u}}{\partial x_d} - \int_{\Omega^h} \nabla u \cdot \nabla \left( (x_d + h) \frac{\partial \bar{u}}{\partial x_d} \right) + \int_{\Omega^h} k^2 n (x_d + h) u \frac{\partial \bar{u}}{\partial x_d}, \end{aligned} \quad (2.39)$$

where

$$\begin{aligned} & \int_{\partial \Omega^h} (x_d + h) \frac{\partial u}{\partial \nu} \frac{\partial \bar{u}}{\partial x_d} = \int_{\Gamma_{\pm h}} (x_d + h) \frac{\partial u}{\partial x_d} \frac{\partial \bar{u}}{\partial x_d} = 2h \int_{\Gamma_h} \left| \frac{\partial u}{\partial x_d} \right|^2 \\ & \nabla \left( (x_d + h) \frac{\partial \bar{u}}{\partial x_d} \right) = (x_d + h) \nabla \frac{\partial \bar{u}}{\partial x_d} + \vec{e}_2 \frac{\partial \bar{u}}{\partial x_d}. \\ & \nabla u \cdot \nabla \left( (x_d + h) \frac{\partial \bar{u}}{\partial x_d} \right) = (x_d + h) \nabla u \cdot \frac{\partial}{\partial x_d} \nabla \bar{u} + \left| \frac{\partial u}{\partial x_d} \right|^2. \end{aligned}$$

Therefore,

$$I = 2h \int_{\Gamma_h} \left| \frac{\partial u}{\partial x_d} \right|^2 - \int_{\Omega^h} \left( \left| \frac{\partial u}{\partial x_d} \right|^2 + (x_d + h) \nabla u \cdot \frac{\partial}{\partial x_d} \nabla \bar{u} \right) + \int_{\Omega^h} k^2 n (x_d + h) u \frac{\partial \bar{u}}{\partial x_d}.$$

Taking the real part of  $I$  yields

$$\begin{aligned} 2 \operatorname{Re}(I) & = 4h \int_{\Gamma_h} \left| \frac{\partial u}{\partial x_d} \right|^2 - \int_{\Omega^h} \left( 2 \left| \frac{\partial u}{\partial x_d} \right|^2 + (x_d + h) \frac{\partial}{\partial x_d} |\nabla u|^2 \right) \\ & + \int_{\Omega^h} k^2 (x_d + h) \left( \operatorname{Re} n \frac{\partial}{\partial x_d} |u|^2 - 2 \operatorname{Im} n \operatorname{Im} \left( u \frac{\partial}{\partial x_d} \bar{u} \right) \right) dx \\ & = 4h \int_{\Gamma_h} \left| \frac{\partial u}{\partial \nu} \right|^2 - \int_{\Omega^h} 2 \left| \frac{\partial u}{\partial x_d} \right|^2 dx - 2h \int_{\Gamma_h} |\nabla u|^2 + \int_{\Omega^h} |\nabla u|^2 \\ & + 2k^2 h \int_{\Gamma_h} |u|^2 - k^2 \int_{\Omega^h} \left( \operatorname{Re} n + (x_d + h) \frac{\partial \operatorname{Re} n}{\partial x_d} \right) |u|^2 - 2k^2 \int_{\Omega^h} (x_d + h) \operatorname{Im} n \operatorname{Im} \left( u \frac{\partial \bar{u}}{\partial x_d} \right) dx \\ & = \int_{\Omega^h} \left( |\nabla u|^2 - k^2 \operatorname{Re} n |u|^2 - 2 \left| \frac{\partial u}{\partial x_d} \right|^2 - k^2 (x_d + h) \frac{\partial \operatorname{Re} n}{\partial x_d} |u|^2 - 2k^2 (x_d + h) \operatorname{Im} n \operatorname{Im} \left( u \frac{\partial \bar{u}}{\partial x_d} \right) \right) dx \\ & - 2h \int_{\Gamma_h} \left( |\nabla u|^2 - 2 \left| \frac{\partial u}{\partial \nu} \right|^2 - k^2 |u|^2 \right). \end{aligned}$$

Taking the real part of (2.6) with  $v$  is replaced by  $u$  implies

$$\int_{\Omega^h} |\nabla u|^2 - k^2 \operatorname{Re} n |u|^2 = \operatorname{Re} \langle T^+ u, u \rangle + \operatorname{Re} \langle T^- u, u \rangle - \operatorname{Re} \int_{\Omega^h} g \bar{u}.$$

Substituting in the expression of  $\operatorname{Re}(I)$  finally gives

$$\begin{aligned} & \int_{\Omega^h} \left( 2 \left| \frac{\partial u}{\partial x_d} \right|^2 dx + k^2 (x_d + h) \frac{\partial \operatorname{Re} n}{\partial x_d} |u|^2 + 2k^2 (x_d + h) \operatorname{Im} n \operatorname{Im} \left( u \frac{\partial}{\partial x_d} \bar{u} \right) \right) dx \\ & + 2h \int_{\Gamma_h} \left( |\nabla u|^2 - 2 \left| \frac{\partial u}{\partial x_d} \right|^2 - k^2 |u|^2 \right) ds = \operatorname{Re} \langle T^+ u, u \rangle + \operatorname{Re} \langle T^- u, u \rangle \\ & - \operatorname{Re} \int_{\Omega^h} g \bar{u} dx - 2 \operatorname{Re} \int_{\Omega^h} (x_d + h) g \frac{\partial \bar{u}}{\partial x_d} dx \end{aligned}$$

which proves the lemma.  $\square$

We now shall exploit the identity of Lemma 2.3.6 to obtain the desired continuity result for solutions of (2.2). We first treat the boundary terms (as in [86]).

**Lemma 2.3.7.** *For  $u \in H^2(\Omega^h)$  satisfying (2.2) we have*

$$\int_{\Gamma_h} \left( -|\nabla u|^2 + 2 \left| \frac{\partial u}{\partial x_d} \right|^2 + k^2 |u|^2 \right) ds \leq 2k \operatorname{Im} \langle T^+(u), u \rangle. \quad (2.40)$$

*Proof.* We recall from (2.2) that

$$u(\bar{x}, x_d) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{i((x_d-h)\sqrt{k^2-\xi^2} + \bar{x}\xi)} \mathcal{F}(u|_{\Gamma_h})(\xi) d\xi, \quad x_d \geq h. \quad (2.41)$$

Then,

$$\mathcal{F} \left( \frac{\partial u}{\partial \bar{x}} \Big|_{\Gamma_h} \right) = i\xi \mathcal{F}(u|_{\Gamma_h})(\xi) \quad \text{and} \quad \mathcal{F} \left( \frac{\partial u}{\partial x_d} \Big|_{\Gamma_h} \right) = i\sqrt{k^2 - \xi^2} \mathcal{F}(u|_{\Gamma_h})(\xi) \quad (2.42)$$

Using the Plancherel identity we get on the one hand

$$\begin{aligned} \int_{\Gamma_h} -|\nabla u|^2 + 2 \left| \frac{\partial u}{\partial x_d} \right|^2 + k^2 |u|^2 &= \int_{\Gamma_h} -|\mathcal{F}(\nabla u)|^2 + 2 \left| \mathcal{F} \left( \frac{\partial u}{\partial x_d} \right) \right|^2 + k^2 |\mathcal{F}(u|_{\Gamma_h})|^2 \\ &= \int_{\Gamma_h} (|k^2 - \xi^2| + k^2 - \xi^2) |\mathcal{F}(u|_{\Gamma_h})|^2 \\ &= \begin{cases} 0 & \text{if } k^2 \leq \xi^2 \\ 2 \int_{\Gamma_h} (k^2 - \xi^2) |\mathcal{F}(u|_{\Gamma_h})|^2 & \text{if } k^2 \geq \xi^2. \end{cases} \end{aligned}$$

On the other hand, we observe that

$$\operatorname{Im} \langle T^+(u), u \rangle = \begin{cases} 0 & \text{if } k^2 < \xi^2 \\ \sqrt{k^2 - \xi^2} |\mathcal{F}(u|_{\Gamma_h})|^2 & \text{if } k^2 \geq \xi^2. \end{cases}$$

Therefore, to prove (2.40) we just need to show that

$$k^2 - \xi^2 \leq k\sqrt{k^2 - \xi^2}, \quad \forall k^2 \geq \xi^2. \quad (2.43)$$

This inequality is equivalent to  $\xi^2(\xi^2 - k^2) \leq 0$  which holds true for all  $\xi^2 \leq k^2$ . This proves the lemma.  $\square$



**Lemma 2.3.8.** *Assume that hypothesis 2.3.1 holds and assume in addition that*

$$\frac{\partial \operatorname{Re} n}{\partial x_d} \geq 0 \text{ in } D \setminus \tilde{D}.$$

*Then, there exists  $\theta > 0$  such that for all  $u \in H^1(\Omega^h)$  satisfying (2.6) and  $g \in L^2(\Omega^h)$*

$$\left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}^2 \leq \theta \|g\|_{L^2(\Omega^h)}^2. \quad (2.44)$$

*Proof.* Combining equation (2.22) with the assumption  $\operatorname{Im} n \geq 0$  in  $\Omega^h$ , we deduce that

$$\operatorname{Im} \langle T^+(u), u \rangle \leq \operatorname{Im} \int_{\Omega^h} g \bar{u} \, dx. \quad (2.45)$$

Combining with Lemmas 2.3.6 and 2.3.7 we then immediately obtain

$$\begin{aligned} \int_{\Omega^h} 2 \left| \frac{\partial u}{\partial x_d} \right|^2 \, dx + k^2 \int_D (x_d + h) \frac{\partial \operatorname{Re} n}{\partial x_d} |u|^2 \, dx + 2k^2 \int_D (x_d + h) \operatorname{Im} n \operatorname{Im} \left( u \frac{\partial}{\partial x_d} \bar{u} \right) \, dx \\ \leq (2k^2 h + 1) \|g\|_{L^2(\Omega^h)} \|u\|_{L^2(\Omega^h)} + 4h \|g\|_{L^2(\Omega^h)} \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}. \end{aligned} \quad (2.46)$$

Let us denote by  $n_r := \sup_{\tilde{D}} \partial \operatorname{Re} n / \partial x_d$ ,  $n_i := \sup_{\tilde{D}} \operatorname{Im} n$ . We first observe that

$$k^2 \int_{\tilde{D}} (x_d + h) \frac{\partial \operatorname{Re} n}{\partial x_d} |u|^2 \leq 2hk^2 n_r \|u\|_{L^2(\tilde{D})}^2 \leq \frac{2hn_r}{c_0} \|g\|_{L^2(\Omega^h)} \|u\|_{L^2(\Omega^h)}$$

and

$$\begin{aligned} 2k^2 \int_D (x_d + h) \operatorname{Im} n \operatorname{Im} \left( u \frac{\partial \bar{u}}{\partial x_d} \right) \leq 2hk^2 n_i \|u\|_{L^2(\tilde{D})} \left\| \frac{\partial \bar{u}}{\partial x_d} \right\|_{L^2(\Omega^h)} \\ \leq 2hk^2 n_i \|g\|_{L^2(\Omega^h)}^{1/2} \|u\|_{L^2(\Omega^h)}^{1/2} \left\| \frac{\partial \bar{u}}{\partial x_d} \right\|_{L^2(\Omega^h)}. \end{aligned} \quad (2.47)$$

Then we deduce from (2.46) that

$$\begin{aligned} \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}^2 \leq C_1 \|g\|_{L^2(\Omega^h)} \|u\|_{L^2(\Omega^h)} + C_2 \|g\|_{L^2(\Omega^h)}^{1/2} \|u\|_{L^2(\Omega^h)}^{1/2} \left\| \frac{\partial \bar{u}}{\partial x_d} \right\|_{L^2(\Omega^h)} \\ + C_3 \|g\|_{L^2(\Omega^h)} \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}, \end{aligned}$$

where,  $C_1 := k^2 h + 1/2 + hn_r/c_0$ ,  $C_2 := hk^2 n_i$  and  $C_3 := 2h$ . Combining with Corollary 2.46 we obtain

$$\left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}^2 \leq \beta_1 \|g\|_{L^2(\Omega^h)}^2 + \beta_2 \|g\|_{L^2(\Omega^h)} \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)} + \beta_3 \|g\|_{L^2(\Omega^h)}^{1/2} \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}^{3/2} \quad (2.48)$$

where,  $\beta_1 := CC_1$ ,  $\beta_2 := CC_1 + \sqrt{C}C_2 + C_3$ ,  $\beta_3 := \sqrt{C}C_2$ .

Let us set  $t := \|g\|_{L^2(\Omega^h)}^{1/2} \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}^{-1/2}$ . We observe that  $F(t) := \beta_1 t^4 + \beta_2 t^2 + \beta_3 t - 1$  is positive valued for all  $t \geq \theta$ , for some  $\theta > 0$ . Therefore, equation (2.48) implies that

$$\left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}^2 \leq \theta \|g\|_{L^2(\Omega^h)}. \quad (2.49)$$

The lemma is proved.  $\square$

We are now in position to prove the main theorem of this section on the well posedness of problem (2.4) or equivalently (2.6).

**Theorem 2.3.9.** *Assume that the assumptions of Lemma 2.3.8 hold. Then problem (2.6) admits a unique solution  $u \in H^1(\Omega^h)$  and there exists a positive constant  $C$  independent from  $u$  and  $g$  such that*

$$\|u\|_{H^1(\Omega^h)} \leq C \|g\|_{L^2(\Omega^h)}. \quad (2.50)$$

*Proof.* Let  $u \in H^1(\Omega^h)$  be satisfying (2.6). From Lemma 2.3.8 and Corollary 2.3.4 we immediately deduce that there exists  $\beta > 0$  such that

$$\|u\|_{L^2(\Omega^h)} \leq \beta \|g\|_{L^2(\Omega^h)}. \quad (2.51)$$

Taking the real part of variational equation (2.6) (with  $v = u$ ) we then get

$$\begin{aligned} \|u\|_{H^1(\Omega^h)}^2 &\leq (k^2 \sup_{\Omega^h} \operatorname{Re} n + 1) \|u\|_{L^2(\Omega^h)}^2 + \|g\|_{L^2(\Omega^h)} \|u\|_{L^2(\Omega^h)} \\ &\leq \left( \beta^2 k^2 \sup_{\Omega^h} \operatorname{Re} n + \beta^2 + \beta \right) \|g\|_{L^2(\Omega^h)}^2, \end{aligned}$$

which implies (2.50) with  $C := \sqrt{\beta^2 k^2 \sup_{\Omega^h} \operatorname{Re} n + \beta^2 + \beta}$ .

To prove the existence of the solution we here prove the inf-sup condition (2.7). Applying Lemma 2.2.2 we shall show that the a priori estimate (2.50) implies the existence of a constant  $C > 0$  independent from  $F \in H^1(\Omega^h)'$  such that if  $u \in H^1(\Omega^h)$  satisfying (2.11) then

$$\|u\|_{H^1(\Omega^h)} \leq C \|F\|_{H^1(\Omega^h)'}$$

Let us define  $a_0 : H^1(\Omega^h) \times H^1(\Omega^h)$  such that

$$a_0(u, v) = \int_{\Omega^h} (\nabla u \cdot \nabla \bar{v} + u \bar{v}) dx - \langle T^+ u, v \rangle - \langle T^- u, v \rangle$$

then  $\operatorname{Re} a_0(u, u) \geq \|u\|_{H^1(\Omega^h)}$ . Therefore for all  $F \in H^1(\Omega^h)'$ , the equation

$$a_0(u, v) = F(v), \quad v \in H^1(\Omega^h)$$

admit unique solution. We denote this solution by  $u_0$ . Moreover,

$$\|u_0\|_{H^1(\Omega^h)} \leq \|F\|_{H^1(\Omega^h)'}$$

Let define  $w := u - u_0$ , from problem (2.6) can be rewritten as:

$$a(w, v) - ((k^2 n + 1)u_0, v) = 0, \quad \forall v \in H^1(\Omega^h). \quad (2.52)$$

Hence  $w$  satisfies (2.6) with  $g = -(k^2 n + 1)u_0$ . We then have

$$\|w\|_{H^1(\Omega^h)} \leq C \left(1 + k^2 \|n\|_{L^\infty(\Omega^h)}\right) \|u_0\|_{L^2(\Omega^h)} \leq C \left(1 + k^2 \|n\|_{L^\infty(\Omega^h)}\right) \|F\|_{H^1(\Omega^h)'}. \quad (2.53)$$

Therefore

$$\|u\|_{H^1(\Omega^h)} \leq \left(1 + C(1 + k^2 \|n\|_{L^\infty(\Omega^h)})\right) \|F\|_{H^1(\Omega^h)'}, \quad (2.54)$$

which proves existence of solutions and finishes the proof.  $\square$

We remark from the proof that the result of Theorem 2.3.9 holds true if  $g \in H^1(\Omega^h)'$  and the a priori estimate reads in this case

$$\|u\|_{H^1(\Omega^h)} \leq C \|g\|_{H^1(\Omega^h)'}. \quad (2.55)$$

## 2.4 Scattering problem from periodic layers

We now consider a special case where the refractive index is given as a periodic function with respect to the first  $d - 1$  variables. To distinguish this case from the more general one treated before, we denote by  $n_p \in L^\infty(\mathbb{R}^d)$  the periodic refractive index of period  $L := (L_1, \dots, L_{d-1}) \in \mathbb{R}^{d-1}$ ,  $L_j > 0$ ,  $j = 1, \dots, d - 1$ .  $n_p$  and consider the Helmholtz equation

$$\Delta u + k^2 n_p u = g \quad \text{in } \mathbb{R}^d \quad (2.56)$$

with the source term  $g \in L^2(\mathbb{R}^d)$  and such that  $u$  satisfies the Rayleigh radiation condition with respect to the  $x_d$  variable. For a point  $x \in \mathbb{R}^d$  we shall use the decomposition  $x = (\bar{x}, x_d)$  with  $x_d \in \mathbb{R}$  and  $\bar{x} = (x_1, \dots, x_{d-1}) \in \mathbb{R}^{d-1}$ . For  $m = (m_1, \dots, m_{d-1}) \in \mathbb{Z}^{d-1}$  we denote by

$$\Omega_m := \left[ \left(m - \frac{1}{2}\right)L, \left(m + \frac{1}{2}\right)L \right] \times \mathbb{R},$$

where we use the notation  $\llbracket a, b \rrbracket := [a_1, b_1] \times \dots \times [a_{d-1}, b_{d-1}]$ . We shall also use the notation

$$\Omega^h := \mathbb{R}^{d-1} \times ] - h, h[ \quad \text{and} \quad \Omega_m^h := \Omega^h \cap \Omega_m.$$

**Definition 2.4.1.** *A function  $u$  is called quasi-periodic with parameter  $\xi = (\xi_1, \dots, \xi_{d-1})$  and period  $L = (L_1, \dots, L_{d-1})$ , with respect to the first  $d - 1$  variables (briefly denoted as  $\xi$ -quasi-periodic with period  $L$ ) if:*

$$u(\bar{x} + (mL), x_d) = e^{i\xi \cdot (mL)} u(\bar{x}, x_d), \quad \forall m \in \mathbb{Z}^{d-1}.$$

Using the Floquet-Bloch transform with respect to transverse coordinates  $\xi \in \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket$ , which will be detailed in Chapter 3, we are led to consider the problem for  $\xi$ -quasi-periodic functions  $u_\xi$  verifying

$$\Delta u_\xi + k^2 n_p u_\xi = g_\xi \quad \text{in } \mathbb{R}^d \quad (2.57)$$

where  $g_\xi \in L^2_{\xi, L}(\mathbb{R}^d)$  and

$$L^2_{\xi, L}(\mathbb{R}^d) := \{u \in L^2(\mathbb{R}^d); u \text{ is } \xi\text{-quasi-periodic with period } L\}$$

denote the set  $L^2$   $\xi$ -quasi-periodic functions with period  $L$ . The function  $u_\xi$  needs also to satisfy a radiation condition that will be specified later.

**Remark 2.4.2.** *Another way to introduce the  $\xi$ -quasi-periodic problem is to consider the scattering problem for plane waves. Consider an incident plane wave*

$$u^i(x) = e^{ikd \cdot x}$$

with  $|d| = 1$ . The scattering problem for  $u^i$  can be formulated as seeking a total field  $u$  solution of

$$\Delta u + k^2 n_p u = 0 \quad \text{in } \mathbb{R}^d$$

such that  $u = u^i + u^s$  and  $u^s$  satisfies the Rayleigh radiation condition. The problem for  $u^s$  can be written as

$$\Delta u^s + k^2 n_p u^s = k^2 (1 - n_p) u^i \quad \text{in } \mathbb{R}^d.$$

Since the function  $(1 - n_p) u^i$  is  $\xi$ -quasi-periodic functions with period  $L$  and  $\xi = kd$ , we are led to consider a problem of the form (2.57).

Indeed considering problem (2.57) in  $\mathbb{R}^d$  or in  $\Omega_0$  is equivalent (see Fig. 2.3). In the following of this section we shall not distinguish  $u_\xi$  and  $D$  from their restrictions to  $\Omega_0$  and set  $\Omega_{0, \pm}^h := \Omega_\pm^h \cap \Omega_0$ . We shall reformulate and consider problem (2.57) in  $\Omega_0^h$  using a DtN operator similar to the one introduced in the previous section. To do so one needs to assume that  $g_\xi$  has also support in  $\Omega^h$ .

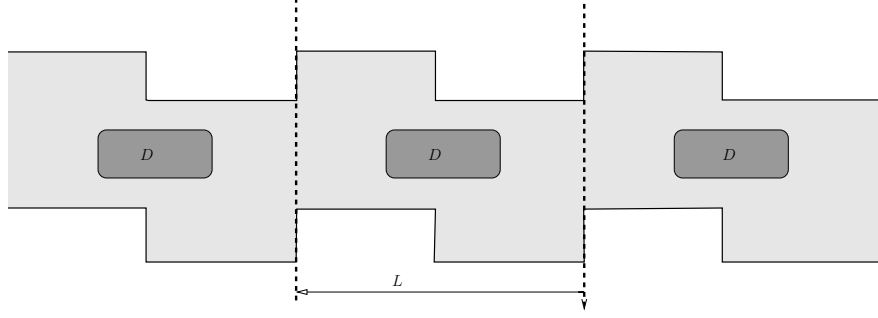


Figure 2.3: Sketch of the geometry in the case of periodic layers

### 2.4.1 The DtN operator for $\xi$ -quasi-periodic problems

A function  $u_\xi \in L^2_{\xi,L}(\mathbb{R}^d)$  can be expanded in the form

$$u_\xi(\bar{x}, x_d) = \sum_{j \in \mathbb{Z}^{d-1}} a_j(x_d) e^{i(\frac{2\pi}{L}j + \xi) \cdot \bar{x}}, \quad x \in \Omega_0, \quad (2.58)$$

where  $a_j(x_d)$  depends only on  $x_d$  and is defined by

$$a_j(x_d) := \frac{1}{\llbracket L \rrbracket} \int_{[-\frac{L}{2}, \frac{L}{2}]} u_\xi(\bar{x}, x_d) e^{-i(\xi + \frac{2\pi}{L}j) \cdot \bar{x}} d\bar{x}, \quad j \in \mathbb{Z}^{d-1}. \quad (2.59)$$

Since  $n = 1$  outside the domain  $\Omega^h$ , then  $u_\xi$  verifies  $\Delta u_\xi + k^2 u_\xi = 0$  in  $\Omega^h_\pm$ . Therefore, expansion (2.58) gives us

$$\sum_{j \in \mathbb{Z}^{d-1}} \left( a_j''(x_d) + \left( k^2 - \left( \frac{2\pi}{L}j + \xi \right)^2 \right) a_j(x_d) \right) e^{i(\frac{2\pi}{L}j + \xi) \cdot \bar{x}} = 0. \quad (2.60)$$

Assuming a time dependence of the form  $e^{-i\omega t}$  and retaining only upward (resp. downward) propagating modes in  $\Omega^h_+$  (resp.  $\Omega^h_-$ ) or exponentially decaying modes yields a representation of  $u_\xi$  in the form

$$\begin{cases} u_\xi(x) = \sum_{j \in \mathbb{Z}^{d-1}} \hat{u}_j^+ e^{i\alpha_\xi(j) \cdot \bar{x} + i\beta_\xi(j)(x_d - h)}, \quad \forall x_d \geq h, \\ u_\xi(x) = \sum_{j \in \mathbb{Z}^{d-1}} \hat{u}_j^- e^{i\alpha_\xi(j) \cdot \bar{x} - i\beta_\xi(j)(x_d + h)}, \quad \forall x_d \leq -h, \end{cases} \quad (2.61)$$

where

$$\alpha_\xi(j) := \xi + \frac{2\pi}{L}j = \left( \xi_1 + \frac{2\pi}{L_1}j_1, \dots, \xi_{d-1} + \frac{2\pi}{L_{d-1}}j_{d-1} \right), \quad (2.62)$$

$$\beta_\xi(j) := \sqrt{k^2 - \|\alpha_\xi(j)\|^2}, \quad \text{Im } \beta_\xi(j) \geq 0, \quad (2.63)$$

and where the coefficients  $\hat{u}_j^\pm$  are the Rayleigh coefficients, defined by

$$\hat{u}_j^\pm = \frac{1}{\llbracket L \rrbracket} \int_{[-\frac{L}{2}, \frac{L}{2}]} u_\xi(\bar{x}, \pm h) e^{-i\alpha_\xi(j) \cdot \bar{x}} d\bar{x}. \quad (2.64)$$

From the representation of  $u_\xi$  as in (2.61) we have that

$$\frac{\partial u}{\partial x_d} \Big|_{\Gamma^{\pm h}} = i \sum_{j \in \mathbb{Z}^{d-1}} \beta_\xi(j) \hat{u}_j^\pm e^{i\alpha_\xi(j) \cdot \bar{x}} \quad (2.65)$$

Let us define  $H_\xi^s(\Gamma^{\pm h})$  as the completion of  $C^\infty$   $\xi$ -quasi-periodic functions of period  $L$  in  $\Gamma^{\pm h}$  with respect to the norm  $\|\cdot\|_{H_\xi^s(\Gamma^{\pm h})}$  defined by

$$\|\varphi\|_{H_\xi^s(\Gamma^{\pm h})} := \left( \sum_{j \in \mathbb{Z}^{d-1}} |k^2 + \|\alpha_\xi(j)\|^2|^s |\widehat{\varphi}_j^\pm|^2 \right)^{1/2} \quad (2.66)$$

where

$$\widehat{\varphi}_j^\pm := \frac{1}{\llbracket L \rrbracket} \int_{\llbracket -\frac{L}{2}, \frac{L}{2} \rrbracket} \varphi(\bar{x} \pm h) e^{-i\alpha_\xi(j) \cdot \bar{x}} d\bar{x}.$$

The DtN operators  $T_\xi^\pm : H_\xi^{1/2}(\Gamma^{\pm h}) \rightarrow H_\xi^{-1/2}(\Gamma^{\pm h})$  are then defined by

$$T_\xi^\pm(\varphi) = i \sum_{j \in \mathbb{Z}^{d-1}} \beta_\xi(j) \widehat{\varphi}_j^\pm e^{i\alpha_\xi(j) \cdot \bar{x}}.$$

Similarly to Lemma 2.2.1 we have the following properties of the operators  $T_\xi^\pm$

**Lemma 2.4.3.** *The operators  $T_\xi^\pm : H_\xi^{1/2}(\Gamma^{\pm h}) \rightarrow H_\xi^{-1/2}(\Gamma^{\pm h})$  are bounded with  $\|T_\xi^\pm\| \leq 1$  and*

$$\operatorname{Re} \langle T_\xi^\pm u, u \rangle \leq 0 \quad \text{and} \quad \operatorname{Im} \langle T_\xi^\pm u, u \rangle \geq 0, \quad \forall u \in H_\xi^{1/2}(\Gamma^{\pm h}). \quad (2.67)$$

Let us define for real numbers  $m$ ,

$$H_\xi^m(\Omega_0^h) := \{u = U|_{\Omega_0^h} \text{ for some function } U \in H_{loc}^m(\Omega^h) \text{ } \xi\text{-quasi periodic with period } L\}.$$

Then problem (2.57) subject to the Rayleigh radiation condition (2.61) can be stated in the domain  $\Omega_0^h$  as:  $u_\xi \in H_\xi^1(\Omega_0^h)$

$$\begin{aligned} \Delta u_\xi + k^2 n_p u_\xi &= g_\xi & \text{in } \Omega_0^h, \\ \frac{\partial u_\xi}{\partial x_d} &= T_\xi^\pm(u) & \text{on } \Gamma^{\pm h}. \end{aligned} \quad (2.68)$$

Problem (2.68) can be variationally written as: Find  $u_\xi \in H_\xi^1(\Omega_0^h)$  such that

$$\int_{\Omega_0^h} (\nabla u_\xi \cdot \nabla \bar{v} - k^2 n_p u_\xi \bar{v}) dx - \langle T_\xi^+(u_\xi), v \rangle - \langle T_\xi^-(u_\xi), v \rangle = - \int_{\Omega_0^h} g_\xi \bar{v} dx \quad (2.69)$$

for all  $v \in H_\xi^1(\Omega_0^h)$ . Here again,  $\langle \cdot, \cdot \rangle$  denote the sesquilinear duality products  $H_\xi^{-1/2}(\Gamma^{\pm h})$ - $H_\xi^{1/2}(\Gamma^{\pm h})$ . Problem (2.69) is of Fredholm type since the sesquilinear form

$$\mathcal{A}_\xi(u_\xi, v) := \int_{\Omega_0^h} \nabla u_\xi \nabla \bar{v} - k^2 n_p u_\xi \bar{v} dx - \langle T_\xi^+(u_\xi), v \rangle - \langle T_\xi^-(u_\xi), v \rangle$$

is continuous on  $H_\xi^1(\Omega_0^h)$  and satisfies the Garding inequality

$$|\mathcal{A}_\xi(u_\xi, u_\xi)| \geq \|u_\xi\|_{H_\xi^1(\Omega_0^h)}^2 - \int_{\Omega_0^h} (k^2 \operatorname{Re} n_p + 1) |u_\xi|^2 dx. \quad (2.70)$$

(using in particular Lemma 2.4.3).

### 2.4.2 Well-posedness of the problem

The uniqueness of solutions to the problem holds if the refractive index  $n_p$  has a positive imaginary part in a sub-domain  $\mathcal{O} \subset \Omega_0^h$ . This is made precise in the following theorem.

**Theorem 2.4.4.** *Assume that the refractive index  $n_p \in L^\infty(\Omega_0^h)$  and that  $\text{Im}(n_p) > 0$  in a sub-domain  $\mathcal{O} \subset \Omega_0^h$ . Then problem (2.69) is well-posed.*

*Proof.* Since problem (2.69) is of Fredholm type, it is sufficient to prove the uniqueness of solution. Assume that  $\mathcal{A}_\xi(u, v) = 0$ , for all  $v \in H_\xi^1(\Omega_0^h)$ , i.e.,

$$\int_{\Omega^h} (\nabla u_\xi \cdot \nabla \bar{v} - k^2 n_p u_\xi \bar{v}) \, dx - \langle T_\xi^+(u_\xi), v \rangle - \langle T_\xi^-(u_\xi), v \rangle = 0 \quad (2.71)$$

Considering the imaginary part of equation (2.71) with  $v$  is replaced by  $u$ , we obtain thanks to Lemma 2.4.3, that

$$\int_{\mathcal{O}} \text{Im} n_p |u_\xi|^2 \, dx \leq 0. \quad (2.72)$$

This implies

$$u_\xi = 0 \quad \text{in } \mathcal{O}.$$

By unique continuation principle (since  $u_\xi$  verifies (2.68)) we arrive that  $u_\xi = 0$  in  $\Omega_0^h$ , which prove the theorem.  $\square$

We remark that if  $n_p$  satisfies the assumption of Theorem 2.4.4, then the solution  $u_\xi$  of (2.69) satisfies

$$\|u_\xi\|_{H^1(\Omega_0^h)} \leq C(\xi) \|g_\xi\|_{L^2(\Omega_0^h)} \quad (2.73)$$

where  $C(\xi)$  is a positive constant independent from  $u_\xi$  and  $g_\xi$  but that may depend on  $\xi$ . Following the same approach as for Theorem 2.3.9 one can prove an a priori estimate for the solution with a constant independent from  $\xi$ . Since the proof is exactly the same as for Theorem 2.3.9 we skip it here and only state the theorem.

**Theorem 2.4.5.** *Assume that  $n = n_p$  is  $L$  periodic with respect to the first  $d - 1$  variables and satisfies the assumptions of Lemma 2.3.8. Let  $u_\xi \in H_\xi^1(\Omega_0^h)$  be the solution of (2.69). Then there exists a constant  $C$  independent from  $\xi$ ,  $u_\xi$  and  $g_\xi$  such that*

$$\|u_\xi\|_{H^1(\Omega_0^h)} \leq C \|g_\xi\|_{L^2(\Omega_0^h)}.$$

Since the volume integral formulation of the  $\xi$ -quasi-periodic problem will occur in a number of occasions in next chapters, we shall present it here for the sake of completeness. Let  $G_\xi(\cdot) \in L_{\xi, L}^2(\mathbb{R}^d)$  be the  $\xi$ -quasi-periodic Green function of the Helmholtz equation, i.e. satisfying,

$$\Delta G_\xi + k^2 G_\xi(\cdot) = -\delta_0 \quad \text{in } \Omega_0 \quad (2.74)$$

and  $G_\xi$  verifies the Rayleigh radiation condition (2.61). If we assume that  $\beta_\xi(j) \neq 0$  for all  $j \in \mathbb{Z}^{d-1}$  (which is generically the case except for special values of the wave number  $k$ , corresponding to the so-called Wood anomalies), then  $G_\xi$  can be expressed as (see [69]),

$$G_\xi(x) = \frac{i}{2\llbracket L \rrbracket} \sum_{j \in \mathbb{Z}^{d-1}} \frac{1}{\beta_\xi(j)} \exp(i\alpha_\xi(j) \cdot \bar{x} + i\beta_\xi(j)|x_d|), \quad (\bar{x}, x_d) \in \mathbb{R}^d, \quad (2.75)$$

We now consider the volume potential  $V_\xi$  on  $L^2(\Omega_0^h)$  with kernel  $G_\xi$  defined by

$$V_\xi g(x) := \int_{\Omega_0^h} G_\xi(x-y)g(y)dy, \quad x \in \mathbb{R}^d. \quad (2.76)$$

Then we have the following classical property that can be found for instance in [85, 46].

**Lemma 2.4.6.** *The volume potential  $V_\xi$  is a linear bounded operator from  $L^2(\Omega_0^h)$  into  $H_\xi^2(\Omega_0)$ . Moreover, for all  $g \in L^2(\Omega_0^h)$ , the potential  $w := V_\xi g \in H_\xi^2(\Omega_0)$  and is the unique solution to (2.69) (or equivalently (2.68)) with  $g_\xi = g$  and  $n_p = 1$ .*

Thus, from Lemma 3.3.1, we see that  $u_\xi \in H_\xi^2(\Omega_0^h)$  satisfies (2.69) if and only if  $u_\xi \in L^2(\Omega_0^h)$  and satisfies

$$u_\xi = k^2 V_\xi((n_p - 1)u_\xi) - V_\xi g_\xi \quad \text{in } L^2(\Omega_0^h). \quad (2.77)$$

We also refer to the next chapter for more details on how to obtain this equivalence. Equation (3.17) is known as the Lippman-Schwinger integral equation. This type of formulation of the problem will be used to derive a discretization scheme for the locally perturbed periodic problem.

## 2.5 The case of infinite layers with local perturbations

We shall show in this section that the well-posedness of problem (2.1) still holds if the infinite layer has a local perturbation that does not satisfies Assumption 2.3.1. We first introduce the following assumption:

**Assumption 2.5.1.** *Assume that  $n_p \in L^\infty(\mathbb{R}^d)$  such that the problem*

$$\begin{aligned} \Delta u + k^2 n_p u &= g \quad \text{in } \Omega^h \\ \frac{\partial u}{\partial x_d} \Big|_{\Gamma_{\pm h}} &= T^\pm u \quad \text{on } \Gamma_{\pm h} \end{aligned} \quad (2.78)$$

for all  $g \in H^{-1}(\mathbb{R}^d)$  with  $\text{supp}(g) \in \Omega^h$  is well-posed. The index of refraction  $n$  is such that

$$n \neq n_p \quad \text{in } \tilde{\mathcal{O}}$$

where  $\tilde{\mathcal{O}}$  is a compact bounded domain in  $\Omega^h$ .

With Assumption 2.5.1, we here prove that Problem (2.4) is well-posed.

**Theorem 2.5.2.** *Assume that Assumption 2.5.1 holds and  $\text{Im } n > 0$  in an open domain  $\mathcal{O}$  such that  $|\mathcal{O}| > 0$ . Then problem (2.4) is well-posed.*

*Proof.* The uniqueness of solution can be obtained similarly as in Theorem 2.4.4. Let us denote by  $\langle \cdot, \cdot \rangle$  the  $H^1(\Omega^h)$  scalar product. We now introduce, using the Riesz representation theorem the operators  $\mathcal{A}, \mathcal{A}_p, \mathcal{T} : H^1(\Omega^h) \rightarrow H^1(\Omega^h)$  such that

$$\begin{aligned} \langle \mathcal{A}u, v \rangle &= \int_{\Omega^h} (\nabla u \cdot \nabla \bar{v} - k^2 n u \bar{v}) dx - \langle T^+ u, v \rangle - \langle T^- u, v \rangle \\ \langle \mathcal{A}_p u, v \rangle &= \int_{\Omega^h} (\nabla u \cdot \nabla \bar{v} - k^2 n_p u \bar{v}) dx - \langle T^+ u, v \rangle - \langle T^- u, v \rangle \\ \langle \mathcal{T}u, v \rangle &= k^2 \int_{\Omega^h} (n_p - n) u \bar{v} dx \end{aligned} \quad (2.79)$$

for all  $u$  and  $v$  in  $H^1(\Omega^h)$  and consider  $\mathcal{G} \in H^1(\Omega^h)$  defined by

$$\langle \mathcal{G}, v \rangle = -k^2 \int_{\Omega^h} g \bar{v} \, dx$$

for all  $v$  in  $H^1(\Omega^h)$  where the integral should be understood as a duality product. We first deduce from the assumption that  $\tilde{\mathcal{O}}$  is compact and from the compact embedding of  $H^1(\tilde{\mathcal{O}})$  into  $L^2(\tilde{\mathcal{O}})$  that the operator  $\mathcal{T}$  is compact. From the variational writing of (2.4), one has to equivalently solve

$$\mathcal{A}_p u + \mathcal{T}u = \mathcal{G} \text{ in } H^1(\Omega^h). \quad (2.80)$$

The well-posedness of problem (2.78) implies that  $\mathcal{A}_p$  is invertible and  $\mathcal{A}_p^{-1} : H^1(\Omega^h) \rightarrow H^1(\Omega^h)$  is bounded. Hence equation (2.80) is equivalent to

$$u + \mathcal{A}_p^{-1} \mathcal{T}u = \mathcal{A}_p^{-1} \mathcal{G} \text{ in } H^1(\Omega^h). \quad (2.81)$$

Since  $\mathcal{A}_p^{-1} \mathcal{T} : H^1(\Omega^h) \rightarrow H^1(\Omega^h)$  is a compact operator and that the solution to (2.80) is unique, we then obtain, using the Fredholm alternative the existence of solution  $u$  to problem (2.81) and hence to problem (2.4). This proves the theorem.  $\square$

In the next Chapter we will be concerned with numerical schemes to solve this problem in the special case where  $n_p$  is periodic.

## 2.6 On the case of the scattering problem for half space with Dirichlet boundary conditions

Using the same technique as above we can also treat the Helmholtz equation in the case of half-space (see also Fig. 2.4)

$$\mathbb{R}_+^d := \mathbb{R}^{d-1} \times [0, +\infty),$$

with Dirichlet boundary conditions on the boundary

$$\Gamma_0 := \mathbb{R}^{d-1} \times \{0\}.$$

In this case, the assumption on the imaginary part of the refractive index can be relaxed thanks to a Poincaré inequality (see 2.6.2). Let us consider the problem

$$\begin{cases} \Delta u + k^2 n u = g & \text{in } \mathbb{R}_+^d \\ u = 0 & \text{on } \Gamma_0, \end{cases} \quad (2.82)$$

with a source term  $g \in L^2(\mathbb{R}_+^d)$  and  $u$  satisfies the Rayleigh radiation condition with respect to direction  $x_d$  (i.e.  $u$  has the Fourier representation (2.2) in  $\Omega_h^+$ ).

In the following we change the notation of  $\Omega^h$  as

$$\Omega^h := \mathbb{R}^{d-1} \times [0, h]$$

but keep the same notations as above for  $\Gamma_h$  and  $\Omega_h^+$ .



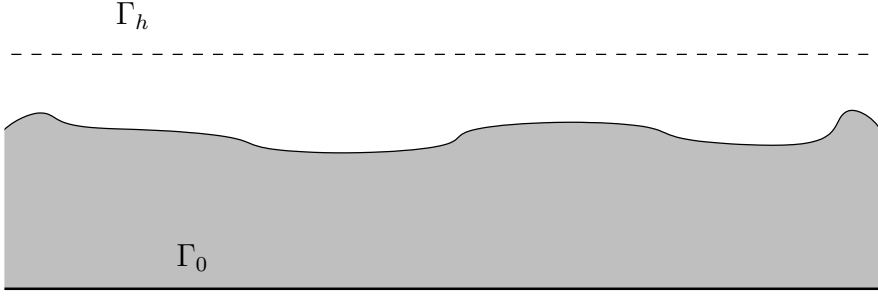


Figure 2.4: Sketch of the configuration for the half-space problem

### 2.6.1 Variational formulation of the problem

We here present the variational formulation of problem in  $H^1(\Omega^h)$  using the explicit formula of the Dirichlet to Neumann map. The problem is equivalently stated as the following:

$$\begin{aligned} \Delta u + k^2 n u &= g & \text{in } \mathbb{R}^d \\ u &= 0 & \text{on } \Gamma_0 \\ \frac{\partial u}{\partial x_d} \Big|_{\Gamma_h} &= T^+ u & \text{on } \Gamma_h \end{aligned} \quad (2.83)$$

where the source term  $g \in L^2(\Omega^h)$  is assumed to have support in  $\Omega^h$ . Let us denote by

$$\dot{H}^1(\Omega^h) := \{u \in H^1(\Omega^h), u|_{\Gamma_0} = 0\}. \quad (2.84)$$

Multiplying the first equation of (2.83) with  $\bar{v} \in H^1(\Omega^h)$  then integrating by parts we obtain the equivalent variational formulation of the scattering problem: Find  $u \in \dot{H}^1(\Omega^h)$  such that

$$\int_{\Omega^h} (\nabla u \cdot \nabla \bar{v} - k^2 n u \bar{v}) - \langle T^+ u, \bar{v} \rangle = - \int_{\Omega^h} g \bar{v}, \quad \forall v \in \dot{H}^1(\Omega^h). \quad (2.85)$$

Let us denote by  $\tilde{a}$  the sesquilinear form on  $\dot{H}^1(\Omega^h) \times \dot{H}^1(\Omega^h)$  defined by

$$\tilde{a}(u, v) = \int_{\Omega^h} (\nabla u \cdot \nabla \bar{v} - k^2 n u \bar{v}) - \langle T^+ u, \bar{v} \rangle \quad (2.86)$$

Similarly to the case of the full space, the sesquilinear form  $\tilde{a}$  is bounded. If we prove in addition that for  $u \in \dot{H}^1(\Omega^h)$  verifying

$$\tilde{a}(u, v) = F(v), \quad \forall v \in \dot{H}^1(\Omega^h)$$

for  $F \in H^1(\Omega^h)'$  the following a priori estimate

$$\|u\|_{H^1(\Omega^h)} \leq C \|F\|_{H^1(\Omega^h)'}$$

holds, where  $C$  is independent from  $F$ , then we can conclude on the existence of solution to (2.85). We already saw in the proof of Theorem 2.3.9 that for our scattering problem it is sufficient to prove the following a priori estimate

$$\|u\|_{H^1(\Omega^h)} \leq C \|g\|_{L^2(\Omega^h)}.$$

for a solution  $u$  verifying (2.85).

### 2.6.2 The Rellich identity for the case without absorption

We first make the following assumption

**Assumption 2.6.1.** *The index of refraction  $n \in L^\infty(\mathbb{R}^d)$  with non negative imaginary part and*

$$D := \text{supp}(n - 1) \subset \Omega^h.$$

Moreover,

$$\frac{\partial \text{Re } n}{\partial x_d} \geq 0, \quad \text{or} \quad n^* := \sup_{\Omega^h} \left| \frac{\partial \text{Re } n}{\partial x_d} \right| < 4/(k^2 h^3).$$

As indicated above, the main ingredient that allows us to relax the assumption on the imaginary part of  $n$  is the following (standard) Poincaré inequality.

**Lemma 2.6.2.** *The following Poincaré inequality holds*

$$\|u\|_{L^2(\Omega^h)} \leq \frac{h}{\sqrt{2}} \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}, \quad \forall u \in \dot{H}^1(\Omega^h). \quad (2.87)$$

*Proof.* For  $x = (\bar{x}, x_d) \in \Omega^h$ :

$$u(x) = u(\bar{x}, 0) + \int_0^{x_d} \frac{\partial u}{\partial x_d}(\bar{x}, s) \, ds = \int_0^{x_d} \frac{\partial u}{\partial x_d}(\bar{x}, s) \, ds. \quad (2.88)$$

Therefore,

$$|u(x)|^2 = \left| \int_0^{x_d} \frac{\partial u}{\partial x_d}(\bar{x}, s) \, ds \right|^2 \leq x_d \int_0^{x_d} \left| \frac{\partial u}{\partial x_d}(\bar{x}, s) \right|^2 \, ds \leq x_d \int_0^h \left| \frac{\partial u}{\partial x_d}(\bar{x}, s) \right|^2 \, ds. \quad (2.89)$$

Integrating inequality (2.89) with respect to  $x_d$  over  $[0, h]$  we get:

$$\int_0^h |u(x)|^2 \, dx_d \leq \int_0^h \left( x_d \int_0^h \left| \frac{\partial u}{\partial x_d}(\bar{x}, s) \right|^2 \, ds \right) \, dx_d = \frac{h^2}{2} \int_0^h \left| \frac{\partial u}{\partial x_d}(\bar{x}, s) \right|^2 \, ds. \quad (2.90)$$

Integrating inequality (2.90) over  $\Omega^h$  we finally obtain:

$$\|u\|_{L^2(\Omega^h)} \leq \frac{h}{\sqrt{2}} \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}. \quad (2.91)$$

which proves the lemma.  $\square$

We are now in position to prove the main Rellich identity for the case of the half plane. This is a similar result to the one in Lemma 2.3.6.

**Lemma 2.6.3.** *Assume that  $u \in \dot{H}^1(\Omega^h)$  satisfies (2.85). Then  $u \in H^2(\Omega^h)$  and satisfies*

$$\begin{aligned} & \int_{\Omega^h} \left( 2 \left| \frac{\partial u}{\partial x_d} \right|^2 \, dx + k^2 x_d \frac{\partial \text{Re } n}{\partial x_d} |u|^2 + 2k^2 x_d \text{Im } n \text{Im} \left( u \frac{\partial \bar{u}}{\partial x_d} \right) \right) \, dx \\ & \quad + h \int_{\Gamma_h} \left( |\nabla u|^2 - 2 \left| \frac{\partial u}{\partial x_d} \right|^2 - k^2 |u|^2 \right) \, ds \\ & \quad = \text{Re} \int_{\Gamma_h} T^+(u) \bar{u} - \text{Re} \int_{\Omega^h} g \bar{u} - 2 \text{Re} \int_{\Omega^h} g x_d \frac{\partial \bar{u}}{\partial x_d} \end{aligned}$$

*Proof.* The fact that  $u \in H^2(\Omega^h)$  can be proved in a very similar way as in Lemma 2.3.5. We therefore skip the proof here. For the Rellich identity we shall follow the same steps as in the proof of Lemma 2.3.6 but with some adaptations to the new Hypothesis.

Multiplying the original equation with  $x_d \frac{\partial \bar{u}}{\partial x_d}$  and then integrating over  $\Omega^h$  we get

$$\int_{\Omega^h} x_d \frac{\partial \bar{u}}{\partial x_d} (\Delta u + k^2 n u) \, dx = \int_{\Omega^h} x_d \frac{\partial \bar{u}}{\partial x_d} g \, dx \quad (2.92)$$

We first have

$$\begin{aligned} \int_{\Omega^h} x_d \frac{\partial \bar{u}}{\partial x_d} \Delta u &= \int_{\partial \Omega^h} x_d \frac{\partial u}{\partial x_d} \frac{\partial \bar{u}}{\partial x_d} - \int_{\Omega^h} \nabla u \cdot \nabla \left( x_d \frac{\partial \bar{u}}{\partial x_d} \right) \\ &= h \int_{\Gamma_h} \left| \frac{\partial u}{\partial x_d} \right|^2 - \int_{\Omega^h} \nabla u \cdot \left( \vec{e}_2 \frac{\partial \bar{u}}{\partial x_d} + x_d \frac{\partial}{\partial x_d} \nabla \bar{u} \right) \end{aligned}$$

Thus,

$$\begin{aligned} I := \int_{\Omega^h} x_d \frac{\partial \bar{u}}{\partial x_d} (\Delta u + k^2 n u) &= h \int_{\Gamma_h} \left| \frac{\partial u}{\partial x_d} \right|^2 - \int_{\Omega^h} \left( \left| \frac{\partial u}{\partial x_d} \right|^2 + x_d \nabla u \cdot \frac{\partial}{\partial x_d} \nabla u \right) \, dx \\ &\quad + \int_{\Omega^h} k^2 x_d \frac{\partial \bar{u}}{\partial x_d} n u \, dx. \end{aligned}$$

Taking the real part of  $I$  yields

$$\begin{aligned} 2\operatorname{Re}(I) &= 2h \int_{\Gamma_h} \left| \frac{\partial u}{\partial x_d} \right|^2 - \int_{\Omega^h} \left( 2 \left| \frac{\partial u}{\partial x_d} \right|^2 + x_d \frac{\partial}{\partial x_d} |\nabla u|^2 \right) + k^2 \int_{\Omega^h} \operatorname{Re} n x_d \frac{\partial}{\partial x_d} |u|^2 \\ &\quad - 2k^2 \int_{\Omega^h} x_d \operatorname{Im} n \operatorname{Im} \left( u \frac{\partial \bar{u}}{\partial x_d} \right) \\ &= 2h \int_{\Gamma_h} \left| \frac{\partial u}{\partial x_d} \right|^2 - \int_{\Omega^h} 2 \left| \frac{\partial u}{\partial x_d} \right|^2 \, dx - h \int_{\Gamma_h} |\nabla u|^2 + \int_{\Omega^h} |\nabla u|^2 + k^2 h \int_{\Gamma_h} |u|^2 \\ &\quad - k^2 \int_{\Omega^h} \left( \operatorname{Re} n + x_d \frac{\partial \operatorname{Re} n}{\partial x_d} \right) |u|^2 - 2k^2 \int_{\Omega^h} x_d \operatorname{Im} n \operatorname{Im} \left( u \frac{\partial \bar{u}}{\partial x_d} \right) \\ &= \int_{\Omega^h} \left( |\nabla u|^2 - 2 \left| \frac{\partial u}{\partial x_d} \right|^2 - k^2 \operatorname{Re} n |u|^2 - k^2 x_d \frac{\partial \operatorname{Re} n}{\partial x_d} |u|^2 \right) \, dx \\ &\quad - h \int_{\Gamma_h} \left( |\nabla u|^2 - 2 \left| \frac{\partial u}{\partial x_d} \right|^2 - k^2 |u|^2 \right) - 2k^2 \int_{\Omega^h} x_d \operatorname{Im} n \operatorname{Im} \left( u \frac{\partial \bar{u}}{\partial x_d} \right) \end{aligned}$$

Taking the real part of equation (2.85) with  $v$  is replaced by  $u$  implies

$$\int_{\Omega^h} |\nabla u|^2 - k^2 \operatorname{Re} n |u|^2 = \operatorname{Re} \int_{\Gamma_h} T^+(u) \bar{u} - \operatorname{Re} \int_{\Omega^h} g \bar{u}. \quad (2.93)$$

Substituting in the expression of  $\operatorname{Re}(I)$  yields

$$\begin{aligned} \int_{\Omega^h} \left( 2 \left| \frac{\partial u}{\partial x_d} \right|^2 \, dx + k^2 x_d \frac{\partial \operatorname{Re} n}{\partial x_d} |u|^2 + 2k^2 x_d \operatorname{Im} n \operatorname{Im} \left( u \frac{\partial \bar{u}}{\partial x_d} \right) \right) \, dx \\ + h \int_{\Gamma_h} \left( |\nabla u|^2 - 2 \left| \frac{\partial u}{\partial x_d} \right|^2 - k^2 |u|^2 \right) \, ds \\ = \operatorname{Re} \int_{\Gamma_h} T^+(u) \bar{u} - \operatorname{Re} \int_{\Omega^h} g \bar{u} - 2 \operatorname{Re} \int_{\Omega^h} g x_d \frac{\partial \bar{u}}{\partial x_d}, \end{aligned}$$

which proves the lemma.  $\square$

The following result can be seen as a straightforward adaptation of Lemma 2.3.7.

**Lemma 2.6.4.** *Assume that Hypothesis 2.6.1 holds. Let  $u \in \dot{H}^1(\Omega^h)$  satisfying (2.85) with  $g \in L^2(\Omega^h)$ . Then*

$$\int_{\Gamma_h} \left( -|\nabla u|^2 + 2 \left| \frac{\partial u}{\partial \nu} \right|^2 + k^2 |u|^2 \right) ds \leq 2k \operatorname{Im} \langle T^+ u, u \rangle. \quad (2.94)$$

Similarly to Lemma 2.3.5 and Lemma 2.40 we can exploit the identity of Lemma 2.6.3 to obtain the desired continuity result for solutions to (2.2).

**Lemma 2.6.5.** *Assume that Hypothesis 2.6.1 holds. Then there exists a positive constant  $C$  such that for all  $u \in \dot{H}^1(\Omega^h)$  satisfying (2.85) and  $g \in L^2(\Omega^h)$*

$$\left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)} \leq C \|g\|_{L^2(\Omega^h)}. \quad (2.95)$$

*Proof.* Combining Lemmas 2.6.3 with the fact that  $\operatorname{Im} n > 0$  we obtain

$$\begin{aligned} & 2 \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}^2 + k^2 \int_{\Omega^h} x_d \frac{\partial \operatorname{Re} n}{\partial x_d} |u|^2 dx + 2k^2 \int_D \operatorname{Im} n \operatorname{Im} \left( u \frac{\partial \bar{u}}{\partial x_d} \right) dx \\ &= h \int_{\Gamma_h} \left( -|\nabla u|^2 + 2 \left| \frac{\partial u}{\partial x_d} \right|^2 + k^2 n |u|^2 \right) + \operatorname{Re} \int_{\Gamma_h} T^+(u) \bar{u} - \operatorname{Re} \int_{\Omega^h} g \bar{u} - 2 \operatorname{Re} \int_{\Omega^h} x_d \frac{\partial \bar{u}}{\partial x_d} g \\ &\leq 2hk \operatorname{Im} \int_{\Omega^h} g \bar{u} - \operatorname{Re} \int_{\Omega^h} g \bar{u} - 2 \operatorname{Re} \int_{\Omega^h} x_d \frac{\partial \bar{u}}{\partial x_d} g \end{aligned}$$

We observe that

$$\begin{aligned} & 2hk \operatorname{Im} \int_{\Omega^h} g \bar{u} - \operatorname{Re} \int_{\Omega^h} g \bar{u} - 2 \operatorname{Re} \int_{\Omega^h} x_d \frac{\partial \bar{u}}{\partial x_d} g \\ &\leq (2kh + 1) \|g\|_{L^2(\Omega^h)} \|u\|_{L^2(\Omega^h)} + 2h \|g\|_{L^2(\Omega^h)} \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)} \\ &\leq \left( \sqrt{2}kh^2 + \frac{2\sqrt{2}+1}{\sqrt{2}}h \right) \|g\|_{L^2(\Omega^h)} \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}. \end{aligned}$$

Taking the imaginary part of equation (2.85), we deduce  $\int_D \operatorname{Im}(n) |u|^2 dx \leq \|g\|_{\Omega_h} \|u\|_{\Omega^h}$ . Thus,

$$\begin{aligned} \left| 2k^2 \int_D \operatorname{Im} n \operatorname{Im} \left( u \frac{\partial \bar{u}}{\partial x_d} \right) dx \right| &\leq 2k^2 \left( \int_D \operatorname{Im} n |u|^2 dx \right)^{1/2} \left( \int_D \operatorname{Im} n \left| \frac{\partial u}{\partial x_d} \right|^2 dx \right)^{1/2} \\ &\leq 2k^2 n_* \|g\|_{\Omega_h}^{1/2} \|u\|_{\Omega^h}^{1/2} \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)} \leq \sqrt{2}hk^2 n_* \|g\|_{\Omega_h}^{1/2} \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}^{3/2}, \quad (2.96) \end{aligned}$$

where,  $n_* = \sup_D \operatorname{Im} n$ . Therefore,

$$\begin{aligned} 2 \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}^2 + k^2 \int_{\Omega^h} x_d \frac{\partial \operatorname{Re} n}{\partial x_d} |u|^2 dx &\leq \left( \sqrt{2}kh^2 + \frac{2\sqrt{2}+1}{\sqrt{2}}h \right) \|g\|_{L^2(\Omega^h)} \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)} \\ &\quad + \sqrt{2}hk^2 n_* \|g\|_{\Omega_h}^{1/2} \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}^{3/2} \quad (2.97) \end{aligned}$$

We now observe that if  $n^* := \sup_{\Omega^h} |\partial \operatorname{Re} n / \partial x_d| < 4/(k^2 h^3)$  then

$$\left| k^2 \int_{\Omega^h} x_d \frac{\partial \operatorname{Re} n}{\partial x_d} |u|^2 dx \right| \leq k^2 h n^* \|u\|_{L^2(\Omega^h)}^2 \leq \frac{k^2 h^3}{2} n^* \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}^2. \quad (2.98)$$

From the three previous equations (2.96), (2.97) and (2.98) we have that

$$\left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)} \leq C_1 \|g\|_{L^2(\Omega^h)} + C_2 \|g\|_{L^2(\Omega^h)}^{1/2} \left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)}^{1/2}, \quad (2.99)$$

where  $C_1 := (\sqrt{2}kh^2 + \frac{2\sqrt{2+1}}{\sqrt{2}}h)/(2 - k^2h^3/2)$ ,  $C_2 := \sqrt{2}k^2hn_*/(2 - k^2h^3/2)$ ,  $n_* = \sup_D \operatorname{Im} n$ .

In the case  $\partial \operatorname{Re} n / \partial x_d \geq 0$  we immediately obtain (2.99) from (2.97) with  $C_1 = (\sqrt{2}kh^2 + \frac{2\sqrt{2+1}}{\sqrt{2}}h)/2$ ,  $C_2 = \sqrt{2}k^2hn_*/2$ . Finally, there exists  $C > 0$  such that

$$\left\| \frac{\partial u}{\partial x_d} \right\|_{L^2(\Omega^h)} \leq C \|g\|_{L^2(\Omega^h)}, \quad (2.100)$$

which proves the lemma.  $\square$

Combining Lemma 2.6.2 with Lemma 2.6.5 we immediately deduce that

$$\|u\|_{L^2(\Omega^h)} \leq C \|g\|_{L^2(\Omega^h)}. \quad (2.101)$$

which help proving the following final result of this chapter:

**Theorem 2.6.6.** *Assume that Hypothesis 2.6.1 holds. Then problem (2.85) admits a unique solution  $u \in \dot{H}^1(\Omega^h)$  and there exists a positive constant  $C$  independent from  $u$  and  $g$  such that*

$$\|u\|_{\dot{H}^1(\Omega^h)} \leq \theta \|g\|_{L^2(\Omega^h)}. \quad (2.102)$$

*Proof.* Assume that  $u \in \dot{H}^1(\Omega^h)$  and satisfies (2.85). Taking the imaginary part of the variational equation (2.85) with  $v$  is replaced by  $u$  and using the property  $\operatorname{Re} \langle T^+ u, u \rangle \leq 0$ , we get

$$\|u\|_{\dot{H}^1(\Omega^h)}^2 \leq (k^2 \sup_{\Omega^h} \operatorname{Re} n + 1) \|u\|_{L^2(\Omega^h)}^2 + \|g\|_{L^2(\Omega^h)} \|u\|_{L^2(\Omega^h)} \leq (k^2 \sup_{\Omega^h} \operatorname{Re} n + C) \|g\|_{L^2(\Omega^h)}^2,$$

which implies

$$\|u\|_{\dot{H}^1(\Omega^h)} \leq \theta \|g\|_{L^2(\Omega^h)}$$

with  $\theta := \sqrt{k^2 \sup_{\Omega^h} \operatorname{Re} n + C}$ . This estimate ensures also the existence of solution as explained at the beginning of this section. The theorem is proved.  $\square$

# A Volume integral method for solving scattering problems from locally perturbed infinite periodic layers

---

## Abstract

This chapter is mainly extracted from the published article [60]. We design a numerical method to compute the scattered wave from locally perturbed periodic layers in  $\mathbb{R}^d$ ,  $d = 2, 3$ . Using the Floquet-Bloch transform in the periodicity direction we reformulate this scattering problem as an equivalent system of coupled volume integral equations. We then apply a spectral method to discretize the obtained system after periodization in the direction orthogonal to the periodicity directions of the medium. The convergence of this method is established only in the case of complex wave numbers with positive imaginary part. Validating numerical results are provided for  $d = 2$ . The only difference with respect to [60], in addition some minor corrections, is the last section dedicated to some numerical examples for real wave numbers that are not covered by the convergence theory but correspond with the configurations studied in Chapter 2.

## Contents

---

<b>3.1</b>	<b>Introduction</b>	<b>41</b>
<b>3.2</b>	<b>Setting of the problem</b>	<b>43</b>
3.2.1	Introduction of the problem and notation	43
3.2.2	Formulation of the problem using the Floquet-Bloch transform	44
<b>3.3</b>	<b>Volume integral formulation of the <math>\xi</math>-quasi-periodic problem</b>	<b>46</b>
3.3.1	Setting of the volume integral equation	46
3.3.2	Periodization of the integral equation	47
3.3.3	Spectral approximation of problem (3.19)	49
<b>3.4</b>	<b>Discretization of the problem and convergence analysis</b>	<b>54</b>
3.4.1	Discretization and convergence in the Floquet-Bloch variable	55
3.4.2	Discretization in the spatial variable	59
<b>3.5</b>	<b>Numerical Algorithm and Experiments</b>	<b>65</b>
3.5.1	Numerical examples for real wave numbers	69

---

## 3.1 Introduction

We are concerned in this chapter by the design of a numerical method to compute the solution of the scattering problem from locally perturbed infinite periodic layers. This problem is

encountered for instance in applications related to photonics or non destructive testing of gratings. Our work can be seen as an extension of the numerical method in [103, 85] where the solution for periodic media is computed based on quasi-periodicity of the incident wave. The principle of our method is similar to the method employed in [42, 44] and relies on the use of the Floquet-Bloch transform in the periodicity directions to transform the problem into coupled quasi-periodic problems. Our methodologies then differs in the way we formulate the problem and discretize it with respect to the space variable. While in [44] an finite element method is employed using a Dirichlet to Neumann map to bound the computation domain as introduced in [68, 54], we shall rely in the present work on a volume integral formulation of the problem. Discretizing the volume integral equation efficiently through the use of spectral method and FFT technique to evaluate the matrix-vector product has been introduced in [103] and studied in a number of papers [103, 84, 63, 64]. We here mainly rely on the numerical algorithm implemented in [85] for quasi-periodic problems but we shall consider TE modes. The TM mode can be treated in a similar manner using the adaptations proposed in [85]. As in [42], the first main difficulty in the convergence analysis is to establish convergence of the discretization of the problem with respect to the Floquet-Bloch variable. Using a trapezoidal rule to discretize the Floquet-Bloch transform induces a semi-discrete problem which is equivalent, up to a change in the source term, to the periodic problem with period  $ML$  where  $M$  is the number of discretization points and  $L$  is the periodicity length. The convergence (with respect to the parameter  $M$ ) then relies on the decay of the solution in the periodicity directions. The latter is indeed not guaranteed in general and usually requires some special assumptions on the material properties, such as strict monotonicity in the direction orthogonal to the periodicity directions [86, 25, 38, 58, 83]. We shall restrict ourselves here to a simplified problem where (small) absorption in the media is assumed through the consideration of complex wave numbers with positive imaginary parts. This assumption has also been done in [42]. Indeed this greatly simplifies the analysis since the underlying differential equation operator becomes coercive. This assumption allows us to focus more on the specificity of coupling spectral discretization of the volume integral method with discretization in the Floquet-Bloch variable. Indeed this step will enable us, in a future work, to treat the case without absorption and focus on the technicality related to establishing Rellich type identities that can be exploited in the convergence analysis. Our assumption allows also to evacuate problems that would occur at the Wood anomalies. Let us notice however that the incorporation of a radiation condition for absorption free problems (as in [83]) would not induce major additional difficulties (for the numerical implementation of the method) since this condition is encoded in the volume integral equation of the problem.

The convergence analysis of the spectral approximation relies on establishing that the obtained discrete system coincides with the one for volume integral equation associated with the  $ML$ -periodic problem. This property is indeed surprising and one would not expect that it holds in general. In fact it is very specific to the choice of a trapezoidal rule to discretize the Floquet-Bloch transform. It has been also observed for finite element discretization in [42]. The optimality of this choice of discretization is not obvious and may for instance not be convenient in the neighborhood of the Wood anomaly (in the absorption free case).

Our discrete system can then be seen as a special re-arranging of the discretization of the  $ML$ -periodic problem (with a slightly different source term). One of the main advantages of this rearranging is that one obtain a linear growth of the computational cost with respect to the number of discretization points in the Floquet-Bloch variable. This is a gain of a logarithmic factor with respect to the application of the method in [103] to the  $ML$  periodic problem. Indeed this would be of interest for large (three dimensional) problems or when the absorption is small (requiring the use of a large discretization points in the Floquet-Bloch variable). We shall discuss

this issue in a future work. In the present one we shall content ourselves with preliminary validating results in 2D setting of the problem.

The chapter is organized as follows. We first introduce the problem and some notation and results for the Floquet-Bloch transform. In order to ease the reading of the technical part treating the full discretization of the problem we introduce in Section 3.3 the method employed in [85] for quasi periodic problem. We complement known results with a uniform convergence result with respect to Fourier-Bloch variable. Section 5.2.2 constitutes the core of our work. We discuss in Section 3.4.1 the semi-discretization in the Floquet-Bloch variable and prove exponential convergence result. Section 5.2.2 is dedicated to a full discretization of the problem and associated convergence result. The last section is concerned with a rapid description of the algorithm and some numerical validating tests for the two dimensional problem. We also present some numerical examples for real wave numbers that are not covered by the convergence theory but correspond with the configurations studied in Chapter 2.

## 3.2 Setting of the problem

### 3.2.1 Introduction of the problem and notation

We investigate the scattering problem for the case of locally perturbed unbounded periodic layers. The problem is formulated as: Given  $f \in L^2(\mathbb{R}^d)$ ,  $d = 2, 3$ , find solution  $u \in H_{loc}^1(\mathbb{R}^d)$  to the Helmholtz equation:

$$\Delta u + k^2 n u = f \quad \text{in } \mathbb{R}^d \quad (3.1)$$

where  $n \in L^\infty(\mathbb{R}^d)$  denotes the index of refraction and is such that  $n = n_p$  outside a compact domain  $D$  where  $n_p \in L^\infty(\mathbb{R}^d)$  is a periodic function with respect to the first  $d - 1$  variables of period  $L := (L_1, \dots, L_{d-1}) \in \mathbb{R}^{d-1}$ ,  $L_j > 0$ ,  $j = 1, \dots, d - 1$ .

If the wave number  $k$  is real, then (3.1) should be supplemented with a radiation condition (see for instance [9]). Although our numerical method can formally be extended to the case of real wave numbers  $k$ , the convergence proof requires some decay properties along the periodicity directions that we are only capable to prove in the case of complex wave numbers. This is why we shall restrict ourselves to the (coercive case) where

$$k^2 = k_0^2 + i\sigma$$

with  $k_0 \geq 0$  and  $\sigma > 0$  and assume that  $\text{Re } n(x) \geq c_0 > 0$  and  $\text{Im } n(x) \geq 0$  in  $\mathbb{R}^d$  for some constant  $c_0$ . In this case one can easily check, using the Lax-Milgram theorem that (3.1) has a unique solution  $u \in H^1(\mathbb{R}^d)$  and that

$$\|u\|_{L^2(\mathbb{R}^d)} \leq \frac{1}{\sigma c_0} \|f\|_{L^2(\mathbb{R}^d)}, \quad \|\nabla u\|_{L^2(\mathbb{R}^d)} \leq \sqrt{\frac{k_0^2 \|n\|_\infty}{\sigma^2 c_0^2} + \frac{1}{\sigma c_0}} \|f\|_{L^2(\mathbb{R}^d)}. \quad (3.2)$$

We also remark that the solution  $u \in H^2(\mathbb{R}^d)$  (since  $\Delta u \in L^2(\mathbb{R}^d)$ ). For the well posedness of the problem in the case of real wave numbers, known results require some additional monotonicity properties of  $n$  with respect to  $x_d$  (see for instance [86]).

For a point  $x \in \mathbb{R}^d$  we shall use the decomposition  $x = (\bar{x}, x_d)$  with  $x_d \in \mathbb{R}$  and  $\bar{x} = (x_1, \dots, x_{d-1}) \in \mathbb{R}^{d-1}$ . For  $m = (m_1, \dots, m_{d-1}) \in \mathbb{Z}^{d-1}$  we denote by

$$\Omega_m := \left[ (m - \frac{1}{2})L, (m + \frac{1}{2})L \right] \times \mathbb{R},$$

where we use the notation  $\llbracket a, b \rrbracket := [a_1, b_1] \times \dots \times [a_{d-1}, b_{d-1}]$ . We shall also use the notation

$$\Omega^h := \mathbb{R}^{d-1} \times ]-h, h[ \quad \text{and} \quad \Omega_m^h := \Omega^h \cap \Omega_m.$$



We shall assume that there exists  $h > 0$  such that

$$\text{supp}(n_p - 1) \subset \Omega^h \text{ and } \text{supp}(f) \subset \Omega^h.$$

Moreover, to further ease the presentation, we suppose that the support of  $n - n_p$  is strictly contained in  $\Omega_0^h$ . Indeed this can always be ensured by increasing the periodicity length. One can also work with a periodicity equal  $\Omega_0^h$  even if the support of  $n - n_p$  is larger  $\Omega_0^h$  but at the cost of additional non essential technicalities (See Remark 3.2.3 below).

**Remark 3.2.1.** For a point  $x_0 \in \mathbb{R}^d$ , the total field  $u$  generated by a point source at  $x_0$  is solution to

$$\Delta u + k^2 n u = -\delta_{x_0} \quad \text{in } \mathbb{R}^d. \quad (3.3)$$

Let  $u^i(\cdot; x_0)$  be the incident wave generated by the point source at  $x_0$ ,

$$u^i(\cdot; x_0) := \begin{cases} \frac{i}{4} H_0^{(1)}(k|\cdot - x_0|) & \text{in } \mathbb{R}^2, \\ \frac{e^{ik|\cdot - x_0|}}{4\pi|\cdot - x_0|} & \text{in } \mathbb{R}^3. \end{cases}$$

The scattering problem for a point source at  $x_0$  can be reformulated as finding  $u^s = u - u^i$  with  $u^s \in H^1(\mathbb{R}^d)$  such that

$$\Delta u^s + k^2 n u^s = -k^2(n - 1)u^i \quad \text{in } \mathbb{R}^d. \quad (3.4)$$

Indeed problem (3.4) has the same form as (3.1) with  $f := -k^2(n - 1)u^i$ .

### 3.2.2 Formulation of the problem using the Floquet-Bloch transform

**Definition 3.2.2.** A regular function  $u$  is called quasi-periodic with parameter  $\xi = (\xi_1, \dots, \xi_{d-1})$  and period  $L = (L_1, \dots, L_{d-1})$ , with respect to the first  $d-1$  variables (briefly denoted as  $\xi$ -quasi-periodic with period  $L$ ) if:

$$u(\bar{x} + mL, x_d) = e^{i\xi \cdot (mL)} u(\bar{x}, x_d), \quad \forall m \in \mathbb{Z}^{d-1}.$$

We now introduce some notation for functional spaces that will be used in the sequel:

- $C_{\xi, L}^\infty(\mathbb{R}^d) := \{u \in C^\infty(\mathbb{R}^d); u \text{ is } \xi\text{-quasi-periodic with period } L\}$ ;
- $C_\xi^\infty(\Omega_0) := \{u|_{\Omega_0}; u \in C_{\xi, L}^\infty(\mathbb{R}^d)\}$ ;
- $H_\xi^s(\Omega_0)$  is the closure of  $C_\xi^\infty(\Omega_0)$  with respect to the  $H^s(\Omega_0)$  norm;
- $H_{\xi, L}^s(\mathbb{R}^d)$  is the extension of  $H_\xi^s(\Omega_0)$  functions to a  $\xi$ -quasi-periodic function with period  $L$  to all of  $\mathbb{R}^d$ .

For a regular function  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ , the Floquet-Bloch transform of  $\varphi$  in the first  $d-1$  variables with period  $L \in \mathbb{R}^{d-1}$  is defined by:

$$\mathcal{F}\varphi(\bar{x}, x_d; \xi) := \sum_{m \in \mathbb{Z}^{d-1}} \varphi(\bar{x} + mL, x_d) e^{-i(mL) \cdot \xi}, \quad \forall \xi \in \left[-\frac{\pi}{L}, \frac{\pi}{L}\right], x \in \mathbb{R}^d. \quad (3.5)$$

$\mathcal{F}\varphi(\cdot; \xi)$  defined in (3.5) is  $\xi$ -quasi-periodic with period  $L$ . The inverse Floquet-Bloch transform  $\mathcal{F}^{-1}$  is given by

$$\varphi(\bar{x} + mL, x_d) = \frac{[L]}{(2\pi)^{d-1}} \int_{[-\frac{\pi}{L}, \frac{\pi}{L}]} \mathcal{F}\varphi(\bar{x}, x_d; \xi) e^{i(mL) \cdot \xi} d\xi, \quad m \in \mathbb{Z}^{d-1}, x \in \mathbb{R}^d, \quad (3.6)$$

where  $\llbracket L \rrbracket := L_1 \cdots L_{d-1}$ . We shall also use the following notations for vector in  $\mathbb{Z}^{d-1}$ :

$$|L| := |L_1| + \cdots + |L_{d-1}|, \quad \|L\| := \sqrt{L_1^2 + \cdots + L_{d-1}^2}.$$

Operators on vectors such as multiplication or division should be understood as component-wise operations. Since  $\mathcal{S}(\mathbb{R}^d)$  is dense in  $L^2(\mathbb{R}^d)$  and

$$\|\mathcal{F}\varphi\|_{L^2(\Omega_0 \times \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket)} = \|\varphi\|_{L^2(\mathbb{R}^d)}$$

for  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ , then the operator  $\mathcal{F}$  extends to an isometry between  $L^2(\mathbb{R}^d)$  and  $L^2(\Omega_0 \times \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket)$ . Furthermore, for all integers  $s > 0$ ,  $\mathcal{F}(H^s(\mathbb{R}^d)) = L^2(\llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket, H_\xi^s(\Omega_0))$  and for all  $u \in H^s(\mathbb{R}^d)$ :

$$\mathcal{F}(\partial_x^j u) = \partial_x^j \mathcal{F}(u), \quad \forall 0 \leq j \leq s. \quad (3.7)$$

This result can be found in [42]. We also refer to [78] for an extensive study of the Floquet-Bloch transform.

We now propose to use the Floquet-Bloch transform in the first  $d-1$  directions to study the spectral numerical approximation of problem (3.1). We first remark that since  $n_p$  is  $L$  periodic, then  $\mathcal{F}(n_p u) = n_p \mathcal{F}(u)$ . Moreover, if a function  $\varphi$  has a support in  $\Omega_0$ , then  $\mathcal{F}\varphi(\bar{x}, x_d; \xi) = \varphi(\bar{x}, x_d)$  for  $(\bar{x}, x_d) \in \Omega_0$ . Consequently, applying the Floquet-Bloch transform to equation (3.1) we get

$$\Delta(\mathcal{F}u)(\bar{x}, x_d; \xi) + k^2 n_p(\mathcal{F}u)(\bar{x}, x_d; \xi) + k^2(n - n_p)u(\bar{x}, x_d) = (\mathcal{F}f)(\bar{x}, x_d; \xi); \quad (\bar{x}, x_d) \in \Omega_0. \quad (3.8)$$

Let us set  $f_\xi := (\mathcal{F}f)(\cdot; \xi)$  and

$$\tilde{u}(\bar{x}, x_d; \xi) := \mathcal{F}u(\bar{x}, x_d; \xi) \quad (x, x_d) \in \Omega_0.$$

Then  $\tilde{u} \in L^2(\llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket, H_\xi^2(\Omega_0))$  and

$$\begin{cases} \Delta \tilde{u}(\cdot; \xi) + k^2 n_p \tilde{u}(\cdot; \xi) + k^2(n - n_p)\mathcal{M}(\tilde{u}) = f_\xi & \text{in } \Omega_0, \xi \in \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket \\ \mathcal{M}(\tilde{u}) := \frac{\llbracket L \rrbracket}{(2\pi)^{d-1}} \int_{\llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket} \tilde{u}(\cdot; \xi) d\xi, \end{cases} \quad (3.9)$$

where the second equation comes from the inverse Floquet-Bloch transform expression

$$u(\bar{x} + mL, x_d) = \frac{\llbracket L \rrbracket}{(2\pi)^{d-1}} \int_{\llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket} \tilde{u}(\bar{x}, x_d; \xi) e^{i(mL)\cdot\xi} d\xi, \quad m \in \mathbb{Z}, (\bar{x}, x_d) \in \Omega_0. \quad (3.10)$$

One can also check that the converse is true, i.e. if  $(\mathcal{F}f) \in L^2(\Omega_0 \times \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket)$ , the  $u$  defined by (3.10) is in  $H^2(\mathbb{R}^2)$  and is solution to (3.1).

**Remark 3.2.3.** *If the support of  $n - n_p$  is contained in  $\Omega_{P^-, P^+}^h := \cup_{m=-P^-}^{P^+} \Omega_m^h$  for  $P^-, P^+ \geq 0$  then the third term of the first equation of system (3.9) has to be modified and the system becomes*

$$\begin{cases} \Delta \tilde{u}(\cdot; \xi) + k^2 n_p \tilde{u}(\cdot; \xi) + k^2 \sum_{m=-P^-}^{P^+} (n - n_p)(\cdot + mL) \mathcal{M}_m(\tilde{u}) e^{-i(mL)\cdot\xi} = f_\xi & \text{in } \Omega_0, \xi \in \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket \\ \mathcal{M}_m(\tilde{u}) := \frac{\llbracket L \rrbracket}{(2\pi)^{d-1}} \int_{\llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket} \tilde{u}(\cdot; \xi) e^{i(mL)\cdot\xi} d\xi. \end{cases} \quad (3.11)$$

*The analysis below can be extended to treat this case without any major additional difficulty.*

### 3.3 Volume integral formulation of the $\xi$ –quasi-periodic problem

Before dealing with problem (3.9) where the equations for different values of  $\xi$  are coupled through  $\mathcal{M}(\bar{u})$ , we shall first recall the principles of the spectral volume integral method for fixed  $\xi$  as in [85] for periodic media. This may be helpful in making the technical details of the general case treated in next section more digest. We shall also supplement the convergence results with a result on uniform convergence properties with respect to  $\xi$ . Although the latter is not necessary to treat problem (3.9), we thought that it has its own interest and would complement the picture for the use of this type of methods in periodic media. Some of the technical results presented will also be useful for Section 4.

#### 3.3.1 Setting of the volume integral equation

We first consider problem (3.9) for the case of periodic media, i.e.  $n = n_p$ . Let us set  $u_\xi := \mathcal{F}u(\cdot; \xi)$  and set  $f_\xi := \mathcal{F}f(\cdot; \xi)$ . Then we are led to consider the problem,  $u_\xi \in H_{\xi, L}^2(\mathbb{R}^d)$  verifying

$$\Delta u_\xi + k^2 n_p u_\xi = f_\xi \quad \text{in } \mathbb{R}^d \quad (3.12)$$

where  $f_\xi \in L_{\xi, L}^2(\mathbb{R}^d)$ . Indeed considering problem (5.4) in  $\mathbb{R}^d$  or in  $\Omega_0$  is equivalent. In the following we shall not distinguish  $u_\xi$  from its restriction to  $\Omega_0$ . We recall that

$$\|u_\xi\|_{H_{\xi, L}^1(\Omega_0)} \leq C \|f_\xi\|_{L^2(\Omega_0^h)}, \quad \text{where } C := \max \left\{ \frac{1}{\sigma c_0}, \sqrt{\frac{k_0^2 \|n\|_\infty}{\sigma^2 c_0^2} + \frac{1}{\sigma c_0}} \right\}. \quad (3.13)$$

Let  $G_\xi(\cdot) \in L_{\xi, L}^2(\mathbb{R}^d)$  be the  $\xi$ –quasi-periodic Green function of the Helmholtz equation, i.e. satisfying,

$$\Delta G_\xi + k^2 G_\xi(\cdot) = -\delta_0 \quad \text{in } \Omega_0. \quad (3.14)$$

Then  $G_\xi$  can be expressed as (see [69]),

$$G_\xi(x) = \frac{i}{2\llbracket L \rrbracket} \sum_{j \in \mathbb{Z}^{d-1}} \frac{1}{\beta_\xi(j)} \exp(i\alpha_\xi(j) \cdot \bar{x} + i\beta_\xi(j)|x_d|), \quad (\bar{x}, x_d) \in \mathbb{R}^d, \quad (3.15)$$

where

$$\alpha_\xi(j) := \xi + \frac{2\pi}{L}j = \left( \xi_1 + \frac{2\pi}{L_1}j_1, \dots, \xi_{d-1} + \frac{2\pi}{L_{d-1}}j_{d-1} \right), \quad \beta_\xi(j)^2 := k^2 - \|\alpha_\xi(j)\|^2, \quad \text{Im } \beta_\xi(j) \geq 0.$$

(Remark that  $\beta_\xi(j) \neq 0$  for all  $j$  and  $\xi$  since  $\sigma > 0$ .) We now consider the volume potential  $V_\xi$  on  $L^2(\Omega_0^h)$  with kernel  $G_\xi$  defined by

$$V_\xi g(x) := \int_{\Omega_0^h} G_\xi(x-y)g(y)dy, \quad x \in \mathbb{R}^d. \quad (3.16)$$

**Lemma 3.3.1.** (See for instance [85, 46]) *The volume potential  $V_\xi$  is a linear bounded operator from  $L^2(\Omega_0^h)$  into  $H_{\xi, L}^2(\mathbb{R}^d)$ . Moreover, for all  $g \in L^2(\Omega_0^h)$ , the potential  $w := V_\xi g \in H_{\xi, L}^2(\mathbb{R}^d)$  and is the unique solution to  $\Delta w + k^2 w = -g$  in  $\mathbb{R}^d$ .*

Thus, if we set  $w := V_k(k^2(n_p - 1)u_\xi - f_\xi)$ , then from Lemma 3.3.1,  $w \in H_{\xi, L}^2(\mathbb{R}^d)$  and satisfies

$$\Delta w + k^2 w = -k^2(n_p - 1)u_\xi + f_\xi \quad \text{in } \mathbb{R}^d.$$

Therefore  $w = u_\xi$  in  $\mathbb{R}^d$  and  $u_\xi \in L^2(\Omega_0^h)$  and satisfies

$$u_\xi = k^2 V_\xi((n_p - 1)u_\xi) - V_\xi f_\xi \quad \text{in } L^2(\Omega_0^h). \quad (3.17)$$

Conversely, if (3.17) is verified, then obviously  $u_\xi$  can be obtained in all  $\mathbb{R}^d$  using the expression of  $w$ . The numerical scheme we shall consider is based on the spectral discretization of (3.17). In order to be efficient, by the use of FFT in evaluating the matrix-vector product, we need to periodize the equation in the direction  $x_d$ . This is the step we shall discuss now.

### 3.3.2 Periodization of the integral equation

Let  $R \in \mathbb{R}$  such that  $R > 2h$ , we define  $G_\xi^R$  as

$$G_\xi^R(x) := G_\xi(x), \quad \forall x = (\bar{x}, x_d) \in \mathbb{R}^{d-1} \times ]-R, R[, \quad (3.18)$$

and extend  $G_\xi^R$  to all  $\mathbb{R}^d$  as a  $2R$ -periodic function with respect to  $x_d$ . We then define the periodized volume potential  $V_\xi^R : L^2(\Omega_0^h) \rightarrow L^2(\Omega_0^R)$  using the same expression as (3.16) where the kernel  $G_\xi$  is replaced by  $G_\xi^R$  and consider the periodized volume integral equation,  $u_\xi^R \in L^2(\Omega_0^R)$ ,

$$u_\xi^R = k^2 V_\xi^R((n_p - 1)u_\xi^R) - V_\xi^R f_\xi \quad \text{in } L^2(\Omega_0^R). \quad (3.19)$$

We then have the following result.

**Lemma 3.3.2.** 1. For all  $g \in L^2(\Omega_0^h)$ ,  $V_\xi^R g = V_\xi g$  in  $\mathbb{R}^d$ .

2. For for all  $f_\xi \in L^2(\Omega_0^h)$ , equation (3.19) has a unique solution  $u_\xi^R \in L^2(\Omega_0^R)$  and we have

$$u_\xi^R = u_\xi \quad \text{in } \Omega_0^h$$

where  $u_\xi \in L^2(\Omega_0^h)$  is the unique solution of (3.17).

*Proof.* The equality  $V_\xi^R g = V_\xi g$  in  $\mathbb{R}^d$  holds for all  $g \in L^2(\Omega_0^h)$  since  $G_\xi^R$  and  $G_\xi$  coincide in  $\Omega^R$ ,  $|x_d - y_d| < R$  for all  $(x, y) \in \Omega_0^h \times \Omega_0^h$  and  $R > 2h$ .

Let  $f_\xi \in L^2(\Omega_0^h)$ . Since the support of  $n_p - 1$  is included in  $\Omega^h$ , we deduce from the first point that a solution  $u_\xi^R \in L^2(\Omega_0^R)$  of (3.19) is such that  $u_\xi^R|_{\Omega_0^h} \in L^2(\Omega_0^h)$  and verifies (3.17). Since  $u_\xi^R|_{\Omega_0^h} = 0$  and  $f_\xi = 0$  imply  $u_\xi^R = 0$  in  $\Omega_0^R$  we easily conclude from the well-posedness of (3.17) that (3.19) has at most one solution. In addition, a solution of (3.19) can be constructed as

$$u_\xi^R = u_\xi \quad \text{in } \Omega_0^h \quad \text{and} \quad u_\xi^R = k^2 V_\xi^R((n_p - 1)u_\xi) - V_\xi^R f_\xi \quad \text{in } \Omega_0^R \setminus \Omega_0^h$$

where  $u_\xi \in L^2(\Omega_0^h)$  is the solution of (3.17). □

The spectral discretization method is based on the Fourier basis defined as

$$\varphi_\xi^j(x) := \frac{1}{\sqrt{2[L]R}} \exp\left(i\left(\xi + \frac{2\pi}{L}j\right) \cdot \bar{x} + i\frac{\pi}{R}j_d x_d\right), \quad j = (j_1, \dots, j_d) = (\bar{j}, j_d) \in \mathbb{Z}^d. \quad (3.20)$$

For all  $u \in L^2(\Omega_0^R)$ ,  $u = \sum_{j \in \mathbb{Z}^d} \widehat{u}(j; \xi) \varphi_\xi^j(x)$  where  $\widehat{u}(j; \xi)$  to denotes the  $j$ -th coefficient of  $u$  with respect to this Fourier basis, which is defined as:

$$\widehat{u}(j; \xi) := \int_{\Omega_0^R} u(x) \overline{\varphi_\xi^j(x)} dx. \quad (3.21)$$

If  $u \in H_\xi^s(\Omega_0^R)$ ,  $0 \leq s < +\infty$  then an equivalent norm in  $H_\xi^s(\Omega_0^R)$  is given by

$$\|u\|_{H_\xi^s(\Omega_0^R)}^2 := \sum_{j \in \mathbb{Z}^d} \left( 1 + \left\| \frac{\bar{j}}{L} \right\|^2 + \left\| \frac{j_d}{R} \right\|^2 \right)^s |\widehat{u}(j; \xi)|^2. \quad (3.22)$$

Remark that since  $\|u\|_{L^2(\Omega_0^R)} = \|e^{i\xi x} u\|_{L^2(\Omega_0^R)}$ , using Plancherel theorem we conclude the norm in  $L^2(\Omega_0^R)$  defined by (3.22) (for  $s = 0$ ) is independent from  $\xi$ , i.e.,

$$\|u\|_{L^2(\Omega_0^R)}^2 = \sum_{j \in \mathbb{Z}^d} |\widehat{u}(j; 0)|^2 = \sum_{j \in \mathbb{Z}^d} |\widehat{u}(j; \xi)|^2 \quad \forall \xi. \quad (3.23)$$

We can easily establish the link between the Fourier coefficients with respect to  $\xi$ -quasi-periodic Fourier basis and periodic Fourier basis ( $\xi \in \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket$ ,  $\xi^\ell \neq 0, \forall \ell = 1, \dots, d$ ) as

$$\widehat{u}(t; 0) = 2R \sum_{j \in \mathbb{Z}^d} \widehat{u}(\bar{j}, t_d; \xi) \prod_{\ell=1}^{d-1} \frac{2 \cos(\pi(j_\ell - t_\ell)) \sin(\xi^\ell \frac{L_\ell}{2})}{\xi^\ell + \frac{2\pi}{L_\ell}(j_\ell - t_\ell)} \quad (3.24)$$

where here  $\xi^\ell$  denotes the component  $\ell$  of  $\xi$ .

Using (3.15), one can easily compute the Fourier coefficients of  $G_\xi^R$  and get

$$\widehat{G}_\xi^R(j; \xi) = \gamma \frac{e^{i\beta_\xi(\bar{j})R} \cos(\pi j_d) - 1}{\beta_\xi(\bar{j})^2 - (\frac{\pi}{R} j_d)^2} = \gamma \frac{e^{i\beta_\xi(\bar{j})R} \cos(\pi j_d) - 1}{k^2 - \|\alpha_\xi(\bar{j})\|^2 - (\frac{\pi}{R} j_d)^2} \quad (3.25)$$

with  $\gamma := \frac{1}{\sqrt{2\llbracket L \rrbracket R}}$ . Since  $k^2 = k_0^2 + i\sigma$  with positive  $\sigma$  and  $\text{Im} \beta_\xi(\bar{j}) \geq 0$ , we have the estimate, for sufficiently small  $\varepsilon$  such that  $\sigma - \varepsilon(k_0^2 + 2|\xi|^2) > 0$ ,

$$|\widehat{G}_\xi^R(j; \xi)| \leq \gamma \frac{2}{\sigma + \varepsilon(-k_0^2 + \|\alpha_\xi(\bar{j})\|^2 + (\frac{\pi}{R} j_d)^2)}.$$

Choosing  $\varepsilon$  such that  $\sigma - \varepsilon(k_0^2 + 2\|\pi/L\|^2) > 0$ , we infer the existence of a constant  $C$  independent from  $\xi$ ,  $L$ ,  $R$  and  $j$  such that

$$|\widehat{G}_\xi^R(j; \xi)| \leq C \frac{\gamma}{\left( 1 + \left\| \frac{\bar{j}}{L} \right\|^2 + \left\| \frac{j_d}{R} \right\|^2 \right)}. \quad (3.26)$$

Since  $V_\xi^R$  is a convolution operator

$$\widehat{V}_\xi^R g(j; \xi) = \frac{1}{\gamma} \widehat{G}_\xi^R(j; \xi) \widehat{g}(j; \xi).$$

We then immediately get from (3.26) the following result.

**Lemma 3.3.3.** *Let  $s \in \mathbb{N}$ . The operator  $V_\xi^R : H_\xi^s(\Omega_0^R) \rightarrow H_\xi^{2+s}(\Omega_0^R)$  is continuous. Moreover, there exists a constant  $C$  independent of  $\xi$  such that*

$$\|V_\xi^R g\|_{H_\xi^{2+s}(\Omega_0^R)} \leq C \|g\|_{H_\xi^s(\Omega_0^R)} \quad \forall g \in H_\xi^s(\Omega_0^R).$$

We then obtain the following uniform regularity result.

**Proposition 3.3.4.** *Let  $f_\xi \in L^2(\Omega_0^h)$  and  $u_\xi^R \in L^2(\Omega_0^R)$  be the solution of (3.19). Then there exists a positive constant  $C$  independent from  $\xi$  and  $f_\xi$  such that*

$$\|u_\xi^R\|_{H^2(\Omega_0^R)} \leq C \|f_\xi\|_{L^2(\Omega_0^h)}. \quad (3.27)$$

*Proof.* Thanks to Lemma 3.3.2,  $u_\xi^R = u_\xi$  in  $\Omega_0^h$  where  $u_\xi$  denotes the solution of (3.17). Then, using (3.13),

$$\|u_\xi^R\|_{L^2(\Omega_0^h)} \leq C_1 \|f_\xi\|_{L^2(\Omega_0^h)}$$

for some constant  $C_1$  independent from  $\xi$ . Using equation (3.19), Lemma 3.3.3 and the fact that  $\text{supp}(1 - n_p) \subset \Omega_0^h$ , we get the existence of a constant  $C_2$  independent from  $\xi$  such that

$$\|u_\xi^R\|_{H^2(\Omega_0^h)} \leq C_2 (\|f_\xi\|_{L^2(\Omega_0^h)} + \|u_\xi^R\|_{L^2(\Omega_0^h)}).$$

The result of the proposition then immediately follows from the two inequalities.  $\square$

Let us introduce for later use  $T_\xi^R : L^2(\Omega_0^R) \rightarrow L^2(\Omega_0^R)$  defined by

$$T_\xi^R g := k^2 V_\xi^R ((n_p - 1)g) \quad (3.28)$$

and  $A_\xi^R : L^2(\Omega_0^R) \rightarrow L^2(\Omega_0^R)$  defined by

$$A_\xi^R g := g - T_\xi^R g. \quad (3.29)$$

Then, as a straightforward consequences of Lemma 3.3.3, we have:

**Lemma 3.3.5.** *The operator  $T_\xi^R : L^2(\Omega_0^R) \rightarrow L^2(\Omega_0^R)$  is compact and there exists a positive constant  $C$  independent of  $\xi$  such that:*

$$\|T_\xi^R g\|_{H^2(\Omega_0^R)} \leq C \|g\|_{L^2(\Omega_0^R)}, \quad \forall g \in L^2(\Omega_0^R). \quad (3.30)$$

An immediate corollary of this Lemma and Lemma 3.3.2 is the following.

**Corollary 3.3.6.** *The operator  $A_\xi^R : L^2(\Omega_0^R) \rightarrow L^2(\Omega_0^R)$  is bounded and injective. Moreover, there exists a positive constant  $C$  independent of  $\xi$  such that*

$$\|A_\xi^R g\|_{L^2(\Omega_0^R)} \leq C \|g\|_{L^2(\Omega_0^R)}.$$

### 3.3.3 Spectral approximation of problem (3.19)

Let  $N = (N_1, \dots, N_d) \in \mathbb{N}^d$  and  $\mathbb{Z}_N^d := \{j = (j_1, \dots, j_d) \in \mathbb{Z}^d \mid -\frac{N_\ell}{2} + 1 \leq j_\ell \leq \frac{N_\ell}{2}, \ell = 1, \dots, d\}$ , we define the discrete space  $\mathcal{T}_\xi^N := \text{span} \{\varphi_\xi^j; j \in \mathbb{Z}_N^d\}$  and the orthogonal projection

$$P_\xi^N : L^2(\Omega_0^R) \longrightarrow \mathcal{T}_\xi^N, \quad u \longmapsto P_\xi^N(u) := \sum_{j \in \mathbb{Z}_N^d} \widehat{u}(j; \xi) \varphi_\xi^j. \quad (3.31)$$

Let  $u_\xi^R$  be the solution of (3.19). We now consider  $u_\xi^{R,N} \in \mathcal{T}_\xi^N$  an approximation of  $u_\xi^R$ , which is solution to following variational equation

$$\left\langle u_\xi^{R,N} - k^2 V_\xi^R \left( (n_p - 1) u_\xi^{R,N} \right), v_\xi^N \right\rangle_{L^2(\Omega_0^R)} = \left\langle -V_\xi^R f_\xi, v_\xi^N \right\rangle_{L^2(\Omega_0^R)}, \quad \forall v_\xi^N \in \mathcal{T}_\xi^N. \quad (3.32)$$

The next theorem shows that the convergence rate of this method using standard convergence result of Galerkin discretization, is independent from  $\xi$ .

**Theorem 3.3.7.** *There exists  $N_0 \in \mathbb{N}^d$  independent from  $\xi$  such that problem (3.32) admits a unique solution for all  $N > N_0$  and  $f_\xi \in L^2(\Omega_0^h)$ . Moreover, there is a positive constant  $C > 0$  independent of  $\xi$  and  $N$  such that,*

$$\|u_\xi^R - u_\xi^{R,N}\|_{L^2(\Omega_0^R)} \leq C \inf_{v_\xi^N \in \mathcal{T}_\xi^N} \|u_\xi^R - v_\xi^N\|_{L^2(\Omega_0^R)}. \quad (3.33)$$

*Proof.* We concentrate on the second claim since the first one follows from the similar arguments. The solution  $u_\xi^R$  of (3.19) satisfies

$$\langle A_\xi^R u_\xi^R, v_\xi \rangle_{L^2(\Omega_0^R)} = \langle -V_\xi^R f_\xi, v_\xi \rangle_{L^2(\Omega_0^R)}, \quad \forall v_\xi \in L^2(\Omega_0^R). \quad (3.34)$$

We then obtain

$$\langle A_\xi^R (u_\xi^R - u_\xi^{R,N}), v_\xi^N \rangle_{L^2(\Omega_0^R)} = 0, \quad \forall v_\xi^N \in \mathcal{T}_\xi^N. \quad (3.35)$$

Since  $u_\xi^{R,N} - v_\xi^N \in \mathcal{T}_\xi^N$ , we get

$$\begin{aligned} \langle A_\xi^R (u_\xi^R - u_\xi^{R,N}), u_\xi^R - u_\xi^{R,N} \rangle_{L^2(\Omega_0^R)} &= \langle A_\xi^R (u_\xi^R - u_\xi^{R,N}), u_\xi^R - v_\xi^N \rangle_{L^2(\Omega_0^R)} \\ &\leq C \|u_\xi^R - u_\xi^{R,N}\|_{L^2(\Omega_0^R)} \|u_\xi^R - v_\xi^N\|_{L^2(\Omega_0^R)}. \end{aligned} \quad (3.36)$$

We shall prove for all  $\varepsilon > 0$ , there exists  $N_0 > 0$  independent from  $\xi$  such that  $\forall N > N_0$ ,  $\forall \xi \in \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket$ :

$$\langle T_\xi^R (u_\xi^R - u_\xi^{R,N}), u_\xi^R - u_\xi^{R,N} \rangle_{L^2(\Omega_0^R)} \leq \varepsilon \|u_\xi^R - u_\xi^{R,N}\|_{L^2(\Omega_0^R)}^2. \quad (3.37)$$

Let us define

$$w_\xi^N := \frac{u_\xi^R - u_\xi^{R,N}}{\|u_\xi^R - u_\xi^{R,N}\|_{L^2(\Omega_0^R)}} \quad (3.38)$$

and suppose that (3.37) is not true. Then there exists a sequence  $N(n) \rightarrow \infty$  such that

$$\langle T_{\xi_n}^R w_{\xi_n}^{N(n)}, w_{\xi_n}^{N(n)} \rangle_{L^2(\Omega_0^R)} \geq \varepsilon. \quad (3.39)$$

for some  $\xi_n \in \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket$ . To shorten the notation, we set  $w_n := w_{\xi_n}^{N(n)}$ . Since  $(\xi_n)_{n \in \mathbb{N}} \subset \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket$  then there is a sub-sequence, also denoted  $(\xi_n)_{n \in \mathbb{N}}$ , that converges to  $\xi \in \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket$ . Furthermore,  $(w_n)$  is bounded in  $L^2(\Omega_0^R)$  then there exist a sub-sequence, also denoted  $(w_n)$  that converges weakly to  $w$  in  $L^2(\Omega_0^R)$ . Let  $v \in L^2(\Omega_0^R)$  and set  $v_n := P_{\xi_n}^{N(n)} v \in \mathcal{T}_{\xi_n}^{N(n)}$ . Then

$$\begin{aligned} \langle A_\xi^R w_n, v \rangle_{L^2(\Omega_0^R)} &= \langle (A_\xi^R - A_{\xi_n}^R) w_n, v \rangle_{L^2(\Omega_0^R)} + \langle A_{\xi_n}^R w_n, v_n \rangle_{L^2(\Omega_0^R)} + \langle A_{\xi_n}^R w_n, v - v_n \rangle_{L^2(\Omega_0^R)} \\ &= \langle (A_\xi^R - A_{\xi_n}^R) w_n, v \rangle_{L^2(\Omega_0^R)} + \langle A_{\xi_n}^R w_n, v - v_n \rangle_{L^2(\Omega_0^R)} \end{aligned}$$

where we used (3.35) for the last equality. Observe that, using Corollary 3.3.6,

$$\left| \langle A_{\xi_n}^R w_n, v - v_n \rangle_{L^2(\Omega_0^R)} \right| \leq \|A_{\xi_n}^R\|_{L^2(\Omega_0^R) \rightarrow L^2(\Omega_0^R)} \|v - v_n\|_{L^2(\Omega_0^R)} \xrightarrow{n \rightarrow \infty} 0 \quad (3.40)$$

and from Lemma 3.3.10 we have,

$$\begin{aligned} \left| \langle (A_\xi^R - A_{\xi_n}^R) w_n, v \rangle_{L^2(\Omega_0^R)} \right| &= \left| \langle (T_{\xi_n}^R - T_\xi^R) w_n, v \rangle_{L^2(\Omega_0^R)} \right| \\ &\leq \|T_{\xi_n}^R - T_\xi^R\|_{L^2(\Omega_0^R) \rightarrow L^2(\Omega_0^R)} \|v\|_{L^2(\Omega_0^R)} \xrightarrow{n \rightarrow \infty} 0. \end{aligned} \quad (3.41)$$

Therefore, taking the limit as  $n \rightarrow +\infty$  we get

$$\langle A_\xi^R w, v \rangle_{L^2(\Omega_0^R)} = 0, \quad \forall v \in L^2(\Omega_0^R). \quad (3.42)$$

Since  $A_\xi^R$  is injective (see Corollary 3.3.6), we deduce that  $w = 0$ . We now observe that

$$\langle T_{\xi_n}^R w_n, w_n \rangle_{L^2(\Omega_0^R)} = \langle (T_{\xi_n}^R - T_\xi^R) w_n, w_n \rangle_{L^2(\Omega_0^R)} + \langle T_\xi^R w_n, w_n \rangle_{L^2(\Omega_0^R)}$$

where

$$\left| \langle (T_{\xi_n}^R - T_\xi^R) w_n, w_n \rangle_{L^2(\Omega_0^R)} \right| \leq \|T_{\xi_n}^R - T_\xi^R\|_{L^2(\Omega_0^R) \rightarrow L^2(\Omega_0^R)} \xrightarrow{n \rightarrow \infty} 0,$$

and

$$\langle T_\xi^R w_n, w_n \rangle_{L^2(\Omega_0^R)} \xrightarrow{n \rightarrow \infty} 0,$$

since  $T_\xi^R$  is compact. We then conclude that

$$\langle T_{\xi_n}^R w_n, w_n \rangle_{L^2(\Omega_0^R)} \xrightarrow{n \rightarrow \infty} \langle T_\xi^R w, w \rangle_{L^2(\Omega_0^R)} = 0, \quad (3.43)$$

which contradicts (3.39) and proves (3.37). Now choose  $\varepsilon = \frac{1}{2}$  and recall that  $A_\xi^R = I - T_\xi^R$ . We obtain from (3.36)

$$\frac{1}{2} \|u_\xi^R - u_\xi^{R,N}\|_{L^2(\Omega_0^R)}^2 \leq \langle A_\xi^R (u_\xi^R - u_\xi^{R,N}), u_\xi^R - u_\xi^{R,N} \rangle \leq C \|u_\xi^R - u_\xi^{R,N}\| \|u_\xi^R - v_\xi^N\| \quad (3.44)$$

which implies

$$\|u_\xi^R - u_\xi^{R,N}\|_{L^2(\Omega_0^R)} \leq C \inf_{v_\xi^N \in \mathcal{T}_\xi^N} \|u_\xi^R - v_\xi^N\|_{L^2(\Omega_0^R)}. \quad (3.45)$$

□

It is possible to characterize the convergence rate due to the regularity result of Proposition 3.3.4 and the following classical interpolation result (See for instance [99]). We here make precise the dependence on  $L$  and  $R$  for later use in the analysis of the discretization of the full problem.

**Lemma 3.3.8.** *If  $u \in H_\xi^\mu(\Omega_0^R)$ , ( $\mu \in \mathbb{R}$ ) then*

$$\|u - P_\xi^N u\|_{H_\xi^\lambda(\Omega_0^R)} \leq \left( \sum_{\ell=1}^{d-1} \frac{1}{\left(1 + \frac{N_\ell^2}{L^2}\right)^{\mu-\lambda}} + \frac{1}{\left(1 + \frac{N_d^2}{R^2}\right)^{\mu-\lambda}} \right)^{1/2} \|u\|_{H_\xi^\mu(\Omega_0^R)}, \quad \text{for all } \lambda \leq \mu. \quad (3.46)$$

*Proof.* We give the proof for the reader's convenience. For  $j \in \mathbb{Z}^d$  let us denote by  $I(j) \in \mathbb{Z}^d$

$$I(j)_\ell = \frac{j_\ell}{L}; \quad \ell = 1, \dots, d-1 \quad \text{and} \quad I(j)_d = \frac{j_d}{R}.$$



$$\begin{aligned}
\|u - P_\xi^N u\|_{H_\xi^\lambda}^2 &= \sum_{j \in \mathbb{Z}^d \setminus \mathbb{Z}_N^d} (1 + \|I(j)\|^2)^\lambda |\widehat{u}(j; \xi)|^2 \\
&= \sum_{\ell=1}^d \sum_{j \in \mathbb{Z}^d, |j_\ell| > N_\ell} (1 + \|I(j)\|^2)^\lambda |\widehat{u}(j; \xi)|^2 \\
&\leq \sum_{\ell=1}^d \frac{1}{(1 + I(N)_\ell^2)^{\mu-\lambda}} \sum_{j \in \mathbb{Z}^d, |j_\ell| > N_\ell} (1 + \|I(j)\|^2)^\mu |\widehat{u}(j; \xi)|^2 \\
&\leq \sum_{\ell=1}^d \frac{1}{(1 + I(N)_\ell^2)^{\mu-\lambda}} \|u\|_{H_\xi^\mu}^2,
\end{aligned}$$

which proves the Lemma.  $\square$

Combining Proposition 3.3.4, Lemma 3.3.8 and Theorem 3.3.7 we obtain the following theorem.

**Theorem 3.3.9.** *There exists  $N_0 \in \mathbb{N}^d$  and a positive constant  $C > 0$  that are independent of  $\xi$  such that,*

$$\|u_\xi^R - u_\xi^{R,N}\|_{L^2(\Omega_0^R)} \leq C \left( \sum_{\ell=1}^d \frac{1}{(1 + N_\ell^2)^2} \right)^{1/2} \|f_\xi\|_{L^2(\Omega_0^b)}. \quad (3.47)$$

for all  $N \geq N_0$ .

To prove Theorem 3.3.7 we used some auxiliary results given in Lemma 3.3.10 and Lemma 3.3.11 below.

**Lemma 3.3.10.** *The operator  $T_\xi^R$  in Lemma 3.3.5 is uniform Lipschitz continuous with respect to  $\xi$ , i.e.*

$$\|T_\xi^R - T_{\xi'}^R\|_{L^2(\Omega_0^R) \rightarrow L^2(\Omega_0^R)} \leq C \|\xi - \xi'\| \quad (3.48)$$

for some constant  $C > 0$  independent of  $\xi, \xi' \in \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket$ .

*Proof.* To prove this lemma we use the norm in  $L^2(\Omega_0^R)$  defined in (3.23). We first estimate the Fourier coefficients of  $G_\xi^R - G_{\xi'}^R$  with respect to the periodic Fourier basis. For  $j = (\bar{j}, j_d)$ , from (3.24) and (3.25) we have

$$\widehat{G}_\xi^R(t; 0) = 2R\gamma \sum_{j \in \mathbb{Z}^d} \frac{e^{i\beta_\xi(\bar{j})R} \cos(\pi j_d) - 1}{k^2 - \|\alpha_\xi(\bar{j})\|^2 - (\frac{\pi}{R}j_d)^2} \prod_{\ell=1}^{d-1} \frac{2 \cos(\pi(j_\ell - t_\ell)) \sin(\xi_\ell \frac{L_\ell}{2})}{\xi_\ell + \frac{2\pi}{L_\ell}(j_\ell - t_\ell)}. \quad (3.49)$$

(In this proof,  $\xi_\ell$  is used to indicate the  $\ell$ -component of  $\xi$ ,  $\ell = 1, \dots, d$ ). Then,

$$\begin{aligned}
\left| \widehat{G}_\xi^R(t; 0) - \widehat{G}_{\xi'}^R(t; 0) \right| &= 2R\gamma \left| \sum_{j \in \mathbb{Z}^d} \frac{e^{i\beta_\xi(\bar{j})R} \cos(\pi j_d) - 1}{k^2 - \|\alpha_\xi(\bar{j})\|^2 - (\frac{\pi}{R}j_d)^2} \prod_{\ell=1}^{d-1} \frac{2 \cos(\pi(j_\ell - t_\ell)) \sin(\xi_\ell \frac{L_\ell}{2})}{\xi_\ell + \frac{2\pi}{L_\ell}(j_\ell - t_\ell)} \right. \\
&\quad \left. - \sum_{j \in \mathbb{Z}^d} \frac{e^{i\beta_{\xi'}(\bar{j})R} \cos(\pi j_d) - 1}{k^2 - \|\alpha_{\xi'}(\bar{j})\|^2 - (\frac{\pi}{R}j_d)^2} \prod_{\ell=1}^{d-1} \frac{2 \cos(\pi(j_\ell - t_\ell)) \sin(\xi'_\ell \frac{L_\ell}{2})}{\xi'_\ell + \frac{2\pi}{L_\ell}(j_\ell - t_\ell)} \right| \quad (3.50)
\end{aligned}$$

To simplify notation we set

$$A_j(\xi) := \frac{e^{i\beta_\xi(\bar{j})R} \cos(\pi j_d) - 1}{k^2 - \|\alpha_\xi(\bar{j})\|^2 - (\frac{\pi}{R}j_d)^2} \quad \text{and} \quad B_j(\xi) := \prod_{\ell=1}^{d-1} \frac{2 \cos(\pi(j_\ell - t_\ell)) \sin(\xi_\ell \frac{L_\ell}{2})}{\xi_\ell + \frac{2\pi}{L_\ell}(j_\ell - t_\ell)}.$$

So we can simply write (3.50) as

$$\left| \widehat{G}_\xi^R(t; 0) - \widehat{G}_{\xi'}^R(t; 0) \right| = 2R\gamma \left| \sum_{j \in \mathbb{Z}^d} A_j(\xi) B_j(\xi) - A_j(\xi') B_j(\xi') \right|.$$

We now can prove that there exists positive constants  $C_1, C_2, C_3, C_4$  independent of  $\xi, \xi'$  such that

$$\begin{aligned} |A_j(\xi) - A_j(\xi')| &\leq C_1 \frac{\|\xi - \xi'\|}{1 + \|j\|^2} \quad \text{and} \quad |A_j(\xi)| \leq \frac{C_3}{1 + \|j\|^2} \\ |B_j(\xi) - B_j(\xi')| &\leq C_2 \frac{\|\xi - \xi'\|}{\prod(-\frac{1}{2} + |j_\ell - t_\ell|)}, \quad \text{and} \quad |B_j(\xi)| \leq \frac{C_4}{\prod(-\frac{1}{2} + |j_\ell - t_\ell|)} \end{aligned}$$

So we finally have

$$\begin{aligned} |A_j(\xi) B_j(\xi) - A_j(\xi') B_j(\xi')| &= |(A_j(\xi) - A_j(\xi')) B_j(\xi) + A_j(\xi') (B_j(\xi) - B_j(\xi'))| \\ &\leq |A_j(\xi) - A_j(\xi')| |B_j(\xi)| + |A_j(\xi')| |B_j(\xi) - B_j(\xi')| \\ &\leq \frac{C \|\xi - \xi'\|}{(1 + \|j\|^2) \prod_{\ell=1}^{d-1} (-\frac{1}{2} + |j_\ell - t_\ell|)} \end{aligned}$$

Since

$$\sum_{j \in \mathbb{Z}_N^d} \frac{1}{(1 + \|j\|^2) \prod_{\ell=1}^{d-1} (-\frac{1}{2} + |j_\ell - t_\ell|)} \leq C_0 \quad (3.51)$$

So,

$$2R\gamma \left| \sum_{j \in \mathbb{Z}_N^d} A_j(\xi) B_j(\xi) - A_j(\xi') B_j(\xi') \right| \leq 2R\gamma \sum_{j \in \mathbb{Z}^d} |A_j(\xi) B_j(\xi) - A_j(\xi') B_j(\xi')| \leq 2R\gamma C C_0 \|\xi - \xi'\|$$

We then obtain

$$\left| \widehat{G}_\xi^R(t; 0) - \widehat{G}_{\xi'}^R(t; 0) \right| \leq C \|\xi - \xi'\| \quad (3.52)$$

for some constant  $C$  independent from  $\xi$  and  $\xi'$ . Since  $V_\xi^R - V_{\xi'}^R$  is a convolution operator

$$(V_\xi^R - V_{\xi'}^R) \widehat{g}(t; 0) = \frac{1}{\gamma} (\widehat{G}_\xi^R(t; 0) - \widehat{G}_{\xi'}^R(t; 0)) \widehat{g}(t; 0).$$

We then have

$$\|(T_\xi^R - T_{\xi'}^R)g\|_{L^2(\Omega_0^R)} = \|(V_\xi^R - V_{\xi'}^R)(n-1)g\|_{L^2(\Omega_0^R)} \leq C \|\xi - \xi'\| \|n\|_{L^\infty} \|g\|_{L^2(\Omega_0^R)}, \quad (3.53)$$

which proves the lemma.  $\square$

**Lemma 3.3.11.** *Let  $\xi \in \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket$  and a sequence  $(\xi_N)_N$  such that  $\xi_N \rightarrow \xi$  as  $N \rightarrow +\infty$  then*

$$\|v - P_{\xi_N}^N v\|_{L^2(\Omega_0^R)} \xrightarrow{n \rightarrow \infty} 0, \quad \forall v \in L^2(\Omega_0^R). \quad (3.54)$$

*Proof.* We first see that

$$\|v - P_{\xi_N}^N v\|_{L^2(\Omega_0^R)} \leq \|v - P_\xi^N v\|_{L^2(\Omega_0^R)} + \|P_\xi^N v - P_{\xi_N}^N v\|_{L^2(\Omega_0^R)}. \quad (3.55)$$

Since  $\|v - P_\xi^N v\|_{L^2(\Omega_0^R)} \rightarrow 0$  as  $N \rightarrow \infty$ , it remains to prove that

$$\|P_\xi^N v - P_{\xi_N}^N v\|_{L^2(\Omega_0^R)} \rightarrow 0 \quad \text{as } N \rightarrow \infty.$$

From (3.31) we have that

$$\begin{aligned} \|P_\xi^N v - P_{\xi_N}^N v\|_{L^2(\Omega_0^R)} &= \left\| \sum_{m \in \mathbb{Z}_N} \widehat{v}(m; \xi) \varphi_\xi^m - \widehat{v}(m; \xi_N) \varphi_{\xi_N}^m \right\|_{L^2(\Omega_0^R)} \\ &\leq \left\| \sum_{m \in \mathbb{Z}_N} (\widehat{v}(m; \xi) - \widehat{v}(m; \xi_N)) \varphi_\xi^m \right\|_{L^2(\Omega_0^R)} + \left\| \sum_{m \in \mathbb{Z}_N} \widehat{v}(m; \xi_N) (\varphi_\xi^m - \varphi_{\xi_N}^m) \right\|_{L^2(\Omega_0^R)}. \end{aligned}$$

We observe that

$$\begin{aligned} \left\| \sum_{m \in \mathbb{Z}_N} (\widehat{v}(m; \xi) - \widehat{v}(m; \xi_N)) \varphi_\xi^m \right\|_{L^2(\Omega_0^R)}^2 &= \sum_{m \in \mathbb{Z}_N} |\widehat{v}(m; \xi) - \widehat{v}(m; \xi_N)|^2 \\ &= \sum_{m \in \mathbb{Z}_N} \left| \int_{\Omega_0^R} (1 - e^{i\bar{x} \cdot (\xi - \xi_N)}) v(x) \varphi_\xi^m dx \right|^2 \\ &\leq \left\| (1 - e^{i\bar{x} \cdot (\xi - \xi_N)}) v \right\|_{L^2(\Omega_0^R)}^2 \leq \left\| 1 - e^{i\bar{x} \cdot (\xi - \xi_N)} \right\|_{L^\infty(\Omega_0^R)}^2 \|v\|_{L^2(\Omega_0^R)}^2 \end{aligned}$$

and

$$\begin{aligned} \left\| \sum_{m \in \mathbb{Z}_N} \widehat{v}(m; \xi_N) (\varphi_\xi^m - \varphi_{\xi_N}^m) \right\|_{L^2(\Omega_0^R)}^2 &= \left\| (e^{i\bar{x} \cdot (\xi - \xi_N)} - 1) \sum_{m \in \mathbb{Z}_N} \widehat{v}(m; \xi_N) \varphi_{\xi_N}^m \right\|_{L^2(\Omega_0^R)}^2 \\ &\leq \left\| e^{i\bar{x} \cdot (\xi - \xi_N)} - 1 \right\|_{L^\infty(\Omega_0^R)}^2 \|P_{\xi_N}^N v\|_{L^2(\Omega_0^R)}^2 \leq \left\| 1 - e^{i\bar{x} \cdot (\xi - \xi_N)} \right\|_{L^\infty(\Omega_0^R)}^2 \|v\|_{L^2(\Omega_0^R)}^2. \end{aligned}$$

We also see that

$$\left\| 1 - e^{i\bar{x} \cdot (\xi - \xi_N)} \right\|_{L^\infty(\Omega_0^R)}^2 = \left\| 2 \sin \frac{\bar{x} \cdot (\xi_N - \xi)}{2} \right\|_{L^\infty(\Omega_0^R)}^2 \xrightarrow{N \rightarrow \infty} 0,$$

which proves the lemma.  $\square$

### 3.4 Discretization of the locally perturbed periodic problem and convergence analysis

We now address the discretization of the original problem (3.9). Let us set again  $f_\xi = \mathcal{F}f(\cdot, \xi)$ . The idea is to perform first a discretization with respect to the Floquet-Bloch variable, then perform the discretization in space as done in the previous section. We recall that  $\tilde{u} \in (L^2(\llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket), H_\xi^2(\Omega_0))$  and

$$\begin{cases} \Delta \tilde{u}(\cdot; \xi) + k^2 n_p \tilde{u}(\cdot; \xi) + k^2 (n - n_p) \mathcal{M}(\tilde{u}) = f_\xi & \text{in } \Omega_0, \quad \xi \in \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket \\ \mathcal{M}(\tilde{u}) := \frac{\llbracket L \rrbracket}{(2\pi)^{d-1}} \int_{\llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket} \tilde{u}(\cdot; \xi) d\xi & \text{in } \Omega_0, \end{cases} \quad (3.56)$$

and that  $u$  can be reconstructed from  $\tilde{u}$  using (3.10).

### 3.4.1 Discretization and convergence in the Floquet-Bloch variable

We discretize the integral  $\mathcal{M}(\tilde{u})$  using the trapezoidal rule. Consider a uniform partition of  $\llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket$  into  $\prod_{\ell=1}^{d-1} M_\ell$  sub-domains of size  $\Delta\xi_1 \times \dots \times \Delta\xi_{d-1}$  with  $\Delta\xi_\ell := 2\pi/(M_\ell L_\ell)$ . Set  $M := (M_1, \dots, M_{d-1}) \in \mathbb{Z}_+^{d-1}$  and define  $\mathbb{Z}_M^{d-1} := \{j = (j_1, \dots, j_{d-1}), [-\frac{M_\ell}{2}] + 1 \leq j_\ell \leq [\frac{M_\ell}{2}], \ell = 1, \dots, d-1\}$ . Here we use the notation  $[\cdot]$  to denote the floor function and  $[a] = ([a_1], \dots, [a_d])$  if  $a = (a_1, \dots, a_d) \in \mathbb{R}^d$ . The discretization points are

$$\xi_j = (j_1 \Delta\xi_1, \dots, j_{d-1} \Delta\xi_{d-1}); \quad j = (j_1, \dots, j_{d-1}) \in \mathbb{Z}_M^{d-1}.$$

Denote by  $\tilde{u}_M(\cdot; \xi_j)$  an approximation of  $\tilde{u}(\cdot, \xi_j)$ . Then, using the fact that  $\tilde{u}(\cdot, -\pi/L) = \tilde{u}(\cdot, \pi/L)$ , we set as discretized equations associated with (3.56) the following system:  $\tilde{u}_M(\cdot; \xi_j) \in H_{\xi_j}^2(\Omega_0)$

$$\begin{cases} \Delta \tilde{u}_M(\cdot; \xi_j) + k^2 n_p \tilde{u}_M(\cdot; \xi_j) + k^2(n - n_p)u_M = f_{\xi_j} & \text{in } \Omega_0, \quad j \in \mathbb{Z}_M^{d-1} \\ u_M := \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} \tilde{u}_M(\cdot; \xi_j), & \text{in } \Omega_0. \end{cases} \quad (3.57)$$

Since  $\text{supp}((n - n_p)u_M) \subset \Omega_0^h$ , one deduces (using similar arguments as in establishing (3.17)) that system (3.57) is equivalent to

$$\begin{cases} \tilde{u}_M(\cdot; \xi_j) - k^2 V_{\xi_j}((n_p - 1)\tilde{u}_M(\cdot; \xi_j)) - k^2 V_{\xi_j}((n - n_p)u_M) = -V_{\xi_j}(f_{\xi_j}) & \text{in } L^2(\Omega_0^h), \\ u_M := \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} \tilde{u}_M(\cdot; \xi_j) & \text{in } \Omega_0. \end{cases} \quad (3.58)$$

We will prove later that the function  $u_M$  constitutes in fact an approximation of  $u$  in  $\Omega_0^h$ . We recall that  $\tilde{u}_M(\cdot; \xi_j)$  can be extended to  $\Omega_0$  using the first equation in (3.58) and then extended to all  $\mathbb{R}^d$  using the quasi-periodicity property as

$$\tilde{u}_M((\bar{x} + mL, x_d); \xi_j) := e^{i(mL) \cdot \xi_j} \tilde{u}_M((\bar{x}, x_d); \xi_j), \quad x = (\bar{x}, x_d) \in \Omega_0, \quad m \in \mathbb{Z}^{d-1}. \quad (3.59)$$

Then we extend  $u_M$  to  $\mathbb{R}^d$  using the same formula as in the second equation of (3.58), namely

$$u_M := \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} \tilde{u}_M(\cdot; \xi_j) \quad \text{in } \mathbb{R}^d.$$

We also define  $f_M$  as

$$f_M := \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} f_{\xi_j} \quad \text{in } \mathbb{R}^d.$$

In particular,

$$u_M(\bar{x} + mL, x_d) = \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} e^{i(mL) \cdot \xi_j} \tilde{u}_M((\bar{x}, x_d); \xi_j), \quad x = (\bar{x}, x_d) \in \Omega_0, \quad m \in \mathbb{Z}^{d-1}. \quad (3.60)$$

The following notation will be used

$$\Omega_M^h := \cup_{m \in \mathbb{Z}_M^{d-1}} \Omega_m^h,$$

$$\Omega_M := \cup_{m \in \mathbb{Z}_M^{d-1}} \Omega_m = \llbracket [-\frac{M}{2}] + \frac{1}{2} \rrbracket L, ([\frac{M}{2}] + \frac{1}{2})L \rrbracket \times \mathbb{R},$$

$$\Gamma_{\ell, M}^- := \{x \in \Omega_M, x_\ell = ([-\frac{M_\ell}{2}] + \frac{1}{2})L_\ell\}, \quad \Gamma_{\ell, M}^+ := \{x \in \Omega_M, x_\ell = ([\frac{M_\ell}{2}] + \frac{1}{2})L_\ell\}, \quad \forall \ell = 1, \dots, d-1$$

and the term  $ML$ -periodic will refer to periodic functions with respect to the first  $d - 1$  variables with period  $ML$ . To distinguish the notation for the space of  $ML$ -periodic we shall set, for  $m \in \mathbb{N}$ ,

$$H_{\#,M}^m(\mathbb{R}^d) := H_{0,ML}^m(\mathbb{R}^d) \quad \text{and} \quad H_{\#}^m(\Omega_M) := \{u|_{\Omega_M}; u \in H_{\#,M}^m(\mathbb{R}^d)\}.$$

We shall make extensive use of the following identity (that somehow occurs naturally from the conjugation of the  $\xi$ -quasi-periodicity with the use of the trapezoidal rule). For  $\ell, \ell' \in \mathbb{Z}_M^{d-1}$

$$\sum_{m \in \mathbb{Z}_M^{d-1}} e^{-i(mL) \cdot (\xi_\ell - \xi_{\ell'})} = \sum_{m \in \mathbb{Z}_M^{d-1}} e^{-i(m \frac{2\pi}{M}) \cdot (\ell - \ell')} = \begin{cases} 0 & \text{if } \ell \neq \ell' \\ \llbracket M \rrbracket & \text{if } \ell = \ell' \end{cases}. \quad (3.61)$$

**Lemma 3.4.1.** *Let us define  $G_{\#,M} := \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} G_{\xi_j}$ . Then  $G_{\#,M} \in L_{\#,M}^2(\mathbb{R}^d)$  and satisfies:*

$$\Delta G_{\#,M} + k^2 G_{\#,M} = -\delta_{x_0} \quad \text{in } \Omega_M. \quad (3.62)$$

*Proof.* This lemma can be proved using the Fourier representation of the  $\xi$ -quasi periodic Green function. Recall that the  $ML$ -periodic Green's function, denoted by  $G_{ML}$ , has the series representation (apply (3.15) with  $\xi = 0$  and  $L = ML$ )

$$G_{ML}(x) = \frac{i}{2\llbracket ML \rrbracket} \sum_{j \in \mathbb{Z}^{d-1}} \frac{1}{\beta_{\#}(j)} \exp(i\alpha_{\#}(j) \cdot \bar{x} + i\beta_{\#}(j)|x_d|), \quad x = (\bar{x}, x_d) \in \mathbb{R}^d \quad (3.63)$$

where  $\alpha_{\#}(j) := \frac{2\pi}{ML}j$ ,  $\beta_{\#}(j) := \sqrt{k^2 - |\alpha_{\#}(j)|^2}$ . Also recall that the  $\xi$ -quasi periodic Green's function  $G_{\xi}$  has the series representation (3.15). Then, forming the expression of  $G_{\#,M}$  we deduce

$$\begin{aligned} G_{\#,M} &= \frac{1}{\llbracket M \rrbracket} \frac{i}{2\llbracket L \rrbracket} \sum_{\ell \in \mathbb{Z}_M^{d-1}} \sum_{j \in \mathbb{Z}^{d-1}} \frac{1}{\sqrt{k^2 - (\xi_\ell + \frac{2\pi}{L}j)^2}} \exp(i(\xi_\ell + \frac{2\pi}{L}j) \cdot \bar{x} + i\sqrt{k^2 - (\xi_\ell + \frac{2\pi}{L}j)^2}|x_d|) \\ &= \frac{i}{2\llbracket ML \rrbracket} \sum_{\ell \in \mathbb{Z}_M^{d-1}} \sum_{j \in \mathbb{Z}^{d-1}} \frac{1}{\sqrt{k^2 - (\frac{2\pi}{L}J(\ell, j))^2}} \exp(i\frac{2\pi}{L}J(\ell, j) \cdot \bar{x} + i\sqrt{k^2 - (\frac{2\pi}{L}J(\ell, j))^2}|x_d|), \end{aligned}$$

where  $J(\ell, j) := \ell + Mj$ . We see that,  $\text{span}\{J(\ell, j), \ell \in \mathbb{Z}_M^{d-1}, j \in \mathbb{Z}^{d-1}\} = \mathbb{Z}^{d-1}$  and  $J(\ell_1, j_1) = J(\ell_2, j_2)$  if and only if  $j_1 = j_2$  and  $\ell_1 = \ell_2$ . Therefore, reordering the summation, we get

$$G_{\#,M} = \frac{i}{2\llbracket ML \rrbracket} \sum_{J \in \mathbb{Z}^{d-1}} \frac{1}{\beta_{\#}(J)} \exp(i\alpha_{\#}(J) \cdot \bar{x} + i\beta_{\#}(J)|x_d|) \quad (3.64)$$

which clearly proves that  $G_{\#,M} = G_{ML}$ . □

We now introduce the potential  $V_{\#,M}$  with kernel  $G_{\#,M}$ :

$$V_{\#,M}f(x) := \int_{\Omega_M^h} G_{\#,M}(x-y)f(y) \, dy. \quad (3.65)$$

We have the following properties for the operator  $V_{\#,M}$ .

**Theorem 3.4.2.** *1. The volume potential  $V_{\#,M}$  is a linear bounded operator from  $L^2(\Omega_M)$  into  $H_{\#,M}^2(\mathbb{R}^2)$  and for all  $g \in L^2(\Omega_M)$ ,  $u := V_{\#,M}g \in H_{\#}^2(\Omega_M)$  is the unique variational solution to:  $\Delta u + k^2 u = -g$  in  $\Omega_M$ .*

2. Let  $\tilde{u}_M(\cdot; \xi_j)$  be the solution to system (3.58). Then the function  $u_M$  defined in (3.60) belongs to  $H_{\#}^1(\Omega_M)$  and verifies the following volume integral equation

$$u_M - k^2 V_{\#,M}((n-1)u_M) = -V_{\#,M}(f_M) \quad \text{in } L^2(\Omega_M). \quad (3.66)$$

*Proof.* Proving the first claim can be done similarly to Lemma (3.3.1). We now prove the second claim. We first introduce the operator  $\tilde{V}_{\xi} : L^2(\Omega_M^h) \rightarrow L^2(\Omega_M^h)$ :

$$\tilde{V}_{\xi} g(x) := \int_{\Omega_M^h} G_{\xi}(x-y)g(y) \, dy. \quad (3.67)$$

Then,  $V_{\#,M} = \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} \tilde{V}_{\xi_j}$  and if  $g \in L^2(\mathbb{R}^2)$  with compact support in  $\Omega_0^h$ ,

$$\tilde{V}_{\xi_{\ell}} g(x) = \int_{\Omega_M^h} G_{\xi_{\ell}}(x-y)g(y) \, dy = \int_{\Omega_0^h} G_{\xi_{\ell}}(x-y)g(y) \, dy = V_{\xi_{\ell}} g(x), \quad \forall x \in \mathbb{R}^2. \quad (3.68)$$

We now observe that for  $\tilde{v}(\cdot; \xi_{\ell'})$  in  $L^2(\Omega_M)$  a  $\xi_{\ell'}$ -quasi-periodic function with period  $L$ , we have

$$\tilde{V}_{\xi_{\ell}}(\tilde{v}(\cdot; \xi_{\ell'})) = \begin{cases} \llbracket M \rrbracket V_{\xi_{\ell}}(\tilde{v}(\cdot; \xi_{\ell})) & \text{if } \ell' = \ell, \\ 0 & \text{if } \ell' \neq \ell. \end{cases} \quad (3.69)$$

To prove this claim we write,

$$\begin{aligned} \tilde{V}_{\xi_{\ell}}(\tilde{v}(\cdot; \xi_{\ell'}))(x) &= \int_{\Omega_M^h} G_{\xi_{\ell}}(x-y)\tilde{v}(y; \xi_{\ell'}) \, dy = \sum_{m \in \mathbb{Z}_M^{d-1}} \int_{\Omega_m^h} G_{\xi_{\ell}}(x-y)\tilde{v}(y; \xi_{\ell'}) \, dy \\ &= \sum_{m \in \mathbb{Z}_M^{d-1}} \int_{\Omega_0^h} G_{\xi_{\ell}}(x-y)\tilde{v}(y; \xi_{\ell'}) e^{-i(mL) \cdot (\xi_{\ell} - \xi_{\ell'})} \, dy \\ &= \int_{\Omega_0^h} G_{\xi_{\ell}}(x-y)\tilde{v}(y; \xi_{\ell'}) \, dy \left( \sum_{m \in \mathbb{Z}_M^{d-1}} e^{-i(mL) \cdot (\xi_{\ell} - \xi_{\ell'})} \right). \end{aligned}$$

Using property (3.61) we then obtain (3.69). Since  $n - n_p$  has compact support in  $\Omega_0^h$ , then from (3.68) we have:

$$V_{\xi_{\ell}}((n - n_p)u_M) = \tilde{V}_{\xi_{\ell}}((n - n_p)u_M). \quad (3.70)$$

Since  $(n_p - 1)\tilde{u}_M(\cdot; \xi_{\ell})$  is  $\xi_{\ell}$ -quasi periodic with period  $L$ , then from (3.69) we have:

$$V_{\xi_{\ell}}((n_p - 1)\tilde{u}_M(\cdot; \xi_{\ell})) = \frac{1}{\llbracket M \rrbracket} \tilde{V}_{\xi_{\ell}}((n_p - 1)\tilde{u}_M(\cdot; \xi_{\ell})) \quad \text{and} \quad \tilde{V}_{\xi_{\ell}}((n_p - 1)\tilde{u}_M(\cdot; \xi_{\ell'})) = 0, \quad \ell \neq \ell'.$$

Therefore,

$$V_{\xi_{\ell}}((n_p - 1)\tilde{u}_M(\cdot; \xi_{\ell})) = \tilde{V}_{\xi_{\ell}}((n_p - 1)u_M). \quad (3.71)$$

Similarly we also get

$$V_{\xi_{\ell}} f_{\xi_{\ell}} = \tilde{V}_{\xi_{\ell}} f_M.$$

Therefore

$$\tilde{u}_M(\cdot; \xi_{\ell}) - k^2 \tilde{V}_{\xi_{\ell}}((n_p - 1)u_M) - k^2 \tilde{V}_{\xi_{\ell}}((n - n_p)u_M) = -\tilde{V}_{\xi_{\ell}} f_M \quad \text{in } \Omega_M,$$

in other words,

$$\tilde{u}_M(\cdot; \xi_\ell) - k^2 \tilde{V}_{\xi_\ell}((n-1)u_M) = -\tilde{V}_{\xi_\ell} f_M, \quad \text{in } \Omega_M. \quad (3.72)$$

Taking the sum with respect to  $\ell \in \mathbb{Z}_M^{d-1}$  we finally obtain

$$u_M - k^2 V_{\#,M}((n-1)u_M) = -V_{\#,M} f_M, \quad (3.73)$$

which proves the theorem.  $\square$

The following Corollary can be immediately deduced from Theorem 3.4.2, Lemma 3.4.3 and estimate (3.2).

**Corollary 3.4.3.** *The function  $u_M \in H_{\#}^2(\Omega_M)$  and is solution to*

$$\Delta u_M + k^2 n u_M = f_M \quad \text{in } \Omega_M. \quad (3.74)$$

Furthermore, there exists a positive constant  $C$  independent from  $M$  such that

$$\|u_M\|_{H_{\#}^2(\Omega_M)} \leq C \|f_M\|_{L^2(\Omega_M^h)}. \quad (3.75)$$

Due to the non vanishing imaginary part of  $k^2$ , we now can formulate the following exponential convergence result with respect to  $M$  that is a consequence of the exponential decay of the solution  $u$ . A proof of this theorem can also be found in [42]. For the reader's convenience, we here give a sketch of the proof.

**Theorem 3.4.4.** *Let  $f \in L^2(\mathbb{R}^d)$  such that  $e^{\tau^* \|x\|} f \in L^2(\mathbb{R}^d)$  for some  $\tau^* > 0$ . Let  $u \in H^1(\mathbb{R}^d)$  be the solution of (3.1) and  $u_M \in H_{\#}^1(\Omega_M)$  be the solution of (3.74). Then there exists  $\tau_0 > 0$  and a constant  $C$  independent of  $M$  and  $f$  and  $\tau$  such that*

$$\|u - u_M\|_{H^1(\Omega_M)} \leq C \left( e^{-\tau \min(ML)/2} \|e^{\tau \|x\|} f\|_{L^2(\mathbb{R}^d)} + \|f - f_M\|_{L^2(\mathbb{R}^d)} \right) \quad \forall 0 \leq \tau \leq \tau_0. \quad (3.76)$$

*Proof.* The first observation is that, if we set  $v = e^{\tau \|x\|} u$ , then  $v$  satisfies

$$\Delta v - 2\tau \frac{x}{\|x\|} \cdot \nabla v + \left( k^2 n + \tau^2 - \tau \frac{d-1}{\|x\|} \right) v = e^{\tau \|x\|} f \quad \text{in } \mathbb{R}^d. \quad (3.77)$$

It is easy to see (similarly to (3.2)) by taking the imaginary part of the variational formulation of the previous equation that

$$\|v\|_{L^2(\mathbb{R}^d)} \leq \frac{1}{\sigma c_0} (\|e^{\tau \|x\|} f\|_{L^2(\mathbb{R}^d)} + 2\tau \|\nabla v\|_{L^2(\mathbb{R}^d)})$$

and using the real part of the variational formulation (observe that  $d-1 \geq 0$ )

$$\|\nabla v\|_{L^2(\mathbb{R}^d)}^2 \leq \left( \|e^{\tau \|x\|} f\|_{L^2(\mathbb{R}^d)} \|v\|_{L^2(\mathbb{R}^d)} + 2\tau \|\nabla v\|_{L^2(\mathbb{R}^d)} \right) \|v\|_{L^2(\mathbb{R}^d)} + |k^2 n + \tau^2| \|v\|_{L^2(\mathbb{R}^d)}^2.$$

Let us define for some positive  $\varepsilon$

$$B_\varepsilon := \{x \in \mathbb{R}^d : \|x\| \leq \varepsilon\}$$

such that  $B_\varepsilon \subset \Omega_M$ . Using again equation (3.77) and choosing  $\tau_0 > 0$  such that  $1 - \frac{2(2\tau_0)^2}{(\sigma c_0)^2} (\sigma c_0 + |k^2 n + \tau_0^2|) > 1/2$  and  $\tau_0 < \tau^*$  we get the existence of a constant  $C(\tau_0)$  independent from  $\tau$  such that

$$\|\Delta v\|_{L^2(\mathbb{R}^d \setminus B_\varepsilon)}^2 + \|v\|_{H^1(\mathbb{R}^d)}^2 \leq C \|e^{\tau \|x\|} f\|_{L^2(\mathbb{R}^d)}^2 \quad \forall 0 \leq \tau \leq \tau_0. \quad (3.78)$$

Let us denote by  $\phi_j^\pm := u|_{\Gamma_{j,M}^\pm}$  and  $\psi_j := \frac{\partial}{\partial x_j} u|_{\Gamma_{j,M}^\pm}$ . Then, from trace theorems and since  $\Gamma_{j,M}^\pm$  can be identified as a part of  $\mathbb{R}^{d-1}$ , there exists a constant  $C$  independent from  $j$  such that

$$e^{\tau(M_j \pm 1)L_j} (\|\phi_j^\pm\|_{H^{1/2}(\Gamma_{j,M}^\pm)}^2 + \|\psi_j^\pm\|_{H^{-1/2}(\Gamma_{j,M}^\pm)}^2) \leq C(\|\Delta v\|_{L^2(\mathbb{R}^d \setminus B_\varepsilon)}^2 + \|v\|_{H^1(\mathbb{R}^d \setminus B_\varepsilon)}^2). \quad (3.79)$$

Since  $w := u - u_M$  solves

$$\begin{cases} \Delta w + k^2 n w = f - f_M & \text{in } \Omega_M, \\ w|_{\Gamma_{j,M}^+} - w|_{\Gamma_{j,M}^-} = \phi_j^+ - \phi_j^-, \\ \frac{\partial}{\partial x_1} w|_{\Gamma_{j,M}^+} - \frac{\partial}{\partial x_1} w|_{\Gamma_{j,M}^-} = \psi_j^+ - \psi_j^-, \end{cases}$$

then, due to  $\sigma > 0$ , there exists constant  $C' > 0$  independent from  $M$  such that

$$\|w\|_{H^1(\Omega_M)}^2 \leq C' \left( \sum_j (\|\phi_j^\pm\|_{H^{1/2}(\Gamma_{j,M}^\pm)}^2 + \|\psi_j^\pm\|_{H^{-1/2}(\Gamma_{j,M}^\pm)}^2) + \|f - f_M\|_{L^2(\Omega_M)} \right). \quad (3.80)$$

The result directly follows from combining (3.78), (3.79) and (3.80).  $\square$

### 3.4.2 Discretization in the spatial variable

The discretization procedure in the spatial variable is the same as the one discussed in Section 3.3.

#### Periodization of the coupled integral equations

To apply our method we first periodize the volume integral equations of system (3.58) in  $x_d$ -direction and get the following system.

$$\begin{cases} \tilde{u}_M^R(\cdot; \xi_j) - k^2 V_{\xi_j}^R((n_p - 1)\tilde{u}_M^R(\cdot; \xi_j)) - k^2 V_{\xi_j}^R((n - n_p)u_M^R) = -V_{\xi_j}^R f_{\xi_j} & \text{in } L^2(\Omega_0^R), \\ u_M^R = \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} \tilde{u}_M^R(\cdot; \xi_j). \end{cases} \quad (3.81)$$

From Lemma 3.3.2, the solution to equation (3.58) and the solution to equation (3.81) coincides in  $\Omega_0^h$ . Therefore, the extension by quasi-periodicity of the solution to equation (3.58) coincide with the extension by quasi-periodicity of the solution to equation (3.81) in  $\Omega_M^h$ . Let

$$G_{\#,M}^R := \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} G_{\xi_j}^R$$

be the periodization of  $G_{\#,M}$  in  $x_d$  direction. We introduce the volume potential  $V_{\#,M}^R$  with kernel  $G_{\#,M}^R$  as:

$$V_{\#,M}^R f(x) := \int_{\Omega_M^h} G_{\#,M}^R(x - y) f(y) dy. \quad (3.82)$$

Then, following the same arguments as in the proof of the second item in Theorem 3.4.2, we obtain that  $u_M^R \in L^2(\Omega_M^R)$  verifies:

$$u_M^R = k^2 V_{\#,M}^R((n - 1)u_M^R) - V_{\#,M}^R f_M \quad \text{in } L^2(\Omega_M^R). \quad (3.83)$$

The link between the solution of the periodized equation (3.83) and the solution of (3.66) can be established in a similar way as for  $\xi$ -quasi-periodic problems (apply the procedure with  $\xi = 0$  and a period =  $ML$ ). We here regroup these results:



**Proposition 3.4.5.** • Equation (3.66) has a unique solution  $u_M \in L^2(\Omega_M^h)$  for all  $f_M \in L^2(\Omega_M^h)$  if and only if equation (3.83) has a unique solution  $u_M^R \in L^2(\Omega_M^R)$  for all  $f_M \in L^2(\Omega_M^h)$ .

- Let  $f_M \in L^2(\Omega_M^h)$ ,  $u_M^R \in L^2(\Omega_M^R)$  be the solution to (3.83) and  $u_M \in L^2(\Omega_M^h)$  be the solution to (3.66). Then

$$u_M^R = u_M \quad \text{in } \Omega_M^h.$$

Moreover  $u_M^R \in H_{\#}^2(\Omega_M^R)$  and there exists a positive constant  $C$  independent of  $M$  and  $f_M$  such that:

$$\|u_M^R\|_{H_{\#}^2(\Omega_M^R)} \leq C \|f_M\|_{L^2(\Omega_M^h)}. \quad (3.84)$$

The definition of  $H_{\#}^s(\Omega_M^R)$  is the same as  $H_{\xi}^s(\Omega_0^R)$  with  $\xi = 0$  and the period  $L$  replaced with  $ML$ . The uniform bound with respect to  $M$  is a consequence of the following lemma that can be proved in the same manner as in Lemma 3.3.3 using (3.26).

**Lemma 3.4.6.** Let  $s \in \mathbb{N}$ . The operator  $V_{\#,M}^R : H_{\#}^s(\Omega_M^R) \rightarrow H_{\#}^{s+2}(\Omega_M^R)$  is continuous. Moreover, there exists a constant  $C$  independent of  $M$  such that

$$\|V_{\#,M}^R g\|_{H_{\#}^{2+s}(\Omega_M^R)} \leq C \|g\|_{H_{\#}^s(\Omega_M^R)}, \quad \forall g \in H_{\#}^s(\Omega_M^R). \quad (3.85)$$

Let us introduce for later use in the convergence analysis  $T_{\#,M}^R : L^2(\Omega_M^R) \rightarrow L^2(\Omega_M^R)$  defined by

$$T_{\#,M}^R g := k^2 V_{\#,M}^R ((n_p - 1)g) \quad (3.86)$$

and  $A_{\#,M}^R : L^2(\Omega_M^R) \rightarrow L^2(\Omega_M^R)$  defined by

$$A_{\#,M}^R g := g - T_{\#,M}^R g. \quad (3.87)$$

Then, as straightforward consequences of Lemma 3.4.6 and the first item in Proposition 3.4.5 we get the following.

**Lemma 3.4.7.** The operator  $T_{\#,M}^R : L^2(\Omega_M^R) \rightarrow L^2(\Omega_M^R)$  is compact. The operator  $A_{\#,M}^R : L^2(\Omega_M^R) \rightarrow L^2(\Omega_M^R)$  is bounded and injective. Moreover, there exists a positive constant  $C$  independent of  $M$  such that

$$\|A_{\#,M}^R g\|_{L^2(\Omega_M^R)} \leq C \|g\|_{L^2(\Omega_M^R)}.$$

### Spectral approximation

Let  $\tilde{u}_M^R(\cdot; \xi_j) \in L^2(\Omega_0^R)$  be the solution of (3.81). We consider  $\tilde{u}_M^{R,N}(\cdot; \xi_j) \in \mathcal{T}_{\xi_j}^N$  an approximation of  $\tilde{u}_M^R(\cdot; \xi_j)$ , which is solution to following variational equations

$$\begin{aligned} & \left\langle \tilde{u}_M^{R,N}(\cdot; \xi_j), v_{\xi_j}^N \right\rangle_{L^2(\Omega_0^R)} - k^2 \left\langle V_{\xi_j}^R((n_p - 1)\tilde{u}_M^{R,N}(\cdot; \xi_j)), v_{\xi_j}^N \right\rangle_{L^2(\Omega_0^R)} \\ & - k^2 \left\langle V_{\xi_j}^R((n - n_p)(\Delta \xi \sum_{\ell}^{R,N}(\cdot; \xi_{\ell}))), v_{\xi_j}^N \right\rangle_{L^2(\Omega_0^R)} = - \left\langle V_{\xi_j}^R(f_{\xi_j}), v_{\xi_j}^N \right\rangle_{L^2(\Omega_0^R)}, \quad \forall v_{\xi_j}^N \in \mathcal{T}_{\xi_j}^N(\Omega_0^R). \end{aligned} \quad (3.88)$$

$u_M^{R,N}$  is defined as

$$u_M^{R,N} := \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} \tilde{u}_M^{R,N}(\cdot; \xi_j). \quad (3.89)$$

We now define  $\mathcal{T}_{\#, \xi_j}^N$  as the space of functions that are extensions by  $\xi_j$ -quasi-periodicity of functions in  $\mathcal{T}_{\xi_j}^N$  to the domain  $\Omega_M^R$ . We then set

$$\mathcal{T}_{\#, M}^N := \bigoplus_{j \in \mathbb{Z}_M^{d-1}} \mathcal{T}_{\#, \xi_j}^N.$$

We extend  $\tilde{u}_M^{R,N}(\cdot; \xi_j)$  by  $\xi_j$ -quasi-periodicity to  $\Omega_M^R$ , which then can be considered as element of  $\mathcal{T}_{\#, \xi_j}^N$ . The function  $u_M^{R,N}$  is then extended to  $\Omega_M^R$  using (3.89). We now prove that  $u_M^{R,N}$  is also an approximation of the solution  $u_M^R$  of equation (3.83) using the Fourier basis on  $L_{\#}^2(\Omega_M^R)$ .

We first observe that

$$\mathcal{T}_{\#, M}^N = \text{span}\{\varphi_{\#, M}^J, J \in \mathbb{Z}_{M,N}^2\} \quad (3.90)$$

where for  $J = (\bar{J}, J_d) \in \mathbb{Z}^{d-1} \times \mathbb{Z}$

$$\varphi_{\#, M}^J(\bar{x}, x_d) := \frac{1}{\sqrt{2R\llbracket ML \rrbracket}} \exp\left(i \left(\frac{2\pi}{ML} \bar{J}\right) \cdot \bar{x} + i \frac{\pi}{R} J_d x_d\right)$$

and where  $\mathbb{Z}_{M,N}^d \subset \mathbb{Z}^d$  is defined for  $N = (\bar{N}, N_d) \in \mathbb{Z}^{d-1} \times \mathbb{Z}$  and  $M \in \mathbb{Z}^{d-1}$  by

$$\mathbb{Z}_{M,N}^2 := \{J = (\bar{J}, J_d) \in \mathbb{Z}^{d-1} \times \mathbb{Z}; [-\frac{M}{2}] - \frac{MN}{2} + M + 1 \leq \bar{J} \leq [\frac{M}{2}] + \frac{MN}{2} \text{ and } -N_d/2 + 1 \leq J_d \leq N_d/2\}.$$

This is a consequence of the fact that for any  $J = (\bar{J}, J_d) \in \mathbb{Z}^{d-1} \times \mathbb{Z}$  we can uniquely associate a couple  $(j, \ell) \in \mathbb{Z}_N^d \times \mathbb{Z}_N^{d-1}$  such that

$$\bar{J} = \bar{j} + \ell M \text{ and } j_d = J_d \quad (3.91)$$

and that if (3.91) holds then

$$\varphi_{\#, M}^J = \frac{1}{\sqrt{\llbracket M \rrbracket}} \varphi_{\xi_\ell}^j \text{ in } \Omega_M^R. \quad (3.92)$$

We now can prove the following.

**Theorem 3.4.8.** *The function  $u_M^{R,N}$  defined in (3.89) belongs to  $\mathcal{T}_{\#, M}^N$  and verifies*

$$\begin{aligned} \left\langle u_M^{R,N}, v_M^N \right\rangle_{L^2(\Omega_M^R)} - k^2 \left\langle V_{\#, M}^R((n-1)u_M^{R,N}), v_M^N \right\rangle_{L^2(\Omega_M^R)} = \\ - \left\langle V_{\#, M}^R f_M, v_M^N \right\rangle_{L^2(\Omega_M^R)}, \forall v_M^N \in \mathcal{T}_{\#, M}^N. \end{aligned} \quad (3.93)$$

*Proof.* Similarly to the proof of Theorem 3.4.2 let us introduce the periodized potential  $\tilde{V}_{\xi_j}^R$  as

$$\tilde{V}_{\xi_j}^R g(x) := \int_{\Omega_M^R} G_{\xi_j}^R(x-y) g(y) \, dy. \quad (3.94)$$

Then, by definition,  $V_{\#, M}^R = \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} \tilde{V}_{\xi_j}^R$ . Since  $(n-n_p)u_M^{R,N}$  has support in  $\Omega_0^R$ , then

$$V_{\xi_j}^R \left( (n-n_p)u_M^{R,N} \right) = \tilde{V}_{\xi_j}^R \left( (n-n_p)u_M^{R,N} \right). \quad (3.95)$$

As in the proof of Theorem 3.4.2, the main ingredient here is that for  $\tilde{v}(\cdot; \xi_{\ell'})$  in  $L^2(\Omega_M^R)$  a  $\xi_{\ell'}$ -quasi-periodic function with period  $L$ , we have

$$\tilde{V}_{\xi_{\ell'}}^R(\tilde{v}(\cdot; \xi_{\ell'})) = \begin{cases} \llbracket M \rrbracket V_{\xi_{\ell'}}^R(\tilde{v}(\cdot; \xi_{\ell})) & \text{if } \ell' = \ell, \\ 0 & \text{if } \ell' \neq \ell. \end{cases}$$

The proof of this property is the same as the proof of (3.69) and relies on identity (3.61). We then get

$$\tilde{V}_{\xi_j}^R\left((n_p - 1)\tilde{u}_M^{R,N}(\cdot; \xi_j)\right) = \tilde{V}_{\xi_j}^R\left((n_p - 1)u_M^{R,N}\right) \quad \text{and} \quad \tilde{V}_{\xi_j}^R(f_{\xi_j}) = \tilde{V}_{\xi_j}^R(f_M). \quad (3.96)$$

Combined with equation (3.88), we obtain that  $\tilde{u}_M^{R,N}(\cdot; \xi_j)$  verifies

$$\begin{aligned} \left\langle \tilde{u}_M^{R,N}(\cdot; \xi_j), v_{\xi_j}^N \right\rangle_{L^2(\Omega_0^R)} - k^2 \left\langle \tilde{V}_{\xi_j}^R((n-1)u_M^{R,N}), v_{\xi_j}^N \right\rangle_{L^2(\Omega_0^R)} \\ = - \left\langle \tilde{V}_{\xi_j}^R f_M, v_{\xi_j}^N \right\rangle_{L^2(\Omega_0)}, \quad \forall v_{\xi_j}^N \in \mathcal{T}_{\xi_j}^N. \end{aligned} \quad (3.97)$$

Thanks to the the quasi-periodicity property we deduce that

$$\begin{aligned} \left\langle \tilde{u}_M^{R,N}(\cdot; \xi_j), v_{\xi_j}^N \right\rangle_{L^2(\Omega_m^R)} - k^2 \left\langle \tilde{V}_{\xi_j}^R((n-1)u_M^{R,N}), v_{\xi_j}^N \right\rangle_{L^2(\Omega_m^R)} \\ = - \left\langle \tilde{V}_{\xi_j}^R f_M, v_{\xi_j, M}^N \right\rangle_{L^2(\Omega_m)}, \quad \forall v_{\xi_j, M}^N \in \mathcal{T}_{\xi_j, M}^N(\Omega_m^R), \quad \forall m \in \mathbb{Z}_M^{d-1}. \end{aligned} \quad (3.98)$$

Taking the sum with respect to  $m$  we then obtain

$$\begin{aligned} \left\langle \tilde{u}_{\#, M}^{R,N}(\cdot; \xi_j), v_{\xi_j}^N \right\rangle_{L^2(\Omega_M^R)} - k^2 \left\langle \tilde{V}_{\xi_j}^R((n-1)u_M^{R,N}), v_{\xi_j}^N \right\rangle_{L^2(\Omega_M^R)} = \\ - \left\langle \tilde{V}_{\xi_j}^R f_M, v_{\xi_j}^N \right\rangle_{L^2(\Omega_M)}, \quad \forall v_{\xi_j}^N \in \mathcal{T}_{\#, \xi_j}^N. \end{aligned} \quad (3.99)$$

Using identity (3.61) we also get that if  $\tilde{v}(\cdot; \xi_j)$  in  $L^2(\Omega_M^R)$  is a  $\xi_j$ -quasi-periodic function with period  $L$  then

$$\left\langle \tilde{v}(\cdot; \xi_j), v_{\xi_{\ell}}^N \right\rangle_{L^2(\Omega_M^R)} = 0 \quad \text{if } \ell \neq j.$$

Therefore we can replace  $v_{\xi_j}^N$  in (3.99) with any test function  $v_M^N \in \mathcal{T}_{\#, M}^N$ . That is,

$$\begin{aligned} \left\langle \tilde{u}_M^{R,N}(\cdot; \xi_j), v_M^N \right\rangle_{L^2(\Omega_M^R)} - k^2 \left\langle \tilde{V}_{\xi_j}^R((n-1)u_M^{R,N}), v_M^N \right\rangle_{L^2(\Omega_M^R)} \\ = - \left\langle \tilde{V}_{\xi_j}^R f_M, v_M^N \right\rangle_{L^2(\Omega_M)}, \quad \forall v_M^N \in \mathcal{T}_{\#, M}^N. \end{aligned}$$

Now, taking the average with respect to  $j \in \mathbb{Z}_M^{d-1}$  we obtain (5.39).  $\square$

### Convergence analysis for spatial discretization

The convergence analysis is based on the interpretation of our discrete system (3.88) as the one obtained with the discretization of the periodic problem with period  $= ML$  (Theorem 3.4.8). Indeed the following convergence results for  $u_M^R$  implies uniform convergence results for each  $\xi_{\ell}$ -quasi-periodic component. The latter can be seen as a consequence of the results of Section 3.3.

**Theorem 3.4.9.** *Let  $u_M^R \in H^1(\Omega_M)$  and  $u_M^{R,N} \in \mathcal{T}_{\#,M}^N$  be solutions to (3.83) and (3.89) respectively. Then for all positive  $M \in \mathbb{N}^{d-1}$ , there exists  $N_0(M) \in \mathbb{N}^d$  such that for all  $N \in \mathbb{N}^d$ ,  $N > N_0$ ,  $u_M^{R,N}$  is uniquely defined and there exists a positive constant  $C$  independent from  $M$  and  $N$  such that*

$$\|u_M^R - u_M^{R,N}\|_{L^2(\Omega_M^R)} \leq C \inf_{v_M^N \in \mathcal{T}_M^N} \|u_M^R - v_M^N\|_{L^2(\Omega_M^R)}. \quad (3.100)$$

*Proof.* The proof of this Theorem can be seen as a consequence of Theorem 4.2.7 in [100]. The solution  $u_M^R$  of (3.83) satisfies

$$\langle A_{\#,M}^R u_M^R, v_M \rangle_{L^2(\Omega_M^R)} = \langle -V_{\#,M}^R f_M, v_M \rangle_{L^2(\Omega_M^R)}, \quad \forall v_M \in L^2(\Omega_M^R). \quad (3.101)$$

We then obtain from (5.39),

$$\langle A_{\#,M}^R (u_M^R - u_M^{R,N}), v_M^N \rangle_{L^2(\Omega_M^R)} = 0, \quad \forall v_M^N \in \mathcal{T}_{\#,M}^N. \quad (3.102)$$

Since  $u_M^{R,N} - v_M^N \in \mathcal{T}_{\#,M}^N$ , then  $\langle A_{\#,M}^R (u_M^R - u_M^{R,N}), u_M^{R,N} - v_M^N \rangle_{L^2(\Omega_M^R)} = 0$ .

$$\begin{aligned} \langle A_{\#,M}^R (u_M^R - u_M^{R,N}), u_M^R - u_M^{R,N} \rangle_{L^2(\Omega_M^R)} &= \langle A_{\#,M}^R (u_M^R - u_M^{R,N}), u_M^R - v_M^N \rangle_{L^2(\Omega_M^R)} \\ &\leq C \|u_M^R - u_M^{R,N}\|_{L^2(\Omega_M^R)} \|u_M^R - v_M^N\|_{L^2(\Omega_M^R)}. \end{aligned} \quad (3.103)$$

For a fixed  $M$ , let us define

$$w_M^{R,N} := \frac{u_M^R - u_M^{R,N}}{\|u_M^R - u_M^{R,N}\|_{L^2(\Omega_M^R)}}, \quad (3.104)$$

For a given  $\varepsilon > 0$ , there exists  $N_0(M) \in \mathbb{N}^d$  such that for all  $N > N_0$

$$|\langle T_{\#,M}^R w_M^{R,N}, w_M^{R,N} \rangle| < \varepsilon. \quad (3.105)$$

This claim is proved by contradiction. Suppose that there exists a sequence  $w_M^{R,N'}$  such that  $|\langle T_{\#,M}^R w_M^{R,N'}, w_M^{R,N'} \rangle| \geq \varepsilon$ . We see that  $(w_M^{R,N'})$  is bounded in  $L^2(\Omega_M^R)$  and therefore there exists a subsequence, which we denoted by  $(w_M^{R,N'})$ , that converges weakly to  $w \in L^2(\Omega_M^R)$ . Observe that (by use of (5.39))

$$\langle A_{\#,M}^R w_M^{R,N'}, v \rangle_{L^2(\Omega_M^R)} = \langle A_{\#,M}^R w_M^{R,N'}, v - P_{\#}^{N'} v \rangle_{L^2(\Omega_M^R)}$$

where  $P_{\#}^{N'}$  denotes the projection operator from  $L^2(\Omega_M^R)$  onto  $\mathcal{T}_{\#,M}^{N'}$ . Taking the limit as  $N' \rightarrow \infty$  implies that  $w = 0$  by the injectivity of  $A_{\#,M}^R$ . The compactness of  $T_{\#,M}^R$  then implies  $\langle T_{\#,M}^R w_M^{R,N'}, w_M^{R,N'} \rangle \rightarrow \langle T_{\#,M}^R w, w \rangle = 0$ . This is a contradiction with  $\varepsilon > 0$ . Now, we choose  $\varepsilon = \frac{1}{2}$  in (3.105) and use that  $A_{\#,M}^R = I - T_{\#,M}^R$  to deduce from (3.103) that:

$$\frac{1}{2} \|u_M^R - u_M^{R,N}\|_{L^2(\Omega_M^R)}^2 \leq C \|u_M^R - u_M^{R,N}\|_{L^2(\Omega_M^R)} \|u_M^R - v_M^N\|_{L^2(\Omega_M^R)}.$$

Thus,

$$\|u_M^R - u_M^{R,N}\|_{L^2(\Omega_M^R)} \leq C \inf_{v_M^N \in \mathcal{T}_M^N} \|u_M^R - v_M^N\|_{L^2(\Omega_M^R)}. \quad (3.106)$$

□

It is now possible to immediately deduce the convergence rate thanks to the regularity result of Proposition 3.3.4 and using the interpolation result of Lemma 3.3.8 with  $\xi = 0$ ,  $L$  replaced with  $ML$  and  $N_\ell$  replaced with  $M_\ell N_\ell$  ( $\ell = 1, d-1$ ).

**Theorem 3.4.10.** *Let  $u_M^R \in H^1(\Omega_M)$  and  $u_M^{R,N} \in \mathcal{T}_{\#,M}^N$  be verifying (3.83) and (3.89) respectively. Then for all positive  $M \in \mathbb{N}^{d-1}$ , there exists  $N_0(M) \in \mathbb{N}^d$  and a positive constant  $C$  independent from  $M$  such that for all  $N \in \mathbb{N}^d$ ,  $N > N_0$ ,*

$$\|u_M^R - u_M^{R,N}\|_{L^2(\Omega_M^R)} \leq C \left( \sum_{\ell=1}^d \frac{1}{(1 + N_\ell^2)^2} \right)^{1/2} \|f_M\|_{L^2(\Omega_M^h)}. \quad (3.107)$$

### A convergence result for a combined discretization

We are now in position to combine the previous convergence results and deduce a convergence result for the discretization of our original problem. Let  $f \in L^2(\mathbb{R}^d)$  with support in  $\Omega^h$  and let  $u \in H^1(\mathbb{R}^d)$  be the solution to (3.1) and  $u_M^{R,N} \in \mathcal{T}_{\#,M}^N$  be verifying (3.89). We shall assume that there exists  $\tau_* > 0$  such that

$$e^{\tau_* \|x\|} f \in L^2(\mathbb{R}^d). \quad (3.108)$$

To obtain the following convergence result, we first write the decomposition

$$u - u_M^{R,N} = (u - u_M) + (u_M - u_M^{R,N}).$$

We apply Theorem 3.4.4 to treat the term  $u - u_M$  and for the second term we first use Proposition 3.4.5 to see that

$$u_M - u_M^{R,N} = u_M^R - u_M^{R,N} \text{ in } \Omega_M^h,$$

then apply Theorem 3.4.10.

**Theorem 3.4.11.** *For all positive  $M \in \mathbb{N}^{d-1}$ , there exists  $N_0(M) \in \mathbb{N}^d$  and two positive constants  $C$  and  $\tau_0$  independent from  $M$  such that for all  $N \in \mathbb{N}^d$ ,  $N > N_0$ ,*

$$\|u - u_M^{R,N}\|_{L^2(\Omega_M^h)} \leq C \left( e^{-\tau_0 \min(ML)/2} \|e^{\tau_0 \|x\|} f\|_{L^2(\mathbb{R}^d)} + \|f - f_M\|_{L^2(\Omega_M)} + \left( \sum_{\ell=1}^d \frac{1}{(1 + N_\ell^2)^2} \right)^{1/2} \|f\|_{L^2(\mathbb{R}^d)} \right).$$

Indeed in this theorem we used that  $\|f_M\|_{L^2(\Omega_M)} \leq \|f\|_{L^2(\mathbb{R}^d)}$ . We also notice that  $\|f - f_M\|_{L^2(\Omega_M)}$  converges exponentially fast to 0 as  $M \rightarrow \infty$  thanks to assumption (3.108). In the cases where  $f$  has compact support, assumption (3.108) is automatically verified and  $f_M = f$  in  $\Omega_M$  as soon as the support of  $f$  is contained in  $\Omega_M$ . Let us also note that the rate of convergence with respect to  $N$  can be higher if additional smoothness on the coefficients and the source term is imposed.

**Theorem 3.4.12.** *Assume in addition that  $n \in W^{m,\infty}(\mathbb{R}^d)$  and  $f \in H_{loc}^m(\mathbb{R}^d)$  for some integer  $m > 0$ . Then, for all positive  $M \in \mathbb{N}^{d-1}$ , there exists  $N_0(M) \in \mathbb{N}^d$  and two positive constants  $C$  and  $\tau_0$  independent from  $M$  such that for all  $N \in \mathbb{N}^d$ ,  $N > N_0$ ,*

$$\|u - u_M^{R,N}\|_{L^2(\Omega_M^h)} \leq C \left( e^{-\tau_0 \min(ML)/2} \|e^{\tau_0 \|x\|} f\|_{L^2(\mathbb{R}^d)} + \|f - f_M\|_{L^2(\Omega_M)} + \left( \sum_{\ell=1}^d \frac{1}{(1 + N_\ell^2)^{m+2}} \right)^{1/2} \|f_M\|_{H_{\#}^m(\Omega_M)} \right).$$

*Proof.* The proof of this theorem follows the same lines as the proof of theorem 3.4.11 and the observation and that under the regularity assumption of the theorem, there exists a constant independent from  $M$  such that

$$\|u_M^R\|_{H^{2+m}(\Omega_M^R)} \leq C \|f_M\|_{H_{\#}^m(\Omega_M)}.$$

The latter is a direct consequence of (3.83) and Lemma 3.4.6.  $\square$

### 3.5 Numerical Algorithm and Experiments

We first notice that for any function in  $L^2(\Omega_0^R)$ ,

$$P_{\xi_j}^N V_{\xi_j}^R v = V_{\xi_j}^R P_{\xi_j}^N v.$$

Therefore, the discrete system (3.88) can be written as

$$\begin{cases} \tilde{u}_M^{R,N}(\cdot; \xi_j) - k^2 V_{\xi_j}^R (P_{\xi_j}^N ((n_p - 1) \tilde{u}_M^{R,N}(\cdot; \xi_j))) - k^2 V_{\xi_j}^R (P_{\xi_j}^N (n - n_p) u_M^{R,N}) = -V_{\xi_j}^R (P_{\xi_j}^N f) \\ u_M^{R,N} = \frac{1}{M} \sum_{\ell \in \mathbb{Z}_M^d} \tilde{u}_M^{R,N}(\cdot; \xi_\ell). \end{cases} \quad (3.109)$$

The Fourier coefficients of  $V_{\xi_j}^R P_{\xi_j}^N v$  can be expressed as

$$\widehat{V_{\xi_j}^R P_{\xi_j}^N v}(\ell; \xi_j) = \frac{1}{\gamma} \sum_{\ell \in \mathbb{Z}_N^d} \widehat{G_{\xi_j}}(\ell; \xi_j) \widehat{v}(\ell; \xi_j). \quad (3.110)$$

The coefficients  $\widehat{G_{\xi_j}}(\ell; \xi_j)$  have explicit expressions given in (3.25). Writing (3.109) in the Fourier domain then leads to simple algebraic equations for the  $\xi_j$ -Fourier coefficients (the term  $\xi_j$ -Fourier coefficients refers to the values  $\widehat{v}(\ell; \xi_j)$  of a function  $v$ ) of  $\tilde{u}_M^{R,N}(\cdot; \xi_j)$ , which constitute the unknowns of our problem. In these equations one needs in particular to evaluate the  $\xi_j$ -Fourier coefficients of

$$(n_p - 1) \tilde{u}_M^{R,N}(\cdot; \xi_j) \text{ and } (n - n_p) u_M^{R,N}.$$

This step is numerically done using (inverse) FFT to evaluate first the nodal points of  $\tilde{u}_M^{R,N}(\cdot; \xi_j)$  (from the  $\xi_j$ -Fourier coefficients) then FFT to compute the  $\xi_j$ -Fourier coefficients of  $(n_p - 1) \tilde{u}_M^{R,N}(\cdot; \xi_j)$  and  $(n - n_p) u_M^{R,N}$ .

As in [85], we use the Matlab version of GMRES algorithm to invert the linear system in the Fourier domain. In the numerical experiments below, we used a tolerance  $= 10^{-9}$ . We also note that we did not use any preconditioning of our system since our objective here is to rather give some preliminary validating tests of the method in a two dimensional setting of the problem.

#### Example 1: Case of $n_p = 1$

We here consider the scattering problem described in Remark 3.2.1 for a locally perturbed homogeneous domain, where the local perturbation is a disk of center  $0 = (0, 0)$  and radius  $r < \min\{\frac{L}{2}, x_d^0\}$ , which has refractive index  $n > 1$ . Thanks to separation of variables we can calculate an analytic expression for the solution as

$$\begin{cases} u = \sum_m \beta_m J_m(k\sqrt{n}|x|) e^{im\theta} & |x| < r, \\ u^s = \sum_m \alpha_m H_m^{(1)}(k|x|) e^{im\theta} & |x| > r, \end{cases} \quad (3.111)$$

where,

$$\begin{cases} \alpha_m = \frac{i}{4} H_m^{(1)}(k|x^0|) e^{-in\theta^0} \frac{\sqrt{n} J_m'(k\sqrt{n}|x|) J_m(k|x|) - J_m(k\sqrt{n}|x|) J_m'(k|x|)}{-\sqrt{n} J_m'(k\sqrt{n}|x|) H_m^{(1)}(k|x|) + J_m(k\sqrt{n}|x|) H_m^{(1)'}(k|x|)}, \\ \beta_m = \frac{i}{4} H_m^{(1)}(k|x^0|) e^{-in\theta^0} \frac{\sqrt{n} H_m^{(1)'}(k|x|) J_m(k|x|) - H_m^{(1)}(k|x|) J_m'(k|x|)}{H_m^{(1)'}(k|x|) J_m(k\sqrt{n}|x|) - \sqrt{n} J_m'(k\sqrt{n}|x|) H_m^{(1)}(k|x|)}. \end{cases} \quad (3.112)$$

Here  $\theta^0$  denotes the angle between  $x$  and  $x^0$  and  $J_m$  and  $H_m^{(1)}$  respectively denote the Bessel function and the Hankel function of the first kind of order  $m$ . Figure 3.1 reports the results obtained using the numerical approximation (right) and the one obtained using the analytic expression (left). These results correspond with a refractive index  $n = 3$ , a wave number  $k = \pi + 0.1i$ , a period length  $L = 2\pi$ ,  $x^0 = (0, 2.5\lambda)$  where  $\lambda := \frac{2\pi}{|k|}$  and  $r = 0.5\lambda$ . For our numerical solution, we used  $h = 1.5\lambda$  and  $R = \pi\lambda$ . We observe that we obtain qualitatively the same solution.

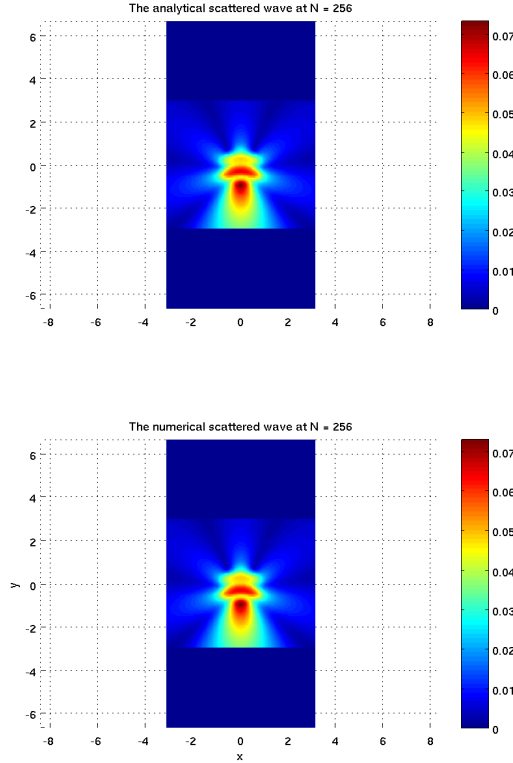


Figure 3.1: The analytic solution in  $\Omega_0^h$  (left) and the numerical solution in  $\Omega_0^h$  (right) for the first example (here  $h = 3$ ).

Figure 3.2-right reports the relative error for fixed  $N = 256$  and varying  $M$ . We observe a fast convergence rate with a rapid saturation of the error, meaning that the approximation error with respect to  $M$  becomes smaller than the one with respect to  $N$ .

Figure 3.2-left indicates the behavior of the  $L^2(\Omega_0^h)$  error with respect to  $N$  (we choose  $N_1 = N_2 = N$ ) for fixed  $M = 9$ . We observe that the slope of the log-log curve is  $-1/2$  less than the predicted one (that should be  $-2$ ). We think that this is due to the fact that we do not evaluate exactly the  $\xi_j$ -Fourier coefficients of  $(n - n_p)u_M^{R,N}$  or those of  $(n - 1)f$ . In our

algorithm, we first project the function on the nodal points then evaluate numerically the FFT of the obtained discrete vector. Indeed in our convergence theorem we did not take into account this type of error.

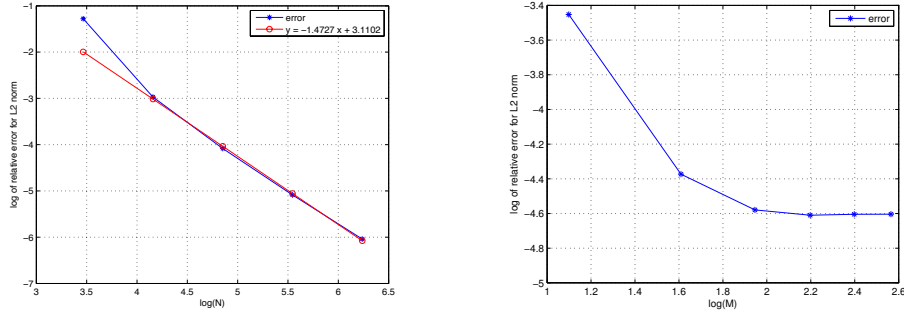


Figure 3.2: Left: The  $L^2$  relative error in  $\Omega_0^h$  with respect to  $N$  when  $M = 9$ . Right: The  $L^2$  relative error in  $\Omega_0^h$  with respect to  $M$  when  $N = 256$ .

To illustrate the effect of the numerical approximation of the Fourier transform of discontinuous functions on the convergence rate, we compare in Figure 3.3 the error in evaluating  $V_\xi^R f$  with different methods in computing the Fourier coefficients of  $f$ . Consider the piecewise constant  $f \in L^2(\Omega_0^R)$  where  $f = 2$  in  $\Omega_0^h$  and  $f = 0$  otherwise. We first compute the convolution operator  $F_N := P_\xi^N V_\xi^R f$  using the exact Fourier coefficients of  $f$ . We second compute  $\tilde{F}_N \approx P_\xi^N V_\xi^R f$  by evaluating the Fourier coefficients of  $f$  using FFT. The Fourier coefficients of  $V_\xi^R$  are exactly computed in both cases. In Figure 3.3 we report the  $L^2(\Omega_0^R)$  norm of  $E_N = F_N - F_{N_{\max}}$  and  $\tilde{E}_N = \tilde{F}_N - \tilde{F}_{N_{\max}}$  for  $N = 2^n, n = 5, \dots, 9$  and with  $N_{\max} = 2^{11}$ . While the rate of convergence is equal to the one predicted by the theory for  $E_N$  (Figure 3.3-left) we observe a non monotone rate of convergence for  $\tilde{E}_N$  (Figure 3.3-right).

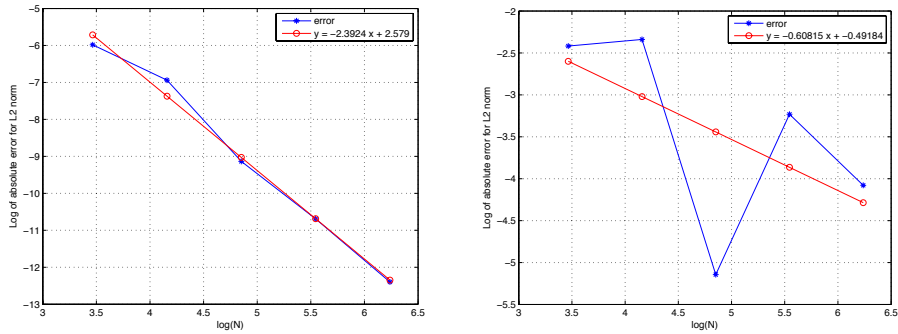


Figure 3.3: Left: The rate of convergence of  $E_N$  with respect to  $N$  when evaluating the convolution operator using exact Fourier coefficients. The rate of convergence of  $\tilde{E}_N$  with respect to  $N$  when evaluating the convolution operator using FFT.

Let us notice that the observation made for Figure 3.3 is less visible for regular functions  $f$  as attested by the following experiment.

In the experiment illustrated by Figure 3.4 we again consider the scattering problem with  $n_p = 1$  but here the local perturbation (the support of  $n - 1$ ) is a square  $\omega$  centered at the origin with an edge length  $r = 0.6$ . We give in Figure 3.4 the convergence rate with respect to  $N$  (for  $M = 9$ ) for three different choices of the refractive index  $n$  inside the domain  $\omega$ :  $n = n_1 := 3$ ,



$n = n_2 := 1 + \frac{1}{2} \cos(\frac{\pi x}{r}) \cos(\frac{\pi y}{r})$  and  $n = n_3 := 1 + 2 \cos^2(\frac{\pi x}{r}) \cos^2(\frac{\pi y}{r})$ . The error is computed as the  $L^2(\Omega_0^h)$  norm of  $u_M^{R,N} - u_M^{R,N_{\max}}$  where  $N_{\max} = 2^{10}$  and  $N = 2^n, n = 5, \dots, 9$ . As we can observe, the correct rate of convergence is only observed for the third choice of  $n$ . For the second choice, the rate is closer to the correct one (that should be 3) for the first values of  $N$ .

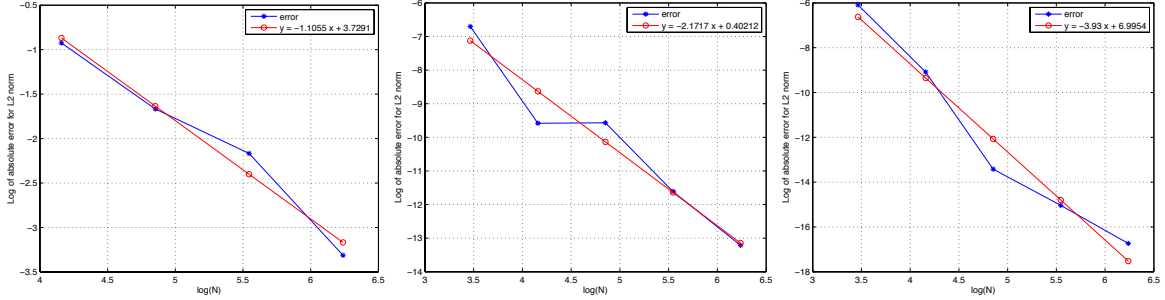


Figure 3.4:  $L^2(\Omega_0^h)$  norm of  $u_M^{R,N} - u_M^{R,N_{\max}}$  where  $N_{\max} = 2^{10}$  and  $N = 2^n, n = 5, \dots, 9$  in the case of  $n = n_1$  (left),  $n = n_2$  (middle) and  $n = n_3$  (right).

### Example 2: A case with $n_p \neq 1$

We end our numerical examples with the case of locally perturbed periodic media where  $n_p \neq 1$ . The experiment corresponds with a scattering from a point source as in the first example. The parameters are the same except for the definition of  $n_p$  and  $n$ . The choice of the refractive index  $n_p$  is such that  $n_p = 3$  in  $B_0(r)$ ,  $n_p = 1$  in  $\Omega_0^h \setminus B_0(r)$  and  $r = 0.5 * \lambda$ . The local perturbation is such that  $n = 1$  in  $\Omega_0^h$ . See Figure 3.5.

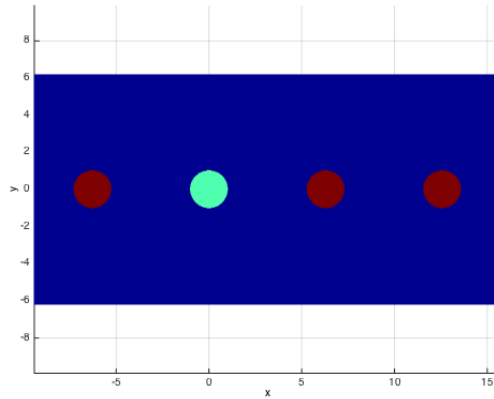


Figure 3.5: Sketch of the geometry for Example 2

We compare in Figure 3.6 the solution obtained by our numerical algorithm and the numerical solution computed on  $\Omega_M^R$  using the finite element code FreeFem++ with a periodicity conditions imposed of the boundary of  $\Omega_M^R$  for  $M = 9$ . We observe a slight difference between the two solutions that indeed comes from difference between the two numerical schemes but also from the difference between  $f_M$  and  $f$  in  $\Omega_M$  where  $f = (n - 1)u_i$ .

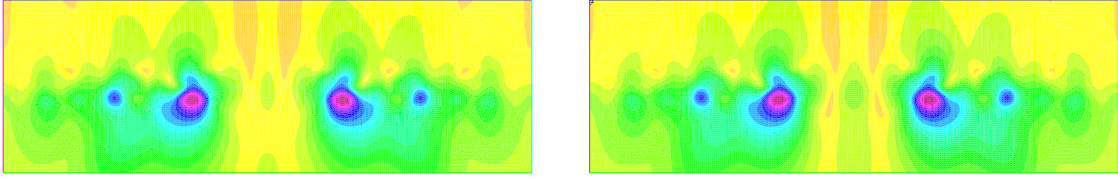


Figure 3.6: Left: Numerical solution obtained by finite element discretization of the periodic problem using Freefem++. Right: Numerical solution obtained by our algorithm.

We illustrate in Figure 3.7 the convergence rate with respect to  $N$  by computing the  $L^2(\Omega_0^h)$  norm of  $u_M^{R,N} - u_M^{R,N_{\max}}$  for  $M = 9$  with  $N_{\max} = 2^{11}$  and  $N = 2^n, n = 3, \dots, 7$ . We observe a rate of convergence compatible with the one observed for the first example.

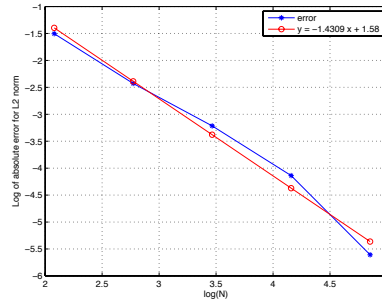


Figure 3.7:  $L^2(\Omega_0^h)$  norm of  $u_M^{R,N} - u_M^{R,N_{\max}}$  where  $N_{\max} = 2^{11}$  and  $N = 2^n, n = 3, \dots, 7$  in the case of locally perturbed periodic media and piecewise constant refractive index.

### 3.5.1 Numerical examples for real wave numbers

We end this chapter with some numerical examples for the case of real wave numbers  $k$ . Indeed our algorithm can be applied to this case but the convergence cannot be guaranteed if there exists propagative modes along the periodicity directions. We showed in Chapter 2 that if the index of refraction has a positive imaginary part on a strip then there are no guided modes and the scattered solution is in  $H^1$ . Our convergence analysis still does not cover this case. However, one can expect convergence of our scheme following for instance arguments as in [87]. One cannot expect anymore exponential convergence with respect to the discretization number  $M$  for the Floquet-Bloch variable. We hereafter report numerical linear convergence results for this parameter. In order to formulate the scattering problem as a source problem with compactly supported source we consider  $\xi_j$ -quasi-periodic incident waves  $u^i$  (that will be chosen as plane waves). We then compute the  $\xi_j$ -quasi-periodic total field  $u_p$  due to the background (i.e. with  $n = n_p$ ). This total field is computed using the scheme described in Section 5.2.2. We then observe that the total field  $u$  (by the perturbed medium) can be decomposed as

$$u = \tilde{u}^s + u_p$$

where  $\tilde{u}^s$  satisfies

$$\Delta \tilde{u}^s + k^2 n \tilde{u}^s = k^2 (n_p - n) u_p$$

and the Rayleigh radiation condition (See Chapter 2). The right hand side of this equation has a compact support by assumption.

**Example 3: real  $k$  and complex  $n_p$** 

We start with a favorable case where the domain setting is similar to the Example 2 (See Figure 3.5) but we replace the index of refraction  $n_p = 3$  by  $n_p = 3 + 1i$  and index of refraction on the support of  $n - n_p$  by  $n = 1 + 1i$ . The wave number is  $k = \pi$ . We here consider an incident plane wave given by

$$u^i(x) = e^{3i\pi/(16L)x_1 + \sqrt{k^2 - (3i\pi/(16L))^2}x_2}. \quad (3.113)$$

For each value of  $M$  (we choose  $M \geq 4$ ), we represent the scattered wave in a truncated domain  $\Omega_{M_0}$  with  $M_0 = 4$  fixed. Let us denote by  $u_M$  the numerical solution corresponding to  $M$  (we fix the number of space discretization and choose it sufficiently large  $N = 256$ ). Let  $M_{max}$  be chosen sufficiently large. We shall represent the relative error  $u_{M_{max}} - u_M$  in the domain  $\Omega_{M_0}$ . Figure 3.8-left illustrates  $u_M$  with  $M = 16$ . Figure 3.8-right shows the relative error between  $u_{M_{max}}$  and  $u_M$  versus  $M$  for  $M_{max} = 64$  fixed and  $M \in [4, 8, 12, 16]$ . We observe linear convergence with respect to  $M$ . This is also confirmed by the next example.

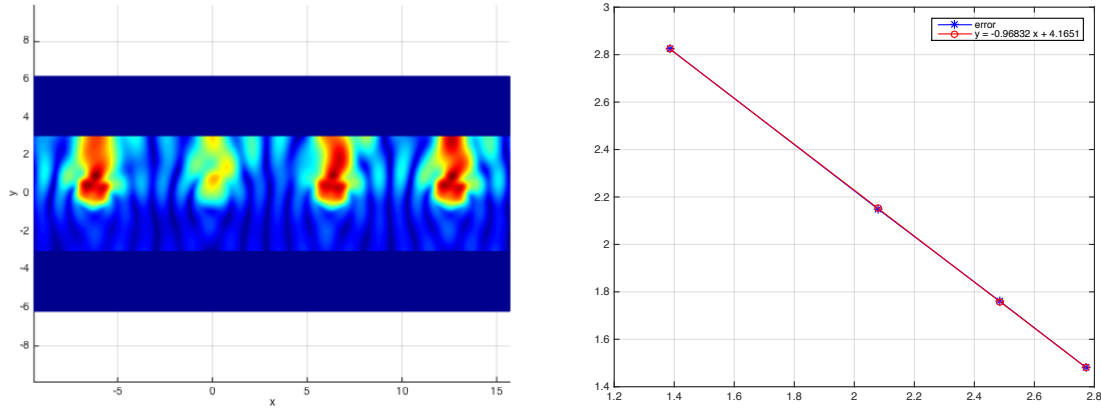


Figure 3.8: Left: Illustration of the scattered wave in  $\Omega_{M_0}$ . Right: The relative error in  $\Omega_{M_0}$  with respect to  $M$ . The geometry of the experiment is shown in Figure 3.5

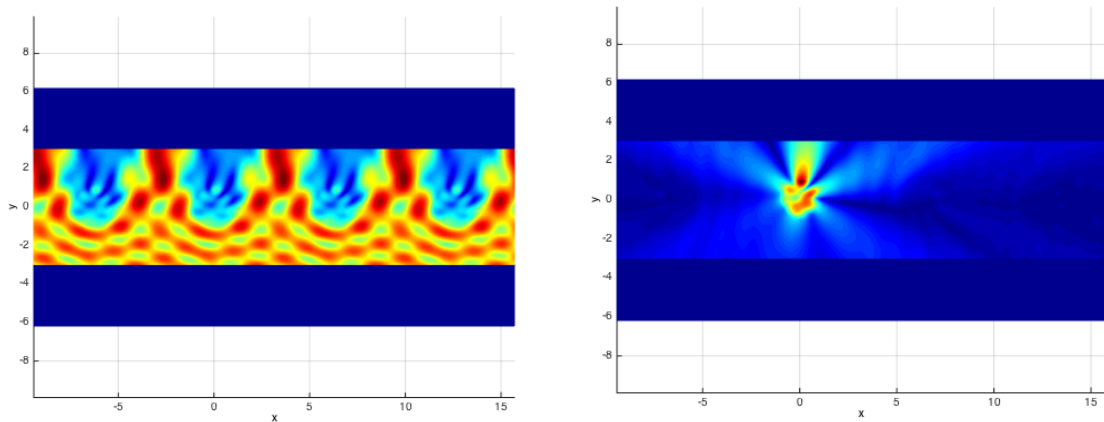


Figure 3.9: Left: The background total field  $u_p$  in  $\Omega_{M_0}$ . Right: The scattered field  $\tilde{u}^s$  in  $\Omega_{M_0}$ . The geometry of the experiment is shown in Figure 3.5

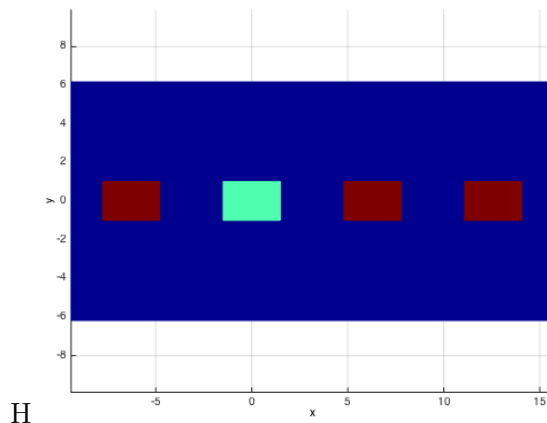


Figure 3.10: Sketch of the geometry for Example 3 with rectangles

For the next example we change the balls by rectangles with side length  $= 1.5\lambda$  as shown in Figure 3.10. The related numerical results are shown in Figures 3.11 and 3.12.

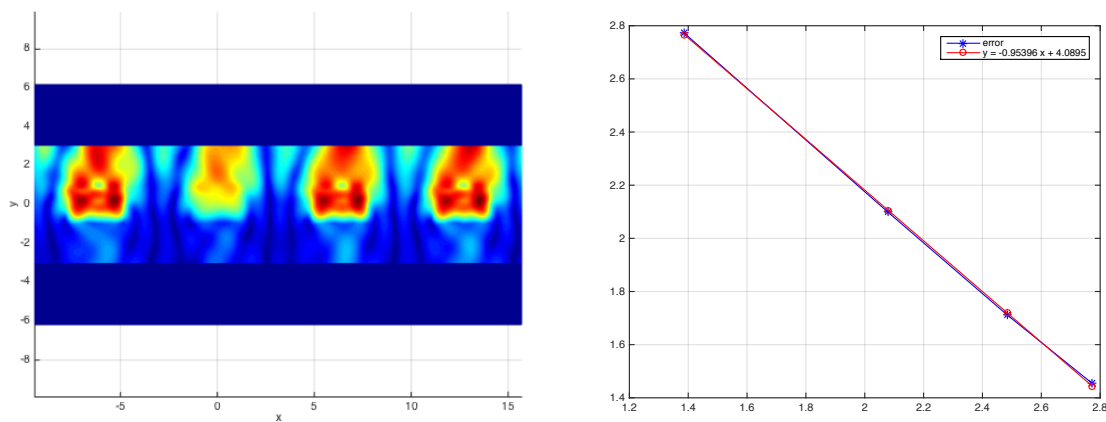


Figure 3.11: Left: Illustration of the scattered wave in  $\Omega_{M_0}$ . Right: The relative error in  $\Omega_{M_0}$  with respect to  $M$ . The geometry of the experiment is shown in Figure 3.10

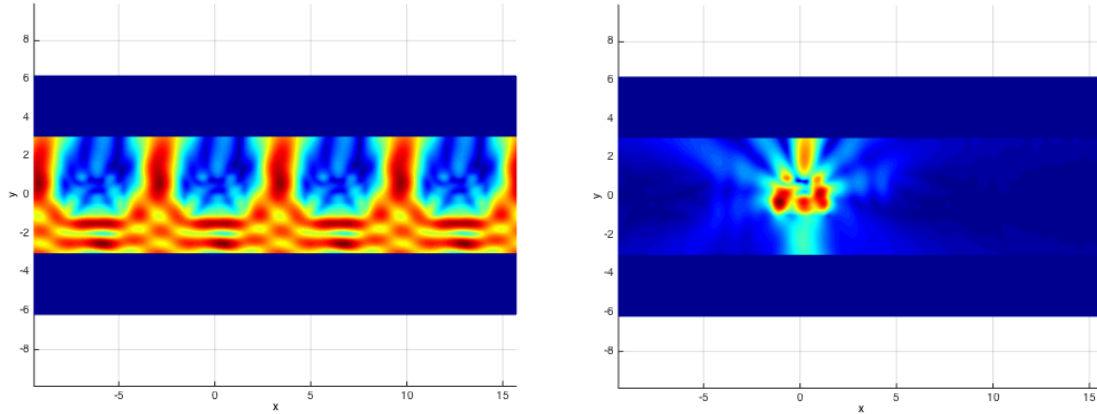


Figure 3.12: Left: The background total field  $u_p$  in  $\Omega_{M_0}$ . Right: The scattered field  $\tilde{u}^s$  in  $\Omega_{M_0}$ . The geometry of the experiment is shown in Figure 3.10

#### Example 4: real $k$ and real $n_p$

We end this chapter with a numerical example for real  $n$  for which we observe convergence, which means that there are no guided modes for the considered configuration. Observing convergence means that we obtain a decrease in the error as  $M$  increases. This is in contrast with observed cases where the error never decreases as  $M$  increases. The example corresponds with Example 2 but we replace the wavenumber  $k = \pi + 0.1i$  by  $k = \pi$ . We then consider the incident plane wave  $u^i$  given by (3.113). For each value of  $M$  (we choose  $M \geq 4$ ), we compute the scattered wave in the domain  $\Omega_{M_0}$  with  $M_0 = 4$  fixed. Let us denote by  $u_M$  the numerical solution corresponding to  $M$  (we fix the number of space discretization and choose it sufficiently large  $N = 256$ ). Figure 3.13-left illustrates  $u_M$  with  $M = 20$ . Figure 3.13-right shows the relative error between  $u_{M_{max}}$  and  $u_M$  versus  $M$  for  $M_{max} = 128$  fixed and  $M \in [4, 8, 12, 16, 20]$ .

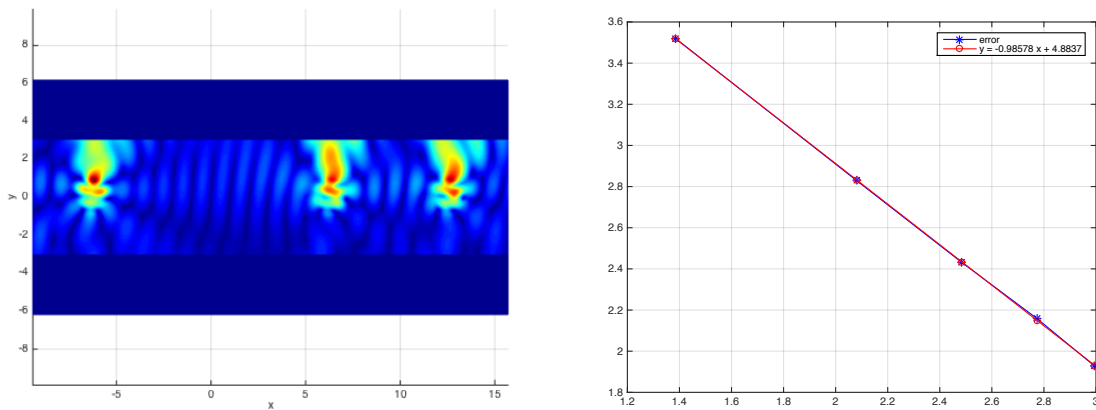


Figure 3.13: Left: Illustration of the scattered wave in  $\Omega_{M_0}$ . Right: The relative error in  $\Omega_{M_0}$  with respect to  $M$ . The geometry of the experiment is shown in Figure 3.5

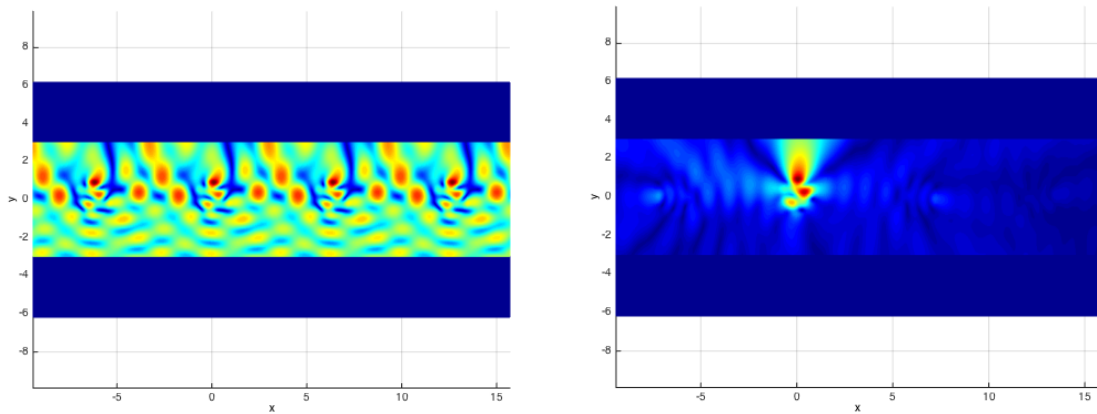


Figure 3.14: Left: The background total field  $u_p$  in  $\Omega_{M_0}$  for real  $n_p$ . Right: The scattered field  $\tilde{u}^s$  in  $\Omega_{M_0}$  for real  $n_p$ . The geometry of the experiment is shown in Figure 3.5



# Inverse scattering problem from locally perturbed periodic media

---

## Abstract

This chapter is dedicated to the design and analysis of sampling methods to recover the shape of a perturbation from measurements of scattered waves at a fixed frequency. We first introduce the model problem that corresponds with the semi-discretized version of the continuous model with respect to the Floquet-Bloch variable. We then present the inverse problem setting where (propagative and evanescent) plane waves are used to illuminate the structure and measurements of the scattered wave at a parallel plane to the periodicity directions are performed. We introduce the near field operator and analyze two possible factorizations of this operator. We then establish sampling methods (LSM, FM and GLSM) to identify the defect and the periodic background geometry from this operator measurement. We also show how one can recover the geometry of the background independently from the defect. Some validating results are given in 2D and simple configurations. We introduce and analyze in the last part the single Floquet-Bloch mode measurement operators and show how one can exploit them to build an indicator function of the defect. Numerical validating results are provided for simple and complex backgrounds.

## Contents

---

<b>4.1</b>	<b>Introduction</b>	<b>76</b>
<b>4.2</b>	<b>Setting of the direct scattering problem</b>	<b>77</b>
<b>4.3</b>	<b>Setting of the inverse problem</b>	<b>80</b>
4.3.1	Definition of the sampling operator	80
4.3.2	Some useful properties for sampling methods	81
<b>4.4</b>	<b>Application to Sampling methods</b>	<b>86</b>
4.4.1	The Linear Sampling Method (LSM)	86
4.4.2	The Factorization Method	87
4.4.3	The Generalized Linear Sampling Method (GLSM)	88
4.4.4	Reconstruction of the periodic domain $D_p$ from $N^\pm$	89
4.4.5	Validating Numerical Experiments	91
<b>4.5</b>	<b>On the Use of Differential measurements</b>	<b>95</b>
4.5.1	Theory of the Differential LSM	95
4.5.2	A numerical example	97
<b>4.6</b>	<b>Sampling methods for a single Floquet-Bloch mode</b>	<b>98</b>
4.6.1	Near field operator for a fixed Floquet-Bloch mode	98
4.6.2	Some properties of the operators $\mathcal{H}_q^\pm$ , $N_q^\pm$ and $G_q^\pm$	99
4.6.3	A new differential imaging functional	106
4.6.4	Numerical validating examples	108

---



## 4.1 Introduction

We investigate in this chapter the inverse problem where one is interested in reconstructing the support of the perturbation of the periodic media from measurements of scattered waves. We are concerned with the design of a sampling method that would reconstruct the support of inhomogeneities without reconstructing the index of refraction. The development of sampling methods has gained a large interest in recent years and many methods have been introduced in the literature to deal with a variety of problems and we refer to [46, 29, 76, 30] for an account of recent developments of these methods. The case of periodic media has been treated in [82, 91, 10, 7, 98, 53, 105, 66]. For the inverse problem in a locally perturbed periodic waveguides we refer to [101, 27] and references therein.

Up to our knowledge, the sampling methods for locally perturbed infinite periodic layers has not been treated in the literature. Even though this problem is the one that motivates our study, we shall consider here a slightly different problem that will be referred to as the  $ML$ -periodic problem: it corresponds with a locally perturbed infinite periodic layer with period  $L$  that has been reduced to a domain of size  $ML$  (with  $M$  a sufficiently large parameter) with periodic boundary conditions. This is mainly for technical reasons since our analysis for the newly introduced differential imaging functional heavily rely on the discrete Floquet-Bloch transform. According to Chapter 3 the  $ML$ -periodic problem can be seen as the discretized problem with respect to the Floquet-Bloch variable using  $M$  discretization points in the trapezoidal rule (see Section 5.2.2 of Chapter 3). We refer to Chapter 2 for the study of the well posedness of  $ML$ -periodic problem.

As an inversion method, we shall employ the recently introduced Generalized version of the Linear Sampling Method GLSM (see [12, 14, 30]) and also consider the Factorization method (see [70, 71, 76]). We consider the case where the data correspond with Rayleigh sequences of the scattered field associated with incident waves that are propagative and evanescent modes. For similar inverse problems in waveguides we refer to [28, 27, 82, 101]. We shall prove in a first part how the GLSM and also the factorization method can be applied to our setting.

The main contribution of our work is the design of a new sampling method that enable the imaging of the defect location without reconstructing the  $L$  periodic background. This method is in the spirit of the so-called Differential LSM that has been introduced in [13] for the imaging of defects in complex backgrounds using differential measurements. However, in our case we shall introduce a method that does not require the measurement operator for the background media. We exploit the  $L$  periodicity of the background and the Floquet-Bloch transform to design a differential criterion between different periods. This criterion is based on the study of sampling methods for the  $ML$ -periodic media where a single Floquet-Bloch mode is used. This study constitutes the main theoretical ingredient for our method. The sampling operator for a single Floquet-Bloch mode somehow plays the role of the measurement operator for the background media. Indeed the main interest for this new sampling method is that it is capable of identifying the defect even though classical sampling methods fail in obtaining high fidelity reconstructions of the (complex) background media.

The performance of the sampling methods are tested against synthetic data generated by the solver developed in Chapter 3.

The outline of this chapter is as the following. We introduce in Section 4.2 the direct  $ML$ -periodic scattering problem and briefly outline the formulation of the Rayleigh radiation condition and the variational formulation of the problem. We present in Section 4.3 the setting of the inverse problem for incident plane waves and measurements constituted by the Rayleigh coefficients of the scattered waves. We then introduce the near field operator as well as the factorizations

of this operator that will be needed for the sampling methods. Some key properties of these operators are then proved as preparations for the analysis of sampling methods. Section 4.4 is dedicated to the theoretical analysis of sampling methods for retrieving the geometry of the background media and the defect. We also explain in this section how these methods can be used to identify the  $L$ -periodic background media. In order to motivate the differential imaging method we briefly outline in Section 4.5 the sampling method that uses differential measurements in order to retrieve only the defect. Section 4.6 is dedicated to the analysis of sampling methods using a single Floquet-Bloch mode. This analysis is the last main ingredient for the differential imaging functional presented in Section 4.6.3. Each of the sections 4.4, 4.5 and 4.6 is ended with validating numerical results. In order to make this work self-contained we included in Appendix A a summary of the main abstract theoretical results that are used for the foundations of the sampling methods.

## 4.2 Setting of the direct scattering problem

Consider a parameter  $L := (L_1, \dots, L_{d-1}) \in \mathbb{R}^{d-1}$ ,  $L_j > 0$ ,  $j = 1, \dots, d-1$  that will refer to the periodicity of the media with respect to the first  $d-1$  variables and  $M := (M_1, \dots, M_{d-1}) \in \mathbb{N}^{d-1}$  that refer to the number of period truncation. A function defined in  $\mathbb{R}^d$  is called  $L$  periodic if it is periodic with respect to the  $d-1$  first variables, with period  $L$ . We consider in the following the  $ML$ -periodic Helmholtz equation (we recall that vector multiplications is to be understood component wise, i.e.  $ML = (M_1L_1, \dots, M_{d-1}L_{d-1})$ ). In this problem, the total field satisfies

$$\begin{cases} \Delta u + k^2 n u = 0 & \text{in } \mathbb{R}^d, \quad d = 2, 3 \\ u \text{ is } ML\text{-periodic} \end{cases} \quad (4.1)$$

where the *wavenumber*  $k$  is positive and real valued. We assume that the index of refraction  $n \in L^\infty(\mathbb{R}^d)$  has a non negative imaginary part and is  $ML$ -periodic. Moreover, we assume that  $n = n_p$  outside a compact domain  $\omega$  where  $n_p \in L^\infty(\mathbb{R}^d)$  is  $L$ -periodic and assume in addition that there exists  $h > 0$  such that  $n = 1$  for  $|x_d| > h$  (see Fig. 4.1).

Thanks to the  $ML$ -periodicity, solving equation (4.1) in  $\mathbb{R}^d$  is equivalent to solving it in one period

$$\Omega_M := \bigcup_{m=[-\frac{M}{2}] + 1}^{m=[\frac{M}{2}]} \Omega_m = \llbracket M_L^-, M_L^+ \rrbracket \times \mathbb{R}$$

where  $M_L^- := ([-\frac{M}{2}] + \frac{1}{2})L$ ,  $M_L^+ := ([\frac{M}{2}] + \frac{1}{2})L$  and  $\Omega_m := \llbracket -\frac{L}{2} + mL, \frac{L}{2} + mL \rrbracket \times \mathbb{R}$ . We denote by  $D$  (respectively  $D_p$ ) a bounded domain composed by simply connected components and such that  $n = 1$  (respectively  $n_p = 1$ ) outside  $D$  (respectively  $D_p$ ) and  $D = D_p \cup \bar{\omega}$ . We shall consider down-to-up or up-to-down incident plane waves of the form:

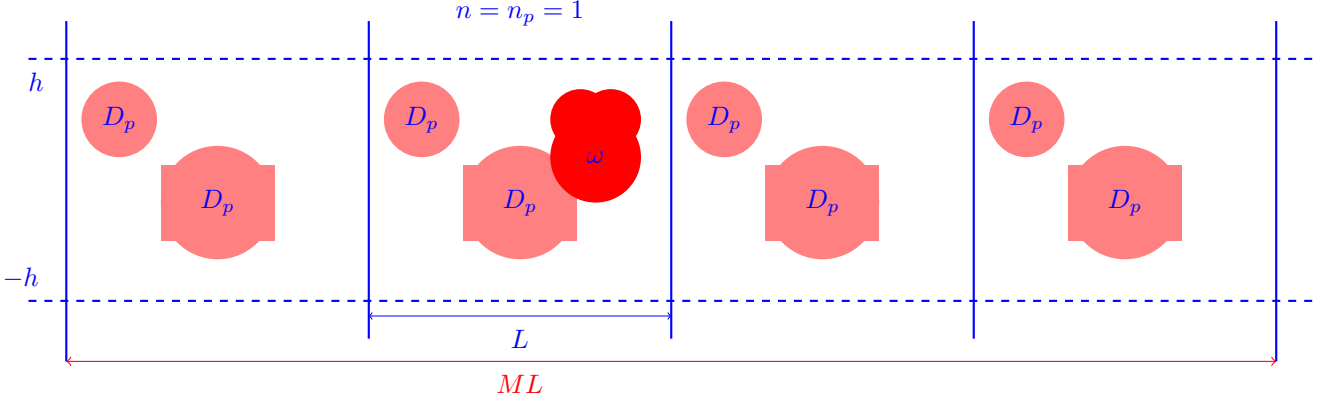
$$u^i(x) = e^{i\alpha_{\#}(j)\bar{x} \pm i\beta_{\#}(j)x_d}$$

where

$$\alpha_{\#}(j) := i\frac{2\pi}{ML}j \quad \text{and} \quad \beta_{\#}(j) := \sqrt{k^2 - \alpha_{\#}^2(j)}, \quad \text{Im}(\beta_{\#}(j)) \geq 0, \quad j \in \mathbb{Z}$$

and where  $x = (\bar{x}, x_d) \in \mathbb{R}^{d-1} \times \mathbb{R}$ . Then the scattered field  $u^s = u - u^i$  verifies

$$\begin{cases} \Delta u^s + k^2 n u^s = -k^2(n-1)u^i & \text{in } \mathbb{R}^d, \\ u^s \text{ is } ML\text{-periodic} \end{cases} \quad (4.2)$$


 Figure 4.1: Sketch of the geometry for the  $ML$ -periodic problem

and we impose as a radiation condition the Rayleigh expansions:

$$\begin{cases} u^s(\bar{x}, x_d) = \sum_{\ell \in \mathbb{Z}} \hat{u}^{s+}(\ell) e^{i(\alpha_{\#}(\ell)\bar{x} + \beta_{\#}(\ell)(x_d - h))}, & \forall x_d > h \\ u^s(\bar{x}, x_d) = \sum_{\ell \in \mathbb{Z}} \hat{u}^{s-}(\ell) e^{i(\alpha_{\#}(\ell)\bar{x} - \beta_{\#}(\ell)(x_d + h))}, & \forall x_d < -h, \end{cases} \quad (4.3)$$

where the Rayleigh coefficients  $\hat{u}^{s,\pm}(\ell)$  are given by

$$\begin{aligned} \hat{u}^{s+}(\ell) &:= \frac{1}{\llbracket ML \rrbracket} \int_{\llbracket M_L^-, M_L^+ \rrbracket} u^s(\bar{x}, h) e^{-i\alpha_{\#}(\ell)\bar{x}} d\bar{x} \\ \hat{u}^{s-}(\ell) &:= \frac{1}{\llbracket ML \rrbracket} \int_{\llbracket M_L^-, M_L^+ \rrbracket} u^s(\bar{x}, -h) e^{-i\alpha_{\#}(\ell)\bar{x}} d\bar{x}. \end{aligned} \quad (4.4)$$

We recall the notation

$$\begin{aligned} \Omega_M^h &:= \llbracket M_L^-, M_L^+ \rrbracket \times ]-h, h[ \\ \Gamma_M^h &:= \llbracket M_L^-, M_L^+ \rrbracket \times \{h\}, \quad \Gamma_M^{-h} := \llbracket M_L^-, M_L^+ \rrbracket \times \{-h\} \end{aligned}$$

and the Sobolev spaces of  $ML$ -periodic functions

$$H_{\#,M}^s(\mathbb{R}^d) := H_{0,ML}^s(\mathbb{R}^d) \quad \text{and} \quad H_{\#}^s(\Omega_M) := \{u|_{\Omega_M}; u \in H_{\#,M}^s(\mathbb{R}^d)\}.$$

For later use, we introduce the following spaces

$$\begin{aligned} H_{\#}^s(\Gamma_M^h) &:= \{\phi \text{ is } ML\text{-periodic such that } \|\phi\|_{H_{\#}^s(\Gamma_M^h)}^2 := \sum_{\ell \in \mathbb{Z}} (1 + |\ell|^2)^{s/2} |\hat{\phi}^+(\ell)|^2 < +\infty\}, \\ H_{\#}^s(\Gamma_M^{-h}) &:= \{\phi \text{ is } ML\text{-periodic such that } \|\phi\|_{H_{\#}^s(\Gamma_M^{-h})}^2 := \sum_{\ell \in \mathbb{Z}} (1 + |\ell|^2)^{s/2} |\hat{\phi}^-(\ell)|^2 < +\infty\}. \end{aligned}$$

Using the the radiation condition (4.3) we can define Dirchlet-to-Neumann operators  $T^{\pm}$  as

$$\begin{aligned} T^+ &: H_{\#}^{1/2}(\Gamma_M^h) \longrightarrow H_{\#}^{-1/2}(\Gamma_M^h) \\ \phi^+ &\longmapsto T^+ \phi^+ = i \sum_{\ell \in \mathbb{Z}} \beta_{\#}(\ell) \hat{\phi}^+(\ell) e^{i\alpha_{\#}(\ell)\bar{x}} \\ T^- &: H_{\#}^{1/2}(\Gamma_M^{-h}) \longrightarrow H_{\#}^{-1/2}(\Gamma_M^{-h}) \\ \phi^- &\longmapsto T^- \phi^- = i \sum_{\ell \in \mathbb{Z}} \beta_{\#}(\ell) \hat{\phi}^-(\ell) e^{i\alpha_{\#}(\ell)\bar{x}} \end{aligned} \quad (4.5)$$

It is easy to check that  $T^\pm$  are bounded operators and:

$$\begin{aligned} \operatorname{Im} \langle T^+ \phi, \phi \rangle &\geq 0, & \operatorname{Im} \langle T^- \phi, \phi \rangle &\geq 0, \\ \operatorname{Re} \langle T^+ \phi, \phi \rangle &\leq 0, & \operatorname{Re} \langle T^- \phi, \phi \rangle &\leq 0, \end{aligned} \quad (4.6)$$

where  $\langle \cdot, \cdot \rangle$  denotes the sesquilinear duality products  $H_{\#}^{-1/2}(\Gamma_M^h) - H_{\#}^{1/2}(\Gamma_M^h)$  (See also Chapter 2). The scattering problem can be reformulated as: Find  $u^s \in H_{\#}^1(\Omega_M^h)$  such that

$$\begin{aligned} \Delta u^s + k^2 n u^s &= -k^2(n-1)u^i \quad \text{in } \Omega_M^h \\ \frac{\partial u^s}{\partial x_d} &= \pm T^\pm(u^s) \quad \text{for } x_d = \pm h. \end{aligned} \quad (4.7)$$

Multiplying equation (4.7) with  $v \in H_{\#}^1(\Omega_M^h)$  and using the Green theorem we arrive at the variational formulation of the problem as

$$\int_{\Omega_M^h} \nabla u^s \nabla \bar{v} - k^2 n u^s \bar{v} \, dx - \langle T^+(u^s), v \rangle - \langle T^-(u^s), v \rangle = k^2 \int_{\Omega_M^h} (n-1)u^i \bar{v} \, dx \quad (4.8)$$

for all  $v \in H_{\#}^1(\Omega_M^h)$ . Problem 4.8 is of Fredholm type since the sesquilinear form

$$\mathcal{A}(u, v) := \int_{\Omega_M^h} \nabla u^s \nabla \bar{v} - k^2 n u^s \bar{v} \, dx - \langle T^+(u^s), v \rangle - \langle T^-(u^s), v \rangle$$

is continuous on  $H_{\#}^1(\Omega_M^h) \times H_{\#}^1(\Omega_M^h)$  and satisfies the Garding inequality

$$|\mathcal{A}(u, u)| \geq \|u\|_{H^1(\Omega_M^h)}^2 - \int_{\Omega_M^h} (k^2 \operatorname{Re} n + 1)|u|^2 \, dx. \quad (4.9)$$

which follows from (4.6). The uniqueness of solutions to this problem can be studied using Rellich type identities as in Chapter 2 or exploiting the fact that the imaginary part of the refractive index is positive in an open ball.

For the purpose of this chapter we shall assume that the index of refraction  $n$  is such that Problem (4.8) is well posed. More precisely, let  $f \in L^2(\Omega_M^h)$  and consider the following variational problem: Find  $w \in H_{\#}^1(\Omega_M^h)$  such that for all  $v \in H_{\#}^1(\Omega_M^h)$ ,

$$\int_{\Omega_M^h} \nabla w \nabla \bar{v} - k^2 n w \bar{v} \, dx - \langle T^+(w), v \rangle - \langle T^-(w), v \rangle = k^2 \int_{\Omega_M^h} (n-1)f \bar{v} \, dx. \quad (4.10)$$

Then we make the following assumption:

**Assumption 4.2.1.** *We assume that  $n$  and  $k$  are such that problem (4.10) is well posed for all  $f \in L^2(\Omega_M^h)$ .*

We remark that the solution  $w \in H_{\#}^1(\Omega_M^h)$  of (4.10) can be extended to a function in  $\Omega_M$  satisfying  $\Delta w + k^2 n w = -k^2(n-1)f$ , using the Rayleigh expansion (4.3). Let  $G_M$  be the  $ML$ -periodic Green function satisfying  $\Delta G_M + k^2 G_M = -\delta$  in  $\Omega_M$  and the Rayleigh radiation condition. Then  $w$  can also be represented as

$$w(x) = k^2 \int_D G_M(x-y)(n-1)(f+w)(y) \, dy. \quad (4.11)$$

This implies in particular that  $w \in H_{\#, \text{loc}}^2(\Omega_M)$ . For later use we recall that  $G_M$  has the representation

$$G_M(x) = \frac{i}{2ML} \sum_{\ell \in \mathbb{Z}} \frac{1}{\beta_{\#}(\ell)} e^{i\alpha_{\#}(\ell)\bar{x} + i\beta_{\#}(\ell)|x_d|}. \quad (4.12)$$

### 4.3 Setting of the inverse problem

We first use as incident waves all down-to-up (scaled) incident plane waves  $u^{i,+}(x; j)$  defined as

$$u^{i,+}(x; j) = \frac{-i}{2\bar{\beta}_{\#}(j)} e^{i\alpha_{\#}(j)\bar{x} + i\bar{\beta}_{\#}(j)(x_d - h)}, \quad j \in \mathbb{Z}. \quad (4.13)$$

Then our measurements (data for the inverse problem) will be formed by the Rayleigh sequences (see (4.4))

$$\widehat{u}^{s+}(\ell; j), \quad (j, \ell) \in \mathbb{Z} \times \mathbb{Z}$$

where  $j$  is related to the incident wave index and  $\ell$  is related to the Rayleigh sequence index. We can also use as incident waves all up-to-down (scaled) incident plane waves  $u^{i,-}(x; j)$  defined as

$$u^{i,-}(x; j) = \frac{-i}{2\bar{\beta}_{\#}(j)} e^{i\alpha_{\#}(j)\bar{x} - i\bar{\beta}_{\#}(j)(x_d + h)}, \quad j \in \mathbb{Z}, \quad (4.14)$$

and as measurements (data for the inverse problem) the Rayleigh sequences (see (4.4))

$$\widehat{u}^{s-}(\ell; j), \quad (j, \ell) \in \mathbb{Z} \times \mathbb{Z}.$$

#### 4.3.1 Definition of the sampling operator

Let us consider the (Herglotz) operators  $\mathcal{H}^+ : \ell^2(\mathbb{Z}) \rightarrow L^2(D)$  and  $\mathcal{H}^- : \ell^2(\mathbb{Z}) \rightarrow L^2(D)$  defined by

$$\mathcal{H}^{\pm} a := \sum_{j \in \mathbb{Z}} a(j) u^{i,\pm}(\cdot; j)|_D, \quad \forall a = \{a(j)\}_{j \in \mathbb{Z}} \in \ell^2(\mathbb{Z}). \quad (4.15)$$

Then  $\mathcal{H}^{\pm}$  is compact, injective (will be proved latter) and its adjoint  $(\mathcal{H}^{\pm})^* : L^2(D) \rightarrow \ell^2(\mathbb{Z})$  is given by

$$(\mathcal{H}^{\pm})^* \varphi := \{\widehat{\varphi}^{\pm}(j)\}_{j \in \mathbb{Z}}, \quad \text{where} \quad \widehat{\varphi}_j^{\pm} := \int_D \varphi(x) \overline{u^{i,\pm}(\cdot; j)}(x) dx. \quad (4.16)$$

Let us denote by  $H_{\text{inc}}^{\pm}(D)$  the closure of the range of  $\mathcal{H}^{\pm}$  in  $L^2(D)$ . We then consider the (compact) operator  $G^{\pm} : H_{\text{inc}}^{\pm}(D) \rightarrow \ell^2(\mathbb{Z})$  defined by

$$G^{\pm}(f) := \{\widehat{w}^{\pm}(\ell)\}_{\ell \in \mathbb{Z}}, \quad (4.17)$$

where  $\{\widehat{w}^{\pm}(\ell)\}_{\ell \in \mathbb{Z}}$  is the Rayleigh sequence of  $w \in H_{\#}^1(\Omega_M^h)$  the solution of (4.10). We now define the sampling operators  $N^{\pm} : \ell^2(\mathbb{Z}) \rightarrow \ell^2(\mathbb{Z})$  by

$$N^{\pm}(a) = G^{\pm} \mathcal{H}^{\pm}(a). \quad (4.18)$$

By linearity of the operators  $G^{\pm}$  and  $\mathcal{H}^{\pm}$  we also get an equivalent definition of  $N^{\pm}$  as

$$[N^{\pm}(a)]_{\ell} = \sum_{j \in \mathbb{Z}} a(j) \widehat{u}^{s\pm}(\ell; j) \quad \ell \in \mathbb{Z}. \quad (4.19)$$

Let us introduce the operator  $T : L^2(D) \rightarrow L^2(D)$  defined by

$$Tf := k^2(n-1)(f + w|_D) \quad (4.20)$$

with  $w$  being the solution of (4.10). We then have the following:

**Lemma 4.3.1.** *The operators  $G^\pm$  defined by (4.17) can be factorized as*

$$G^\pm = (\mathcal{H}^\pm)^* \mathsf{T}.$$

*Proof.* Let  $f \in L^2(D)$  and  $w \in H_{\#}^1(\Omega_M^h)$  be solution to (4.10). By definition of the Rayleigh coefficients and combining with (4.12) we have

$$\begin{aligned} \widehat{w}^\pm(j) &= \frac{i}{2ML} \int_{x_d=\pm h} e^{-i\alpha_{\#}(j)\bar{x}} \int_D \sum_{\ell \in \mathbb{Z}} \frac{1}{\beta_{\#}(\ell)} e^{i\alpha_{\#}(\ell)(\bar{x}-\bar{y}) + i\beta_{\#}(\ell)|h \mp y_d|} \mathsf{T}(f)(y) \, dy \, d\bar{x} \\ &= \int_D \frac{ie^{i\beta_{\#}(j)h}}{2\beta_{\#}(j)} e^{-i\alpha_{\#}(j)\bar{y} \mp i\beta_{\#}(j)y_d} \mathsf{T}f(y) \, dy \end{aligned} \quad (4.21)$$

Observing that  $\frac{ie^{i\beta_{\#}(j)h}}{2\beta_{\#}(j)} e^{-i\alpha_{\#}(j)y_1 \mp i\beta_{\#}(j)y_2} = \overline{u^{i,\pm}(y; j)}$ , we then have

$$\widehat{w}^\pm(j) = \int_D \mathsf{T}f(y) \overline{u^{i,\pm}(y; j)} \, dy,$$

which proves the lemma.  $\square$

Using Lemma 4.3.1 we end up with

$$N^\pm = (\mathcal{H}^\pm)^* \mathsf{T} \mathcal{H}^\pm. \quad (4.22)$$

The justification of the Sampling Methods that will be introduced later uses the solvability of the so-called *interior transmission problem* defined as: Seek  $(u, v) \in L^2(D) \times L^2(D)$  such that  $u - u_0 \in H_{\#}^2(D)$  and

$$\begin{cases} \Delta u + k^2 n u = 0 & \text{in } D, \\ \Delta v + k^2 v = 0 & \text{in } D, \\ u - v = \varphi & \text{on } \partial D, \\ \partial(u - v)/\partial\nu = \psi & \text{on } \partial D, \end{cases} \quad (4.23)$$

for given  $(\varphi, \psi) \in H^{3/2}(\partial D) \times H^{1/2}(\partial D)$  where  $\nu$  denotes the outward normal on  $\partial D$ . Values of  $k$  for which this problem is not well posed are referred to as *transmission eigenvalues*. For a detailed discussion on the solution to this problem we refer to [30, 31, 32, 102]. For our purpose we shall assume that this problem is well posed.

**Assumption 4.3.2.** *We assume that the refractive index  $n$  and the real wave number  $k$  are such that (4.23) defines a well posed problem.*

We recall that the well-posedness of (4.23) requires at least that  $n \neq 1$  in a neighborhood of  $\partial D$  and that  $k$  is outside a countable set without finite accumulation points.

### 4.3.2 Some useful properties for sampling methods

Most of our results are based on the assumption that

$$\Omega_M \setminus D \text{ is connected and } \partial\Omega_m \cap \overline{D} = \emptyset \text{ for all } m \in \mathbb{Z}_M.$$

The last assumption can be avoided with minor adaptations by changing the structure of the interior transmission problem. This assumption also justifies the use of  $N^+$  or  $N^-$  and not both

of them. We made the choice to adopt this assumption in order to avoid unnecessary additional technicalities.

A first step towards the justification of the sampling methods is the characterization of the closure of the range of  $\mathcal{H}^\pm$ .

**Lemma 4.3.3.** *The operator  $\mathcal{H}^\pm$  is compact and injective. Let  $H_{\text{inc}}^\pm(D)$  be the closure of the range of  $\mathcal{H}^\pm$  in  $L^2(D)$ . Then*

$$H_{\text{inc}}^\pm(D) = H_{\text{inc}}(D) := \{v \in L^2(D) : \Delta v + k^2 v = 0 \text{ in } D\}. \quad (4.24)$$

*Proof.* We shall prove this lemma only for  $\mathcal{H}^+$  since the proof for  $\mathcal{H}^-$  is similar. Let  $a = \{a(j)\}_{j \in \mathbb{Z}} \in \ell^2(\mathbb{Z})$  and assume that  $\mathcal{H}^+ a = 0$  in  $D$ . Since,

$$\Delta(\mathcal{H}^+ a) + k^2(\mathcal{H}^+ a) = 0 \quad \text{in } \mathbb{R}^3$$

then by the unique continuation principle,  $\mathcal{H}^+ a = 0$  in  $\mathbb{R}^3$ . This implies that

$$0 = (\mathcal{H}^+ a)(\bar{x}, h) = -\frac{i}{2} \sum_{j \in \mathbb{Z}} \frac{a(j)}{\bar{\beta}_\#(j)} e^{i\alpha_\#(j) \cdot \bar{x}}$$

for all  $\bar{x} \in \mathbb{R}^{d-1}$ . This implies, using the inverse Fourier transform that  $a_j = 0$  for all  $j \in \mathbb{Z}$ , which proves the injectivity of  $\mathcal{H}^+$ .

We now prove identity (4.24). We first obviously see that  $H_{\text{inc}}^+(D) \subset H_{\text{inc}}(D)$ . To prove the identity (4.24) it is then sufficient to prove that the adjoint  $(\mathcal{H}^+)^*$  is injective on  $H_{\text{inc}}(D)$ . Let  $f \in H_{\text{inc}}(D)$  and assume that  $(\mathcal{H}^+)^*(f) = 0$ . We set

$$u(x) := \int_D G_M(x, y) f(y) dy, \quad x \in \mathbb{R}^3, \quad (4.25)$$

where  $G_M$  is the  $ML$ -periodic Green function that has the expansion (4.12). From the regularity properties of volume potentials, we infer that  $u \in H_{\#, \text{loc}}^2(\Omega_M)$  and satisfies

$$\begin{cases} \Delta u + k^2 u = -f & \text{in } D, \\ \Delta u + k^2 u = 0 & \text{in } \Omega_M \setminus \bar{D}. \end{cases} \quad (4.26)$$

From expansion (4.12) and the definition of  $u$  in (4.25) we have that

$$\begin{aligned} \widehat{u}^+(j) &= \int_{\llbracket M_L^-, M_L^+ \rrbracket} \int_D \frac{i}{2ML} \sum_{\ell \in \mathbb{Z}} \frac{1}{\beta_\#(\ell)} e^{i\alpha_\#(\ell)(\bar{x}-\bar{y}) + i\beta_\#(\ell)(h-y_d)} f(y) dy e^{-i\alpha_\#(j) \cdot \bar{x}} d\bar{x} \\ &= \int_D f(y) \frac{i}{2\beta_\#(j)} e^{i\alpha_\#(j) - i\beta_\#(j)(x_d - h)} = ((\mathcal{H}^+)^*(f))(j). \end{aligned} \quad (4.27)$$

i.e.,  $(\mathcal{H}^+)^*(f) = \{\widehat{u}^+(j)\}_{j \in \mathbb{Z}}$ , the Rayleigh sequence of  $u$ . Therefore, the assumption  $(\mathcal{H}^+)^*(f) = 0$  implies that  $\widehat{u}^+(j) = 0$  for all  $j \in \mathbb{Z}$  and therefore  $u = 0$  for all  $x_d > h$ . By the unique continuation principal and since  $\Omega_M \setminus \bar{D}$  is connected, we have that  $u = 0$  in  $\Omega_M \setminus D$ . The regularity  $u \in H_{\#, \text{loc}}^2(\Omega_M)$  then implies  $u \in H_0^2(D)$ . Taking the  $L^2(D)$  scalar product of the first equation in (4.26) with  $f$  we obtain

$$\int_D (\Delta u + k^2 u) \bar{f} dx = \|f\|_{L^2(D)}^2.$$

Since  $\Delta f + k^2 f = 0$  in  $D$  in the sense of distributions and since  $u \in H_0^2(D)$ , then

$$\int_D (\Delta u + k^2 u) \bar{f} dx = 0,$$

which proves that  $f = 0$  and yields the injectivity of  $(\mathcal{H}^+)^*$  on  $H_{\text{inc}}(D)$ .  $\square$

Lemma 4.3.3 shows in particular that the closure of the range of  $\mathcal{H}^\pm$  are identical and will be denoted in the sequel by  $H_{\text{inc}}(D)$ . The following reciprocity lemma will also be useful.

**Lemma 4.3.4.** *Let  $f_0, f_1 \in L^2(D)$  and let  $w_0$  and  $w_1 \in H_{\#}^1(\Omega_M^h)$  be the corresponding solutions satisfying (4.10) with  $f = f_0, f = f_1$  respectively. Then*

$$\int_D (1-n)w_0 f_1 dx = \int_D (1-n)w_1 f_0 dx. \quad (4.28)$$

*Proof.* Taking  $v = w_0$  and  $v = w_1$  in the variational formulation (4.10) satisfied by  $w_1$  and  $w_0$  respectively then taking the difference yields

$$\begin{aligned} \int_{\Gamma_M^h} \frac{\partial w_0}{\partial x_2} w_1 - \frac{\partial w_1}{\partial x_2} w_0 ds(x) - \int_{\Gamma_M^{-h}} \frac{\partial w_0}{\partial x_2} w_1 - \frac{\partial w_1}{\partial x_2} w_0 ds(x) \\ = k^2 \int_D (1-n)f_0 w_1 - (1-n)f_1 w_0 dx. \end{aligned} \quad (4.29)$$

Obviously  $T^\pm$  are symmetric, since (using Parceval's theorem)

$$\pm \int_{\Gamma_M^{\pm h}} \frac{\partial w_0}{\partial x_2} w_1 ds(x) = \sum_{\ell \in \mathbb{Z}} i\beta_{\#}(\ell) \widehat{w}_0^\pm(\ell) \widehat{w}_1^\pm(-\ell) \quad (4.30)$$

and  $\beta_{\#}(\ell) = \beta_{\#}(-\ell)$  for all  $\ell \in \mathbb{Z}$ . Therefore,

$$\pm \int_{\Gamma_M^{\pm h}} \left( \frac{\partial w_0}{\partial x_2} w_1 - \frac{\partial w_1}{\partial x_2} w_0 \right) ds(x) = 0,$$

which proves the lemma.  $\square$

We now prove one of the main ingredients for the justification of the inversion methods discussed below. From now on, for  $z \in \Omega_M^h$ , we denote  $\Phi(\cdot; z) := G_M(\cdot - z)$  which has the Rayleigh sequences  $\widehat{\Phi}^\pm(\cdot; z) := \{\widehat{\Phi}^\pm(\ell; z)\}_{\ell \in \mathbb{Z}}$  with

$$\widehat{\Phi}^\pm(\ell; z) := \frac{i}{2ML\beta_{\#}(\ell)} e^{-i(\alpha_{\#}(\ell)\bar{z} - \beta_{\#}(\ell)|z_d \mp h|)}.$$

**Theorem 4.3.5.** *Assume that Assumptions 4.2.1 and 4.3.2 hold. Then the operator  $G^\pm : H_{\text{inc}}(D) \rightarrow \ell^2(\mathbb{Z})$  defined by (4.17) is injective with dense range. Moreover  $\widehat{\Phi}^\pm(\cdot; z)$  belongs to  $\mathcal{R}(G^\pm)$  if and only if  $z \in D$ .*

*Proof.* We only give here the proof of theorem for the operator  $G^+$  since the proof for the operator  $G^-$  is similar. We start by proving that  $G^+ : H_{\text{inc}}(D) \rightarrow \ell^2(\mathbb{Z})$  is injective with dense range. Let  $f \in H_{\text{inc}}(D)$  and let  $w \in H_{\#}^1(\Omega_M^h)$  be the associated scattered field via (4.10). As observed earlier,  $w \in H_{\#, \text{loc}}^2(\Omega_M)$ . Assume that  $G^+(f) = 0$ . Then  $w = 0$  for  $x_d > h$ . By unique continuation principal we then deduce that

$$w = 0 \quad \text{in } \Omega_M \setminus D,$$

and therefore  $w \in H_0^2(D)$ . We now set,  $u := w + f$ , then the pair  $(u, f)$  is a solution to (4.23) with zero data. Assumption 4.3.2 then ensures that  $f = 0$ , which proves the injectivity of  $G^+$ . We now prove the denseness of the range of  $G^+$ . Let  $g \in \overline{\mathcal{R}(G^+)}^\perp$ . Then

$$(G^+(f), g)_{\ell^2(\mathbb{Z})} = 0, \quad \forall f \in H_{\text{inc}}(D).$$



Consider  $f$  of the form  $f = \overline{\mathcal{H}^+(a)}$  for some  $a \in \ell^2(\mathbb{Z})$ . Since  $G^+ = (\mathcal{H}^+)^*T$ , we then have

$$\langle T(\overline{\mathcal{H}^+(a)}), \mathcal{H}^+(g) \rangle_{L^2(D)} = 0, \quad \forall a \in \ell^2(\mathbb{Z}). \quad (4.31)$$

Let  $w(a)$  and  $w(g)$  solution to (4.10) with respect to  $\overline{\mathcal{H}^+(a)}$  and  $\overline{\mathcal{H}^+(g)}$ . From Lemma 4.3.4 we get

$$\begin{aligned} \left( T(\overline{\mathcal{H}^+(a)}), \mathcal{H}^+(g) \right)_{L^2(D)} &= k^2 \int_D (n-1)(\overline{\mathcal{H}^+a} + w(a)) \overline{\mathcal{H}^+g} \, dx \\ &= k^2 \int_D (n-1)(\overline{\mathcal{H}^+g} + w(g)) \overline{\mathcal{H}^+g} \, dx. \end{aligned}$$

Therefore,

$$\left( T(\overline{\mathcal{H}^+(g)}), \mathcal{H}^+(a) \right)_{L^2(D)} = \left( T(\overline{\mathcal{H}^+(a)}), \mathcal{H}^+(g) \right)_{L^2(D)}, \quad \forall a \in \ell^2(\mathbb{Z}).$$

We deduce from (4.31) that

$$\left( G^+(\overline{\mathcal{H}^+(g)}), a \right)_{\ell^2(\mathbb{Z})} = 0, \quad \forall a \in \ell^2(\mathbb{Z}),$$

which implies  $G^+(\overline{\mathcal{H}^+(g)}) = 0$ . The injectivity of  $G^+$  ensures that  $\mathcal{H}^+g = 0$  in  $D$  and consequently  $g = 0$  (by Lemma 4.3.3). This proves the denseness of the range of  $G^+$ .

We now prove the last part of the theorem. We first observe that  $\widehat{\Phi}^+(\cdot; z)$  is the Rayleigh sequence of  $\Phi(\cdot; z)$  satisfying  $\Delta\Phi(\cdot; z) + k^2\Phi(\cdot; z) = -\delta_z$  in  $\Omega_M$  and the Rayleigh radiation condition. Let  $z \in D$ . We consider  $(u, v) \in L^2(D) \times L^2(D)$  as being the solution to (4.23) with

$$\varphi(x) = \Phi(x; z) \text{ and } \psi(x) = \partial\Phi(x; z)/\partial\nu(x) \text{ for } x \in \partial D. \quad (4.32)$$

We then define  $w$  by

$$\begin{aligned} w(x) &= u(x) - v(x) \quad \text{in } D, \\ w(x) &= \Phi(x; z) \quad \text{in } \Omega_M \setminus D. \end{aligned}$$

Due to (4.32), we have that  $w \in H_{\#, \text{loc}}^2(\Omega_M)$  and satisfies (4.10). Hence  $G^+w = \widehat{\Phi}^+(\cdot; z)$ .

Now let  $z \in \Omega_M \setminus D$ . Assume that there exists  $f \in H_{\text{inc}}(D)$  such that  $G^+f = \widehat{\Phi}^+(\cdot; z)$ . This implies that  $w = \Phi(\cdot; z)$  in  $\{x \in \Omega_M, \pm x_d \geq h\}$  where  $w$  is the solution to (4.10). By the unique continuation principle we deduce that  $w = \Phi(\cdot; z)$  in  $\Omega_M \setminus D$ . This gives a contradiction since  $w \in H_{\#, \text{loc}}^2(\Omega_M \setminus D)$  while  $\Phi(\cdot; z) \notin H_{\#, \text{loc}}^2(\Omega_M \setminus D)$ .  $\square$

**Lemma 4.3.6.** *Let  $O$  be an open domain such that  $\overline{O} \subset D$  and assume that there exists a real valued function  $n_0 \in L^\infty(D)$  such that  $(n_0(x) - 1) \geq \sigma > 0$ ,  $x \in D$  (respectively  $(1 - n_0(x)) \geq \sigma > 0$ ,  $x \in D$ ) and  $\text{Re } n = n_0$  in  $D \setminus O$ . Then the operator  $T : L^2(D) \rightarrow L^2(D)$  defined by (4.20) satisfies*

$$\text{Im}(T\phi, \phi) \geq 0, \quad \forall \phi \in H_{\text{inc}}(D) \quad (4.33)$$

and  $\text{Re } T = T_0 + T_1$ , where  $T_0$  (respectively  $-T_0$ ) is self-adjoint and coercive and  $T_1$  is compact on  $H_{\text{inc}}(D)$ . Moreover, assume in addition that Assumption (4.3.2) holds or  $(n-1)^{-1} \in L^\infty(D)$ . Then  $T$  is injective on  $H_{\text{inc}}(D)$

*Proof.* Let  $\phi \in L^2(D)$  and  $w_\phi$  be solution to (4.10) associated with  $f = \phi$ . By definition of the operator  $T$  we have

$$\langle T\phi, \phi \rangle_{L^2(D)} = k^2 \int_D (n-1)(\phi + w_\phi)\bar{\phi} = k^2 \int_D (n-1)|\phi + w_\phi|^2 - k^2 \int_D (n-1)(w_\phi + \phi)\bar{w}_\phi \quad (4.34)$$

where, using the variational formulation (4.10),

$$\begin{aligned} -k^2 \int_D (n-1)(w_\phi + \phi)\bar{w}_\phi &= \int_D (\Delta w_\phi + w_\phi)\bar{w}_\phi \\ &= \langle T^+ w_\phi, w_\phi \rangle + \langle T^- w_\phi, w_\phi \rangle - \int_D |\nabla w_\phi|^2 - k^2 |w_\phi|^2 dx \end{aligned} \quad (4.35)$$

Therefore,

$$\begin{aligned} \langle T\phi, \phi \rangle &= k^2 \int_D (n-1)|\phi + w_\phi|^2 \\ &\quad - \int_D |\nabla w_\phi|^2 - k^2 |w_\phi|^2 + \langle T^+ w_\phi, w_\phi \rangle + \langle T^- w_\phi, w_\phi \rangle \end{aligned} \quad (4.36)$$

Thanks to the non-negative sign of the imaginary part of  $T^\pm$  and the assumption  $\text{Im}(n) \geq 0$  we deduce that

$$\text{Im} \langle T\phi, \phi \rangle = k^2 \int_D \text{Im}(n)|w_\phi + \phi|^2 + \text{Im} \langle T^+ w_\phi, w_\phi \rangle + \text{Im} \langle T^- w_\phi, w_\phi \rangle \geq 0. \quad (4.37)$$

Define  $T_0 : L^2(D) \rightarrow L^2(D)$  by

$$T_0\phi := k^2(n_0 - 1)\phi. \quad (4.38)$$

Clearly,  $T_0$  (respectively  $-T_0$ ) is self-adjoint and coercive. Moreover,  $T_1 = T - T_0$  satisfies,  $T_1\phi = k^2(1-n)w_\phi + k^2(n_0 - \text{Re } n)\phi$ . The application  $\phi \mapsto w_\phi$  is continuous from  $L^2(D)$  into  $H^2(D)$  and since for  $\phi \in H_{\text{inc}}(D)$ ,  $\Delta\phi + k^2\phi = 0$  in  $D$ , the application that  $\phi \mapsto \phi|_O$  is continuous from  $H_{\text{inc}}(D)$  into  $H^2(O)$ . Therefore, the operator  $T_1 : H_{\text{inc}}(D) \rightarrow L^2(D)$  is compact thanks to the Rellich compact embedding theorem.

Assume that  $\phi \in H_{\text{inc}}(D)$  and  $T\phi = k^2(n-1)(\phi + w_\phi) = 0$ . This implies that  $w_\phi = 0$  by uniqueness of solutions to problem (4.10) with  $n = 1$ .

If we assume that in addition  $(n-1)^{-1} \in L^\infty(D)$ . Then  $T\phi = 0$  also implies  $\phi + w_\phi = 0$  in  $D$  and therefore  $\phi = 0$ .

The injectivity of  $T$  is remains true if Assumption 4.3.2 holds. With  $\phi \in H_{\text{inc}}(D)$  verifying  $\Delta\phi + k^2\phi = 0$  in  $D$  we get that  $u := \phi + w_\phi$  and  $v := \phi$  are such that  $u - v \in H_0^2(D)$  and satisfy the interior transmission problem (4.23) with  $\varphi = \psi = 0$ . We then deduce that  $u = v = 0$ . This concludes the proof of the injectivity of the operator  $T$ .  $\square$

**Lemma 4.3.7.** *Assume that Assumptions (4.2.1) and (4.3.2) hold. Then the operators  $N^\pm$  are injective with dense ranges.*

*Proof.* The injectivity and the denseness of the ranges of  $N^\pm$  directly follow from the same properties satisfied by  $\mathcal{H}^\pm$  (Lemma 4.3.3) and  $G^\pm$  (Theorem 4.3.5).  $\square$

## 4.4 Application to Sampling methods

We shall provide here the theoretical justifications of three sampling methods : the Linear Sampling Method, the Factorization Method and the Generalized Linear Sampling Method, to reconstruct the domain  $D$  from one of the near field operators  $N^\pm$ . These justifications are mainly a direct application of the results of the previous section and the abstract theoretical framework of these methods that is recalled in the appendix. This section is preparatory to the next section where we propose a new algorithm capable of reconstructing directly the domain  $\omega$  from  $N^\pm$ .

### 4.4.1 The Linear Sampling Method (LSM)

We give here the classical justification for the use of so-called Linear Sampling Method (LSM). This justification is a consequence of Theorem 4.3.5 and the Lemma 4.3.3.

**Theorem 4.4.1.** *Assume that Assumptions (4.2.1) and (4.3.2) hold. Then:*

- If  $z \in D$  then there exists a sequence  $a_\alpha^\pm(z) \in \ell^2(\mathbb{Z})$  such that  $\lim_{\alpha \rightarrow 0} \|N^\pm(a_\alpha^\pm(z)) - \widehat{\Phi}^\pm(\cdot; z)\|_{\ell^2(\mathbb{Z})} = 0$  and  $\lim_{\alpha \rightarrow 0} \|\mathcal{H}^\pm a_\alpha^\pm(z)\|_{L^2(D)} < \infty$ .
- If  $z \notin D$  then for all  $a_\alpha^\pm(z) \in \ell^2(\mathbb{Z})$  such that  $\lim_{\alpha \rightarrow 0} \|N^\pm(a_\alpha^\pm(z)) - \widehat{\Phi}^\pm(\cdot; z)\|_{\ell^2(\mathbb{Z})} \rightarrow 0$ ,  $\lim_{\alpha \rightarrow 0} \|\mathcal{H}^\pm a_\alpha^\pm(z)\|_{L^2(D)} = \infty$ .

*Proof.* The proof is classical but we give it here for the reader's convenience.

If  $z \in D$  then let  $f \in H_{\text{inc}}(D)$  be such that  $G^\pm f = \widehat{\Phi}^\pm(\cdot; z)$  which exists by Theorem 4.3.5. From Lemma 4.3.3 there exists a sequence  $a_z^\alpha \in \ell^2(\mathbb{Z})$  such that  $\mathcal{H}^\pm a_z^\alpha \rightarrow f$  as  $\alpha \rightarrow 0$ , and the first statement follows from the fact that  $N^\pm = G^\pm \mathcal{H}^\pm$ .

Let  $z \notin D$  and  $a_\alpha^\pm(z) \in \ell^2(\mathbb{Z})$  be such that  $\lim_{\alpha \rightarrow 0} \|N^\pm a_\alpha^\pm(z) - \widehat{\Phi}^\pm(\cdot; z)\|_{\ell^2(\mathbb{Z})} \rightarrow 0$ . Assume that  $\|\mathcal{H}^\pm a_\alpha^\pm(z)\|_{L^2(D)}$  is bounded as  $\alpha \rightarrow 0$ . We can assume that  $\mathcal{H}^\pm a_\alpha^\pm(z)$  weakly converges to some  $f \in H_{\text{inc}}(D)$ . Since  $G^\pm \mathcal{H}^\pm = N^\pm$  then we get as  $\alpha \rightarrow 0$ ,  $G^\pm f = \widehat{\Phi}^\pm(\cdot; z)$  which contradicts the last part of Theorem 4.3.5.  $\square$

This theorem does not indicate how to construct the sequence  $a_\alpha^\pm(z)$  when  $z \in D$ . In practice one relies on the use of Tikhonov regularization and considers  $\tilde{a}_\alpha^\pm(z) \in \ell^2(\mathbb{Z})$  satisfying

$$(\alpha + (N^\pm)^* N^\pm) \tilde{a}_\alpha^\pm(z) = (N^\pm)^* \left( \widehat{\Phi}^\pm(\cdot; z) \right). \quad (4.39)$$

Since  $N^\pm$  has dense range,  $\lim_{\alpha \rightarrow 0} \|N^\pm \tilde{a}_\alpha^\pm(z) - \widehat{\Phi}^\pm(\cdot; z)\|_{\ell^2(\mathbb{Z})} = 0$ . However, one cannot guarantee in general that  $\lim_{\alpha \rightarrow 0} \|\mathcal{H}^\pm \tilde{a}_\alpha^\pm(z)\|_{L^2(D)} < \infty$  if  $z \in D$ . In addition, one cannot compute  $\|\mathcal{H}^\pm a_\alpha^\pm(z)\|_{L^2(D)}$  since  $D$  is not known. In practice one uses  $z \rightarrow \|a_\alpha^\pm(z)\|_{\ell^2(\mathbb{Z})}$  as an indicator function for  $D$ . A possible method to fix the Tikhonov regularization parameter  $\alpha$  in (4.39) is to use the Morozov discrepancy principle. Assume that  $N^{\pm, \delta}$  is the noisy operator corresponding to noisy measurements, i.e

$$\|N^{\pm, \delta} - N^\pm\| \leq \delta.$$

Then for each sampling point  $z$ , the parameter  $\alpha$  is chosen such that

$$\|N^{\pm, \delta} a_\alpha^\pm(z) - \widehat{\Phi}^\pm(\cdot; z)\|_{\ell^2(\mathbb{Z})} = \delta \|a_\alpha^\pm(z)\|_{\ell^2(\mathbb{Z})}.$$

This leads to a nonlinear equation that determines  $\alpha$  in terms of the noise level  $\delta$  [47].

### 4.4.2 The Factorization Method

We here proceed with the justification of the Factorization method that has been introduced in [70] and that was applied in a number of papers to various configurations [76, 56, 26, 59, 11, 10, 71, 77, 79, 80, 81, 98, 40, 61]. Our setting is similar to the case of guided waves that has been treated in [27] or the case of periodic media in [82] where the half space problem was considered. We include the analysis of this method here since we shall prove the counterpart for a single Floquet-Bloch mode later. Let us define the operator

$$N_{\#}^{\pm} := |\operatorname{Re}(N^{\pm})| + |\operatorname{Im}(N^{\pm})| \quad (4.40)$$

where  $\operatorname{Re}(N^{\pm}) := \frac{1}{2}(N^{\pm} + (N^{\pm})^*)$ ,  $\operatorname{Im}(N^{\pm}) := \frac{1}{2i}(N^{\pm} - (N^{\pm})^*)$ . Then we have the following theorem:

**Theorem 4.4.2.** *Under the hypothesis of Lemma 4.3.6, the following factorization holds:*

$$N_{\#}^{\pm} = (\mathcal{H}^{\pm})^* T_{\#}^{\pm} \mathcal{H}^{\pm}, \quad (4.41)$$

where  $T_{\#}^{\pm} : L^2(D) \rightarrow L^2(D)$  is self-adjoint and coercive on  $H_{\text{inc}}(D)$ . Moreover,  $z \in D$  if and only if  $\widehat{\Phi}^{\pm}(\cdot; z) \in \mathcal{R}\left((N_{\#}^{\pm})^{1/2}\right)$ .

*Proof.* The proof of this theorem is a direct application of the abstract framework given in Theorem A.1.2 using the results of Lemma 4.3.6 and Lemma 4.4.3 below.  $\square$

**Lemma 4.4.3.** *For  $z \in \Omega_M$ ,  $z \in D$  if and only if  $\widehat{\Phi}^{\pm}(\cdot; z)$  belongs to the range of  $(\mathcal{H}^{\pm})^*$ .*

*Proof.* For  $z \in D$  choose a cut-off function  $\rho \in C^{\infty}(\Omega_M)$  which vanishes near  $z$  and equals one in  $\Omega_M \setminus D$ . We define  $v(x) := \rho(x)G_{M,z}$ . Then the Rayleigh sequence of  $v(x)$  are  $\widehat{\Phi}^{\pm}(\cdot; z)$ . We observe that  $f := -(\Delta v + k^2 v)$  has compact support in  $D$  and  $f \in L^2(D)$ . Since  $v$  satisfies the Rayleigh radiation condition, then

$$v(x) = \int_D G_M(x-y)f(y)dy. \quad (4.42)$$

We hence have from expansion (4.12) and the fact the Rayleigh sequences of  $v$  and  $\Phi(\cdot; z)$  are identical that

$$\widehat{\Phi}^{\pm}(\cdot; z) = (\mathcal{H}^{\pm})^* f.$$

We now assume that  $z \notin D$  (without loss of generality we can assume that  $D \cup \{z\} \subset \Omega_M^h$ ) and  $\widehat{\Phi}^{\pm}(\cdot; z) = (\mathcal{H}^{\pm})^* f$  for some  $f \in L^2(D)$ . We also consider  $v$  which is defined by (4.42). Since  $\Phi(\cdot; z)$  and  $v$  satisfies the Rayleigh radiation condition then  $\Phi(\cdot; z) = v$  in domain  $\pm x_d > h$ . By unique continuation principal we deduce that  $\Phi(\cdot; z) = v$  in the exterior of  $D \cup \{z\}$ . This gives a contradiction since  $v$  is smooth near  $z$  but  $\Phi(\cdot; z)$  is singular at  $z$ .  $\square$

We can also reformulate the second part of Theorem 4.4.2 by using Picard's criterion:  $z \in D$  if and only if the series

$$\sum_{m=1}^{\infty} \frac{\left| (\widehat{\Phi}^{\pm}(\cdot; z), \Psi_m^{\pm})_{\ell^2(\mathbb{Z})} \right|^2}{\lambda_m^{\pm}} \quad (4.43)$$

converges, where  $\{\lambda_m^{\pm}, \Psi_m^{\pm}\}$  is eigensystems of the self-adjoint and positive defined  $N_{\#}^{\pm}$ . This criterion can also be used in the numerical implementation of the factorization method with a suitable choice of the cut-off parameter with respect to the noise level. One can also rely on the use of Tikhonov regularization as explained above for the linear sampling method.

### 4.4.3 The Generalized Linear Sampling Method (GLSM)

This section is dedicated to the third family of sampling methods that has been introduced in the literature [14, 12, 30] and that somehow combines the benefits from the two previously presented sampling methods. The GLSM constructs a nearby solution as predicted by the LSM theorem by considering minimizing sequences of a cost functional with data fidelity the LSM residual and a penalty term the norm of the Herglotz function. The latter is constructed exploiting the factorization method. The first advantage of the GLSM is indeed to have a more convincing theoretical justification than LSM. Compared to the factorization method, the GLSM keep the link with the so-called interior transmission problem as for LSM which will be exploited later for the design of the new imaging functional capable of directly identifying a defect in a periodic background. We here restrict ourselves to the simplest version of GLSM that exploits “symmetric factorizations” of the data operator. For the treatment of other type of factorizations we refer the reader to [12].

We first present the noise free version of GLSM. We denote by  $(\cdot, \cdot)$  the  $\ell^2(\mathbb{Z})$  scalar product and by  $\|\cdot\|$  the associated norm. Let  $\alpha > 0$  be a given parameter and  $\phi \in \ell^2(\mathbb{Z})$ . We introduce functional  $J_\alpha(\phi; \cdot) : \ell^2(\mathbb{Z}) \rightarrow \mathbb{R}$  where

$$J_\alpha^\pm(\phi; a) := \alpha(N_\#^\pm a, a) + \|N^\pm a - \phi\|^2, \quad \forall a \in \ell^2(\mathbb{Z}), \quad (4.44)$$

and define

$$j_\alpha(\phi) := \inf_{a \in \ell^2(\mathbb{Z})} J_\alpha^\pm(\phi; a). \quad (4.45)$$

Let  $c(\alpha) > 0$  be such that  $c(\alpha)/\alpha \rightarrow 0$  as  $\alpha \rightarrow 0$ .

**Theorem 4.4.4.** *Assume that Assumptions 4.2.1, 4.3.2 and the hypothesis of Lemma 4.3.6 hold. Let  $z \in \Omega_M$  and define  $a_{\alpha, G}^\pm(z) \in \ell^2(\mathbb{Z})$  such that*

$$J_\alpha^\pm(\widehat{\Phi}^\pm(\cdot; z); a_{\alpha, G}^\pm(z)) \leq j_\alpha(\widehat{\Phi}^\pm(\cdot; z)) + c(\alpha). \quad (4.46)$$

Then  $z \in D$  if and only if  $\lim_{\alpha \rightarrow 0} (N_\#^\pm a_{\alpha, G}^\pm(z), a_{\alpha, G}^\pm(z)) < \infty$ . Moreover, if  $z \in D$  then  $\mathcal{H}^\pm a_{\alpha, G}^\pm(z) \rightarrow v(\cdot; z)$  in  $L^2(D)$  where  $(u(\cdot; z), v(\cdot; z)) \in L^2(D) \times L^2(D)$  is the solution of problem (4.23) with  $\varphi = \Phi(\cdot; z)$  and  $\psi = \partial\Phi(\cdot; z)/\partial\nu$  on  $\partial D$ .

*Proof.* The proof of this theorem is a direct application of the abstract framework given in Theorem A.2.2 in combination with Theorem A.1.2  $\square$

For the case with noise in the operators, one has to change the functional  $J_\alpha^\pm$ . More precisely, consider the noisy operators  $N^{\pm, \delta} : \ell^2(\mathbb{Z}) \rightarrow \ell^2(\mathbb{Z})$  and  $N_\#^{\pm, \delta} : \ell^2(\mathbb{Z}) \rightarrow \ell^2(\mathbb{Z})$  such that

$$\|N^{\pm, \delta} - N^\pm\| \leq \delta \|N^{\pm, \delta}\| \quad \text{and} \quad \|N_\#^{\pm, \delta} - N_\#^\pm\| \leq \delta \|N_\#^{\pm, \delta}\|, \quad (4.47)$$

for some  $\delta > 0$  and assume that the operators  $N^{\pm, \delta}$  and  $N_\#^{\pm, \delta}$  are compact. We then consider for  $\alpha > 0$  and  $\phi > 0$  the functional  $J_\alpha^\delta(\phi; \cdot) : \ell^2(\mathbb{Z}) \rightarrow \mathbb{R}$  defined by

$$J_\alpha^{\pm, \delta}(\phi; a) := \alpha \left( (N_\#^{\pm, \delta} a, a) + \delta \alpha^{-\eta} \|N_\#^{\pm, \delta}\| \|a\|^2 \right) + \|N^{\pm, \delta} a - \phi\|^2, \quad \forall a \in \ell^2(\mathbb{Z}), \quad (4.48)$$

with  $0 < \eta < 1$  a fixed parameter. Then we have the following result:

**Theorem 4.4.5.** *Assume that Assumptions 4.2.1, 4.3.2 and the assumptions of Lemma 4.3.6 hold. For  $z \in \Omega_M$  denote by  $a_{\alpha,\delta}^{\pm}(z)$  the minimizer of  $J_{\alpha}^{\pm,\delta}(\widehat{\Phi}^{\pm}(\cdot; z); \cdot)$  over  $\ell^2(\mathbb{Z})$ . Then,*

$$z \in D \text{ if and only if } \lim_{\alpha \rightarrow 0} \limsup_{\delta \rightarrow 0} \left( \left| (\mathbb{N}_{\#}^{\pm,\delta} a_{\alpha,\delta}^{\pm}(z), a_{\alpha,\delta}^{\pm}(z)) \right| + \delta \alpha^{-\eta} \|\mathbb{N}_{\#}^{\pm,\delta}\| \left\| a_{\alpha,\delta}^{\pm}(z) \right\|^2 \right) < \infty.$$

Moreover, if  $z \in D$  then  $\lim_{\alpha \rightarrow 0} \limsup_{\delta \rightarrow 0} \|\mathcal{H}^{\pm} a_{\alpha,\delta}^{\pm}(z) - v(\cdot; z)\|_{L^2(D)} = 0$  where  $(u(\cdot; z), v(\cdot; z)) \in L^2(D) \times L^2(D)$  is the solution to problem (4.23) with  $\varphi = \Phi(\cdot; z)$  and  $\psi = \partial\Phi(\cdot; z)/\partial\nu$  on  $\partial D$ .

*Proof.* This is a direct consequence of Theorem A.2.4.  $\square$

We remark that (for numerics) the minimizer  $a_{\alpha,\delta}^{\pm}(z)$  in Theorem 4.4.5 can be computed as the solution of

$$(\alpha \mathbb{N}_{\#}^{\pm,\delta} + \alpha^{1-\eta} \delta \|\mathbb{N}_{\#}^{\pm,\delta}\| I + (\mathbb{N}^{\pm,\delta})^* \mathbb{N}^{\pm,\delta}) a_{\alpha,\delta}^{\pm}(z) = (\mathbb{N}^{\pm,\delta})^* \widehat{\Phi}^{\pm}(\cdot; z). \quad (4.49)$$

Unfortunately, there is no a posteriori rule for the choice of the parameter  $\alpha$  as for the LSM method. In practice, one can follow the empirical rule proposed in [14] by taking

$$\alpha \equiv \alpha(\delta) / (\|\mathbb{N}_{\#}^{\pm,\delta}\| (1 + \delta)) \quad (4.50)$$

where  $\alpha(\delta)$  is the regularization parameter used in the LSM equation (4.39) and determined by the Morozov principle. The parameter  $\eta$  has little incidence on the numerics and can be set to 0. Therefore an approximate indicator function of the domain  $D$  is given by the function

$$z \mapsto \frac{1}{\left( (\mathbb{N}_{\#}^{\pm,\delta} a_{\alpha,\delta}^{\pm}(z), a_{\alpha,\delta}^{\pm}(z)) + \delta \|\mathbb{N}_{\#}^{\pm,\delta}\| \left\| a_{\alpha,\delta}^{\pm}(z) \right\|^2 \right)}. \quad (4.51)$$

#### 4.4.4 Reconstruction of the periodic domain $D_p$ from $\mathbb{N}^{\pm}$

We here explain how one can reconstruct the periodic background  $D_p$  from  $\mathbb{N}^{\pm}$  without knowing the refractive indices  $n$  and  $n_p$  using the same sampling operators as above. To reconstruct  $D_p$  we do not need the local perturbation  $\omega$  to be located in only one period but we need to assume that for all  $z \in \omega$  there exists  $m \in \mathbb{Z}^{d-1}$  such that  $z + m\mathbf{e} \in \Omega_M \setminus D$ . Here and in the following we denote by  $\mathbf{e}$  one of the vectors  $L_i \mathbf{e}_i$ ,  $i = 1, \dots, d-1$  where  $\mathbf{e}_1, \dots, \mathbf{e}_d$  is the canonical basis of  $\mathbb{R}^d$ .

We shall exploit the decomposition of the  $ML$ -periodic fundamental solution into  $L$  quasi-periodic fundamental solutions. We remark from using discrete Floquet-Bloch transform that an  $ML$ -periodic function can be decomposed into the sum of  $M$  quasi periodic functions with periods  $L$  and quasi-periodicity factors  $\alpha_q = 2\pi q/(ML)$ ,  $q = [-\frac{M}{2}] + 1, \dots, [\frac{M}{2}]$  as

$$u = \frac{1}{M} \sum_{[-\frac{M}{2}] + 1}^{[\frac{M}{2}]} u_q \quad (4.52)$$

where  $u_q$  has a Fourier expansion only on  $q + Mj$  Fourier modes. More precisely, for  $u \in L^2(\Omega_M^h)$  that is decomposed as

$$u(\bar{x}, x_d) = \sum_{j \in \mathbb{Z}^{d-1}} \tilde{u}(j, x_d) e^{i\alpha_{\#}(j)\bar{x}}$$

where

$$\tilde{u}(j, x_d) := \frac{1}{\llbracket ML \rrbracket} \int_{\llbracket M_L^-, M_L^+ \rrbracket} u(\bar{x}, x_d) e^{-i\alpha_{\#}(j)\bar{x}} d\bar{x},$$

we define  $u_q$  as:

$$u_q(\bar{x}, x_d) := M \sum_{j \in \mathbb{Z}^{d-1}} \tilde{u}(q + Mj, x_d) e^{i\alpha_{\#}(q+Mj)\bar{x}}.$$

Writing the decomposition

$$\Phi(\cdot; z) := \frac{1}{M} \sum_{q \in \mathbb{Z}_M} \Phi_q(\cdot; z),$$

the functions  $\Phi_q(\cdot; z)$  satisfies  $\Delta \Phi_q(\cdot; z) + k^2 \Phi_q(\cdot; z) = -\delta_z$  in  $\Omega_0$  and is  $\alpha_q$  quasi-periodic with period  $L$ . Moreover,

$$\Phi_q(\cdot; z + \mathbf{e}) = e^{-i\alpha_q \cdot \mathbf{e}} \Phi_q(\cdot; z), \quad z \in \mathbb{R}^d.$$

The Rayleigh coefficients  $\widehat{\Phi}_q^{\pm}(\cdot; z)$  of  $\Phi_q(\cdot; z)$  are given by

$$\widehat{\Phi}_q^{\pm}(j; z) = \begin{cases} \frac{i}{2L\beta_{\#}(q+M\ell)} e^{-i(\alpha_{\#}(q+M\ell)\bar{z} - \beta_{\#}(q+M\ell)|z_d \mp h|)} & \text{if } j = q + M\ell, \ell \in \mathbb{Z}, \\ 0 & \text{if } j \neq q + M\ell, \forall \ell \in \mathbb{Z}. \end{cases}$$

**Lemma 4.4.6.** *Assume that Assumptions (4.2.1) and (4.3.2) hold. Then  $\widehat{\Phi}_q^{\pm}(\cdot; z)$  belongs to  $\mathcal{R}(G^{\pm})$  if and only if  $z \in D_p$ .*

*Proof.* Consider first the case  $z \in D_p$  and let  $(u_q(\cdot; z), v_q(\cdot; z)) \in L^2(D) \times L^2(D)$  be the solution to the interior transmission problem (4.23) with

$$\varphi(x) = \Phi_q(x; z) \text{ and } \psi(x) = \partial \Phi_q(x; z) / \partial \nu(x) \quad \text{for } x \in \partial D. \quad (4.53)$$

We then define  $w_q$  by:

$$w_q(x) = \begin{cases} u_q(x; z) - v_q(x; z) & \text{in } D, \\ \Phi_q(x; z) & \text{in } \Omega_M \setminus D. \end{cases}$$

Due to (4.53),  $w_q \in H_{\#, \text{loc}}^2(\Omega_M)$  and since  $\Phi_q(\cdot; z)$  satisfies  $\Delta \Phi_q(\cdot; z) + k^2 \Phi_q(\cdot; z) = -\delta_z$  in  $\Omega_M \setminus D_p$  and the Rayleigh radiation condition (4.3), then  $w_q$  satisfies (4.10). Moreover, we get that  $\widehat{\Phi}_q^{\pm}(\cdot; z)$  are the Rayleigh coefficients of  $w_q$  and consequently  $G^{\pm} v_q(x; z) = \widehat{\Phi}_q^{\pm}(\cdot; z)$ .

Consider now the case  $z \in \Omega_M \setminus D_p$ . Assume that there exists  $f \in H_{\text{inc}}(D)$  such that  $G^{\pm} f = \widehat{\Phi}_q^{\pm}(\cdot; z)$ . This implies that  $w_q = \Phi_q(\cdot; z)$  in  $\{x \in \Omega_M, \pm x_d \geq h\}$  where  $w_q$  is the solution to (4.10). By the unique continuation principal we deduce that  $w_q = \Phi_q(\cdot; z)$  in  $\Omega_M \setminus D$ . Since there exists  $m \in \mathbb{Z}$  such that  $z + m\mathbf{e} \in \Omega_M \setminus D$  then  $\Phi_q(\cdot; z) \notin H_{\#, \text{loc}}^2(\Omega_M \setminus D)$  (since  $\Phi_q(\cdot; z)$  is singular at all points  $z + m\mathbf{e}$ ). This contradicts the fact that  $w \in H_{\#, \text{loc}}^2(\Omega_M \setminus D)$ .  $\square$

**Lemma 4.4.7.** *For  $z \in \Omega_M$ ,  $z \in D_p$  if and only if  $\widehat{\Phi}_q^{\pm}(\cdot; z)$  belongs to the range of  $(\mathcal{H}^{\pm})^*$ .*

*Proof.* For  $z \in D_p$ , we choose a cut-off  $L$ -periodic function  $\rho \in C^{\infty}(\Omega_M)$  that vanishes in a neighborhood of  $z$  and equals one in  $\Omega_M \setminus D_p$ . We define  $v(x) := \rho(x) \Phi_q(\cdot; z)$ . Then  $v(x)$  is  $\alpha_q$ -quasi-periodic and the Rayleigh sequence of  $v(x)$  is equal to  $\widehat{\Phi}_q^{\pm}(\cdot; z)$ . We observe that  $f := -(\Delta v + k^2 v)$  has compact support in  $D_p$  and  $f \in L^2(D)$ . Since  $v$  satisfies the Rayleigh radiation condition and is  $ML$ -periodic, then

$$v(x) = \int_{D_p} \Phi_q(x - y) f(y) dy = M \int_{D_p} \Phi(x - y) f(y) dy = M \int_D \Phi(x - y) f(y) dy. \quad (4.54)$$

For the first equality we used the fact  $f$  is  $\alpha_q$ -quasi-periodic function and therefore

$$v(x) = \int_{D_p} \Phi_{q'}(x-y)f(y)dy = 0, \quad \forall q' \in \mathbb{Z}_M, q' \neq q \quad (4.55)$$

while for the second equality we simply used  $f = 0$  in  $D \setminus D_p$ . From the expansion

$$\Phi_q(x) = \frac{i}{2L} \sum_{\ell \in \mathbb{Z}} \frac{1}{\beta_{\#}(q+M\ell)} e^{i\alpha_{\#}(q+M\ell)\bar{x} + i\beta_{\#}(q+M\ell)|x_d|} \quad (4.56)$$

and the fact that the Rayleigh sequences of  $v$  and  $\Phi_q(\cdot; z)$  are identical we obtain that

$$\widehat{\Phi}_q^{\pm}(\cdot; z) = (\mathcal{H}^{\pm})^*(Mf).$$

We now assume that  $z \notin D_p$  (without loss of generality we can assume that  $D_p \cup \{z\} \subset \Omega_M^h$ ) and  $\widehat{\Phi}_q^{\pm}(\cdot; z) = (\mathcal{H}^{\pm})^*f$  for some  $f \in L^2(D)$ . We also consider  $v$  which is defined by (4.54). Since  $\Phi_q(\cdot; z)$  and  $v$  satisfies the Rayleigh radiation condition then  $\Phi_q(\cdot; z) = v$  in the domain  $\pm x_d > h$ . By the unique continuation principle, we deduce that  $\Phi_q(\cdot; z) = v$  in the exterior of  $D \cup \{z+mL\}$ ,  $m \in \mathbb{Z}_M$ . This gives a contradiction since  $v$  is a smooth function in the neighborhood of the points  $\{z+mL\}$ ,  $m \in \mathbb{Z}_M$  but  $\Phi_q(\cdot; z) \notin H^2(\Omega_M \setminus \{z+mL, m \in \mathbb{Z}_M\})$ .  $\square$

The previous two results allow us to conclude that one can reconstruct the domain  $D_p$  using the sampling methods introduced above by replacing the sampling function  $\widehat{\Phi}^{\pm}(\cdot; z)$  with  $\widehat{\Phi}_q^{\pm}(\cdot; z)$ .

**Theorem 4.4.8.** *Let  $q$  be a fixed integer between  $[-\frac{M}{2}] + 1$  and  $[\frac{M}{2}]$ . Assume that Assumptions 4.2.1, 4.3.2 and the assumptions of Lemma 4.3.6 hold. Then Theorems 4.4.1, 4.4.2, 4.4.4 and 4.4.5 hold true if  $D$  is replaced by  $D_p$  and  $\widehat{\Phi}^{\pm}(\cdot; z)$  is replaced by  $\widehat{\Phi}_q^{\pm}(\cdot; z)$ .*

#### 4.4.5 Validating Numerical Experiments

We here give some numerical examples for testing the sampling methods introduced above. Some complementary results will be given later. We only consider here two dimensional examples. In order to obtain the data for the inverse problem, we consider  $N_{inc}$  incident down-to-up or up-to-down plane-waves given by formula (4.13) or (4.14). The synthetic data has been generated by solving the forward scattering problem using the spectral discretization scheme of the volume integral formulation of the problem presented in Chapter 3.

We denote by

$$\mathbb{Z}_{inc} := \{j = q + M\ell, [-\frac{M}{2}] + 1 \leq q \leq [\frac{M}{2}], -N_{min} \leq \ell \leq N_{max}\}$$

the set of indices for the incident waves (which is also the set of indices for measured Rayleigh coefficients). The values of all parameters will be indicated below. The discrete version of the operators  $N^{\pm}$  are given by the matrices of size  $N_{inc} \times N_{inc}$ ,

$$N^{\pm} := \left( \widehat{u}^{s^{\pm}}(\ell; j) \right)_{\ell, j \in \mathbb{Z}_{inc}}. \quad (4.57)$$

We add some noise to the data as:

$$N^{\pm, \delta}(j, \ell) := N^{\pm}(j, \ell)(1 + \delta A(j, \ell)), \quad \forall (j, \ell) \in \mathbb{Z}_{inc} \times \mathbb{Z}_{inc} \quad (4.58)$$

where  $A = (A(j, \ell))_{N_{inc} \times N_{inc}}$  is a matrix of uniform complex random variables with real and imaginary parts in  $[-1, 1]^2$  and  $\delta > 0$  is the noise level. We use in all following examples  $\delta = 1\%$ .



In this section we restrict ourselves to a simple setting for the geometry represented in Figure 4.2. The background domain geometry  $D_p$  is constituted by discs of radii  $r$  and the geometry of the perturbation is a disc of radius  $r_w$ .

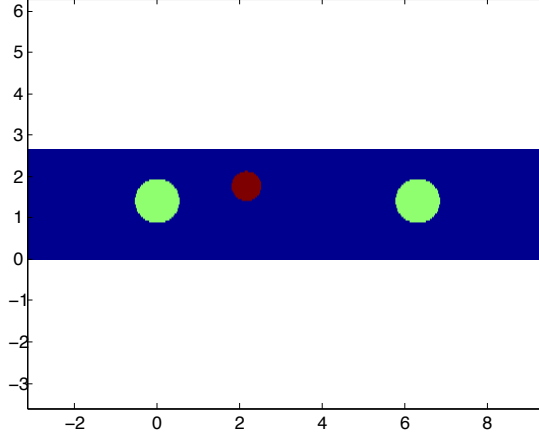


Figure 4.2: Exact geometry for the numerical examples.

The physical parameters are set as

$$k = \pi/3.14, n_p = 2 \text{ in } D_p \text{ and } n = 4 \text{ in } \omega. \quad (4.59)$$

Set  $\lambda := 2\pi/k$  as the wavelength. Then the geometrical parameters are

$$L = \pi\lambda, h = 1.5\lambda, r = 0.3\lambda \text{ and } r_w = 0.2\lambda. \quad (4.60)$$

The coordinate system is chosen such that the center of one disc of  $D_p$  is  $(0, 0.8\lambda)$  and the center of  $\omega$  is  $(1.2\lambda, \lambda)$ . Finally we choose as parameters for the discrete model

$$M = 2, N_{min} = 5 \text{ and } N_{max} = 5. \quad (4.61)$$

### Reconstruction using LSM

Let  $a_\alpha^\pm(z)$  be the Tikhonov regularized solution of (4.39) with respect to the operators  $N^{\pm, \delta}$  where the regularization parameter is computed using the Morozov discrepancy principle. We define the indicator functions

$$\mathcal{I}_+^{\text{LSM}}(z) := \frac{1}{\|a_\alpha^+(z)\|^2} \quad \text{and} \quad \mathcal{I}_-^{\text{LSM}}(z) := \frac{1}{\|a_\alpha^-(z)\|^2}. \quad (4.62)$$

In the following numerical example, illustrated by Figure 4.3, we also show the reconstruction of domain  $D$  using the indicator function

$$\mathcal{I}^{\text{LSM}}(z) := \mathcal{I}_+^{\text{LSM}}(z) + \mathcal{I}_-^{\text{LSM}}(z).$$

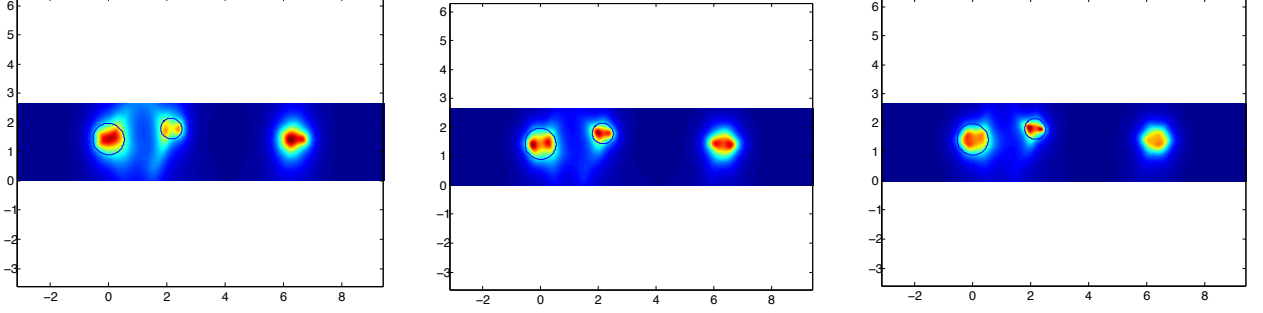


Figure 4.3: Reconstruction of the domain  $D = D_p \cup \omega$  using, from left to right,  $\mathcal{I}_+^{\text{LSM}}(z)$ ,  $\mathcal{I}_-^{\text{LSM}}(z)$  and  $\mathcal{I}^{\text{LSM}}(z)$ .

### Reconstruction using the Factorization Method

Let  $a_{\alpha, \#}^{\pm}(z)$  be the Tikhonov regularized solution of (4.39) with  $N^{\pm, \delta}$  is replaced by  $(N_{\#}^{\pm, \delta})^{1/2}$  and the regularization parameter is computed using the Morozov discrepancy principle. We define

$$\mathcal{I}_+^{\text{F}\#}(z) := \frac{1}{\|a_{\alpha, \#}^+(z)\|^2} \quad \text{and} \quad \mathcal{I}_-^{\text{F}\#}(z) := \frac{1}{\|a_{\alpha, \#}^-(z)\|^2}.$$

We also show the reconstruction of domain  $D$  using the indicator function

$$\mathcal{I}^{\text{F}\#}(z) := \mathcal{I}_+^{\text{F}\#}(z) + \mathcal{I}_-^{\text{F}\#}(z).$$

The numerical results of the Factorization method are given in Figure 4.4.

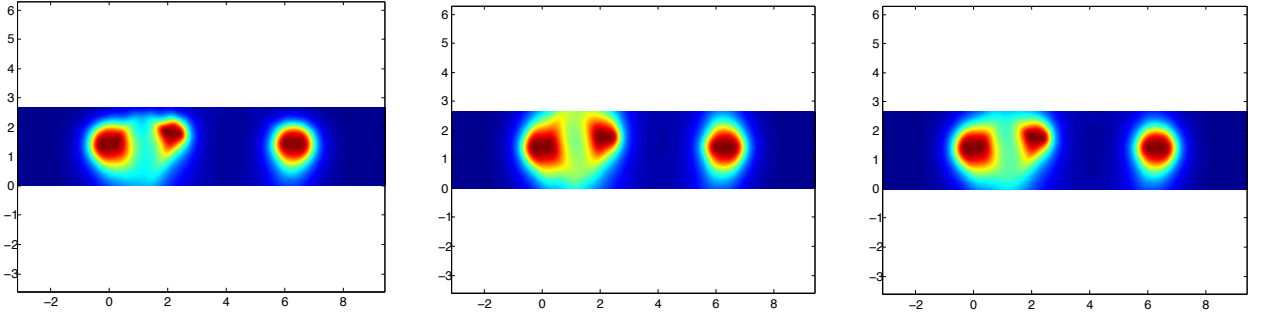


Figure 4.4: Reconstruction of the domain  $D = D_p \cup \omega$  using, from left to right,  $\mathcal{I}_+^{\text{F}\#}(z)$ ,  $\mathcal{I}_-^{\text{F}\#}(z)$  and  $\mathcal{I}^{\text{F}\#}(z)$ .

### Reconstruction using GLSM

Let  $a_{\alpha, G}^{\pm}(z) \in \ell^2(\mathbb{Z})$  be solution of (4.49) with the parameter  $\alpha$  defined in (4.50). We define

$$\mathcal{I}_+^{\text{GLSM}}(z) = \frac{1}{|(N_{\#}^{+, \delta} a_{\alpha, G}^+(z), a_{\alpha, G}^+(z))| + \delta \|N_{\#}^{+, \delta}\| \|a_{\alpha, G}^+(z)\|^2}$$

$$\mathcal{I}_-^{\text{GLSM}}(z) = \frac{1}{|(N_{\#}^{-, \delta} a_{\alpha, G}^-(z), a_{\alpha, G}^-(z))| + \delta \|N_{\#}^{-, \delta}\| \|a_{\alpha, G}^-(z)\|^2}.$$

We also show the reconstruction of domain  $D$  using the indicator function

$$\mathcal{I}^{\text{GLSM}}(z) := \mathcal{I}_+^{\text{GLSM}}(z) + \mathcal{I}_-^{\text{GLSM}}(z)$$

The numerical results of the GLSM are given in Figure 4.5

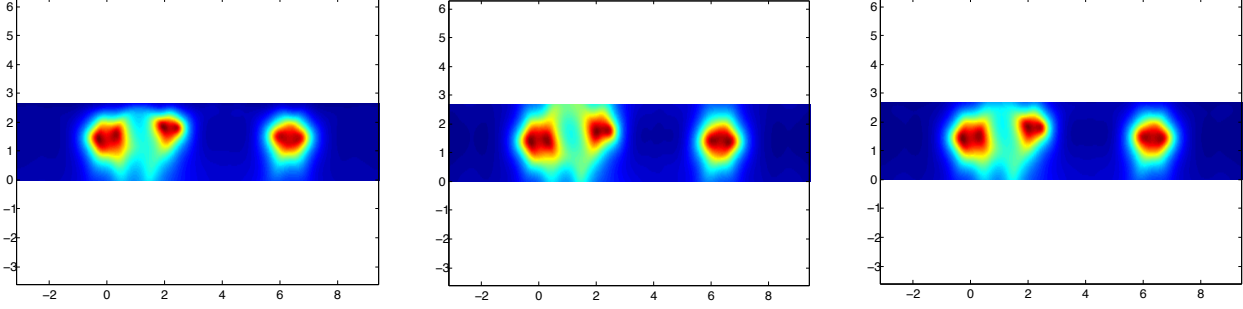


Figure 4.5: Reconstruction of the domain  $D = D_p \cup \omega$  using, from left to right,  $\mathcal{I}_+^{\text{GLSM}}(z)$ ,  $\mathcal{I}_-^{\text{GLSM}}(z)$  and  $\mathcal{I}^{\text{GLSM}}(z)$

### Reconstruction of the periodic domain $D_p$ from $\mathbb{N}^\pm$

Let  $q \in \mathbb{Z}_M$ . Let  $a_{\alpha,q}^\pm(z) \in \ell^2(\mathbb{Z})$  be the Tikhonov regularized solution of (4.39) with  $\widehat{\Phi}^\pm(\cdot; z)$  is replaced by  $\widehat{\Phi}_q^\pm(\cdot; z)$ .

We then define  $\mathcal{I}_\pm^{\text{LSMq}}(z)$  as in (4.62) where  $a_\alpha^\pm(z)$  is replaced by  $a_{\alpha,q}^\pm(z)$ . We similarly define  $\mathcal{I}_\pm^{\text{F}\#,q}(z)$  and  $\mathcal{I}_\pm^{\text{GLSMq}}(z)$ .

Figure 4.6 indicates the obtained reconstructions of the periodic domain  $D_p$  using the following indicator functions

$$\mathcal{I}^{\text{LSMq}}(z) := \mathcal{I}_+^{\text{LSMq}}(z) + \mathcal{I}_-^{\text{LSMq}}(z),$$

$$\mathcal{I}^{\text{F}\#,q}(z) := \mathcal{I}_+^{\text{F}\#,q}(z) + \mathcal{I}_-^{\text{F}\#,q}(z)$$

and

$$\mathcal{I}^{\text{GLSMq}}(z) := \mathcal{I}_+^{\text{GLSMq}}(z) + \mathcal{I}_-^{\text{GLSMq}}(z).$$

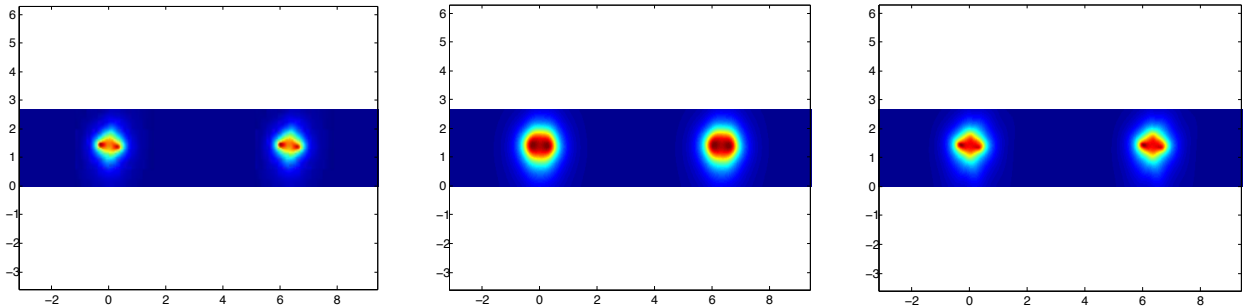


Figure 4.6: Reconstruction of the periodic domain  $D_p$ . From left to right:  $\mathcal{I}^{\text{LSMq}}(z)$ ,  $\mathcal{I}^{\text{F}\#,q}(z)$  and  $\mathcal{I}^{\text{GLSMq}}(z)$ .

As we can observe from all previous examples the three methods are capable of reconstructing the full background (in the the simple configuration that we have chosen) but with different

precisions. The LSM tends to under estimate the size of the objects while the Factorization method tends to over estimate the size (which would result in a less resolution for closely spaced objects). The GLSM method somehow provides a trade-off between the two methods.

## 4.5 On the Use of Differential measurements for reconstructing the inclusion support

Before introducing our new algorithm to directly reconstruct the inclusion we shall present in the following the approach that has been proposed in [13] based on so-called differential measurements. This approach assume that one has access to the measurement operator for the background media (i.e. here the periodic media without the presence of the defect  $\omega$ ) in addition to the measurement operator for the configuration with defect. As indicated in the introduction, our method will remove the need for measuring the operator for the background. However, the idea of our method is inspired from the method in [13] since we will construct an operator that mimics the measurement operator for the background. This is also one of the reasons we choose to present this method.

### 4.5.1 Theory of the Differential LSM

We here present the theory behind the application of the algorithm proposed in [13], called Differential LSM (DLSM), to our setting of the periodic layer with local perturbation. As indicated above, in the DLSM method, one tries to reconstruct the local perturbation of the periodic media without knowing the refractive indices  $n$  and  $n_p$  from the knowledge of differential measurements. We use here the same notation as above for the operators  $N^\pm$ ,  $\mathcal{H}^\pm$  and  $G^\pm$  associated with the index of refraction  $n$  and which obey the following factorizations.

$$N_{\#}^\pm = (\mathcal{H}^\pm)^* T_{\#}^\pm \mathcal{H}^\pm \quad \text{and} \quad N^\pm = G^\pm \mathcal{H}^\pm. \quad (4.63)$$

Let us then denote by  $N_p^\pm$ ,  $\mathcal{H}_p^\pm$  and  $G_p^\pm$  the respective operators associated with  $n = n_p$  (i.e. the operators associated with the background media), and for which the following factorization hold

$$N_{p,\#}^\pm = (\mathcal{H}_p^\pm)^* T_{p,\#}^\pm \mathcal{H}_p^\pm \quad \text{and} \quad N_p^\pm = G_p^\pm \mathcal{H}_p^\pm \quad (4.64)$$

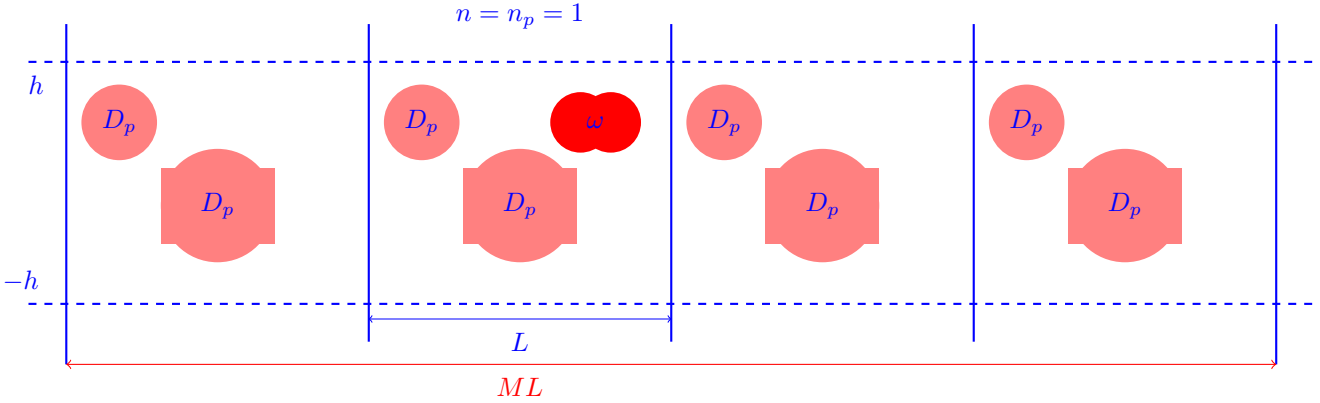
where  $T_{p,\#}^\pm$  is self-adjoint and coercive. We are interested in identifying the perturbation  $\omega := D \setminus D_p$  from the knowledge of  $N^\pm$  and  $N_p^\pm$  (without reconstructing  $n$  and  $n_p$  or  $D$  and  $D_p$ ). We shall present the method in the following simple geometric setting: we assume that the perturbation  $\omega$  is located in a period  $\Omega_0$  and  $D_p \cap \omega = \emptyset$ . We also recall that we assume that  $\Omega_M \setminus D$  is connected (see Figure 4.7)

Denoting by  $\text{ITP}(n, D, \Phi)$  the interior transmission problem (4.23) with  $\varphi = \Phi$  and  $\psi = \frac{\partial \Phi}{\partial \nu}$ , we here assume that  $\text{ITP}(n, D, \Phi)$  is well-posed (which implies that  $\text{ITP}(n_p, D_p, \Phi)$  is also well posed since  $\omega \cap D_p = \emptyset$ ). Let us introduce

$$J_\alpha(z; a) := \alpha(N_{\#}^\pm a, a)_{\ell^2(\mathbb{Z})} + \left\| N^\pm a - \widehat{\Phi}^\pm(\cdot; z) \right\|_{\ell^2(\mathbb{Z})}^2$$

and

$$J_{p,\alpha}(z; a) := \alpha(N_{p,\#}^\pm a, a)_{\ell^2(\mathbb{Z})} + \left\| N_p^\pm a - \widehat{\Phi}^\pm(\cdot; z) \right\|_{\ell^2(\mathbb{Z})}^2.$$


 Figure 4.7: Sketch of the geometry in the case where  $D_p \cap \omega = \emptyset$ 

Let  $a_\alpha^z$  and  $a_{p,\alpha}^z$  be in  $\ell^2(\mathbb{Z})$  such that

$$J_\alpha(z; a_\alpha^z) \leq \inf_{a \in \ell^2(\mathbb{Z})} J_\alpha(z; a) + c(\alpha)$$

and

$$J_{p,\alpha}(z; a_{p,\alpha}^z) \leq \inf_{a \in \ell^2(\mathbb{Z})} J_{p,\alpha}(z; a) + c(\alpha)$$

with  $0 < \frac{c(\alpha)}{\alpha} \rightarrow 0$  as  $\alpha \rightarrow 0$ . Let  $(u(\cdot; z), v(\cdot; z)) \in L^2(D) \times L^2(D)$  and  $(u_p(\cdot; z), v_p(\cdot; z)) \in L^2(D_p) \times L^2(D_p)$  be the solutions to  $\text{ITP}(n, D, \Phi(\cdot; z))$  and  $\text{ITP}(n_p, D_p, \Phi(\cdot; z))$  respectively. We observe that if  $z \in D_p$ , then  $\Phi(\cdot; z)$  satisfies  $\Delta \Phi(\cdot; z) + k^2 \Phi(\cdot; z) = 0$  in  $\omega$ . By the assumption of the well-posedness of  $\text{ITP}(n, D, \Phi(\cdot; z))$  we deduce that

$$u(\cdot; z) = \begin{cases} u_p(\cdot; z) & \text{in } D_p \\ 0 & \text{in } \omega \end{cases} \quad \text{and} \quad v(\cdot; z) = \begin{cases} v_p(\cdot; z) & \text{in } D_p \\ -\Phi(\cdot; z) & \text{in } \omega \end{cases}$$

This shows that the solutions of  $\text{ITP}(n, D, \Phi(\cdot; z))$  and  $\text{ITP}(n_p, D_p, \Phi(\cdot; z))$  coincide in  $D_p$  if  $z \in D_p$ . Applying Theorem 4.4.4 to  $N^\pm$  and  $N_p^\pm$  we have that if  $z$  is in  $D_p$  then

$$(N_{p,\sharp}^\pm(a_\alpha^z - a_{p,\alpha}^z), (a_\alpha^z - a_{p,\alpha}^z))_{\ell^2(\mathbb{Z})} \leq C \| \mathcal{H}_p^\pm a_\alpha^z - \mathcal{H}_p^\pm a_{p,\alpha}^z \|_{L^2(D_p)}^2 \rightarrow 0 \quad (4.65)$$

as  $\alpha \rightarrow 0$ . Let us introduce the indicator function

$$\mathcal{I}(a, a_p) := \frac{1}{(N_\sharp^\pm a, a)_{\ell^2(\mathbb{Z})} \left( 1 + \frac{(N_\sharp^\pm a, a)_{\ell^2(\mathbb{Z})}}{(N_{p,\sharp}^\pm(a - a_p), a - a_p)_{\ell^2(\mathbb{Z})}} \right)}. \quad (4.66)$$

**Theorem 4.5.1.** *Let  $z \in \Omega_M$  and assume that the hypothesis of Theorem 4.4.4 hold. Then  $z \in \omega$  if and only if  $\lim_{\alpha \rightarrow 0} \mathcal{I}(a_\alpha^z, a_{p,\alpha}^z) > 0$ .*

*Proof.* If  $z \notin D$  then from Theorem 4.4.4 applied to  $N^\pm$  we get that  $(N_\sharp^\pm a_\alpha^z, a_\alpha^z) \rightarrow +\infty$  as  $\alpha \rightarrow 0$  and therefore  $\lim_{\alpha \rightarrow 0} \mathcal{I}(a_\alpha^z, a_{p,\alpha}^z) = 0$ .

If  $z \in D_p$ . Applying Theorem 4.4.4 to the operator  $N^\pm$  we deduce that  $(N_\sharp^\pm a_\alpha^z, a_\alpha^z)$  is bounded and converges to  $(T_\sharp^\pm v(\cdot; z), v(\cdot; z))_{L^2(D)} > 0$ . Combining this fact with (4.65) implies  $\lim_{\alpha \rightarrow 0} \mathcal{I}(g_\alpha^z, g_{0,\alpha}^z) =$

0.

We now treat the case of  $z \in \omega$ . Theorem 4.4.4 implies that  $(N_{\#}^{\pm} a_{\alpha}^z, a_{\alpha}^z)$  is bounded. We also get from Theorem 4.4.4 applied to  $N_p^{\pm}$  that  $\|\mathcal{H}_p^{\pm} a_{p,\alpha}^z\|$  is unbounded as  $\alpha \rightarrow 0$  and from Theorem 4.4.4 applied to  $N^{\pm}$  that  $\|\mathcal{H}^{\pm} a_{\alpha}^z\|$  is bounded as  $\alpha \rightarrow 0$ , which implies  $\|\mathcal{H}_p^{\pm} a_{\alpha}^z\|$  is bounded (since  $\|\mathcal{H}_p^{\pm} a_{\alpha}^z\|_{L^2(D_p)} \leq \|\mathcal{H}^{\pm} a_{\alpha}^z\|_{L^2(D)}$ ). Combining with the coercivity of the operator  $T_{p,\#}^{\pm}$  we then have,

$$\begin{aligned} (N_{p,\#}^{\pm}(a_{\alpha}^z - a_{p,\alpha}^z), a_{\alpha}^z - a_{p,\alpha}^z)_{\ell^2(\mathbb{Z})} &\geq C \|\mathcal{H}_p^{\pm}(a_{\alpha}^z - a_{p,\alpha}^z)\|_{L^2(D_p)}^2 \\ &\geq C \|\mathcal{H}_p^{\pm} a_{p,\alpha}^z\|_{L^2(D_p)}^2 - C \|\mathcal{H}_p^{\pm} a\|_{L^2(D_p)}^2 \rightarrow \infty \end{aligned} \quad (4.67)$$

as  $\alpha \rightarrow 0$ . We then get  $\lim_{\alpha \rightarrow 0} \mathcal{I}(a_{\alpha}^z, a_{p,\alpha}^z) > 0$  which finishes the proof.  $\square$

We remark that in the case of noisy operators and according to the GLSM method for noisy operators one has to change the indicator function (4.66) as in (4.51). More precisely one has to replace the term  $(N_{\#}^{\pm} a, a)_{\ell^2(\mathbb{Z})}$  in the expression (4.66) with  $\left( (N_{\#}^{\pm, \delta} a, a) + \delta \|N_{\#}^{\pm, \delta}\| \|a\|^2 \right)$  where  $N_{\#}^{\pm, \delta}$  indicates the noisy version of  $N_{\#}^{\pm}$ .

### 4.5.2 A numerical example

We consider the same settings and parameters as in the numerical example of Section 4.4.5. Let  $a_{\alpha}^{z,\pm}$  and  $a_{p,\alpha}^{z,\pm}$  be the Tikhonov regularized solution of (4.49) with respect to the operators  $N^{\pm, \delta}$  and  $N_p^{\pm, \delta}$  respectively where the parameter  $\alpha$  is given by (4.50). We recall that the noisy operators  $N^{\pm, \delta}$  are such that

$$\|N^{\pm, \delta} - N^{\pm}\| \leq \delta \|N^{\pm, \delta}\| \quad \text{and} \quad \|N_p^{\pm, \delta} - N_p^{\pm}\| \leq \delta \|N_p^{\pm, \delta}\|, \quad (4.68)$$

where  $\delta$  is the noise level. We then define

$$\begin{aligned} \mathcal{A}^+(z) &= |(N_{\#}^{+, \delta} a_{\alpha}^{z,+}, a_{\alpha}^{z,+})| + \delta \|N_{\#}^{+, \delta}\| \|a_{\alpha}^{z,+}\|^2, \\ \mathcal{A}^-(z) &= |(N_{\#}^{-, \delta} a_{\alpha}^{z,-}, a_{\alpha}^{z,-})| + \delta \|N_{\#}^{-, \delta}\| \|a_{\alpha}^{z,-}\|^2 \\ \mathcal{D}^+(z) &= |(N_{p,\#}^{+, \delta} (a_{\alpha}^{z,+} - a_{p,\alpha}^{z,+}, a_{\alpha}^{z,+} - a_{p,\alpha}^{z,+})| \\ \mathcal{D}^-(z) &= |(N_{p,\#}^{-, \delta} (a_{\alpha}^{z,-} - a_{p,\alpha}^{z,-}, a_{\alpha}^{z,-} - a_{p,\alpha}^{z,-})| \end{aligned} \quad (4.69)$$

and the following indicator functions

$$\mathcal{I}_+^{\text{DLSM}}(z) = \frac{1}{\mathcal{A}^+(z)(1 + \frac{\mathcal{A}^+(z)}{\mathcal{D}^+(z)})} \quad \text{and} \quad \mathcal{I}_-^{\text{DLSM}}(z) = \frac{1}{\mathcal{A}^-(z)(1 + \frac{\mathcal{A}^-(z)}{\mathcal{D}^-(z)})} \quad (4.70)$$

We also test the reconstruction obtained using the indicator function

$$\mathcal{I}^{\text{DLSM}}(z) := \mathcal{I}_+^{\text{DLSM}}(z) + \mathcal{I}_-^{\text{DLSM}}(z).$$

The numerical results are summarized in Figure 4.8 where one clearly observe that the contribution of the background is discarded in the used indicator functions.

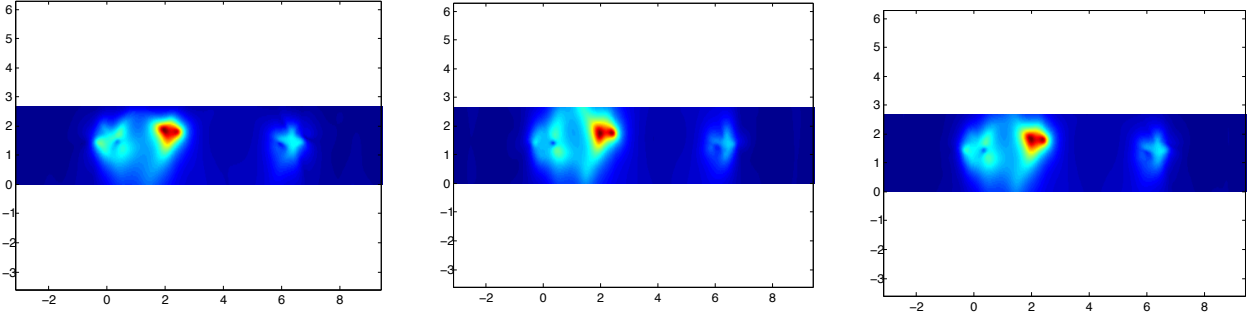


Figure 4.8: Reconstruction of the domain  $\omega$  using, from left to right,  $\mathcal{I}_+^{\text{DLSM}}(z)$ ,  $\mathcal{I}_-^{\text{DLSM}}(z)$  and  $\mathcal{I}^{\text{DLSM}}(z)$ .

## 4.6 Sampling methods for a single Floquet-Bloch mode

The first step toward the construction of our indicator function of  $\omega$  without the need for a measurement operator for the background is to construct a sampling operator that would roughly speaking plays the role of the background sampling operator. A natural candidate for this operator is the one obtained from the operator  $\mathcal{H}$  by restricting to single Floquet-Bloch modes. This choice comes from the fact that in the periodic case, i.e.  $n = n_p$ , this type of operators have been used to reconstruct the background medium (see [91]). As we can shall see later, for the case of perturbed periodic media, this operator reconstruct the background media and the perturbation with periodic copies of period  $L$ .

### 4.6.1 Near field operator for a fixed Floquet-Bloch mode

We shall present in this section how to define the sampling operator for one fixed Floquet-Bloch mode. Let  $a \in \ell^2(\mathbb{Z})$ , we define for  $q \in \mathbb{Z}_M$ , the element  $a_q \in \ell^2(\mathbb{Z})$  by

$$a_q(j) := a(q + jM).$$

We then define the operator  $\mathbf{I}_q : \ell^2(\mathbb{Z}) \rightarrow \ell^2(\mathbb{Z})$ , which transforms  $a \in \ell^2(\mathbb{Z})$  to  $\tilde{a} \in \ell^2(\mathbb{Z})$  such that

$$\tilde{a}_q = a \quad \text{and} \quad \tilde{a}_{q'} = 0 \quad \text{if} \quad q \neq q'.$$

We remark that  $\mathbf{I}_q^*(a) = a_q$ , where  $\mathbf{I}_q^* : \ell^2(\mathbb{Z}) \rightarrow \ell^2(\mathbb{Z})$  is the dual of the operator  $\mathbf{I}_q$ . The single Floquet-Bloch mode Herglotz operator  $\mathcal{H}_q^\pm : \ell^2(\mathbb{Z}) \rightarrow L^2(D)$  is defined by

$$\mathcal{H}_q^\pm a := \mathcal{H}^\pm \mathbf{I}_q a = \sum_j a(j) u^{i,\pm}(\cdot; q + jM)|_D \quad (4.71)$$

and the single Floquet-Bloch mode sampling operator  $\mathbf{N}_q^\pm : \ell^2(\mathbb{Z}) \rightarrow \ell^2(\mathbb{Z})$  is defined by

$$\mathbf{N}_q^\pm a = \mathbf{I}_q^* \mathbf{N}^\pm \mathbf{I}_q a. \quad (4.72)$$

We remark that  $\mathcal{H}_q^\pm a$  is an  $\alpha_q$ -quasi-periodic function with period  $L$ . The sequence  $\mathbf{N}_q^\pm a$  corresponds with the Fourier coefficients of the  $\alpha_q$ -quasi-periodic component of the scattered field in the decomposition (4.52). This operator is then somehow associated with  $\alpha_q$ -quasi-periodicity.

One immediately see from the factorization  $N^\pm = (\mathcal{H}^\pm)^* \mathbb{T} \mathcal{H}^\pm$  that the following factorization holds.

$$N_q^\pm = (\mathcal{H}_q^\pm)^* \mathbb{T} \mathcal{H}_q^\pm. \quad (4.73)$$

For later use we also define the operator  $G_q^\pm : \overline{\mathcal{R}(\mathcal{H}_q^\pm)} \rightarrow \ell^2(\mathbb{Z})$  by

$$G_q^\pm = (\mathcal{H}_q^\pm)^* \mathbb{T}|_{\overline{\mathcal{R}(\mathcal{H}_q^\pm)}} \quad (4.74)$$

where the operator  $\mathbb{T}$  is defined by (4.20).

#### 4.6.2 Some properties of the operators $\mathcal{H}_q^\pm$ , $N_q^\pm$ and $G_q^\pm$

We prove in this sections the needed properties in order to establish the theoretical justifications of sampling methods. This Section is somehow the equivalent of Section 4.3.2 but for the case single Floquet-Bloch mode operator. We recall that we make the assumption that

$$\Omega_M \setminus D \text{ is connected and } \partial\Omega_m \cap \overline{D} = \emptyset \text{ for all } m \in \mathbb{Z}_M.$$

For the support of the perturbation  $\omega$  we assume that

$$\omega \subset \Omega_{m_0}$$

for some  $m_0 \in \mathbb{Z}_M$  and

$$D_p \cap \omega = \emptyset.$$

The latter assumption is important in many aspects of the following proofs and indeed a main perspective of the current work is extend our results to the cases  $D_p \cap \omega \neq \emptyset$ . The first assumption on  $\omega$  is not restrictive on the size of the perturbation (since one can increase  $L$ ) but implies for the inverse problem that one has a priori knowledge on the size of  $\omega$ .

We define for later use the domain

$$\omega_p := \bigcup_{m-m_0 \in \mathbb{Z}_M} \omega + mL \quad (4.75)$$

which is the union of  $L$  periodic copies of  $\omega$ . In all the following of this section

$$q \in \mathbb{Z}_M$$

is a fixed parameter.

**Lemma 4.6.1.** *The operator  $\mathcal{H}_q^\pm$  defined by (4.71) is injective and*

$$\overline{\mathcal{R}(\mathcal{H}_q^\pm)} = H_{\text{inc}}^q(D) := \{v \in L^2(D), \quad \Delta v + k^2 v = 0 \text{ in } D \text{ and } v|_{D_p} \text{ is } \alpha_q\text{-quasi-periodic}\} \quad (4.76)$$

*Proof.*  $\mathcal{H}_q^\pm$  is injective since  $\mathcal{H}^\pm$  is injective and  $\mathbb{I}_q$  is injective. We now prove that  $(\mathcal{H}_q^\pm)^*$  is injective on  $H_{\text{inc}}^q(D)$ . Let  $\varphi \in H_{\text{inc}}^q(D)$  and assume  $(\mathcal{H}_q^\pm)^*(\varphi) = 0$ . We define

$$u(x) := \frac{1}{M} \int_D \Phi_q(x-y) \varphi(y) \, dy.$$

From the expansion of  $\Phi_q(x)$  as in (4.56) and using the same calculations as for (4.27) we have that  $\widehat{u}^\pm(j) = 0$  for all  $j \neq q + M\ell$  and  $\widehat{u}^\pm(q + M\ell) = ((\mathcal{H}^\pm)^*(\varphi))(q + M\ell) = ((\mathcal{H}_q^\pm)^*(\varphi))(\ell) = 0$ . Therefore  $u$  has all Rayleigh coefficients equal 0, which implies that

$$u = 0, \quad \text{for } \pm x_d > h.$$



We now observe that for all  $y \in D$ ,  $\Delta\Phi_q(\cdot; y) + k^2\Phi_q(\cdot; y) = 0$  in the complement of  $D_p \cup \omega_p$ . This implies that

$$\Delta u + k^2 u = 0 \quad \text{in } \mathbb{R}^d \setminus \{D_p \cup \omega_p\}.$$

Using a unique continuation argument we infer that  $u = 0$  in  $\Omega \setminus D_p \cup \omega_p$ . Therefore,  $u \in H_0^2(D_p \cup \omega_p)$  by the regularity of volume potentials.

Since  $\varphi|_{D_p}$  and  $\Phi_q$  are  $\alpha_q$ -quasi-periodic functions with period  $L$ , then for  $m \in \mathbb{Z}_M$ ,

$$u(x) := \frac{1}{M} \int_{\omega} \Phi_q(x; y) \varphi(y) \, dy + \int_{D_p \cap \Omega_m} \Phi_q(x; y) \varphi(y) \, dy \quad x \in D_p \cap \Omega_m.$$

We recall that  $\Delta\Phi_q(\cdot; y) + k^2\Phi_q(\cdot; y) = -\delta_y$  in  $\Omega_m$ . Using in addition the fact that  $D_p \cap \omega = \emptyset$ , we obtain for  $m \in \mathbb{Z}_M$ ,

$$\Delta u(x) + k^2 u(x) = -\varphi(x) \quad \text{in } D_p \cap \Omega_m. \quad (4.77)$$

Let us set for  $m - m_0 \in \mathbb{Z}_M$

$$\varphi_m(x) := e^{i\alpha_q m L} \varphi(x - mL) \quad \text{for } x \in \omega + mL.$$

Then we have, using the  $\alpha_q$ -quasi-periodicity of  $\Phi_q$

$$u(x) := \frac{1}{M} \int_{\omega + mL} \Phi_q(x; y) \varphi_m(y) \, dy + \frac{1}{M} \int_D \Phi_q(x; y) \varphi(y) \, dy \quad x \in \omega + mL$$

with  $m - m_0 \in \mathbb{Z}_M$ . Using again the fact that  $D_p \cap \omega_p = \emptyset$  and  $\Delta\Phi_q(\cdot; y) + k^2\Phi_q(\cdot; y) = -\delta_y$  in  $\Omega_m$  we get

$$\Delta u(x) + k^2 u(x) = -\varphi_m \quad \text{in } \omega + mL. \quad (4.78)$$

Now define the function  $\tilde{\varphi}$  by

$$\tilde{\varphi} = \varphi \quad \text{in } D_p \quad \text{and} \quad \tilde{\varphi} = \varphi_m \quad \text{in } \omega + mL.$$

Clearly

$$\Delta \tilde{\varphi} + k^2 \tilde{\varphi} = 0 \quad \text{in } D_p \cup \omega_p \quad \text{with } m - m_0 \in \mathbb{Z}_M.$$

Since  $u \in H_0^2(D_p \cup \omega_p)$  we then have

$$\int_D (\Delta u + k^2 u) \bar{\tilde{\varphi}} = 0.$$

This implies according to (4.77) and (4.78) that

$$\int_{D_p} |\varphi|^2 \, dx + M \int_{\omega} |\varphi|^2 = 0,$$

which implies  $\varphi = 0$  in  $D_p \cup \omega$ . This proves the injectivity of  $(\mathcal{H}^\pm)^*$  on  $H_{\text{inc}}^q(D)$  and hence proves the Lemma.  $\square$

The following two Lemmas are probably the most important results. They form two of the important cornerstones for the justification of the GLSM and the differential imaging functional that we shall define later.

**Lemma 4.6.2.** *Assume that Assumptions 4.2.1 and 4.3.2 hold. Assume in addition that Assumption 4.2.1 hold with  $n = n_p$ . Then the operator  $G_q^\pm : H_{\text{inc}}^q(D) \rightarrow \ell^2(\mathbb{Z})$  is injective with dense range.*

*Proof.* We here give the proof of the lemma for the operator  $G_q^+$  since the proof for the operator  $G_p^-$  is similar.

Assume that  $G_q^+(f) = 0$  for  $f \in H_{\text{inc}}^q(D)$ . Let  $w \in H_{\#}^1(\Omega_M^h)$  be the associated scattered field via (4.10) and consider the decomposition  $w$  into

$$w = \frac{1}{M} \sum_{q' \in \mathbb{Z}_M} w_{q'}, \quad (4.79)$$

where  $w_{q'}$  is  $\alpha_{q'}$ -quasi-periodic with period  $L$ . We recall that

$$\Delta w + k^2 n_p w + k^2(n - n_p)w = k^2(n_p - n)f + k^2(1 - n_p)f. \quad (4.80)$$

Since  $(1 - n_p)f$  is  $\alpha_q$ -quasi-periodic with period  $L$  and  $\omega \subset \Omega_{m_0}$  we obtain that (projecting the latter equation on the  $L^2(\Omega_{m_0})$  Fourier basis which is  $\alpha_q$ -quasi-periodic with period  $L$ )

$$\Delta w_q + k^2 n_p w_q + k^2(n - n_p)w = k^2(1 - n)f \quad \text{in } \Omega_{m_0}. \quad (4.81)$$

In particular we have that  $\Delta w_q + k^2 w_q = 0$  in  $\Omega_M \setminus \{D_p \cup \omega_p\}$ . Using a similar unique continuation argument as the one at the beginning of the proof of Lemma 4.6.1 we deduce that

$$w_q = 0 \quad \text{in } \Omega_M \setminus \{D_p \cup \omega_p\}.$$

This implies in particular (since  $D_p \cap \omega = \emptyset$ )  $w_q \in H_0^2(D_p)$  and

$$\begin{cases} \Delta w_q + k^2 n_p w_q = k^2(1 - n_p)f & \text{in } D_p \\ \Delta f + k^2 f = 0 & \text{in } D_p \end{cases}$$

Assumption 4.3.2 implies in particular that  $\text{ITP}(n_p, D_p, 0)$  is well posed. We then obtain that  $f = 0$  in  $D_p$  and  $w_q = 0$  in  $D_p$ . On the other hand, using the fact that  $n_p = 1$  in  $\omega$  we also have  $w_q \in H_0^2(\omega)$  and

$$\{ \Delta w_q + k^2 w_q + k^2(n - 1)w = k^2(1 - n)f \quad \text{in } \omega. \quad (4.82)$$

The fact that  $w_q \in H_0^2(\omega)$  gives for instance that

$$\int_{\omega} (\Delta w_p + k^2 w_p)\theta = 0, \quad (4.83)$$

for all  $\theta \in H_{\text{inc}}(\omega) := \{\theta \in L^2(\omega), \Delta \theta + k^2 \theta = 0\}$ . Now, taking the  $L^2$  scalar product of (4.82) with  $\theta$  we arrive at

$$\int_{\omega} (k^2(1 - n)f + k^2(1 - n)w)\bar{\theta} = 0 \quad (4.84)$$

for all  $\theta \in H_{\text{inc}}(\omega)$ . Recall that the solution  $w$  of (4.10) can be represented as

$$w(x) = \int_{D_p \cup \omega} k^2(n - 1)(w + f)(y)\Phi(x; y) dy \quad x \in \Omega_M. \quad (4.85)$$

Since  $y \mapsto \Phi(x; y) \in H_{\text{inc}}(\omega)$  for  $x \notin \omega$  and  $f = 0$  in  $D_p$ , we then obtain from (4.84) that

$$w(x) = \int_{D_p} k^2(1 - n_p)w(y)\Phi(x; y) dy \quad \text{for } x \notin \omega. \quad (4.86)$$

Let us define  $\tilde{w} \in H^2(\omega)$  by

$$\tilde{w}(x) := \int_{D_p} k^2(1 - n_p)w(y)\Phi(x; y) dy \quad x \in \omega.$$

Then obviously  $u = w + f$  and  $f$  satisfy

$$\begin{cases} \Delta u + k^2 n u = 0 & \text{in } \omega \\ \Delta f + k^2 f = 0 & \text{in } \omega \\ u - f = \tilde{w} & \text{on } \partial\omega \\ \frac{\partial(u - f)}{\partial\nu} = \frac{\partial\tilde{w}}{\partial\nu} & \text{on } \partial\omega \end{cases} \quad (4.87)$$

Since

$$\Delta\tilde{w} + k^2\tilde{w} = 0 \text{ in } \omega$$

and  $\text{ITP}(n, \omega, \tilde{w})$  is well posed (by Assumption 4.3.2 and the fact that  $D_p \cap \omega_p = \emptyset$ ) we get that

$$f = -\tilde{w} \quad \text{and} \quad u = 0 \quad \text{in } \omega.$$

We then conclude that  $w + f = 0$  in  $\omega$ , i.e.  $w = \tilde{w}$  in  $\omega$ . This implies for instance that

$$w(x) = \int_{D_p} k^2(1 - n_p)w(y)\Phi(x; y) dy \quad \text{for } x \notin \Omega_M \quad (4.88)$$

which means in particular that  $w$  is solution to (4.10) with  $n = n_p$  and  $f = 0$ . We then infer that  $w = 0$  by uniqueness of the solution to the scattering problem for  $n = n_p$  (third assumption of our Lemma). This proves that  $f = 0$  in  $\omega$  and finishes the arguments for the injectivity of the operator  $G_q^+$ .

Using the injectivity we now prove the denseness of the range of  $G_q^\pm$ . The proof can be done as in Theorem 4.3.5. Let  $g \in \overline{\mathcal{R}(G_q^+)}^\perp$ . Then

$$(G_q^+(f), g)_{\ell^2(\mathbb{Z})} = 0, \quad \forall f \in H_{\text{inc}}^q(D).$$

Consider  $f$  of the form  $f = \overline{\mathcal{H}_q^+(a)}$  for some  $a \in \ell^2(\mathbb{Z})$ . Since  $G_q^+ = (\mathcal{H}_q^+)^*T$ , we then have

$$\langle T(\overline{\mathcal{H}_q^+(a)}), \mathcal{H}_q^+(g) \rangle_{L^2(D)} = 0, \quad \forall a \in \ell^2(\mathbb{Z}). \quad (4.89)$$

Let  $w(a)$  and  $w(g)$  solution to (4.10) with respect to  $\overline{\mathcal{H}_q^+(a)}$  and  $\overline{\mathcal{H}_q^+(g)}$ . From Lemma 4.3.4 we get

$$\begin{aligned} \left( T(\overline{\mathcal{H}_q^+(a)}), \mathcal{H}_q^+(g) \right)_{L^2(D)} &= k^2 \int_D (n-1) \overline{(\mathcal{H}_q^+ a + w(a))} \overline{\mathcal{H}_q^+ g} dx \\ &= k^2 \int_D (n-1) \overline{(\mathcal{H}_q^+ g + w(g))} \overline{\mathcal{H}_q^+ g} dx. \end{aligned}$$

Therefore,

$$\left( T(\overline{\mathcal{H}_q^+(g)}), \mathcal{H}_q^+(a) \right)_{L^2(D)} = \left( T(\overline{\mathcal{H}_q^+(a)}), \mathcal{H}_q^+(g) \right)_{L^2(D)}, \quad \forall a \in \ell^2(\mathbb{Z}).$$

We deduce from (4.89) that

$$\left( G_q^+(\overline{\mathcal{H}_q^+(g)}), a \right)_{\ell^2(\mathbb{Z})} = 0, \quad \forall a \in \ell^2(\mathbb{Z}),$$

which implies  $G_q^+(\overline{\mathcal{H}_q^+(g)}) = 0$ . The injectivity of  $G_q^+$  ensure that  $\mathcal{H}_q^+g = 0$  in  $D$  and consequently  $g = 0$  (using Lemma 4.6.1). This proves the denseness of the range of  $G_q^+$ .  $\square$

Before continuing with the characterization of  $D_p \cup \omega_p$  in terms of the range of  $G_q^\pm$  let us make the following simple observations. Let  $f \in H_{\text{inc}}(D)$  and let  $w$  be the solution of problem (4.10) associated with  $f$  and consider the decomposition

$$w = \frac{1}{M} \sum_{q \in \mathbb{Z}_M} w_q,$$

as in Lemma 4.6.2. Let us denote by  $a^\pm(w) \in \ell^2(\mathbb{Z})$  be the Rayleigh sequences of  $w$ . Then for fixed  $q$

$$\begin{cases} \frac{1}{M} a^\pm(w_q)(q + jM) = a^\pm(w)(q + jM), & \forall j \in \mathbb{Z} \\ \frac{1}{M} a^\pm(w_q)(\ell) = 0, & \forall \ell \neq q + jM. \end{cases}$$

Therefore,

$$I_q^*(a^\pm(w)) = \frac{1}{M} I_q^*(a^\pm(w_q))$$

(and in particular  $I_q^* \widehat{\Phi}_q^\pm(\cdot; z) = \frac{1}{M} I_q^* \widehat{\Phi}_q^\pm(\cdot; z)$ ).

**Lemma 4.6.3.** *Under the same assumptions as in Lemma 4.6.2 we have that  $I_q^* \widehat{\Phi}_q^\pm(\cdot; z) \in \mathcal{R}(G_q^\pm)$  if and only if  $z \in D_p \cup \omega_p$ .*

*Proof.* We recall that  $\widehat{\Phi}_q^\pm(\cdot; z)$  is the Rayleigh sequence of  $\Phi_q(\cdot; z)$ .

Let  $z \in D_p$ . We the consider  $(u_q, v_q)$  solution of ITP( $n, D, \Phi_q(\cdot; z)$ ). Let us define

$$w = \begin{cases} u_q - v_q & \text{in } D_p. \\ \Phi_q & \text{in } \Omega_M \setminus D_p. \end{cases}$$

Then  $w \in H_{\#, \text{loc}}^2(\Omega_M)$  and satisfies the scattering problem (4.10) with  $f = v_q$  (remark that  $v_q = -\Phi_q$  in  $\omega$ ). Therefore  $G^\pm(v_q) = \widehat{\Phi}_q^\pm(\cdot; z)$ . Applying  $I_q^*$  to the two sides of the equality we get  $G_q^\pm(v_q) = I_q^* \widehat{\Phi}_q^\pm(\cdot; z)$ .

Now let  $z \in \omega$  and consider  $v \in H_{\text{inc}}(D)$  is such that  $G(v) = \widehat{\Phi}^\pm(\cdot; z)$  as in Lemma. By construction

$$I_q^* G(Mv) = M I_q^* \widehat{\Phi}^\pm(\cdot; z) = I_q^* \widehat{\Phi}_q^\pm(\cdot; z) \quad (4.90)$$

However the function  $v \notin H_{\text{inc}}^q(D)$  since the function  $v$  is constructed such that  $(u, v) \in L^2(D) \times L^2(D)$  is a solution of ITP( $n, D, \Phi(\cdot; z)$ ) and therefore  $v = -\Phi(\cdot; z)$  in  $D_p$  which is not an  $\alpha_q$ -quasi-periodic function with period  $L$ . We describe in the following how one can construct from  $v$  a function  $\tilde{v} \in H_{\text{inc}}^q(D)$  such that  $G_q(v) = G_q(\tilde{v})$ . We first decompose  $v|_{D_p}$  into

$$v = \frac{1}{M} \sum_{q' \in \mathbb{Z}_M} v_{q'}$$

where  $v_{q'}$  is  $\alpha_{q'}$ -quasi-periodic function with period  $L$  as in 4.52. Each of the  $v_{q'}$  is also a solution to  $\Delta v_{q'} + k^2 v_{q'} = 0$  in  $D_p$ . Remark that for our case  $v_{q'} = \Phi_{q'}(\cdot, z)$ . We set  $\tilde{v} = v - \frac{1}{M} v_q$ . Now consider the solution  $\tilde{w} \in H_{\#, \text{loc}}^2(\Omega_M)$  to

$$\Delta \tilde{w} + k^2 n_p \tilde{w} = k^2(1 - n_p) \tilde{v} \quad \text{in } \Omega_M$$

with the Rayleigh radiation condition. Using the  $L^2(\Omega_M)$  orthogonality of  $(1 - n_p) \tilde{v}$  with  $\alpha_q$ -quasi-periodic functions and uniqueness the solution of the scattering problem for we simply get that

$$\tilde{w}_q = 0.$$

If we denote by  $\tilde{a}^\pm$  the Rayleigh sequences associated with  $\tilde{w}$ , then the latter ensures that  $\mathbb{I}_q^*(\tilde{a}^\pm) = 0$ . Now observe that we can also write the equation for  $\tilde{w}$  as

$$\Delta \tilde{w} + k^2 n \tilde{w} = k^2(1 - n_p) \tilde{v} + k^2(n_p - n)(-\tilde{w}) \quad \text{in } \Omega_M$$

and since  $\omega \cap D_p = \emptyset$ , we have  $\Delta \tilde{w} + k^2 \tilde{w} = 0$  in  $\omega$ . Therefore the function  $\hat{v}$  defined by  $\hat{v} = \tilde{v}$  in  $D_p$  and  $\hat{v} = -\tilde{w}$  in  $\omega$  satisfies  $\hat{v} \in H_{\text{inc}}(D)$  and

$$G^\pm(\hat{v}) = \tilde{a}^\pm.$$

Therefore  $\mathbb{I}_q^* G^\pm(\hat{v}) = 0$ ,  $(v - \hat{v}) \in H_{\text{inc}}^q(D)$  and

$$\mathbb{I}_q^* G^\pm(v - \hat{v}) = \mathbb{I}_q^* G^\pm(v).$$

(Remark that if  $\hat{v} \neq 0$  then  $G^\pm(v - \hat{v}) \neq G^\pm(v)$ ). Applying this procedure to the  $v$  in (4.90) we obtain that,  $M(v - \hat{v}) \in H_{\text{inc}}^q(D)$  and

$$G_q(M(v - \hat{v})) = \mathbb{I}_q^* \widehat{\Phi}_q^\pm(\cdot; z).$$

Now consider the case where  $z \in \omega + mL$  with  $m - m_0 \in \mathbb{Z}_M$ . We recall that  $\widehat{\Phi}_q^\pm(\cdot; z) = e^{imL \cdot \alpha_q} \widehat{\Phi}_q^\pm(\cdot; z - mL)$ . Therefore, if we consider  $v \in H_{\text{inc}}^q(D)$  such that  $G_q(v) = \mathbb{I}_q^* \widehat{\Phi}_q^\pm(\cdot; z - mL)$ , which is possible by the previous step since  $z - mL \in \omega$ , then

$$G_q(e^{imL \cdot \alpha_q} v) = \mathbb{I}_q^* (\widehat{\Phi}_q^\pm(\cdot; z)).$$

We finally consider the case  $z \notin D_p \cup \omega_p$ . If  $G_q(v) = \mathbb{I}_q^* \widehat{\Phi}_q^\pm(\cdot; z)$ , then using the same unique continuation argument as in the proof of Lemma 4.6.2 we get that  $w_q = \Phi_q$  in  $\Omega_M \setminus \{D_p \cup \omega_p\}$  where  $w$  is the solution of (4.10) with  $f = v$  and  $w_q$  is defined by (4.79). This gives a contradiction since  $w_q \in H_{\text{loc}}^2(\Omega_M \setminus \{D_p \cup \omega_p\})$  while  $\Phi_q(\cdot; z) \notin H_{\text{loc}}^2(\Omega_M \setminus \{D_p \cup \omega_p\})$ .  $\square$

The following Lemma allows us to infer some reconstruction results for the domain  $D_p \cup \omega_p$  using the Factorization method but is not important for the differential imaging functional we shall introduce later.

**Lemma 4.6.4.**  $\mathbb{I}_q^* \widehat{\Phi}_q^\pm(\cdot; z) \in \mathcal{R}((\mathcal{H}_q^\pm)^*)$  if and only if  $z \in D_p \cup \omega_p$ .

*Proof.* We first consider  $z \in D_p \cup \omega_p$ . We define  $\rho$  a  $C^\infty$  cut-off function such that

$$\begin{aligned} \rho &= 1 && \text{in } \Omega \setminus D_p \cup \omega_p \\ \rho &= 0 && \text{in a neighborhood of } \{z + mL, m \in \mathbb{Z}_M\} \end{aligned} \tag{4.91}$$

and  $\rho$  is periodic with period  $L$ . We set  $v := \rho \Phi_q(\cdot; z)$ . Obviously  $v$  is  $\alpha_q$ -quasi-periodic with period  $L$  and

$$\widehat{v}^\pm(j) = \widehat{\Phi}_q^\pm(j; z).$$

Moreover, since  $f := -(\Delta v + k^2 v)$  is  $\alpha_q$ -quasi-periodic and with support in  $D_p \cup \omega_p$  then

$$v = \frac{1}{M} \int_{D_p \cup \omega_p} f(y) \Phi_q(x - y) dy$$

Obviously,

$$\widehat{v}^\pm(q + Mj) = \frac{1}{M} \int_{D_p \cup \omega_p} f(y) \overline{u^{i,\pm}}(y; q + Mj) dy.$$

We see that

$$\int_{\omega_p} f(y) \overline{u^{i,\pm}}(y; q + Mj) dy = M \int_{\omega} f(y) \overline{u^{i,\pm}}(y; q + Mj) dy$$

since  $f$  and  $u^{i,\pm}(\cdot; q + Mj)$  are  $\alpha_q$ -quasi-periodic functions with period  $L$ . Therefore,

$$\widehat{v}^\pm(q + Mj) = \frac{1}{M} \int_{D_p} f(y) \overline{u^{i,\pm}}(y; q + Mj) dy + \int_{\omega} f(y) \overline{u^{i,\pm}}(y; q + Mj) dy.$$

We then obtain  $\mathbb{I}_q^* \widehat{v}^\pm = \mathbb{I}_q^* (\mathcal{H}^\pm)^* \varphi$  with

$$\varphi := \begin{cases} \frac{1}{M} f & \text{in } D_p \\ f & \text{in } \omega. \end{cases}$$

Therefore

$$(\mathcal{H}_q^\pm)^* \varphi = \mathbb{I}_q^* \widehat{\Phi}_q^\pm(\cdot; z)$$

If  $z \notin D_p \cup \omega_p$ , using the same unique continuation argument as in the proof of Lemma 4.6.1 we see that if  $(\mathcal{H}_q^\pm)^* \varphi = \mathbb{I}_q^* \widehat{\Phi}_q^\pm(\cdot; z)$  then the function defined by

$$w(x) = \int_D \Phi(x - y) \varphi(y) dy$$

verifies  $w_q = \Phi_q(\cdot; z)$  in  $\Omega \setminus D_p \cup \omega_p \cup \{z\}$  where  $w_q$  is defined as in (4.79). This gives a contradiction since  $w_q$  is smooth in a neighborhood of  $z$  but  $\Phi_q(\cdot; z)$  is singular at  $z$ . We then get the contradiction if  $z \notin D_p \cup \omega_p$ .  $\square$

We now end this section with direct applications of the sampling methods that have been introduced in Section 4.4 to the case where one uses the operator  $\mathbb{N}_q^\pm$ . The previous technical results allow us to phrase theorems related to reconstructing the domain  $D_p \cup \omega_p$ . The first one is related to the factorization method that employs the operator  $\mathbb{N}_{q,\#}^\pm$  that we shall define here as

$$\mathbb{N}_{q,\#}^\pm := \mathbb{I}_q^* \mathbb{N}_\#^\pm \mathbb{I}_q$$

The proof of the following theorem is based on the same arguments as the proof of Theorem 4.4.2 and is a consequence of the factorization 4.73, the properties of the operator  $T$  formulated in Lemma 4.3.6 and the technical Lemma 4.6.4.

**Theorem 4.6.5.** *Under the same assumptions of Lemma 4.3.6 we have that*

$$\mathbb{N}_{q,\#}^\pm = (\mathcal{H}_q^\pm)^* \mathbb{T}_\#^\pm \mathcal{H}_q^\pm, \quad (4.92)$$

where  $\mathbb{T}_\#^\pm : L^2(D) \rightarrow L^2(D)$  is the same as in Theorem 4.4.2. In particular it is self-adjoint and coercive on  $H_{\text{inc}}^q(D)$ . Moreover,  $\mathbb{I}_q^* \widehat{\Phi}_q^\pm(\cdot; z) \in \mathcal{R}((\mathbb{N}_{q,\#}^\pm)^{1/2})$  if and only if  $z \in D_p \cup \omega_p$ .

*Proof.* The proof of the factorization (4.92) follows directly from (4.41). The second part of the theorem follows from the technical Lemma 4.6.4 and Theorem A.1.2.  $\square$

We remark that we could also have defined  $N_{q,\sharp}^\pm$  as in the definition of  $N_\sharp^\pm$  and derived the same results. However, proving in this way that the middle operator  $T_\sharp^\pm$  is the same as for  $N_\sharp^\pm$  would have been less straightforward.

The second useful corollary is the application of the GLSM algorithm. Indeed the following theorem can be proved in a similar way as Theorem 4.4.4 using Theorem 4.6.5, and the technical Lemmas 4.6.1, 4.6.2 and 4.6.3.

**Theorem 4.6.6.** *Assume that Assumptions of Lemmas 4.3.6 and 4.6.2 hold. Then the results of 4.4.4 are 4.4.5 still true if  $D$  is replaced by  $D_p \cup \omega_p$ , the operators  $\mathcal{H}^\pm$ ,  $G^\pm$ ,  $N^\pm$  and  $N_\sharp^\pm$  are replaced by the operators  $\mathcal{H}_q^\pm$ ,  $G_q^\pm$ ,  $N_q^\pm$  and  $N_{q,\sharp}^\pm$  respectively and  $\widehat{\Phi}^\pm(\cdot; z)$  is replaced by  $I_q^* \widehat{\Phi}_q^\pm(\cdot; z)$ .*

### 4.6.3 A new differential imaging functional

We now have all the ingredients to introduce the new differential imaging functional for retrieving the domain  $\omega$  using the measurement operator  $N^+$  or  $N^-$ . Compared to the algorithm presented in Section 4.5, there is no need here for a measure of the background operator. The latter will be replaced by using the operator  $N_q^+$  (or respectively  $N_q^-$ ) for a fixed  $q \in \mathbb{Z}_M$ . In addition to the assumptions on the geometry that was made in the previous section we need to include those of Section 4.5 by assuming that for all  $z \in \omega$  there exists  $m \in \mathbb{Z}^{d-1}$  such that  $z + mL \in \Omega_M \setminus D$ . This means in particular that  $M > 1$  and therefore there is a least one period that does not contain a defect.

We hereafter assume that the hypothesis of Theorems 4.6.5 and 4.6.6 are verified. In order to simplify the notations we only present the results for the operator  $N^+$ . We obtain exactly the same results by changing the exponent  $+$  with the exponent  $-$ . We then consider for  $\phi$  and  $a$  in  $\ell^2(\mathbb{Z})$  the cost functionals

$$\begin{aligned} J_\alpha^+(\phi, a) &:= \alpha(N_\sharp^+ a, a) + \|N^+ a - \phi\|^2, \\ J_{\alpha,q}^+(\phi, a) &:= \alpha(N_{q,\sharp}^+ a, a) + \|N_q^+ a - \phi\|^2. \end{aligned} \quad (4.93)$$

We remark that we also have

$$J_{\alpha,q}^+(\phi, a) = \alpha(N_\sharp^+ I_q a, I_q a) + \|N_q^+ a - \phi\|^2.$$

Let  $a^{\alpha,z}$ ,  $a_q^{\alpha,z}$  and  $\tilde{a}_q^{\alpha,z}$  be in  $\ell(\mathbb{Z})$  verifying

$$\begin{aligned} J_\alpha^+(\widehat{\Phi}^\pm(\cdot; z), a^{\alpha,z}) &\leq \inf_{a \in \ell^2(\mathbb{Z})} J_\alpha^+(\widehat{\Phi}^\pm(\cdot; z), a) + c(\alpha) \\ J_\alpha^+(\widehat{\Phi}_q^\pm(\cdot; z), a_q^{\alpha,z}) &\leq \inf_{a \in \ell^2(\mathbb{Z})} J_\alpha^+(\widehat{\Phi}_q^\pm(\cdot; z), a) + c(\alpha) \\ J_{\alpha,q}^+(I_q^* \widehat{\Phi}_q^\pm(\cdot; z), \tilde{a}_q^{\alpha,z}) &\leq \inf_{a \in \ell^2(\mathbb{Z})} J_{\alpha,q}^+(I_q^* \widehat{\Phi}_q^\pm(\cdot; z), a) + c(\alpha) \end{aligned} \quad (4.94)$$

with  $\frac{c(\alpha)}{\alpha} \rightarrow 0$  as  $\alpha \rightarrow 0$ . We then consider the following imaging functional to characterize  $\omega$ :

$$\mathcal{I}_\alpha^+(z) = \left( (N_\sharp^+ a^{\alpha,z}, a^{\alpha,z}) \left( 1 + \frac{(N_\sharp^+ a^{\alpha,z}, a^{\alpha,z})}{D^+(a_q^{\alpha,z}, \tilde{a}_q^{\alpha,z})} \right) \right)^{-1} \quad (4.95)$$

where for  $a$  and  $b$  in  $\ell^2(\mathbb{Z})$ ,

$$D^+(a, b) := \left( N_\sharp^+(a - I_q b), (a - I_q b) \right).$$

**Theorem 4.6.7.** *Under the assumptions of Theorem 4.6.6, we have that*

$$z \in \omega \text{ if and only if } \lim_{\alpha \rightarrow 0} \mathcal{I}_\alpha^+(z) > 0.$$

*Proof.* If  $z \notin D$  then from Theorem 4.4.4 applied to  $\mathbb{N}^+$  we get that  $(\mathbb{N}_\#^+ a^{\alpha,z}, a^{\alpha,z}) \rightarrow +\infty$  as  $\alpha \rightarrow 0$  and therefore  $\lim_{\alpha \rightarrow 0} \mathcal{I}_\alpha^+(z) = 0$ .

If  $z \in D_p$ , we remark that according to Theorems 4.4.8 and 4.6.6, when  $\alpha \rightarrow 0$ , the sequences  $\mathcal{H}^+ a_q^{\alpha,z}$  and  $\mathcal{H}_q^+ \tilde{a}_q^{\alpha,z}$  respectively converge in  $L^2(D)$  to  $v \in H_{\text{inc}}(D)$  and  $\tilde{v} \in H_{\text{inc}}^q(D)$  such that

$$G(v) = \widehat{\Phi}_q^+(\cdot; z) \text{ and } G_q(\tilde{v}) = \mathbb{I}_q^* \widehat{\Phi}_q^+(\cdot; z).$$

According to the proof of Theorem 4.3.5 and Lemma 4.6.3, the functions  $v$  and  $\tilde{v}$  are solutions to ITP( $n, D, \Phi_q(\cdot; z)$ ) and therefore are the same. From the factorization of  $\mathbb{N}_\#^+$  and the definition of  $\mathcal{H}_q^+$  we get

$$D^+(a_q^{\alpha,z}, \tilde{a}_q^{\alpha,z}) = (\mathbb{T}_\#(\mathcal{H}^+ a_q^{\alpha,z} - \mathcal{H}_q^+ \tilde{a}_q^{\alpha,z}), \mathcal{H}^+ a_q^{\alpha,z} - \mathcal{H}_q^+ \tilde{a}_q^{\alpha,z})$$

and therefore

$$D^+(a_q^{\alpha,z}, \tilde{a}_q^{\alpha,z}) \leq \|\mathbb{T}_\#\| \|\mathcal{H}^+ a_q^{\alpha,z} - \mathcal{H}_q^+ \tilde{a}_q^{\alpha,z}\|_{L^2(D)}^2 \rightarrow 0, \quad \text{as } \alpha \rightarrow 0.$$

From Theorem 4.4.4,  $(\mathbb{N}_\# a^{\alpha,z}, a^{\alpha,z})$  converges as  $\alpha \rightarrow 0$  to  $(\mathbb{T}_\# v_0, v_0) < \infty$  where  $v_0 \neq 0$  is the solution of  $G(v_0) = \widehat{\Phi}_q^+(\cdot; z)$ . We then conclude that

$$\lim_{\alpha \rightarrow 0} \mathcal{I}_\alpha^+(z) = 0 \text{ if } z \in D_p.$$

Finally, if  $z \in \omega$  then again by Theorem 4.4.4,  $(\mathbb{N}_\# a^{\alpha,z}, a^{\alpha,z}) < +\infty$  converges as  $\alpha \rightarrow 0$  to  $(\mathbb{T}_\# v_0, v_0) < \infty$  where  $v_0 \neq 0$  is the solution of  $G(v_0) = \widehat{\Phi}_q^+(\cdot; z)$ . However,

$$D^+(a_q^{\alpha,z}, \tilde{a}_q^{\alpha,z}) \rightarrow \infty, \quad \alpha \rightarrow 0$$

since by Theorem 4.4.8  $(\mathbb{N}_\#^+ a_q^{\alpha,z}, a_q^{\alpha,z}) \rightarrow +\infty$  while  $(\mathbb{N}_\#^+ \mathbb{I}_q \tilde{a}_q^{\alpha,z}, \mathbb{I}_q \tilde{a}_q^{\alpha,z}) \rightarrow (T\tilde{v}, \tilde{v}) < +\infty$  where  $\tilde{v} \in H_{\text{inc}}^q(D)$  is such that  $G_q(\tilde{v}) = \mathbb{I}_q^* \widehat{\Phi}_q^+(\cdot; z)$ . Therefore,

$$\lim_{\alpha \rightarrow 0} \mathcal{I}_\alpha^+(z) \neq 0 \text{ if } z \in \omega$$

and the Theorem is proved.  $\square$

Let us conclude this section by indicating that indeed, following Theorem 4.4.5 in the case of noisy measurements, one has to redefine the functionals  $J_\alpha^+$  and  $J_{\alpha,q}^+$  as

$$\begin{aligned} J_\alpha^{+,\delta}(\phi, a) &:= \alpha \left( (\mathbb{N}_\#^{+,\delta} a, a) + \delta \alpha^{-\eta} \|\mathbb{N}_\#^{+,\delta}\| \|a\|^2 \right) + \|\mathbb{N}^{+,\delta} a - \phi\|^2, \\ J_{\alpha,q}^{+,\delta}(\phi, a) &:= \alpha \left( (\mathbb{N}_\#^{+,\delta} \mathbb{I}_q a, \mathbb{I}_q a) + \delta \alpha^{-\eta} \|\mathbb{N}_\#^{+,\delta}\| \|a\|^2 \right) + \|\mathbb{N}_q^{+,\delta} a - \phi\|^2 \end{aligned} \quad (4.96)$$

where  $\eta < 1$  is a fixed positive parameter and  $\delta$  the relative noise level. We then consider  $a_\delta^{\alpha,z}$ ,  $a_{q,\delta}^{\alpha,z}$  and  $\tilde{a}_{q,\delta}^{\alpha,z}$  in  $\ell(\mathbb{Z})$  as the minimizers of respectively

$$J_\alpha^{+,\delta}(\widehat{\Phi}^+(\cdot; z), a), \quad J_\alpha^{+,\delta}(\widehat{\Phi}_q^+(\cdot; z), a) \text{ and } J_{\alpha,q}^{+,\delta}(\widehat{\Phi}_q^+(\cdot; z), a).$$



We define the indicator function in the noisy case as

$$\mathcal{I}_\alpha^{+, \delta}(z) = \left( \mathcal{G}^{+, \delta}(a_\delta^{\alpha, z}) \left( 1 + \frac{\mathcal{G}^{+, \delta}(a_\delta^{\alpha, z})}{D^{+, \delta}(a_{q, \delta}^{\alpha, z}, \tilde{a}_{q, \delta}^{\alpha, z})} \right) \right)^{-1} \quad (4.97)$$

where for  $a$  and  $b$  in  $\ell^2(\mathbb{Z})$ ,

$$D^{+, \delta}(a, b) := \left( N_{\#}^{+, \delta}(a - I_q b), (a - I_q b) \right)$$

and

$$\mathcal{G}^{+, \delta}(a) := (N_{\#}^{+, \delta} a, a) + \delta \alpha^{-\eta} \|N_{\#}^{+, \delta}\| \|a\|^2.$$

Then following the lines of the previous proof one can prove (thanks to Theorem 4.4.5 and equivalent version for  $N_q^+$ )

**Theorem 4.6.8.** *Under the assumptions of Theorem 4.6.6, we have that*

$$z \in \omega \text{ if and only if } \lim_{\alpha \rightarrow 0} \liminf_{\delta \rightarrow 0} \mathcal{I}_\alpha^{+, \delta}(z) > 0.$$

#### 4.6.4 Numerical validating examples

We here give some numerical examples in 2D, where we reconstruct the local perturbation  $\omega$  using the indicator function that we introduced above. The numerical setting for the discrete model is the same as in Section 4.4.5. In addition to the indicator function  $\mathcal{I}^{+, \delta}(z)$  defined by (4.97) and the similarly defined indicator function  $\mathcal{I}^{-, \delta}(z)$ . We also define

$$\mathcal{I}^\delta(z) := \mathcal{I}^{+, \delta}(z) + \mathcal{I}^{-, \delta}(z).$$

In order to give idea on the influence of the different components in the definition on  $\mathcal{I}^{\pm, \delta}(z)$  we shall also show the behavior of the function

$$\mathcal{D}^{+, \delta}(z) := D^{+, \delta}(a_{q, \delta}^{\alpha, z}, \tilde{a}_{q, \delta}^{\alpha, z})$$

and the function  $\mathcal{D}^{-, \delta}(z)$  defined similarly. These two functions are supposed to have small values in  $D_p$  (and therefore discard  $D_p$  in the indicator functions  $\mathcal{I}^{\pm, \delta}(z)$ ).

#### Example 1.

In the first example, we consider again the domain represented in Figure 4.2 in Section 4.4.5 and with numerical parameters given in (4.59)-(4.61). The reconstruction of the domain  $\omega$  is given in Figure 4.9 where one observes that it is indeed isolated from the background structure. Notice that we obtain comparable results to the one obtained by DLSM (see Figure 4.8).

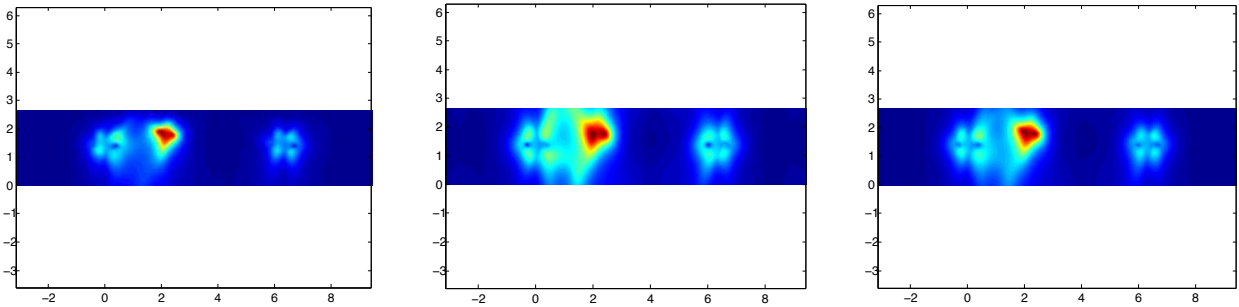


Figure 4.9: Reconstruction the local perturbation  $\omega$ . The exact domain configuration is shown in Figure 4.2. From left to right:  $z \mapsto \mathcal{I}^{+, \delta}(z)$ ,  $z \mapsto \mathcal{I}^{-, \delta}(z)$  and  $z \mapsto \mathcal{I}^\delta(z)$

**Example 2.**

In this example we modify the geometry of  $D_p$  and  $\omega$  and replace the discs by L-shaped domains: see Figure 4.10.

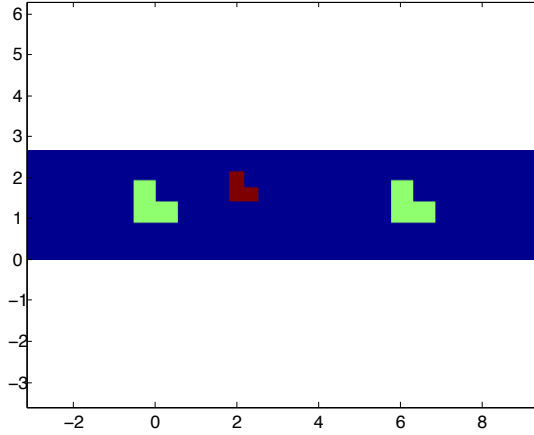


Figure 4.10: Exact geometry for example 2.

We keep the physical parameters as in (4.59)-(4.61). We show in the following the reconstruction of the full domain using GLSM through the indicator function  $z \mapsto \mathcal{I}_{\pm}^{\text{GLSM}}(z)$  (see Figure 4.11), the reconstruction of local perturbation using  $z \mapsto \mathcal{I}_{\pm}^{\text{q}}(z)$  criterion (see Figure 4.12) and the behavior of the functions  $z \mapsto 1/\mathcal{D}^{\pm, \delta}(z)$  (see Figure 4.13).

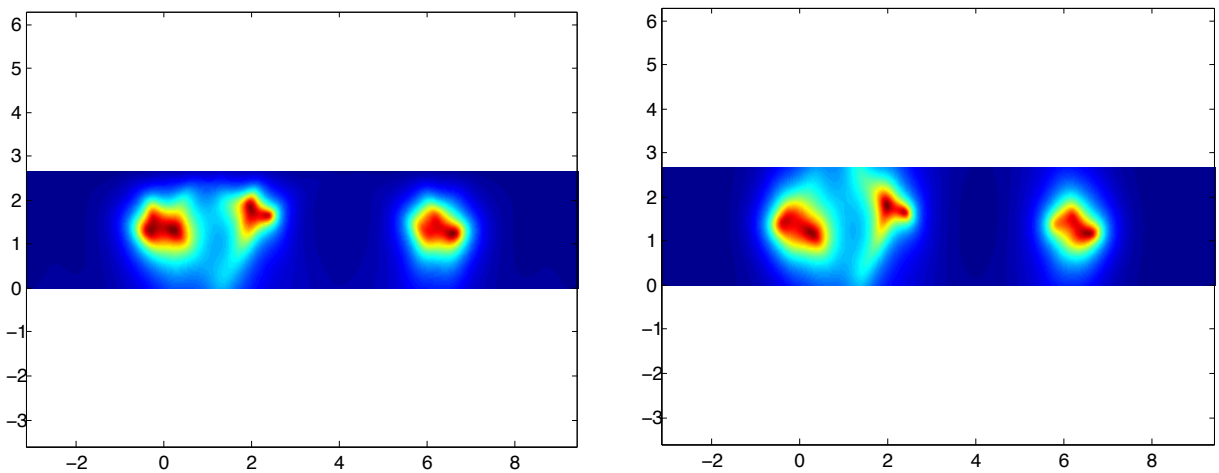


Figure 4.11: Reconstructions of the full domain  $D_p \cup \omega$  using GLSM. The exact geometry is shown in Figure 4.10. From left to right:  $z \mapsto \mathcal{I}_{+}^{\text{GLSM}}(z)$  and  $z \mapsto \mathcal{I}_{-}^{\text{GLSM}}(z)$

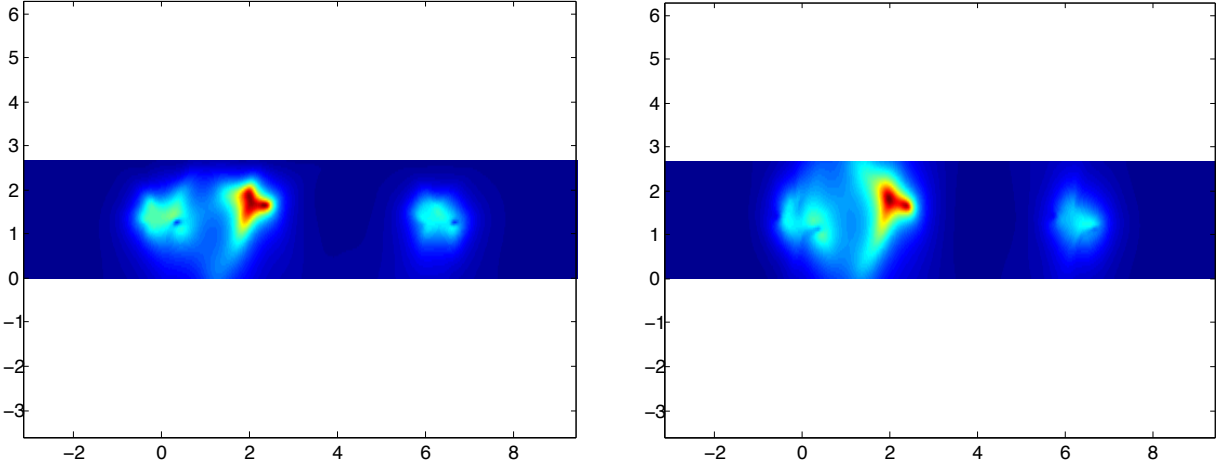


Figure 4.12: Reconstruction the local perturbation  $\omega$  for the case of the geometry depicted in Figure 4.10. From left to right:  $z \mapsto \mathcal{I}^{+,\delta}(z)$  and  $z \mapsto \mathcal{I}^{-,\delta}(z)$ .

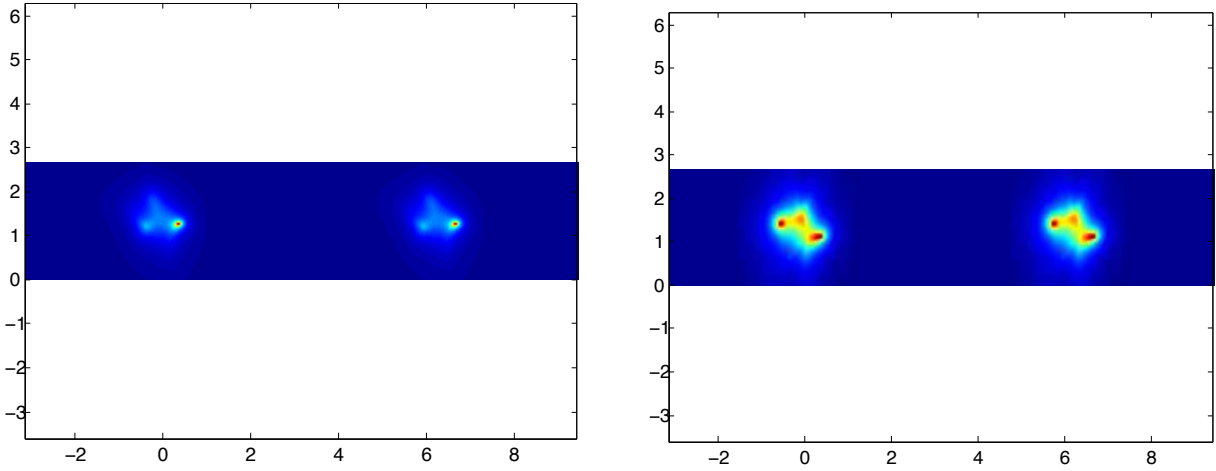


Figure 4.13: Behaviour of  $z \mapsto 1/\mathcal{D}^{+,\delta}$  (left) and  $z \mapsto 1/\mathcal{D}^{-,\delta}$  (right) for Example 2 (Figure 4.10)

We clearly observe a numerical behaviour that corroborates the theoretical predictions. We observe that the resolution of the reconstructed geometry is not good but this is not surprising since the sizes of the domains are smaller than the half of the wavelength.

For the next two examples, we shall just show the configurations of the exact domains, the reconstruction of the full domain using the criterion

$$\mathcal{I}^{\text{GLSM}}(z) := \mathcal{I}_+^{\text{GLSM}}(z) + \mathcal{I}_-^{\text{GLSM}}(z),$$

the reconstruction of local perturbation using the criterion

$$\mathcal{I}^\delta(z) := \mathcal{I}^{+,\delta}(z) + \mathcal{I}^{-,\delta}(z).$$

and the behavior of the function

$$1/\mathcal{D}^\delta(z) := 1/\mathcal{D}^{+,\delta}(z) + 1/\mathcal{D}^{-,\delta}(z).$$

We choose complex configurations for  $D_p$  to illustrate how the indicator function  $\mathcal{I}^\delta(z)$  is capable of visualizing the defect although the GLSM indicator function is not capable of separating the components of the background media. This is indeed well illustrated by Figures 4.14-4.17. For these examples again the parameters as the same as in (4.59)-(4.61).

**Example 3:**

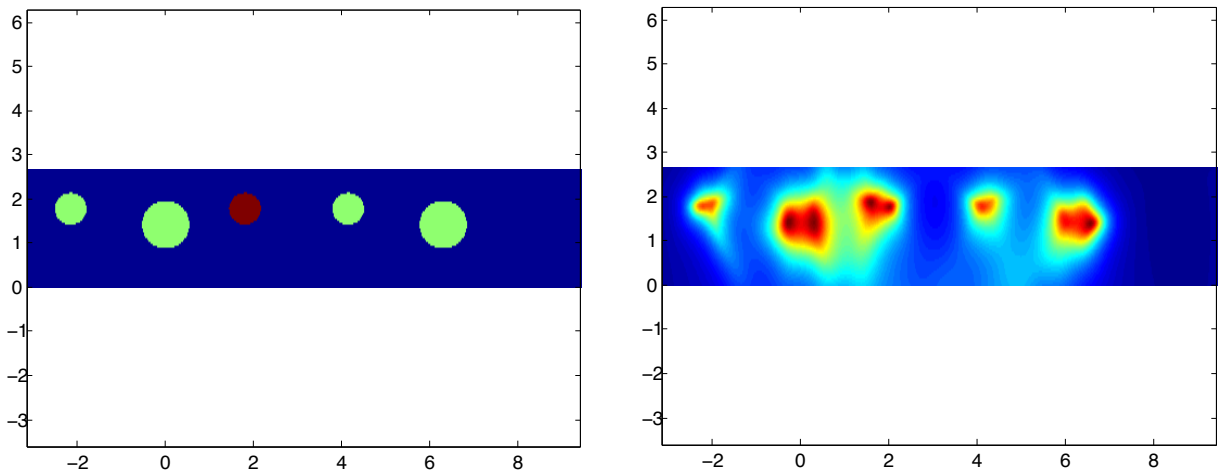


Figure 4.14: Left: The exact geometry for Example 3. Right: the reconstruction of this domain using  $z \mapsto \mathcal{I}^{\text{GLSM}}(z)$ .

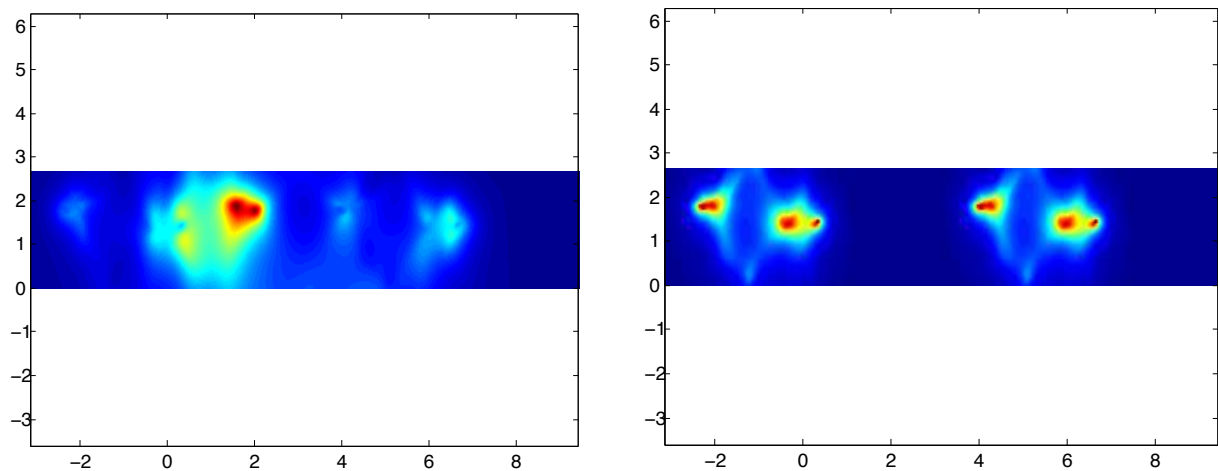


Figure 4.15: Exact geometry is given in Figure 4.14. Left: The reconstruction of the local perturbation using  $z \mapsto \mathcal{I}^\delta(z)$ . Right: behaviour of  $z \mapsto 1/\mathcal{D}^\delta(z)$ .

Example 4.

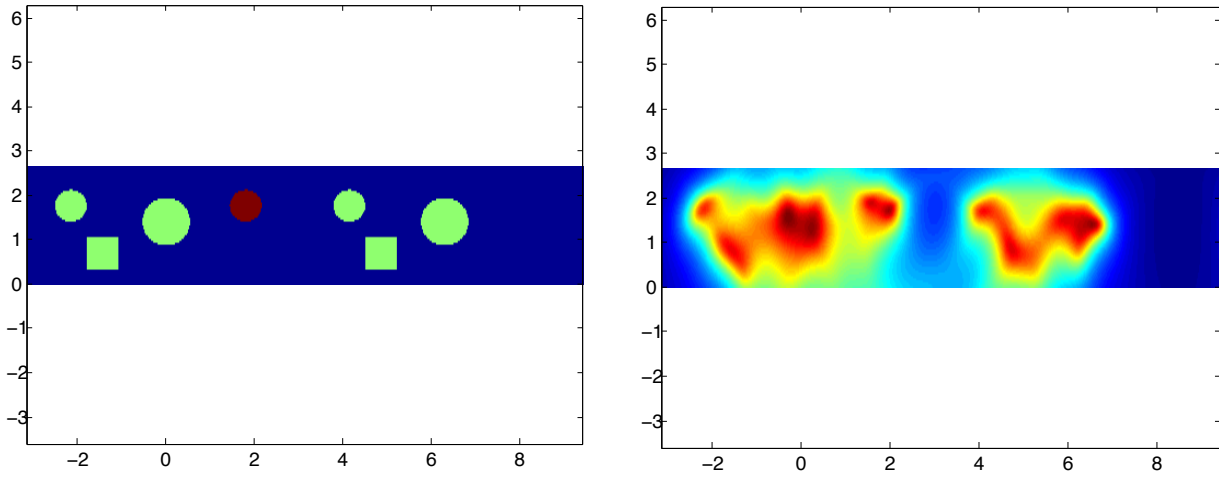


Figure 4.16: Left: The exact geometry for Example 4. Right: the reconstruction of this domain using  $z \mapsto \mathcal{I}^{\text{GLSM}}(z)$ .

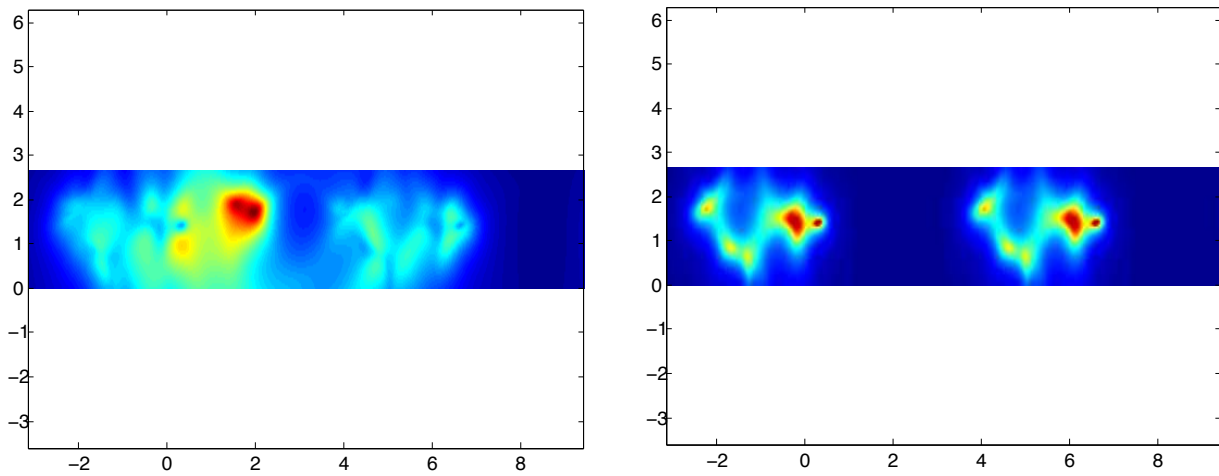


Figure 4.17: Exact geometry is given in Figure 4.16. Left: The reconstruction of the local perturbation using  $z \mapsto \mathcal{I}^\delta(z)$ . Right: behaviour of  $z \mapsto 1/\mathcal{D}^\delta(z)$ .

# Extensions to TM-mode like equations

## Abstract

This chapter is dedicated to the extension of the results of Chapters 3 and 4 to the case where the inhomogeneity induces a contrast in the main operator of the Helmholtz equation. This configuration introduces some additional technical difficulties for the setting of the spectral method that we shall outline in the first part of this chapter. For the inverse problem, except some technical differences in the proofs and the assumptions, the main results remain the same. We shall only give the proofs of the technical lemmas since the proofs for the main results are exactly the same as in Chapter 4. In order to read this chapter, one should first read Chapter 3 for the section on the numerical method and read Chapter 4 for the section on the inverse problem.

## Contents

<b>5.1</b>	<b>Introduction</b>	<b>113</b>
<b>5.2</b>	<b>The spectral volumetric method for solving the direct scattering problem</b>	<b>114</b>
5.2.1	Volume integral formulation of the $\xi$ -quasi-periodic problem	115
5.2.2	Discretization of the locally perturbed periodic problem	117
<b>5.3</b>	<b>The sampling methods for the TM-mode</b>	<b>123</b>
5.3.1	Setting of the inverse problem and definition of the sampling operator	125
5.3.2	Some key properties of the introduced operators	127
5.3.3	Sampling methods for a single Floquet-Bloch mode	132

## 5.1 Introduction

It is known (see [48]) that if the media is invariant along one direction and if the direction of the incident plane wave is chosen to be perpendicular to the invariance axis, Maxwell's equations are decoupled into two scalar Helmholtz equations. They are known as transverse electric (TE) mode and transverse magnetic (TM) mode. In Chapter 3 we considered the Helmholtz equation for the TM mode when the magnetic permeability is constant and studied in Chapter 4 the related inverse problem. We here complement the study of scattering problems from locally perturbed infinite periodic layers by considering the TM mode where the magnetic permeability can vary, which has the form

$$\operatorname{div}(\mu^{-1}\nabla u) + k^2 n u = f \quad \text{in } \mathbb{R}^d, \quad (5.1)$$

where  $f \in L^2(\mathbb{R}^d)$  and  $\mu \in L^\infty(\mathbb{R}^d)$  denotes the magnetic permeability. This equation can also model the TE mode when the electric permittivity is not constant. Although the equation is meaningful, in the case of electromagnetic scattering problems, only for  $d = 2$ , we keep for sake of generality  $d = 2$  or  $3$ . The case  $d = 3$  can also model acoustic scattering problems for media with inhomogeneous mass density.

Although not “exactly” physically true, in order to distinguish the problem studied in Chapter 3 from the problem studied here, we shall refer to the case  $\mu = 1$  as the TE mode and the case  $\mu \neq 1$  as the TM mode. Studying the forward and inverse problems for the TM mode is quasi-similar to the TE mode but is technically more complicated since the magnetic permeability affects the principal part of the Helmholtz operator. In the following we shall only sketch the main results and include the main proofs indicating the major (technical) differences. The material of this chapter thus heavily relies on the material of Chapter 3 for the numerical method to solve the scattering problem and the material of Chapter 4 for the inverse part.

## 5.2 The spectral volumetric method for solving the direct scattering problem

This section is dedicated to the discretization of equation (5.1) using the method introduced in Chapter 3. As we shall see, the only difference comes from the lack of regularity of the solution if the  $\mu$  is not regular enough. This causes difficulty in the periodization of the volume integral equation where a suitable cutoff function has to be incorporated. The latter technique was proposed in [91] for the quasi-periodic scattering problem and we here mainly indicate why this technique can also be applied to the coupled problem that we are considering using the framework of Chapter 3.

We shall use in this section the same notation as in Chapter 3. As in the case of the TE mode we restrict ourselves for the convergence analysis to the case of complex valued wave numbers such that

$$k^2 = k_0^2 + i\sigma$$

with  $k_0 \geq 0$  and  $\sigma > 0$ . We also make the following assumptions on  $\mu$  and  $n$ .

**Assumption 5.2.1.** *We assume that the physical parameters are such that:*

1. *The magnetic permeability  $\mu \in L^\infty(\mathbb{R}^d)$  is real valued and is such that  $\mu^{-1} \in W^{1,\infty}(\mathbb{R}^d)$  and  $\mu = \mu_p$  outside a compact domain  $\omega$  where  $\mu_p$  is a periodic function with respect to the first  $d - 1$  variables of period  $L := (L_1, \dots, L_{d-1}) \in \mathbb{R}^{d-1}$ ,  $L_j > 0$ ,  $j = 1, \dots, d - 1$ . Moreover  $\mu^{-1} \geq c_\mu$  in  $\mathbb{R}^d$  for some constants  $c_\mu > 0$ .*
2. *The electric permittivity  $n \in L^\infty(\mathbb{R}^d)$  is complex valued with non negative imaginary part and  $n = n_p$  outside the compact domain  $D$  with  $n_p$  a periodic function with respect to the first  $d - 1$  variables of period  $L$ . Moreover,  $\operatorname{Re} n \geq c_n$  in  $\mathbb{R}^d$  for some constants  $c_n > 0$ .*

Using the Lax-Milgram theorem and Assumption 5.2.1, equation (5.1) admits a unique solution  $u \in H^1(\mathbb{R}^d)$  and

$$\|u\|_{L^2(\mathbb{R}^d)} \leq \frac{1}{c_n \sigma} \|f\|_{L^2(\mathbb{R}^d)} \quad \text{and} \quad \|\nabla u\|_{L^2(\mathbb{R}^d)} \leq \sqrt{\frac{1}{c_\mu} \left( \frac{k_0^2 \|\operatorname{Re}(n)\|_{L^\infty}}{c_n^2 \sigma^2} + \frac{1}{c_n \sigma} \right)} \|f\|_{L^2(\mathbb{R}^d)} \quad (5.2)$$

Considering the domains  $\Omega^h$ ,  $\Omega_m$  and  $\Omega_m^h$  as in Chapter 3, we also assume that there exists  $h > 0$  such that

$$\operatorname{supp}(\mu^{-1} - 1) \subset \Omega^h, \quad \operatorname{supp}(n - 1) \subset \Omega^h \quad \text{and} \quad \operatorname{supp}(f) \subset \Omega^h.$$

Applying the Floquet-Bloch transform  $\mathcal{F}$  to equation (5.1) and recalling that since  $\mu_p^{-1}$  is  $L$  periodic, then

$$\mathcal{F}(\mu_p^{-1} \nabla u) = \mu_p^{-1} \nabla(\mathcal{F}u)$$

## 5.2. The spectral volumetric method for solving the direct scattering problem 115

and if a function  $\varphi$  has a support in  $\Omega_0$ , then  $\mathcal{F}\varphi(\bar{x}, x_d; \xi) = \varphi(\bar{x}, x_d)$  for  $(\bar{x}, x_d) \in \Omega_0$  we transform the problem (5.1) into seeking  $\tilde{u} \in L^2(\llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket, H_{\xi}^1(\Omega_0))$  such that

$$\begin{aligned} \operatorname{div}(\mu_p^{-1} \nabla \tilde{u})(\cdot; \xi) + \operatorname{div}((\mu^{-1} - \mu_p^{-1}) \nabla \mathcal{M}(\tilde{u}))(\bar{x}, x_d) + k^2 n_p \tilde{u}(\cdot; \xi) + k^2(n - n_p) \mathcal{M}(\tilde{u}) = f_{\xi} \\ \text{in } \Omega_0, \quad \xi \in \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket \end{aligned} \quad (5.3)$$

where  $f_{\xi} := (\mathcal{F}f)(\cdot; \xi)$  and

$$\mathcal{M}(\tilde{u}) := \frac{\llbracket L \rrbracket}{(2\pi)^{d-1}} \int_{\llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket} \tilde{u}(\cdot; \xi) \, d\xi.$$

The latter definition comes from the inverse Floquet-Bloch transform expression (3.10). One can check that the converse is true, i.e. if  $(\mathcal{F}f) \in H^1(\Omega_0 \times \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket)$ , the  $u$  defined by (3.10) is in  $H^1(\mathbb{R}^2)$  and is solution to (5.1). Following the outline of Chapter 3, before dealing with problem (5.3) where the equations for different values of  $\xi$  are coupled through  $\mathcal{M}(\tilde{u})$ , we shall first recall the principles of the spectral volume integral method for fixed  $\xi$  as in [85] for periodic media (and the TM-mode).

### 5.2.1 Volume integral formulation of the $\xi$ -quasi-periodic problem

For a fixed  $\xi \in \llbracket -\frac{\pi}{L}, \frac{\pi}{L} \rrbracket$ , the studying of the  $\xi$ -quasi-periodic problem using the volume integral formulation has been done in [91]. In the following, we just recall the main points for later use. Since studying the  $\xi$ -quasi-periodic problem in  $\mathbb{R}^d$  or in one period  $\Omega_0$  is equivalent, we shall not distinguish  $u_{\xi}$  from its restriction to  $\Omega_0$ . We consider the  $\xi$ -quasi-periodic Helmholtz equation given by: Seek  $u_{\xi} \in H_{\xi, L}^1(\mathbb{R}^d)$  verifying

$$\operatorname{div}(\mu_p^{-1} \nabla u_{\xi}) + k^2 n_p u_{\xi} = f_{\xi} \quad \text{in } \mathbb{R}^d \quad (5.4)$$

where  $f_{\xi} \in L_{\xi, L}^2(\mathbb{R}^d)$ . We recall from (5.2) that

$$\|u_{\xi}\|_{H_{\xi}^1(\mathbb{R}^d)} \leq C \|f_{\xi}\|_{L^2(\mathbb{R}^d)}, \quad \text{where } C := \max \left\{ \frac{1}{c_n \sigma}, \sqrt{\frac{1}{c_{\mu}} \left( \frac{k_0^2 \|\operatorname{Re}(n)\|_{L^{\infty}}}{c_n^2 \sigma^2} + \frac{1}{c_n \sigma} \right)} \right\}. \quad (5.5)$$

Using the volume potential  $V_{\xi}$  in (3.16), we define the volume potential  $L_{\xi}$  as

$$L_{\xi} g(x) := \operatorname{div} V_{\xi} g(x), \quad \forall x \in \mathbb{R}^d. \quad (5.6)$$

As a consequence of Lemma 3.3.1 we have that the operator  $L_{\xi}$  is a linear bounded operator from  $L^2(\Omega_0^h)$  into  $H_{\xi, L}^1(\mathbb{R}^d)$ . Moreover, for all  $g \in L^2(\Omega_0^h)^d$ , the potential  $w := L_{\xi} g \in H_{\xi, L}^1(\mathbb{R}^d)$  and is the unique solution to

$$\Delta w + k^2 w = -\operatorname{div} g \quad \text{in } \mathbb{R}^d.$$

Therefore it is deduced, in combination with Lemma 3.3.1, that if we set

$$w := L_{\xi}((\mu_p^{-1} - 1) \nabla u_{\xi}) + V_{\xi}(k^2(n_p - 1)u_{\xi} - f_{\xi})$$

then  $w \in H_{\xi, L}^1(\mathbb{R}^d)$  and satisfies

$$\Delta w + k^2 w = -\operatorname{div}((\mu_p^{-1} - 1) \nabla u_{\xi}) - k^2(n_p - 1)u_{\xi} + f_{\xi} \quad \text{in } \mathbb{R}^d.$$



This proves  $w = u_\xi$  in  $\mathbb{R}^d$ ,  $u_\xi \in H^1(\Omega_0^h)$  and satisfies

$$u_\xi = L_\xi \left( (\mu_p^{-1} - 1) \nabla u_\xi \right) + V_\xi \left( k^2 (n_p - 1) u_\xi \right) - V_\xi f_\xi \quad \text{in } H^1(\Omega_0^h). \quad (5.7)$$

Conversely, if (5.7) is verified, then obviously  $u_\xi$  can be obtained in all  $\mathbb{R}^d$  using the expression of  $w$ . For the same purpose as for the TE-mode, we need to periodize the equation in the direction  $x_d$ . However, as opposed to the TE-mode the periodized Green function  $G_\xi^R$  defined in (3.18) has no longer sufficient smoothness at the boundaries  $\Gamma_{\pm h}$  (due to the divergence operator in front of the integral operator) so that the periodized equation remains Fredholm. This is why, following [91], we additionally need to smooth the periodized kernel. This is the step we shall discuss now.

### Periodization of the integral equation

For  $R \in \mathbb{R}$  such that  $R > 2h$ , let  $\chi_R \in C^\infty(\mathbb{R})$  which is  $2R$ - periodic function and such that

$$0 \leq \chi_R \leq 1, \quad \chi_R(x_d) = 1, \quad \forall |x_d| \leq 2R \quad \text{and} \quad \frac{d^j \chi_R}{dx_d} = 0, \quad \text{for } j = 0, 1, 2, \quad (5.8)$$

we define the operator  $\mathcal{G}_\xi^R$  as

$$\mathcal{G}_\xi^R(x) := \chi_R(x_d) G_\xi^R(x), \quad \forall x = (\bar{x}, x_d) \in \mathbb{R}^{d-1} \times ]-R, R[, \quad (5.9)$$

where  $G_\xi^R$  is defined in (3.18). We then define the periodized volume potential  $\mathcal{V}_\xi^R : L^2(\Omega_0^h) \rightarrow H^1(\Omega_0^R)$  using the same expression as (3.16) where  $G_\xi$  is replaced by  $\mathcal{G}_\xi^R(x)$  and the periodized volume potential  $\mathcal{L}_\xi$  as

$$\mathcal{L}_\xi := \text{div } \mathcal{V}_\xi.$$

We then consider the periodized volume integral equation,  $u_\xi^R \in H^1(\Omega_0^R)$ ,

$$u_\xi^R = \mathcal{L}_\xi^R \left( (\mu_p^{-1} - 1) \nabla u_\xi^R \right) + \mathcal{V}_\xi^R \left( k^2 (n_p - 1) u_\xi^R \right) - \mathcal{V}_\xi^R f_\xi \quad \text{in } H^1(\Omega_0^R). \quad (5.10)$$

Similarly to Lemma 3.3.2 we have the following equivalence.

**Lemma 5.2.2.** 1. For all  $(g_1, g_2) \in L^2(\Omega_0^h)^d \times L^2(\Omega_0^h)$ ,  $\mathcal{L}_\xi^R g_1 = \mathcal{L}_\xi g_1$  and  $\mathcal{V}_\xi^R g_2 = \mathcal{V}_\xi g_2$  in  $\mathbb{R}^d$ .

2. For all  $f_\xi \in L^2(\Omega_0^h)$ , equation (5.10) has a unique solution  $u_\xi^R \in H^1(\Omega_0^R)$  and we have

$$u_\xi^R = u_\xi \quad \text{in } \Omega_0^h$$

where  $u_\xi \in L^2(\Omega_0^h)$  is the unique solution of (5.7).

Moreover one can prove the following continuity result (see for instance [91] and remark that for this case the constant  $C$  may depend on  $\xi$ ).

**Lemma 5.2.3.** Let  $s \in \mathbb{N}$ . The operator  $\mathcal{L}_\xi^R : H_\xi^s(\Omega_0^R) \rightarrow H_\xi^{1+s}(\Omega_0^R)$  is continuous., i.e, there exists a constant  $C$  such that

$$\|\mathcal{L}_\xi^R g\|_{H_\xi^{1+s}(\Omega_0^R)} \leq C \|g\|_{H_\xi^s(\Omega_0^R)} \quad \forall g \in H_\xi^s(\Omega_0^R).$$

Remark that, as a consequence of Lemma 5.2.3 the operator  $\mathcal{V}_\xi^R : H_\xi^s(\Omega_0^R) \rightarrow H_\xi^{2+s}(\Omega_0^R)$  is continuous. We now can prove the following for the solution to the periodized equation.

## 5.2. The spectral volumetric method for solving the direct scattering problem 117

**Proposition 5.2.4.** *Let  $f_\xi \in L^2(\Omega_0^h)$  and  $u_\xi^R \in H^1(\Omega_0^R)$  be the solution of (5.10). Then there exists a positive constant  $C$  independent from  $\xi$  and  $f_\xi$  such that*

$$\|u_\xi^R\|_{H^1(\Omega_0^R)} \leq C \|f_\xi\|_{L^2(\Omega_0^h)}. \quad (5.11)$$

*Proof.* Thanks to Lemma 5.2.2,  $u_\xi^R = u_\xi$  in  $\Omega_0^h$  where  $u_\xi$  denotes the solution of (5.7). Then, using (5.5),

$$\|u_\xi^R\|_{H^1(\Omega_0^h)} \leq C_1 \|f_\xi\|_{L^2(\Omega_0^h)}$$

for some constant  $C_1$ . Using equation (5.10), Lemma (5.2.3) and the fact that  $\text{supp}(1 - \mu_p^{-1}) \cup \text{supp}(1 - n) \subset \Omega_0^h$ , we get the existence of a constant  $C_2$  such that

$$\|u_\xi^R\|_{H^1(\Omega_0^R)} \leq C_2 \left( \|f_\xi\|_{L^2(\Omega_0^h)} + \|u_\xi^R\|_{H^1(\Omega_0^h)} + \|u_\xi^R\|_{L^2(\Omega_0^h)} \right).$$

The result of the proposition then immediately follows from the two inequalities.  $\square$

Let us finally observe that, as a straightforward consequence of Lemma 5.2.3, we have:

**Corollary 5.2.5.** *The operator  $\mathcal{A}_\xi^R : H^1(\Omega_0^R) \rightarrow H^1(\Omega_0^R)$  defined by*

$$\mathcal{A}_\xi^R g := g - \mathcal{L}_\xi^R((\mu_p^{-1} - 1)\nabla g) - \mathcal{V}_\xi^R(k^2(n_p - 1)g). \quad (5.12)$$

*is bounded and injective, i.e, there exists a positive constant  $C$  such that*

$$\|\mathcal{A}_\xi^R g\|_{H^1(\Omega_0^R)} \leq C \|g\|_{H^1(\Omega_0^R)}.$$

### Spectral approximation of problem (5.10)

We define for  $N = (N_1, \dots, N_d) \in \mathbb{N}^d$  the set of index  $\mathbb{Z}_N^d$ , the Fourier basic  $\varphi_\xi^i$ , the discrete space  $\mathcal{T}_\xi^N$  and the orthogonal projection  $P_\xi^N$  as in Section 3.3.3. Let  $u_\xi^R$  be the solution of (5.10). We now consider  $u_\xi^{R,N} \in \mathcal{T}_\xi^N$  an approximation of  $u_\xi^R$ , which is solution to following variational equation

$$\begin{aligned} \left\langle u_\xi^{R,N} - \mathcal{L}_\xi^R((\mu_p^{-1} - 1)\nabla u_\xi^{R,N}), v_\xi^N - \mathcal{V}_\xi^R(k^2(n_p - 1)u_\xi^{R,N}), v_\xi^N \right\rangle_{H^1(\Omega_0^R)} \\ = \left\langle -\mathcal{V}_\xi^R f_\xi, v_\xi^N \right\rangle_{H^1(\Omega_0^R)}, \quad \forall v_\xi^N \in \mathcal{T}_\xi^N. \end{aligned} \quad (5.13)$$

The next theorem indicates that the convergence rate of this method is proportional to the interpolation error.

**Theorem 5.2.6.** *There exists  $N_0 \in \mathbb{N}^d$  such that problem (5.13) admits a unique solution for all  $N > N_0$  and  $f_\xi \in L^2(\Omega_0^h)$ . Moreover, there is a positive constant  $C > 0$  such that,*

$$\|u_\xi^R - u_\xi^{R,N}\|_{H^1(\Omega_0^R)} \leq C \inf_{v_\xi^N \in \mathcal{T}_\xi^N} \|u_\xi^R - v_\xi^N\|_{H^1(\Omega_0^R)}. \quad (5.14)$$

*Proof.* The proof of this theorem is similar to the proof of theorem 3.3.7 with the help of Lemma 5.2.5. However, we no longer have the guarantee that the constant  $C$  is independent from  $\xi$ .  $\square$

### 5.2.2 Discretization of the locally perturbed periodic problem

We now address the discretization of the original problem (5.3). Let us set again  $f_\xi = \mathcal{F}f(\cdot, \xi)$ . The idea is to perform first a discretization with respect to the Floquet-Bloch variable, then perform the discretization in space as done in the previous section.

### Discretization and convergence in the Floquet-Bloch variable

Using the same procedure as in Section 3.4.1, we discretize the integral  $\mathcal{M}(\tilde{u})$  using the trapezoidal rule and obtain as semi-discrete system:

$$\begin{aligned} \tilde{u}_M(\cdot; \xi_j) - L_{\xi_j}((\mu_p^{-1} - 1)\nabla\tilde{u}_M(\cdot; \xi_j)) - L_{\xi_j}((\mu^{-1} - \mu_p^{-1})\nabla u_M) - V_{\xi_j}(k^2(n_p - 1)\tilde{u}_M(\cdot; \xi_j)) \\ - V_{\xi_j}((n - n_p)u_M) = -V_{\xi_j}(f_{\xi_j}) \quad \text{in } H^1(\Omega_0^h), \text{ for } j \in \mathbb{Z}_M^{d-1}, \end{aligned} \quad (5.15)$$

and

$$u_M := \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} \tilde{u}_M(\cdot; \xi_j) \quad \text{in } \Omega_0. \quad (5.16)$$

We recall that the function  $u_M$  constitutes an approximation of  $u$  in  $\Omega_0^h$ . The functions  $\tilde{u}_M(\cdot; \xi_j)$  can be extended to  $\Omega_0$  using equation (5.15) and then extended to all  $\mathbb{R}^d$  using the quasi-periodicity property as in (3.59). Then we extend  $u_M$  to  $\mathbb{R}^d$  using the same formula as in equation (5.16), namely

$$u_M := \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} \tilde{u}_M(\cdot; \xi_j) \quad \text{in } \mathbb{R}^d.$$

We also define  $f_M$  as

$$f_M := \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} f_{\xi_j} \quad \text{in } \mathbb{R}^d.$$

In particular,

$$u_M(\bar{x} + mL, x_d) = \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} e^{i(mL) \cdot \xi_j} \tilde{u}_M((\bar{x}, x_d); \xi_j), \quad x = (\bar{x}, x_d) \in \Omega_0, \quad m \in \mathbb{Z}^{d-1}. \quad (5.17)$$

Using the potential  $V_{\#,M}$  defined in (3.65), we then introduce the potential  $L_{\#,M}$  as:

$$L_{\#,M}f(x) = \operatorname{div} V_{\#,M}f(x), \quad \forall f \in L^2(\Omega_M^h)^d. \quad (5.18)$$

Then as a consequence of Theorem 3.4.2, the volume potential  $L_{\#,M}$  is a linear bounded operator from  $L^2(\Omega_M^h)^d$  into  $H_{\#,M}^1(\Omega_M)$ . Moreover, let  $\tilde{u}_M(\cdot; \xi_j)$  be the solution to system (5.15). Then the function  $u_M$  defined in (5.17) belongs to  $H_{\#,M}^1(\Omega_M)$  and verifies the following volume integral equation

$$u_M - L_{\#,M}((\mu^{-1} - 1)\nabla u_M) - k^2 V_{\#,M}((n - 1)u_M) = -V_{\#,M}(f_M) \quad \text{in } H^1(\Omega_M). \quad (5.19)$$

**Corollary 5.2.7.** *The function  $u_M \in H_{\#,M}^1(\Omega_M)$  and is solution to*

$$\operatorname{div}(\mu^{-1}\nabla u_M) + k^2 n u_M = f_M \quad \text{in } \Omega_M. \quad (5.20)$$

Furthermore, under Assumption 5.2.1, there exists a positive constant  $C$  independent from  $M$  such that

$$\|u_M\|_{H_{\#,M}^2(\Omega_M)} \leq C \|f_M\|_{L^2(\Omega_M^h)}. \quad (5.21)$$

We remark that the  $H^2$  regularity of the solution comes from the fact that  $\mu^{-1} \in W^{1,\infty}(\mathbb{R}^d)$  and the uniform bound with respect to  $M$  comes from the uniform bound for the  $H^1$  norm and the fact that

$$\Delta u_M = \mu(f_M - k^2 n u_M - \nabla \mu^{-1} \cdot \nabla u_M).$$

Due to the non vanishing imaginary part of  $k^2$ , we now can formulate the following exponential convergence result with respect to  $M$  that is a consequence of the exponential decay of the solution  $u$ . The proof is similar to the TE case and we give a sketch of it in the following.

## 5.2. The spectral volumetric method for solving the direct scattering problem 119

**Theorem 5.2.8.** *Assume that  $\mu$  and  $n$  satisfy Assumption 5.2.1. Let  $f \in L^2(\mathbb{R}^d)$  such that  $e^{\tau^* \|x\|} f \in L^2(\mathbb{R}^d)$  for some  $\tau^* > 0$ . Let  $u \in H^1(\mathbb{R}^d)$  be the solution of (5.1) and  $u_M \in H^1_{\#}(\Omega_M)$  be the solution of (5.20). Then there exists  $\tau_0 > 0$  and a constant  $C$  independent of  $M$  and  $f$  and  $\tau$  such that*

$$\|u - u_M\|_{H^1(\Omega_M)} \leq C \left( e^{-\tau \min(ML)/2} \|e^{\tau \|x\|} f\|_{L^2(\mathbb{R}^d)} + \|f - f_M\|_{L^2(\mathbb{R}^d)} \right) \quad \forall 0 \leq \tau \leq \tau_0. \quad (5.22)$$

*Proof.* If we set  $v = e^{\tau \|x\|} u$ , then  $v$  satisfies

$$\operatorname{div}(\mu^{-1} \nabla v) - 2\mu^{-1} \tau \frac{x}{\|x\|} \cdot \nabla v + \left( k^2 n + \mu^{-1} \tau^2 - \mu^{-1} \frac{\tau(d-1)}{\|x\|} - \nabla(\mu^{-1}) \cdot \frac{x}{\|x\|} \right) v = e^{\tau \|x\|} f \quad \text{in } \mathbb{R}^d. \quad (5.23)$$

Similarly to (5.2) we have that

$$\|v\|_{L^2(\mathbb{R}^d)} \leq \frac{1}{c_n \sigma} \left( \|e^{\tau \|x\|} f\|_{L^2(\mathbb{R}^d)} + 2\mu^{-1} \tau \|\nabla v\|_{L^2(\mathbb{R}^d)} \right)$$

and

$$\|\nabla v\|_{L^2(\mathbb{R}^d)}^2 \leq \frac{1}{c_\mu} \left( \frac{\theta}{c_n^2 \sigma^2} + \frac{1}{c_n \sigma} \right) \left( \|e^{\tau \|x\|} f\|_{L^2(\mathbb{R}^d)} + 2\mu^{-1} \tau \|\nabla v\|_{L^2(\mathbb{R}^d)} \right)^2. \quad (5.24)$$

where,

$$\theta := k_0^2 \|\operatorname{Re}(n)\|_{L^\infty} + \mu^{-1} \tau^2 + \|\nabla(\mu^{-1})\|_{L^\infty(\mathbb{R}^d)}$$

Let us denote by

$$B_\varepsilon := \{x \in \mathbb{R}^d : \|x\| \leq \varepsilon\}$$

for  $\varepsilon > 0$  such that  $B_\varepsilon \subset \Omega_M$ . Choosing  $\tau_0 > 0$  such that  $1 - \frac{2(2\mu^{-1}\tau)^2}{c_\mu} \left( \frac{\theta}{c_n^2 \sigma^2} + \frac{1}{c_n \sigma} \right) > 1/2$  and  $\tau_0 < \tau^*$  we get the existence of a constant  $C(\tau_0)$  independent from  $\tau$  and  $\varepsilon$  such that

$$\|\operatorname{div}(\mu^{-1} \nabla v)\|_{L^2(\mathbb{R}^d \setminus B_\varepsilon)}^2 + \|\mu^{-1} \nabla v\|_{L^2(\mathbb{R}^d)} + \|v\|_{H^1(\mathbb{R}^d)}^2 \leq C \|e^{\tau \|x\|} f\|_{L^2(\mathbb{R}^d)} \quad \forall 0 \leq \tau \leq \tau_0. \quad (5.25)$$

Let us denote by  $\phi_j^\pm := u|_{\Gamma_{j,M}^\pm}$  and  $\psi_j := \mu^{-1} \frac{\partial}{\partial x_j} u|_{\Gamma_{j,M}^\pm}$ . Then, from trace theorems and since  $\Gamma_{j,M}^\pm$  can be identified as a part of  $\mathbb{R}^{d-1}$ , there exists a constant  $C$  independent from  $j$  such that

$$e^{\tau(M_j \pm 1)L_j} \left( \|\phi_j^\pm\|_{H^{1/2}(\Gamma_{j,M}^\pm)}^2 + \|\psi_j^\pm\|_{H^{-1/2}(\Gamma_{j,M}^\pm)}^2 \right) \leq C \left( \|\operatorname{div}(\mu^{-1} \nabla v)\|_{L^2(\mathbb{R}^d \setminus B_\varepsilon)}^2 + \|\mu^{-1} \nabla v\|_{L^2(\mathbb{R}^d \setminus B_\varepsilon)} + \|v\|_{H^1(\mathbb{R}^d \setminus B_\varepsilon)}^2 \right). \quad (5.26)$$

Since  $w := u - u_M$  solves

$$\begin{cases} \operatorname{div}(\mu^{-1} \nabla w) + k^2 n w = f - f_M & \text{in } \Omega_M, \\ w|_{\Gamma_{j,M}^+} - w|_{\Gamma_{j,M}^-} = \phi_j^+ - \phi_j^-, \\ \mu^{-1} \frac{\partial}{\partial x_1} w|_{\Gamma_{j,M}^+} - \mu^{-1} \frac{\partial}{\partial x_1} w|_{\Gamma_{j,M}^-} = \psi_j^+ - \psi_j^-, \end{cases}$$

then, due to  $\sigma > 0$ , there exists constant  $C' > 0$  independent from  $M$  such that

$$\|w\|_{H^1(\Omega_M)}^2 \leq C' \left( \sum_j \left( \|\phi_j^\pm\|_{H^{1/2}(\Gamma_{j,M}^\pm)}^2 + \|\psi_j^\pm\|_{H^{-1/2}(\Gamma_{j,M}^\pm)}^2 \right) + \|f - f_M\|_{L^2(\Omega_M)} \right). \quad (5.27)$$

The result directly follows from combining (5.25), (5.26) and (5.27).  $\square$

### Discretization in the spatial variable

The discretization procedure in the spatial variable is the same as the one discussed in Section 5.2.1.

### Periodization of the coupled integral equations

To apply the method we first periodize the volume integral equations of system (5.15) in  $x_d$ -direction and get the following system.

$$\begin{aligned} \tilde{u}_M^R(\cdot; \xi_j) - \mathcal{L}_{\xi_j}^R((\mu_p^{-1} - 1)\nabla \tilde{u}_M^R(\cdot; \xi_j)) - \mathcal{L}_{\xi_j}^R((\mu^{-1} - \mu_p^{-1})\nabla u_M^R) - \mathcal{V}_{\xi_j}^R(k^2(n_p - 1)\tilde{u}_M^R(\cdot; \xi_j)) \\ - \mathcal{V}_{\xi_j}^R(k^2(n - n_p)u_M^R) = -\mathcal{V}_{\xi_j}^R f_{\xi_j} \quad \text{in } H^1(\Omega_0^R), j \in \mathbb{Z}^{d-1} \end{aligned} \quad (5.28)$$

where

$$u_M^R := \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} \tilde{u}_M^R(\cdot; \xi_j).$$

Analogously to the results in Chapter 3, we have that the solution to equation (5.15) and the solution to equation (5.28) coincide in  $\Omega_0^h$ . Therefore, the extension by quasi-periodicity of the solution to equation (5.15) coincide with the extension by quasi-periodicity of the solution to equation (5.28) in  $\Omega_M^h$ . Let

$$\mathcal{G}_{\#,M}^R(x) := \chi(x_d)G_{\#,M}^R(x)$$

be the smoothed version of  $G_{\#,M}^R$ . We introduce the volume potential  $\mathcal{V}_{\#,M}^R$  as:

$$\mathcal{V}_{\#,M}^R f(x) := \int_{\Omega_M^h} \mathcal{G}_{\#,M}^R(x-y)f(y) dy \quad (5.29)$$

and the volume potential  $\mathcal{L}_{\#,M}^R := \text{div } \mathcal{V}_{\#,M}^R$ . Then  $u_M^R \in H^1(\Omega_M^R)$  verifies:

$$u_M^R = \mathcal{L}_{\#,M}^R((\mu^{-1} - 1)\nabla u_M^R) + \mathcal{V}_{\#,M}^R(k^2(n-1)u_M^R) - \mathcal{V}_{\#,M}^R f_M \quad \text{in } H^1(\Omega_M^R). \quad (5.30)$$

The link between the solution of the periodized equation (5.30) and the solution of (5.19) can be established in a similar way as for  $\xi$ -quasi-periodic problems (apply the procedure with  $\xi = 0$  and a period =  $ML$ ). We here regroup these results:

**Proposition 5.2.9.** *Let  $\mu$  and  $n$  be satisfying Assumption 5.2.1.*

- Equation (5.19) has a unique solution  $u_M \in H_{\#}^1(\Omega_M^h)$  for all  $f_M \in L^2(\Omega_M^h)$  if and only if equation (5.30) has a unique solution  $u_M^R \in H_{\#}^1(\Omega_M^R)$  for all  $f_M \in L^2(\Omega_M^h)$ .
- Let  $f_M \in L^2(\Omega_M^h)$ ,  $u_M^R \in H_{\#}^1(\Omega_M^R)$  be the solution to (5.30) and  $u_M \in H_{\#}^1(\Omega_M^h)$  be the solution to (5.19). Then

$$u_M^R = u_M \quad \text{in } \Omega_M^h. \quad (5.31)$$

Moreover there exists a positive constant  $C$  independent of  $M$  and  $f_M$  such that:

$$\|u_M^R\|_{H_{\#}^2(\Omega_M^R)} \leq C \|f_M\|_{L^2(\Omega_M^h)}. \quad (5.32)$$

The uniform bound with respect to  $M$  is a consequence of the following lemma that can be proved in the same manner as in Lemma 3.3.3 using (3.26) since the cut-off function  $\chi_R$  depends only on the  $x_d$  variable. The derivation of the  $H^2$  bound is a consequence of twice applications of the Lemma 5.2.10 using (5.30), the fact that  $\mu^{-1} \in W^{1,\infty}(\mathbb{R}^d)$ , the support of  $\mu^{-1} - 1$  is in  $\Omega_M^h$ , equality (5.31) and the regularity result of Corollary 5.2.7.

## 5.2. The spectral volumetric method for solving the direct scattering problem 121

**Lemma 5.2.10.** *Let  $s \in \mathbb{N}$ . The operator  $\mathcal{L}_{\#,M}^R : H_{\#}^s(\Omega_M^R) \rightarrow H_{\#}^{s+1}(\Omega_M^R)$  is continuous. Moreover, there exists a constant  $C$  independent of  $M$  such that*

$$\|\mathcal{L}_{\#,M}^R g\|_{H_{\#}^{1+s}(\Omega_M^R)} \leq C \|g\|_{H_{\#}^s(\Omega_M^R)}, \quad \forall g \in H_{\#}^s(\Omega_M^R). \quad (5.33)$$

Let us introduce for later use in the convergence analysis  $\mathcal{A}_{\#,M}^R : H^1(\Omega_M^R) \rightarrow H^1(\Omega_M^R)$  defined by

$$\mathcal{A}_{\#,M}^R g := g - \mathcal{L}_{\#,M}^R((\mu^{-1} - 1)\nabla g) - \mathcal{V}_{\#,M}^R(k^2(n-1)g). \quad (5.34)$$

**Lemma 5.2.11.** *The operator  $\mathcal{A}_{\#,M}^R : H^1(\Omega_M^R) \rightarrow H^1(\Omega_M^R)$  is bounded and injective. Moreover,  $\mathcal{A}_{\#,M}^R$  is a Fredholm operator of index 0, i.e.,  $\mathcal{A}_{\#,M}^R$  has a decomposition*

$$\mathcal{A}_{\#,M}^R = \mathcal{A}_0 + \mathcal{T}_0 \quad (5.35)$$

where the operator  $\mathcal{A}_0 : H^1(\Omega_M^R) \rightarrow H^1(\Omega_M^R)$  is coercive and  $\mathcal{T}_0 : H^1(\Omega_M^R) \rightarrow H^1(\Omega_M^R)$  is a compact operator.

### Spectral approximation

We consider  $\tilde{u}_M^{R,N}(\cdot; \xi_j) \in \mathcal{T}_{\xi_j}^N$  an approximation of  $\tilde{u}_M^R(\cdot; \xi_j)$ , which is solution to following variational equations

$$\begin{aligned} & \left\langle \tilde{u}_M^{R,N}(\cdot; \xi_j), v_{\xi_j}^N \right\rangle_{H^1(\Omega_0^R)} - \left\langle \mathcal{L}_{\xi_j}^R((\mu_p^{-1} - 1)\nabla \tilde{u}_M^{R,N}(\cdot; \xi_j)), v_{\xi_j}^N \right\rangle_{H^1(\Omega_0^R)} \\ & - \left\langle \mathcal{L}_{\xi_j}^R((\mu^{-1} - \mu_p^{-1})(\Delta \xi \sum_{\ell} \nabla u_M^{R,N}(\cdot; \xi_{\ell}))), v_{\xi_j}^N \right\rangle_{H^1(\Omega_0^R)} \end{aligned} \quad (5.36)$$

$$\begin{aligned} & - \left\langle \mathcal{V}_{\xi_j}^R(k^2(n_p - 1)\tilde{u}_M^{R,N}(\cdot; \xi_j)), v_{\xi_j}^N \right\rangle_{H^1(\Omega_0^R)} - \left\langle \mathcal{V}_{\xi_j}^R(k^2(n - n_p)(\Delta \xi \sum_{\ell} u_M^{R,N}(\cdot; \xi_{\ell}))), v_{\xi_j}^N \right\rangle_{H^1(\Omega_0^R)} \\ & = - \left\langle \mathcal{V}_{\xi_j}^R(f_{\xi_j}), v_{\xi_j}^N \right\rangle_{H^1(\Omega_0^R)}, \quad \forall v_{\xi_j}^N \in \mathcal{T}_{\xi_j}^N(\Omega_0^R). \end{aligned} \quad (5.37)$$

where  $u_M^{R,N}$  is defined as

$$u_M^{R,N} := \frac{1}{\llbracket M \rrbracket} \sum_{j \in \mathbb{Z}_M^{d-1}} \tilde{u}_M^{R,N}(\cdot; \xi_j). \quad (5.38)$$

We now define  $\mathcal{T}_{\#, \xi_j}^N$  as the space of functions that are extensions by  $\xi_j$ -quasi-periodicity of functions in  $\mathcal{T}_{\xi_j}^N$  to the domain  $\Omega_M^R$ . We then set

$$\mathcal{T}_{\#,M}^N := \bigoplus_{j \in \mathbb{Z}_M^{d-1}} \mathcal{T}_{\#, \xi_j}^N.$$

We extend  $\tilde{u}_M^{R,N}(\cdot; \xi_j)$  by  $\xi_j$ -quasi-periodicity to  $\Omega_M^R$ , which then can be considered as element of  $\mathcal{T}_{\#, \xi_j}^N$ . The function  $u_M^{R,N}$  is then extended to  $\Omega_M^R$  using (5.38). We now prove that  $u_M^{R,N}$  is also an approximation of the solution  $u_M^R$  of equation (5.30) using the Fourier basis on  $L_{\#}^2(\Omega_M^R)$ . Using the discrete Fourier space  $\mathcal{T}_{\#,M}^N$  based on the basic functions  $\varphi_{\#,M}^J$ ,  $J \in \mathbb{Z}_{M,N}^2$  defined in (3.90), we then can prove the following.

**Theorem 5.2.12.** *The function  $u_M^{R,N}$  defined in (5.38) belongs to  $\mathcal{T}_{\#,M}^N$  and verifies*

$$\begin{aligned} & \left\langle u_M^{R,N}, v_M^N \right\rangle_{H^1(\Omega_M^R)} - \left\langle \mathcal{L}_{\#,M}^R((\mu^{-1} - 1)\nabla u_M^{R,N}), v_M^N \right\rangle_{H^1(\Omega_M^R)} \\ & - \left\langle \mathcal{V}_{\#,M}^R(k^2(n-1)u_M^{R,N}), v_M^N \right\rangle_{H^1(\Omega_M^R)} = - \left\langle \mathcal{V}_{\#,M}^R f_M, v_M^N \right\rangle_{H^1(\Omega_M^R)}, \forall v_M^N \in \mathcal{T}_{\#,M}^N. \end{aligned} \quad (5.39)$$

### Convergence analysis for spatial discretization

The convergence analysis is based on the interpretation of our discrete system (5.36) as the one obtained with the discretization of the periodic problem with period =  $ML$  (Theorem 5.2.12). Indeed the following convergence results for  $u_M^R$  implies uniform convergence results for each  $\xi_\ell$ -quasi-periodic component. The latter was not guaranteed by the analysis in Section 5.2.1.

**Theorem 5.2.13.** *Let  $u_M^R \in H^1(\Omega_M)$  and  $u_M^{R,N} \in \mathcal{T}_{\#,M}^N$  be solutions to (5.30) and (5.38) respectively. Then for all positive  $M \in \mathbb{N}^{d-1}$ , there exists  $N_0(M) \in \mathbb{N}^d$  such that for all  $N \in \mathbb{N}^d$ ,  $N > N_0$ ,  $u_M^{R,N}$  is uniquely defined and there exists a positive constant  $C$  independent from  $M$  and  $N$  such that*

$$\|u_M^R - u_M^{R,N}\|_{H^1(\Omega_M^R)} \leq C \inf_{v_M^N \in \mathcal{T}_{\#,M}^N} \|u_M^R - v_M^N\|_{H^1(\Omega_M^R)}. \quad (5.40)$$

*Proof.* The proof of this theorem can be done similar to the proof of Theorem 3.4.9 using in addition the decomposition of  $\mathcal{A}_{\#,M}^R$  given by (5.35).  $\square$

It is now possible to immediately deduce the convergence rate thanks to the regularity result of Corollary 5.2.9 and using the interpolation result of Lemma 3.3.8 with  $\xi = 0$ ,  $L$  replaced with  $ML$  and  $N_\ell$  replaced with  $M_\ell N_\ell$  ( $\ell = 1, d-1$ ).

**Theorem 5.2.14.** *Assume that  $\mu$  and  $n$  satisfy Assumption 5.2.1. Let  $u_M^R \in H^1(\Omega_M)$  and  $u_M^{R,N} \in \mathcal{T}_{\#,M}^N$  be verifying (5.30) and (5.38) respectively. Then for all positive  $M \in \mathbb{N}^{d-1}$ , there exists  $N_0(M) \in \mathbb{N}^d$  and a positive constant  $C$  independent from  $M$  such that for all  $N \in \mathbb{N}^d$ ,  $N > N_0$ ,*

$$\|u_M^R - u_M^{R,N}\|_{H^1(\Omega_M^R)} \leq \left( \sum_{\ell=1}^d \frac{1}{(1+N_\ell^2)} \right)^{1/2} \|u\|_{H^2(\mathbb{R}^d)} \leq C \left( \sum_{\ell=1}^d \frac{1}{(1+N_\ell^2)} \right)^{1/2} \|f_M\|_{L^2(\Omega_M^h)}. \quad (5.41)$$

### A convergence result for a combined discretization

We now recall how one deduces a convergence result for a full discretization of the problem. Let  $f \in L^2(\mathbb{R}^d)$  with support in  $\Omega^h$  and let  $u \in H^1(\mathbb{R}^d)$  be the solution to (5.1) and  $u_M^{R,N} \in \mathcal{T}_{\#,M}^N$  be verifying (5.38). We shall assume that there exists  $\tau^* > 0$  such that

$$e^{\tau^* \|x\|} f \in L^2(\mathbb{R}^d). \quad (5.42)$$

To obtain the following convergence result, we first write the decomposition

$$u - u_M^{R,N} = (u - u_M) + (u_M - u_M^{R,N}).$$

We apply Theorem 5.2.8 to treat the term  $u - u_M$  and for the second term we first use Proposition 5.2.9 to see that

$$u_M - u_M^{R,N} = u_M^R - u_M^{R,N} \text{ in } \Omega_M^h,$$

then apply Theorem 5.2.14.

**Theorem 5.2.15.** *Assume that  $\mu$  and  $n$  satisfy Assumption 5.2.1 and that  $f$  satisfies (5.42). For all positive  $M \in \mathbb{N}^{d-1}$ , there exists  $N_0(M) \in \mathbb{N}^d$  and two positive constants  $C$  and  $\tau_0$  independent from  $M$  such that for all  $N \in \mathbb{N}^d$ ,  $N > N_0$ ,*

$$\|u - u_M^{R,N}\|_{H^1(\Omega_M^h)} \leq C \left( e^{-\tau_0 \min(ML)/2} \|e^{\tau_0 \|x\|} f\|_{L^2(\mathbb{R}^d)} + \|f - f_M\|_{L^2(\Omega_M)} + \left( \sum_{\ell=1}^d \frac{1}{(1+N_\ell^2)} \right)^{1/2} \|f\|_{L^2(\mathbb{R}^d)} \right).$$

We recall that we used the fact that  $\|f_M\|_{L^2(\Omega_M)} \leq \|f\|_{L^2(\mathbb{R}^d)}$  (in the cases where  $f$  has a compact support, assumption (5.42) is automatically verified and  $f_M = f$  in  $\Omega_M$  as soon as the support of  $f$  is contained in  $\Omega_M$ ). Let us also note that the rate of convergence with respect to  $N$  can be higher if additional smoothness on the coefficients and the source term is imposed. We refer to Theorem 3.4.12 for details on how to obtain higher order convergence results with more regular coefficients.

### 5.3 The sampling methods for the TM-mode

We here present an outline of possible extensions of the results of Chapter 4 to the TM-mode. We consider the semi-discretized version in the Floquet-Bloch variable of (5.1) with  $M$  discretization points which is then solution of

$$\begin{cases} \operatorname{div}(\mu^{-1} \nabla u) + k^2 n u = 0 & \text{in } \mathbb{R}^d, \quad d = 2, 3 \\ u \text{ is } ML\text{-periodic} \end{cases} \quad (5.43)$$

where here the wavenumber  $k$  is positive and real valued and the parameters  $\mu$  and  $n$  satisfy Assumption 5.2.1. We assume in addition that there exists  $h > 0$  such that  $\mu = 1$  and  $\mu^{-1} = 1$  for  $|x_d| > h$ . Indeed, thanks to the  $ML$ -periodicity, solving equation (5.43) in  $\mathbb{R}^d$  is equivalent to solving it in one period  $\Omega_M$ . We denote by  $D$  (respectively  $D_p$ ) a bounded domain composed by simply connected components and such that  $\mu^{-1} = 1$  and  $n = 1$  (respectively  $\mu_p^{-1} = 1$  and  $n_p = 1$ ) outside  $D$  (respectively  $D_p$ ) and  $D = D_p \cup \bar{\omega}$ . We consider down-to-up or up-to-down incident plane waves of the form:

$$u^i(x) = e^{i\alpha_{\#}(j)\bar{x} \pm i\beta_{\#}(j)x_d}$$

where

$$\alpha_{\#}(j) := i\frac{2\pi}{ML}j \quad \text{and} \quad \beta_{\#}(j) := \sqrt{k^2 - \alpha_{\#}^2(j)}, \quad \operatorname{Im}(\beta_{\#}(j)) \geq 0, \quad j \in \mathbb{Z}$$

and where  $x = (\bar{x}, x_d) \in \mathbb{R}^{d-1} \times \mathbb{R}$ . Then the scattered field  $u^s = u - u^i$  verifies

$$\begin{cases} \operatorname{div}(\mu^{-1} \nabla u^s) + k^2 n u^s = -\operatorname{div}(q \nabla u^i) - k^2 p u^i & \text{in } \mathbb{R}^d, \\ u^s \text{ is } ML\text{-periodic} \end{cases} \quad (5.44)$$

where  $q$  and  $p$  are the contrasts defined by

$$q := \mu^{-1} - 1 \quad \text{and} \quad p := n - 1.$$



We impose that  $u^s$  satisfies the Rayleigh radiation condition as defined in (4.3). Using the Dirichlet-to-Neumann operators  $T^\pm$  defined by (4.5), we reformulate the scattering problem as: Find  $u^s \in H_{\#}^1(\Omega_M^h)$  such that (remark that  $\mu^{-1} = 1$  on  $\Gamma_{\pm h}$ )

$$\begin{aligned} \operatorname{div}(\mu^{-1}\nabla u^s) + k^2 n u^s &= -\operatorname{div}(q\nabla u^i) - k^2 p u^i \quad \text{in } \Omega_M^h \\ \frac{\partial u^s}{\partial x_d} &= \pm T^\pm(u^s) \quad \text{for } x_d = \pm h. \end{aligned} \quad (5.45)$$

Multiplying equation (5.45) with  $v \in H_{\#}^1(\Omega_M^h)$  and using the Green theorem we arrive at the variational formulation of the problem as

$$\int_{\Omega_M^h} \mu^{-1}\nabla u^s \cdot \nabla \bar{v} - k^2 n u^s \bar{v} \, dx - \langle T^+(u^s), v \rangle - \langle T^-(u^s), v \rangle = \int_{\Omega_M^h} -q\nabla u^i \cdot \nabla \bar{v} + k^2 p u^i \bar{v} \, dx \quad (5.46)$$

for all  $v \in H_{\#}^1(\Omega_M^h)$ . Problem 5.46 is of Fredholm type since the sesquilinear form

$$\mathcal{A}(u^s, v) := \int_{\Omega_M^h} \mu^{-1}\nabla u^s \cdot \nabla \bar{v} - k^2 n u^s \bar{v} \, dx - \langle T^+(u^s), v \rangle - \langle T^-(u^s), v \rangle$$

is continuous on  $H_{\#}^1(\Omega_M^h) \times H_{\#}^1(\Omega_M^h)$  and satisfies a Garding type inequality. We shall assume that the parameters  $\mu$  and  $n$  are such that Problem (5.46) is well posed. More precisely, let  $f = (f_1, f_2) \in L^2(\Omega_M^h)^d \times L^2(\Omega_M^h)$  and consider the following variational problem: Find  $w \in H_{\#}^1(\Omega_M^h)$  such that for all  $v \in H_{\#}^1(\Omega_M^h)$ ,

$$\int_{\Omega_M^h} \mu^{-1}\nabla w \cdot \nabla \bar{v} - k^2 n w \bar{v} \, dx - \langle T^+(w), v \rangle - \langle T^-(w), v \rangle = \int_{\Omega_M^h} -q f_1 \cdot \nabla \bar{v} + k^2 p f_2 \bar{v} \, dx. \quad (5.47)$$

Then we make the following assumption:

**Assumption 5.3.1.** *We assume that  $\mu$ ,  $n$  and  $k$  are such that problem (5.47) is well posed for all  $f = (f_1, f_2) \in L^2(\Omega_M^h)^d \times L^2(\Omega_M^h)$ .*

We remark that the solution  $w \in H_{\#}^1(\Omega_M^h)$  of (5.47) can be extended to a function in  $\Omega_M$  satisfying  $\operatorname{div}(\mu^{-1}\nabla w) + k^2 n w = -\operatorname{div}(q f_1) - k^2 p f_2$  using the Rayleigh expansion (4.3). Let  $G_M$  be the  $ML$ -periodic Green function satisfying  $\Delta G_M + k^2 G_M = -\delta$  in  $\Omega_M$  and the Rayleigh radiation condition. Then  $w$  can also be represented as

$$w(x) = \operatorname{div} \int_D G_M(x-y) q(y) (f_1 + \nabla w)(y) \, dy + k^2 \int_D G_M(x-y) p(y) (f_2 + w)(y) \, dy. \quad (5.48)$$

We recall that  $G_M$  can be expressed as (4.12)

$$G_M(x) = \frac{i}{2ML} \sum_{\ell \in \mathbb{Z}} \frac{1}{\beta_{\#}(\ell)} e^{i\alpha_{\#}(\ell)\bar{x} + i\beta_{\#}(\ell)|x_d|}. \quad (5.49)$$

### 5.3.1 Setting of the inverse problem and definition of the sampling operator

We still use as incidents waves all down-to-up (scaled) incident plane waves  $u^{i,+}(x; j)$  (res. up-to-down incident plane waves  $u^{i,-}(x; j)$ ) defined by

$$u^{i,+}(x; j) = \frac{-i}{2\bar{\beta}_{\#}(j)} e^{i\alpha_{\#}(j)\bar{x} + i\bar{\beta}_{\#}(j)(x_d-h)}, \quad j \in \mathbb{Z}. \quad (5.50)$$

$$\left( \text{res. } u^{i,-}(x; j) = \frac{-i}{2\bar{\beta}_{\#}(j)} e^{i\alpha_{\#}(j)\bar{x} - i\bar{\beta}_{\#}(j)(x_d+h)}, \quad j \in \mathbb{Z} \right) \quad (5.51)$$

Our measurements (data for the inverse problem) will be formed by the Rayleigh sequences (see (4.4))

$$\widehat{u}^{s+}(\ell; j), \quad (j, \ell) \in \mathbb{Z} \times \mathbb{Z}, \quad \left( \text{res. } \widehat{u}^{s-}(\ell; j), \quad (j, \ell) \in \mathbb{Z} \times \mathbb{Z}. \right)$$

where  $j$  is related to the incident wave index and  $\ell$  is related to the Rayleigh sequence index.

Let us consider the (Herglotz) operators  $H^+ : \ell^2(\mathbb{Z}) \rightarrow L^2(D)^d \times L^2(D)$  and  $H^- : \ell^2(\mathbb{Z}) \rightarrow L^2(D)^d \times L^2(D)$  defined by

$$H^{\pm}a := \left( \sum_{j \in \mathbb{Z}} a(j) \nabla u^{i,\pm}(\cdot; j)|_D, \sum_{j \in \mathbb{Z}} a(j) u^{i,\pm}(\cdot; j)|_D \right), \quad \forall a = \{a(j)\}_{j \in \mathbb{Z}} \in \ell^2(\mathbb{Z}). \quad (5.52)$$

Then  $H^{\pm}$  is compact, injective (will be proved later) and its adjoint  $(H^{\pm})^* : L^2(D)^d \times L^2(D) \rightarrow \ell^2(\mathbb{Z})$  is given by

$$(H^{\pm})^* \varphi := \{\widehat{\varphi}^{\pm}(j)\}_{j \in \mathbb{Z}}, \quad \forall \varphi = (\varphi_1, \varphi_2) \in L^2(D)^d \times L^2(D), \quad (5.53)$$

where

$$\widehat{\varphi}_j^{\pm} := \int_D \left( \varphi_1(x) \cdot \nabla \overline{u^{i,\pm}(\cdot; j)}(x) + \varphi_2(x) \cdot \overline{u^{i,\pm}(\cdot; j)}(x) \right) dx.$$

Let us denote by  $H_{\text{inc}}^{\pm}(D)$  the closure of the range of  $H^{\pm}$  in  $L^2(D)^d \times L^2(D)$ . We then consider the (compact) operator  $G^{\pm} : H_{\text{inc}}^{\pm}(D) \rightarrow \ell^2(\mathbb{Z})$  defined by

$$G^{\pm}(f) := \{\widehat{w}^{\pm}(\ell)\}_{\ell \in \mathbb{Z}}, \quad (5.54)$$

where  $\{\widehat{w}^{\pm}(\ell)\}_{\ell \in \mathbb{Z}}$  is the Rayleigh sequence of  $w \in H_{\#}^1(\Omega_M^h)$  the solution of (5.47). We now define the sampling operators  $N^{\pm} : \ell^2(\mathbb{Z}) \rightarrow \ell^2(\mathbb{Z})$  by

$$N^{\pm}(a) = G^{\pm} H^{\pm}(a). \quad (5.55)$$

By linearity of the operators  $G^{\pm}$  and  $H^{\pm}$  we also get an equivalent definition of  $N^{\pm}$  as

$$[N^{\pm}(a)]_{\ell} = \sum_{j \in \mathbb{Z}} a(j) \widehat{u}^{s\pm}(\ell; j) \quad \ell \in \mathbb{Z}. \quad (5.56)$$

Let us introduce the operator  $T : L^2(D)^d \times L^2(D) \rightarrow L^2(D)^d \times L^2(D)$  defined by

$$Tf := \left( -q(f_1 + \nabla w|_D), k^2 p(f_2 + w|_D) \right), \quad \forall f = (f_1, f_2) \in L^2(D)^d \times L^2(D) \quad (5.57)$$

with  $w$  being the solution of (5.47). We then have the following:

**Lemma 5.3.2.** *The operators  $G^\pm$  defined by (5.54) can be factorized as*

$$G^\pm = (H^\pm)^* T.$$

*Proof.* Let  $f = (f_1, f_2) \in L^2(D)^d \times L^2(D)$  and  $w \in H_{\#}^1(\Omega_M^h)$  be solution to (5.47). Let us write  $T_1(f) := -q(f_1 + \nabla w|_D)$  and  $T_2(f) := k^2 p(f_2 + w|_D)$ . Then, by definition of the Rayleigh coefficients and combining with the representation of  $G_M$  in (4.12) and the writing of  $w$  as in (5.48) we have

$$\begin{aligned} \widehat{w}^\pm(j) &= \frac{1}{2ML} \int_{x_d=\pm h} e^{-i\alpha_{\#}(j)\bar{x}} \int_D \sum_{\ell \in \mathbb{Z}} \frac{\alpha_{\#}(\ell)}{\beta_{\#}(\ell)} e^{i\alpha_{\#}(\ell)\cdot(\bar{x}-\bar{y}) + i\beta_{\#}(\ell)|h \mp y_d|} \cdot (T_1(f)(y)) \, dy \, d\bar{x} \\ &\quad + \frac{i}{2ML} \int_{x_d=\pm h} e^{-i\alpha_{\#}(j)\bar{x}} \int_D \sum_{\ell \in \mathbb{Z}} \frac{1}{\beta_{\#}(\ell)} e^{i\alpha_{\#}(\ell)(\bar{x}-\bar{y}) + i\beta_{\#}(\ell)|h \mp y_d|} T_2(f)(y) \, dy \, d\bar{x} \\ &= \int_D \frac{\alpha_{\#}(j)e^{i\beta_{\#}(j)h}}{2\beta_{\#}(j)} e^{-i\alpha_{\#}(j)\bar{y} \mp i\beta_{\#}(j)y_d} \cdot T_1(f)(y) \, dy + \int_D \frac{ie^{i\beta_{\#}(j)h}}{2\beta_{\#}(j)} e^{-i\alpha_{\#}(j)\bar{y} \mp i\beta_{\#}(j)y_d} T_2 f(y) \, dy \end{aligned}$$

Observing that

$$\frac{\alpha_{\#}(j)e^{i\beta_{\#}(j)h}}{2\beta_{\#}(j)} e^{-i\alpha_{\#}(j)\cdot dy \mp i\beta_{\#}(j)y_d} = \nabla \overline{u^{i,\pm}(y; j)} \quad \text{and} \quad \frac{ie^{i\beta_{\#}(j)h}}{2\beta_{\#}(j)} e^{-i\alpha_{\#}(j)y_1 \mp i\beta_{\#}(j)y_2} = \overline{u^{i,\pm}(y; j)}$$

we then have

$$\widehat{w}^\pm(j) = \int_D T_1 f(y) \cdot \nabla \overline{u^{i,\pm}(y; j)} + T_2 f(y) \overline{u^{i,\pm}(y; j)} \, dy,$$

which proves the lemma.  $\square$

Using Lemma 5.3.2 we end up with

$$N^\pm = (H^\pm)^* T H^\pm. \quad (5.58)$$

The justification of the Sampling Methods uses the solvability of the interior transmission problem defined in the present case as: Seek  $(u, v) \in H^1(D) \times H^1(D)$  such that

$$\begin{cases} \operatorname{div}(\mu^{-1} \nabla u) + k^2 n u = 0 & \text{in } D, \\ \Delta v + k^2 v = 0 & \text{in } D, \\ u - v = \varphi & \text{on } \partial D, \\ \mu^{-1} \frac{\partial u}{\partial \nu} - \mu^{-1} \frac{\partial v}{\partial \nu} = \psi & \text{on } \partial D, \end{cases} \quad (5.59)$$

for given  $(\varphi, \psi) \in H^{1/2}(\partial D) \times H^{-1/2}(\partial D)$  where  $\nu$  denotes the outward normal on  $\partial D$ . Values of  $k$  for which this problem is not well posed are referred to as transmission eigenvalues. For our purpose we shall assume that this problem is well posed.

**Assumption 5.3.3.** *We assume that the parameters  $\mu$ ,  $n$  and the real wave number  $k$  are such that (5.59) defines a well posed problem.*

Remark that we here specified that the solution of the interior transmission problem is in  $H^1(D) \times H^1(D)$  which is the natural space of solutions in the cases  $\mu^{-1} \neq 1$  in a neighborhood of  $\partial D$  [21].

### 5.3.2 Some key properties of the introduced operators

We still keep the assumption (that is not essential but simplifies some of the arguments, and justifies the use of  $N^+$  or  $N^-$  and not both of them)

$$\Omega_M \setminus D \text{ is connected.}$$

The first step towards the justification of the sampling methods is the characterization of the closure of the range of  $H^\pm$ .

**Lemma 5.3.4.** *The operator  $H^\pm$  is compact and injective. Let  $H_{\text{inc}}^\pm(D)$  be the closure of the range of  $H^\pm$  in  $L^2(D)^d \times L^2(D)$ . Then*

$$H_{\text{inc}}^\pm(D) = H_{\text{inc}}(D) := \{(\varphi_1, \varphi_2) = (\nabla v, v) \mid v \in H^1(D) : \Delta v + k^2 v = 0 \text{ in } D\}. \quad (5.60)$$

*Proof.* The compactness and the injectivity of the operators  $H$  directly follow from Lemma 4.3.3. We now prove identity (5.60). Let us denote by

$$\tilde{H}_{\text{inc}}(D) := \{(\varphi_1, \varphi_2) = (\nabla v, v) \mid v \in H^1(D) : \Delta v + k^2 v = 0 \text{ in } D\}. \quad (5.61)$$

We see that  $H_{\text{inc}}(D) \subset \tilde{H}_{\text{inc}}(D)$ . Then to prove the lemma, it remains to prove that  $(H^\pm)^*$  is injective in  $\tilde{H}_{\text{inc}}(D)$ . Let  $\varphi = (\nabla f, f) \in \tilde{H}_{\text{inc}}(D)$  and assume that  $(H^\pm)^*(\varphi) = 0$ . We set,

$$u = \text{div} \int_D G_M(x-y)(-\nabla f(y)) + \int_D G_M(x-y)f(y) \, dy, \quad x \in \mathbb{R}^d \quad (5.62)$$

Then  $u \in H_{\#}^1(\Omega_M^h)$  verifies

$$\int_D -\nabla u \cdot \nabla \bar{v} + k^2 u \bar{v} = - \int_D \nabla f \cdot \nabla \bar{v} + f \bar{v} \, dx \quad (5.63)$$

for all  $v \in H_{\#}^1(\Omega_M^h)$ . From the definition of  $u$  we have that

$$\hat{u}^+(j) = ((H^+)^*(\varphi))(j). \quad (5.64)$$

This implies from the assumption  $((\mathcal{H}^+)^*(\varphi))(j) = 0$  that  $\hat{u}^+(j) = 0$  for all  $j \in \mathbb{Z}$ , hence  $u = 0$  for all  $x_d > h$ . By the unique continuation principal and since  $\Omega_M \setminus D$  connected we have that  $u = 0$  in  $\Omega_M \setminus D$ . The regularity of  $u \in H_{\#, \text{loc}}^1(\Omega_M)$  then implies that  $u \in H_0^1(D)$ . Equation (5.63) then gives

$$\int_D -\nabla u \cdot \nabla \bar{v} + k^2 u \bar{v} = - \int_D \nabla f \cdot \nabla v + f v \, dx \quad (5.65)$$

for all  $v \in H^1(D)$ . Taking  $v = \bar{f}$  we then have

$$\|f\|_{H^1(D)}^2 = - \int_D -\nabla u \cdot \nabla \bar{f} + k^2 u \bar{f} \, dx = 0. \quad (5.66)$$

The last equation comes from  $\Delta f + k^2 f = 0$  and  $u \in H_0^1(D)$ .  $\square$

The following reciprocity lemma will be useful to relate injectivity to the denseness of the range of the operators  $G^\pm$ .

**Lemma 5.3.5.** *Let  $f = (f_1, f_2)$ ,  $\tilde{f} = (\tilde{f}_1, \tilde{f}_2) \in L^2(D)^d \times L^2(D)$  and let  $w, \tilde{w} \in H_{\#}^1(\Omega_M^h)$  be the corresponding solutions satisfying (5.47) with  $f = f, \tilde{f} = \tilde{f}$  respectively. Then*

$$\int_D q \nabla w \cdot \tilde{f}_1 - k^2 p w \tilde{f}_2 dx = \int_D q \nabla \tilde{w} \cdot f_1 - k^2 p \tilde{w} f_2 dx. \quad (5.67)$$

*Proof.* Taking  $v = w$  and  $v = \tilde{w}$  in the variational formulation (5.47) satisfied by  $\tilde{w}$  and  $w$  respectively then taking the difference yields (using the assumption that  $\text{supp}(f) \subset \Omega_M^h$ )

$$\begin{aligned} \int_{\Gamma_M^h} \frac{\partial w}{\partial x_2} \tilde{w} - \frac{\partial \tilde{w}}{\partial x_2} w ds(x) - \int_{\Gamma_M^{-h}} \frac{\partial w}{\partial x_2} \tilde{w} - \frac{\partial \tilde{w}}{\partial x_2} w ds(x) \\ = \int_D (q f_1 \cdot \nabla \tilde{w} - k^2 p f_2 \tilde{w}) - (q \tilde{f}_1 \cdot \nabla w - p \tilde{f}_2 w) dx. \end{aligned} \quad (5.68)$$

Obviously  $T^{\pm}$  are symmetric, since (using Parseval's theorem)

$$\pm \int_{\Gamma_M^{\pm h}} \frac{\partial w_0}{\partial x_2} w_1 ds(x) = \sum_{\ell \in \mathbb{Z}} i \beta_{\#}(\ell) \widehat{w_0}^{\pm}(\ell) \widehat{w_1}^{\pm}(-\ell) \quad (5.69)$$

and  $\beta_{\#}(\ell) = \beta_{\#}(-\ell)$  for all  $\ell \in \mathbb{Z}$ . Therefore,

$$\pm \int_{\Gamma_M^{\pm h}} \left( \frac{\partial w_0}{\partial x_2} w_1 - \frac{\partial w_1}{\partial x_2} w_0 \right) ds(x) = 0,$$

which proves the lemma.  $\square$

We now prove one of the main ingredients for the justification of the inversion methods discussed below.

**Theorem 5.3.6.** *Assume that Assumptions 5.3.1 and 5.3.3 hold. Then the operator  $G^{\pm} : H_{\text{inc}}(D) \rightarrow \ell^2(\mathbb{Z})$  defined by (5.54) is injective with dense range. Moreover,  $\widehat{\Phi}^{\pm}(\cdot; z)$  belongs to  $\mathcal{R}(G^{\pm})$  if and only if  $z \in D$  and  $\widehat{\Phi}_q^{\pm}(\cdot; z)$  belongs to  $\mathcal{R}(G^{\pm})$  if and only if  $z \in D_p$ , where  $q$  is a fixed parameter in  $\mathbb{Z}_M$ .*

*Proof.* We only give here the proof of the theorem for the operator  $G^+$  since the proof for the operator  $G^-$  is similar. We start by proving that  $G^+ : H_{\text{inc}}(D) \rightarrow \ell^2(\mathbb{Z})$  is injective with dense range. Let  $\varphi = (\nabla f, f) \in H_{\text{inc}}(D)$  and let  $w \in H_{\#}^1(\Omega_M^h)$  be the associated scattered field via (5.47). Assume that  $G^+(f) = 0$ . Then  $w = 0$  for  $x_d > h$ . By unique continuation principle we then deduce that

$$w = 0 \quad \text{in } \Omega_M \setminus D,$$

and therefore  $w \in H_0^1(D)$ . We now set,  $u := w + f$ , then the pair  $(u, f)$  is a solution to (5.59) with zero data. Assumption 5.3.3 then ensures that  $f = 0$ , which proves the injectivity of  $G^+$ . We now prove the denseness of the range of  $G^+$ . Let  $g \in \overline{\mathcal{R}(G^+)}^{\perp}$ . Then

$$(G^+(\varphi), g)_{\ell^2(\mathbb{Z})} = 0, \quad \forall \varphi = (\nabla f, f) \in H_{\text{inc}}(D).$$

Consider  $\varphi$  of the form  $\varphi = \overline{H^+(a)}$  for some  $a \in \ell^2(\mathbb{Z})$ . Since  $G^+ = (H^+)^* T$ , we then have

$$\langle T(\overline{H^+(a)}), H^+(g) \rangle_{L^2(D)^d} = 0, \quad \forall a \in \ell^2(\mathbb{Z}). \quad (5.70)$$

In this proof, we denote  $H_1^+ : a \mapsto \sum_j a_j \nabla u^{i,+}(\cdot; j)$  and  $H_2^+ : a \mapsto \sum_j a_j u^{i,+}(\cdot; j)$ . Then,  $H^+ = (H_1^+, H_2^+)$ . Let  $w(a)$  and  $w(g)$  solution to (5.47) with respect to  $\overline{H^+(a)}$  and  $\overline{H^+(g)}$ . From Lemma 5.3.5 we get

$$\begin{aligned} \left( T(\overline{H^+(a)}), H^+(g) \right)_{L^2(D)^d \times L^2(D)} &= \int_D q(\overline{H_1^+ a} + \nabla w(a)) \cdot \overline{H_1^+ g} \, dx + k^2 \int_D p(\overline{H_2^+ a} + w(a)) \overline{H_2^+ g} \, dx \\ &= \int_D q(\overline{H_1^+ g} + \nabla w(g)) \cdot \overline{H_1^+ g} + k^2 \int_D p(\overline{H_2^+ g} + w(g)) \overline{H_2^+ a} \, dx \, dx. \end{aligned}$$

Therefore,

$$\left( T(\overline{H^+(g)}), H^+(a) \right)_{L^2(D)} = \left( T(\overline{H^+(a)}), H^+(g) \right)_{L^2(D)}, \quad \forall a \in \ell^2(\mathbb{Z}). \quad (5.71)$$

We deduce from (5.102) that

$$\left( G^+(\overline{H^+(g)}), a \right)_{\ell^2(\mathbb{Z})} = 0, \quad \forall a \in \ell^2(\mathbb{Z}),$$

which implies  $G^+(\overline{H^+(g)}) = 0$ . The injectivity of  $G^+$  ensures that  $H^+g = 0$  in  $D$  and consequently  $g = 0$  (by Lemma 5.3.4). This proves the denseness of the range of  $G^+$ .

We now prove that  $\widehat{\Phi}^\pm(\cdot; z)$  belongs to  $\mathcal{R}(G^\pm)$  if and only if  $z \in D$ . We first observe that  $\widehat{\Phi}^+(\cdot; z)$  is the Rayleigh sequence of  $\Phi(\cdot; z)$  satisfying  $\Delta \Phi(\cdot; z) + k^2 \Phi(\cdot; z) = -\delta_z$  in  $\Omega_M$  and the Rayleigh radiation condition. Let  $z \in D$ . We consider  $(u, v) \in H^1(D) \times H^1(D)$  as being the solution to (5.59) with

$$\varphi(x) = \Phi(x; z) \text{ and } \psi(x) = \partial \Phi(x; z) / \partial \nu(x) \text{ for } x \in \partial D. \quad (5.72)$$

We then define  $w$  by

$$\begin{aligned} w(x) &= u(x) - v(x) \quad \text{in } D, \\ w(x) &= \Phi(x; z) \quad \text{in } \Omega_M \setminus D. \end{aligned}$$

Due to (5.72), we have that  $w \in H_{\#, \text{loc}}^1(\Omega_M)$  and satisfies (5.47). Hence  $G^+v = \widehat{\Phi}^+(\cdot; z)$ .

Now let  $z \in \Omega_M \setminus D$ . Assume that there exists  $\varphi = (\nabla f, f) \in H_{\text{inc}}(D)$  such that  $G^+\varphi = \widehat{\Phi}^+(\cdot; z)$ . This implies that  $w = \Phi(\cdot; z)$  in  $\{x \in \Omega_M, \pm x_d \geq h\}$  where  $w$  is the solution to (5.47). By the unique continuation principle we deduce that  $w = \Phi(\cdot; z)$  in  $\Omega_M \setminus D$ . This gives a contradiction since  $w \in H_{\#, \text{loc}}^1(\Omega_M \setminus D)$  while  $\Phi(\cdot; z) \notin H_{\#, \text{loc}}^1(\Omega_M \setminus D)$ .

The proof of the statement  $\widehat{\Phi}_q^\pm(\cdot; z)$  belongs to  $\mathcal{R}(G^\pm)$  if and only if  $z \in D_p$  follows the same lines as above replacing  $\Phi(\cdot; z)$  by  $\Phi_q(\cdot; z)$ . The reader can also refer to the proof of Lemma 4.4.6.  $\square$

**Lemma 5.3.7.** *Assume that Assumptions (5.2.1) and (5.3.3) hold. Then the operator  $T : L^2(D) \rightarrow L^2(D)$  defined by (5.57) satisfies*

$$\text{Im}(T\phi, \phi) \geq 0, \quad \forall \phi \in H_{\text{inc}}(D). \quad (5.73)$$

*Assume in addition that  $q(x) \geq \sigma_n > 0$ ,  $x \in D$  (respectively  $-q(x) \geq \sigma_n > 0$ ,  $x \in D$ ) and assume that either  $p^{-1} \in L^\infty(D)$  or  $k$  is not a transmission eigenvalue. Then  $-\text{Re} T = T_0 + T_1$ , where  $T_0$  (respectively  $-T_0$ ) is self-adjoint and coercive and  $T_1$  is compact on  $H_{\text{inc}}(D)$ . Moreover,  $T$  is injective on  $H_{\text{inc}}(D)$ .*

*Proof.* Let  $\varphi = (\varphi_1, \varphi_2) \in L^2(D)^d \times L^2(D)$  and  $w_\varphi$  be solution to (5.47) associated with  $f = \varphi$ . By definition of the operator  $T$  we have

$$\begin{aligned} (T\varphi, \varphi)_{L^2(D)^d \times L^2(D)} &= \int_D -q(\varphi_1 + \nabla w_\varphi) \cdot \overline{\varphi_1} + k^2 p(\varphi_2 + w_\varphi) \overline{\varphi_2} \, dx \\ &= \int_D -q|\varphi_1 + \nabla w_\varphi|^2 + k^2 p|\varphi_2 + w_\varphi|^2 \, dx + \int_D q(\nabla w_\varphi + \varphi_1) \cdot \nabla \overline{w_\varphi} - p(\varphi_2 + w_\varphi) \overline{w_\varphi} \, dx \end{aligned} \quad (5.74)$$

where, using the variational formulation (5.47)

$$\begin{aligned} \int_D q(\nabla w_\varphi + \varphi_1) \cdot \nabla \overline{w_\varphi} - p(w_\varphi + \varphi_2) \overline{w_\varphi} \, dx \\ = \langle T^+ w_\varphi, w_\varphi \rangle + \langle T^- w_\varphi, w_\varphi \rangle - \int_{\Omega_M^h} |\nabla w_\varphi|^2 - k^2 |w_\varphi|^2 \, dx. \end{aligned} \quad (5.75)$$

Therefore, substituting (5.75) into (5.74) we end up with:

$$\begin{aligned} (T\varphi, \varphi)_{L^2(D)^d \times L^2(D)} &= \int_D -q|\varphi_1 + \nabla w_\varphi|^2 + k^2 p|\varphi_2 + w_\varphi|^2 \, dx \\ &\quad - \int_{\Omega_M^h} (|\nabla w_\varphi|^2 - k^2 |w_\varphi|^2) + \langle T^+ w_\varphi, w_\varphi \rangle + \langle T^- w_\varphi, w_\varphi \rangle \end{aligned} \quad (5.76)$$

Thanks to the non-negative sign of the imaginary part of  $T^\pm$  and the assumption  $\text{Im}(n) \geq 0$  we deduce that

$$\text{Im} (T\varphi, \varphi) = \int_D \text{Im}(n) |\varphi_2 + w_\varphi|^2 \, dx + \text{Im} \langle T^+ w_\varphi, w_\varphi \rangle + \langle T^- w_\varphi, w_\varphi \rangle \geq 0.$$

For the case  $q$  positive definite on  $D$  one can define  $T_0 : L^2(D)^d \times L^2(D) \rightarrow L^2(D)^d \times L^2(D)$  by

$$(T_0\varphi, \psi)_{L^2(D)^d \times L^2(D)} := \int_D q(\varphi_1 + \nabla w_\varphi) \cdot \overline{(\psi_1 + \nabla w_\psi)} + \varphi_2 \overline{\psi_2} \, dx + \int_{\Omega_M^h} (\nabla w_\varphi \cdot \overline{\nabla w_\psi}) \, dx$$

which is indeed a selfadjoint and coercive operator. Using (5.76) one then deduces that  $-T + T_0 : H_{\text{inc}}(D) \rightarrow L^2(D)^d \times L^2(D)$  is compact by the  $H^2$  regularity outside  $D$  of  $w_\varphi$  and the Rellich compact embedding theorem. Observe that we used that the operator is restricted to  $H_{\text{inc}}(D)$  to infer compactness of the terms involving  $\varphi_2$  in the expression of  $(-T + T_0)(\varphi)$ .

For the case  $q$  negative definite on  $D$  we first observe that (5.74) and (5.75) also lead to

$$\begin{aligned} (T\varphi, \varphi)_{L^2(D)^d \times L^2(D)} &= \int_D -q|\varphi_1|^2 + \int_{\Omega_M^h} \mu^{-1} |\nabla w_\varphi|^2 + 2i \int_D -q \text{Im}(\nabla w_\varphi \cdot \varphi_1) \\ &\quad + \int_D k^2 p(\varphi_2 + w_\varphi) (\overline{\varphi_2} - \overline{w_\varphi}) \, dx - \int_{\Omega_M^h} k^2 |w_\varphi|^2 \, dx - \langle T^+ w_\varphi, w_\varphi \rangle - \langle T^- w_\varphi, w_\varphi \rangle. \end{aligned} \quad (5.77)$$

We then define  $T_0 : L^2(D)^d \times L^2(D) \rightarrow L^2(D)^d \times L^2(D)$  by

$$(T_0\varphi, \psi)_{L^2(D)^d \times L^2(D)} := \int_D -q\varphi_1 \overline{\psi_1} + \varphi_2 \overline{\psi_2} \, dx + \int_{\Omega_M^h} \mu^{-1} (\nabla w_\varphi \cdot \overline{\nabla w_\psi}) \, dx$$

which is also selfadjoint and coercive. Using (5.77) one deduces using the same arguments as in the previous case that  $T - T_0 : H_{\text{inc}}(D) \rightarrow L^2(D)^d \times L^2(D)$  is compact.

In the case  $k$  is not a transmission eigenvalue, the injectivity of  $T^+$  is implied for instance by Assumption 5.3.3 and the factorization  $G^+ = (H^+)^* T$ : Assume that  $\varphi = (\nabla f, f) \in H_{\text{inc}}(D)$  and  $T\varphi = \left( -q(\nabla f + \nabla w_\varphi), k^2 p(f + w_\varphi) \right) = 0$ . This implies, using the factorization  $G^+ = (H^+)^* T$  that  $\widehat{w}_\varphi^+(j) = 0$  for all  $j \in \mathbb{Z}$  and therefore  $w_\varphi = 0$  in  $\Omega_M \setminus D$  (by unique continuation principal). With  $\varphi = (\nabla f, f) \in H_{\text{inc}}(D)$  and  $f$  verifying  $\Delta f + k^2 f = 0$  in  $D$  we get that  $u := f + w_\varphi$  and  $v := f$  satisfying the interior transmission problem (5.59) with  $\varphi = \psi = 0$ . We then deduce that  $u = v = 0$ . This proves the injectivity of the operator  $T$ . If  $p^{-1} \in L^\infty(D)$  the injectivity obviously follows from the expression of  $T$ .  $\square$

**Lemma 5.3.8.** *Assume that Assumptions (5.3.1) and (5.3.3) hold. Then the operators  $N^\pm$  are injective with dense ranges.*

*Proof.* The injectivity and the denseness of the ranges of  $N^\pm$  directly follow from the same properties satisfied by  $H^\pm$  (Lemma 5.3.4) and  $G^\pm$  (Theorem 5.3.6).  $\square$

**Lemma 5.3.9.** *For  $z \in \Omega_M$ ,  $z \in D$  if and only if  $\widehat{\Phi}^\pm(\cdot; z)$  belongs to the range of  $(\mathcal{H}^\pm)^*$  and  $z \in D_p$  if and only if  $\widehat{\Phi}_q^\pm(\cdot; z)$  belongs to the range of  $(\mathcal{H}^\pm)^*$ , where  $q$  is a (fixed) integer between  $[-\frac{M}{2}] + 1$  and  $[\frac{M}{2}]$ .*

*Proof.* Let us consider first the proof of the first statement. For  $z \in D$  choose a cut-off function  $\rho \in C^\infty(\Omega_M)$  which vanishes near  $z$  and equals one in  $\Omega_M \setminus D$ . We define  $v(x) := \rho(x)G_{M,z}$ . Then the Rayleigh sequence of  $v(x)$  are  $\widehat{\Phi}^\pm(\cdot; z)$ . We observe that  $f := -(\Delta v + k^2 v)$  has compact support in  $D$  and  $f \in L^2(D)$ . Let  $\varphi = (\varphi_1, \varphi_2) \in L^2(D)^d \times L^2(D)$  such that  $\varphi_1 = 0$  and  $\varphi_2 = f$ . Since  $v$  satisfies the Rayleigh radiation condition, then

$$v(x) = \int_D G_M(x-y)f(y)dy \quad \left( = \operatorname{div} \int_D G_M(x-y) \cdot \varphi_1(y)dy + \int_D G_M(x-y)\varphi_2(y)dy \right) \quad (5.78)$$

We hence have from expansion (4.12) and the fact the Rayleigh sequences of  $v$  and  $\Phi(\cdot; z)$  are identical that

$$\widehat{\Phi}^\pm(\cdot; z) = (H^\pm)^* \varphi.$$

We now assume that  $z \notin D$  (without loss of generality we can assume that  $D \cup \{z\} \subset \Omega_M^h$ ) and  $\widehat{\Phi}^\pm(\cdot; z) = (H^\pm)^* \varphi$  for some  $\varphi \in L^2(D)^d \times L^2(D)$ . We also consider  $v$  which is defined by (5.78). Since  $\Phi(\cdot; z)$  and  $v$  satisfies the Rayleigh radiation condition then  $\Phi(\cdot; z) = v$  in domain  $\pm x_d > h$ . By unique continuation principal we deduce that  $\Phi(\cdot; z) = v$  in the exterior of  $D \cup \{z\}$ . This gives a contradiction since  $v$  is smooth near  $z$  but  $\Phi(\cdot; z)$  is singular at  $z$ .

The proof of the second statement uses the same arguments as before and the adaptations as in Lemma 4.4.7 replacing  $G_M(x-y)$  by  $\Phi_q(x; y)$ .  $\square$

### Theorems for the sampling methods to reconstruct $D$ and $D_p$

We now have all the ingredients for the proofs of Theorems 4.4.2-4.4.4 that we summarize hereafter.

Concerning the factorization method, let us recall the definition of the operator

$$N_{\sharp}^\pm := |\operatorname{Re}(N^\pm)| + |\operatorname{Im}(N^\pm)| \quad (5.79)$$

where  $\operatorname{Re}(N^\pm) := \frac{1}{2}(N^\pm + (N^\pm)^*)$ ,  $\operatorname{Im}(N^\pm) := \frac{1}{2i}(N^\pm - (N^\pm)^*)$ . Then we have the following theorem:



**Theorem 5.3.10.** *Under the hypothesis of Lemma 5.3.7, the following factorization holds:*

$$N_{\sharp}^{\pm} = (\mathcal{H}^{\pm})^* T_{\sharp}^{\pm} \mathcal{H}^{\pm}, \quad (5.80)$$

where  $T_{\sharp}^{\pm} : L^2(D) \rightarrow L^2(D)$  is self-adjoint and coercive on  $H_{\text{inc}}(D)$ . Moreover,  $z \in D$  if and only if  $\widehat{\Phi}^{\pm}(\cdot; z) \in \mathcal{R}\left((N_{\sharp}^{\pm})^{1/2}\right)$  and  $z \in D_p$  if and only if  $\widehat{\Phi}_q^{\pm}(\cdot; z) \in \mathcal{R}\left((N_{\sharp}^{\pm})^{1/2}\right)$  where  $q \in \mathbb{Z}_M$ .

*Proof.* The proof of this theorem is a direct application of the abstract framework given in Theorem A.1.2 using the results of Lemma 5.3.7 and Lemma 5.3.9 above.  $\square$

Denote by  $(\cdot, \cdot)$  the  $\ell^2(\mathbb{Z})$  scalar product and by  $\|\cdot\|$  the associated norm. Let  $\alpha > 0$  be a given parameter and  $\phi \in \ell^2(\mathbb{Z})$ . We introduce functional  $J_{\alpha}(\phi; \cdot) : \ell^2(\mathbb{Z}) \rightarrow \mathbb{R}$  where

$$J_{\alpha}^{\pm}(\phi; a) := \alpha(N_{\sharp}^{\pm}a, a) + \|N^{\pm}a - \phi\|^2, \quad \forall a \in \ell^2(\mathbb{Z}), \quad (5.81)$$

and define

$$j_{\alpha}(\phi) := \inf_{a \in \ell^2(\mathbb{Z})} J_{\alpha}^{\pm}(\phi; a). \quad (5.82)$$

Let  $c(\alpha) > 0$  be such that  $c(\alpha)/\alpha \rightarrow 0$  as  $\alpha \rightarrow 0$ .

**Theorem 5.3.11.** *Assume that Assumptions 5.3.1, 5.3.3 and the hypothesis of Lemma 5.3.7 hold. Let  $z \in \Omega_M$  and define  $a^{\alpha, z} \in \ell^2(\mathbb{Z})$  and  $a_q^{\alpha, z} \in \ell^2(\mathbb{Z})$  such that*

$$\begin{aligned} J_{\alpha}^{\pm}(\widehat{\Phi}^{\pm}(\cdot; z); a^{\alpha, z}) &\leq j_{\alpha}(\widehat{\Phi}^{\pm}(\cdot; z)) + c(\alpha) \\ J_{\alpha}^{\pm}(\widehat{\Phi}_q^{\pm}(\cdot; z); a_q^{\alpha, z}) &\leq j_{\alpha}(\widehat{\Phi}_q^{\pm}(\cdot; z)) + c(\alpha) \end{aligned} \quad (5.83)$$

Then  $z \in D$  (respectively  $z \in D_p$ ) if and only if  $\lim_{\alpha \rightarrow 0} (N_{\sharp}^{\pm} a^{\alpha, z}, a^{\alpha, z}) < \infty$  (respectively  $\lim_{\alpha \rightarrow 0} (N_{\sharp}^{\pm} a_q^{\alpha, z}, a_q^{\alpha, z}) < \infty$ ). Moreover, if  $z \in D$  (respectively  $z \in D_p$ ) then  $\mathcal{H}^{\pm} a^{\alpha, z} \rightarrow v(\cdot; z)$  in  $H^1(D)$  (respectively  $\mathcal{H}^{\pm} a_q^{\alpha, z} \rightarrow v(\cdot; z)$  in  $H^1(D)$ ) where  $(u(\cdot; z), v(\cdot; z)) \in H^1(D) \times H^1(D)$  is the solution of problem (5.59) with  $\varphi = \Phi(\cdot; z)$  and  $\psi = \partial\Phi(\cdot; z)/\partial\nu$  on  $\partial D$  (respectively  $\varphi = \Phi_q(\cdot; z)$  and  $\psi = \partial\Phi_q(\cdot; z)/\partial\nu$  on  $\partial D$ ).

*Proof.* The proof of this theorem is a direct application of the abstract framework given in Theorem A.2.2 in combination with Theorem A.1.2  $\square$

For the case with noise in the operators, we refer to Theorem 4.4.5 for which the result also hold in the present setting assuming that the hypothesis of Theorem 5.3.11 are satisfied.

### 5.3.3 Sampling methods for a single Floquet-Bloch mode

We here pursue the analysis of sampling methods with one of the essential steps toward the construction of our indicator function for  $\omega$ . Using the notation of Section 4.6 we recall that the near field operator  $N_q^{\pm} : \ell^2(\mathbb{Z}) \rightarrow \ell^2(\mathbb{Z})$  for a single Floquet-Bloch mode is defined by

$$N_q^{\pm} a = I_q^* N^{\pm} I_q a \quad (5.84)$$

where  $q \in \mathbb{Z}_M$  is fixed. The operator  $I_q : \ell^2(\mathbb{Z}) \rightarrow \ell^2(\mathbb{Z})$  transforms  $a \in \ell^2(\mathbb{Z})$  into  $\tilde{a} \in \ell^2(\mathbb{Z})$  such that

$$\tilde{a}_q = a \quad \text{and} \quad \tilde{a}_{q'} = 0 \quad \text{if} \quad q \neq q'$$

and  $I_q^*(a) = a_q$ . We recall that for  $a \in \ell^2(\mathbb{Z})$ , we define for  $q \in \mathbb{Z}_M$ , the element  $a_q \in \ell^2(\mathbb{Z})$  by

$$a_q(j) := a(q + jM).$$

The single Floquet-Bloch mode Herglotz operator  $H_q^\pm : \ell^2(\mathbb{Z}) \rightarrow L^2(D)^d \times L^2(D)$  is defined by

$$H_q^\pm a := H^\pm \mathbf{I}_q a = \left( \sum_j a(j) \nabla u^{i,\pm}(\cdot; q + jM)|_D, \sum_j a(j) u^{i,\pm}(\cdot; q + jM)|_D \right) \quad (5.85)$$

so that

$$N_q^\pm = (H_q^\pm)^* T H_q^\pm. \quad (5.86)$$

We define the operator  $G_q^\pm : \overline{\mathcal{R}(H_q^\pm)} \rightarrow \ell^2(\mathbb{Z})$  by

$$G_q^\pm = (H_q^\pm)^* T|_{\overline{\mathcal{R}(H_q^\pm)}} \quad (5.87)$$

where the operator  $T$  is defined by (5.57).

### Key properties of the operators $H_q^\pm$ , $N_q^\pm$ and $G_q^\pm$

We prove in this section the needed properties in order to establish the theoretical justifications of sampling methods. This Section is somehow the equivalent of Section 5.3.2 but for the case of single Floquet-Bloch mode operator. We recall that we make the assumption that

$$\Omega_M \setminus D \text{ is connected and } \partial\Omega_m \cap \overline{D} = \emptyset \text{ for all } m \in \mathbb{Z}_M.$$

For the support of the perturbation  $\omega$  we assume that

$$\omega \subset \Omega_{m_0}$$

for some  $m_0 \in \mathbb{Z}_M$  and

$$\mathcal{D}_p \cap \omega = \emptyset.$$

We define for later use the domain

$$\omega_p := \bigcup_{m-m_0 \in \mathbb{Z}_M} \omega + mL \quad (5.88)$$

which is the union of  $L$  periodic copies of  $\omega$ . In all the following of this section

$$q \in \mathbb{Z}_M$$

is a fixed parameter.

**Lemma 5.3.12.** *The operator  $H_q^\pm$  defined by (5.85) is injective and*

$$\overline{\mathcal{R}(H_q^\pm)} = H_{\text{inc}}^q(D) := \{(\nabla v, v) \mid v \in H^1(D) : \Delta v + k^2 v = 0 \text{ in } D \text{ and } v|_{D_p} \text{ is } \alpha_q\text{-quasi-periodic}\}$$

*Proof.*  $H_q^\pm$  is injective since  $H^\pm$  is injective and  $\mathbf{I}_q$  is injective. We now prove that  $(H_q^\pm)^*$  is injective on  $H_{\text{inc}}^q(D)$ . Let  $\varphi = (\nabla f, f) \in H_{\text{inc}}^q(D)$  and assume  $(H_q^\pm)^*(\varphi) = 0$ . We define

$$u(x) := \frac{1}{M} \operatorname{div} \int_D \Phi_q(x-y) (-\nabla f(y)) dy + \frac{1}{M} \int_D \Phi_q(x-y) f(y) dy.$$

From the expansion of  $\Phi_q(x)$  as in (4.56) and using the same calculations as in the proof of Lemma 5.3.2 we have that  $\widehat{u}^\pm(j) = 0$  for all  $j \neq q + M\ell$  and  $\widehat{u}^\pm(q + M\ell) = ((H^\pm)^*(\varphi))(q + M\ell) = ((H_q^\pm)^*(\varphi))(\ell) = 0$ . Therefore  $u$  has all Rayleigh coefficients equal 0, which implies that

$$u = 0, \quad \text{for } \pm x_d > h.$$

We now observe that for all  $y \in D$ ,  $\Delta\Phi_q(\cdot; y) + k^2\Phi_q(\cdot; y) = 0$  in the complement of  $D_p \cup \omega_p$ . This implies that

$$\Delta u + k^2 u = 0 \quad \text{in } \mathbb{R}^d \setminus \{D_p \cup \omega_p\}.$$

Using a unique continuation argument we infer that  $u = 0$  in  $\Omega \setminus D_p \cup \omega_p$ . Therefore,  $u \in H_0^1(D_p \cup \omega_p)$  by the regularity of volume potentials.

Since  $f|_{D_p}$  and  $\Phi_q$  are  $\alpha_q$ -quasi-periodic functions with period  $L$ , then for  $m \in \mathbb{Z}_M$ ,

$$\begin{aligned} u(x) &:= \frac{1}{M} \operatorname{div} \int_{\omega} \Phi_q(x; y) (-\nabla f(y)) \, dy + \operatorname{div} \int_{D_p \cap \Omega_m} \Phi_q(x; y) (-\nabla f(y)) \, dy \quad x \in D_p \cap \Omega_m \\ &+ \frac{1}{M} \int_{\omega} \Phi_q(x; y) f(y) \, dy + \int_{D_p \cap \Omega_m} \Phi_q(x; y) f(y) \, dy \quad x \in D_p \cap \Omega_m \end{aligned} \quad (5.89)$$

We recall that  $\Delta\Phi_q(\cdot; y) + k^2\Phi_q(\cdot; y) = -\delta_y$  in  $\Omega_m$ . Using in addition the fact that  $D_p \cap \omega = \emptyset$ , we obtain for  $m \in \mathbb{Z}_M$ ,

$$\Delta u(x) + k^2 u(x) = \Delta f(x) - f(x) \quad \text{in } D_p \cap \Omega_m. \quad (5.90)$$

Let us set for  $m - m_0 \in \mathbb{Z}_M$

$$f_m(x) := e^{i\alpha_q m L} \varphi(x - mL) \quad \text{for } x \in \omega + mL.$$

Then we have, using the  $\alpha_q$ -quasi-periodicity of  $\Phi_q$

$$\begin{aligned} u(x) &:= \frac{1}{M} \operatorname{div} \int_{\omega + mL} \Phi_q(x; y) (-\nabla f_m(y)) \, dy + \frac{1}{M} \operatorname{div} \int_D \Phi_q(x; y) (-\nabla f(y)) \, dy \\ &+ \frac{1}{M} \int_{\omega + mL} \Phi_q(x; y) f_m(y) \, dy + \frac{1}{M} \int_D \Phi_q(x; y) f(y) \, dy \quad x \in \omega + mL \end{aligned}$$

with  $m - m_0 \in \mathbb{Z}_M$ . Using again the fact that  $D_p \cap \omega_p = \emptyset$  and  $\Delta\Phi_q(\cdot; y) + k^2\Phi_q(\cdot; y) = -\delta_y$  in  $\Omega_m$  we get

$$\Delta u(x) + k^2 u(x) = +\Delta f_m - f_m \quad \text{in } \omega + mL. \quad (5.91)$$

Now define the function  $\tilde{f}$  by

$$\tilde{f} = f \quad \text{in } D_p \quad \text{and} \quad \tilde{f} = f_m \quad \text{in } \omega + mL.$$

Clearly

$$\Delta \tilde{f} + k^2 \tilde{f} = 0 \quad \text{in } D_p \cup \omega_p \quad \text{with } m - m_0 \in \mathbb{Z}_M.$$

Since  $u \in H_0^1(D_p \cup \omega_p)$  we then have

$$\int_D (\Delta u + k^2 u) \tilde{f} = 0.$$

This implies according to (5.90) and (5.91) that

$$\int_{D_p} (-\Delta f + f) \tilde{f} \, dx + M \int_{\omega} (-\Delta f + f) \tilde{f} = 0,$$

which implies  $f = 0$  in  $D_p \cup \omega$  since  $f$  satisfies  $\Delta f + k^2 f = 0$  in  $D = D_p \cup \omega$  (of course  $k^2 \neq -1$ ). Therefore  $\varphi = 0$ . This proves the injectivity of  $(H^\pm)^*$  on  $H_{\text{inc}}^q(D)$  and hence proves the Lemma.  $\square$

The following two Lemmas are the important cornerstones for the justification of the GLSM and the differential imaging functional that we shall define later.

**Lemma 5.3.13.** *Assume that Assumptions 5.3.1 and 5.3.3 hold. Assume in addition that Assumption 5.3.1 hold with  $\mu^{-1} = \mu_p^{-1}$  and  $n = n_p$ . Then the operator  $G_q^\pm : H_{\text{inc}}^q(D) \rightarrow \ell^2(\mathbb{Z})$  is injective with dense range.*

*Proof.* We here give the proof of the lemma for the operator  $G_q^+$  since the proof for the operator  $G_p^-$  is similar.

Assume that  $G_q^+(\varphi) = 0$  for  $\varphi = (\nabla f, f) \in H_{\text{inc}}^q(D)$ . Let  $w \in H_{\#}^1(\Omega_M^h)$  be the associated scattered field via (5.47) and consider the decomposition  $w$  into

$$w = \frac{1}{M} \sum_{q' \in \mathbb{Z}_M} w_{q'}, \quad (5.92)$$

where  $w_{q'}$  is  $\alpha_{q'}$ -quasi-periodic with period  $L$ . We recall that

$$\begin{aligned} \operatorname{div}(\mu_p^{-1} \nabla w) + k^2 n_p w + \operatorname{div}((\mu^{-1} - \mu_p^{-1}) \nabla w) + k^2(n - n_p)w \\ = \operatorname{div}((\mu_p^{-1} - \mu^{-1}) \nabla f) + \operatorname{div}((1 - \mu_p^{-1}) \nabla f) + k^2(n_p - n)f + k^2(1 - n_p)f. \end{aligned} \quad (5.93)$$

Since  $(1 - \mu_p^{-1})f$  and  $(1 - n_p)f$  is  $\alpha_q$ -quasi-periodic with period  $L$  and  $\omega \subset \Omega_{m_0}$  we obtain that (projecting the latter equation on the  $H^1(\Omega_{m_0})$  Fourier basis which is  $\alpha_q$ -quasi-periodic with period  $L$ )

$$\begin{aligned} \operatorname{div}(\mu_p^{-1} \nabla w_q) + k^2 n_p w_q + \operatorname{div}((\mu^{-1} - \mu_p^{-1}) \nabla w_q) + k^2(n - n_p)w \\ = \operatorname{div}((1 - \mu^{-1}) \nabla f) + k^2(1 - n)f \quad \text{in } \Omega_{m_0}. \end{aligned} \quad (5.94)$$

In particular we have that  $\Delta w_q + k^2 w_q = 0$  in  $\Omega_M \setminus \{D_p \cup \omega_p\}$ . Using a similar unique continuation argument as the one at the beginning of the proof of Lemma 5.3.12 we deduce that

$$w_q = 0 \quad \text{in } \Omega_M \setminus \{D_p \cup \omega_p\}.$$

This implies in particular (since  $D_p \cap \omega = \emptyset$ )  $w_q \in H_0^2(D_p)$  and

$$\begin{cases} \operatorname{div}(\mu_p^{-1} \nabla w_q) + k^2 n_p w_q & = \operatorname{div}((1 - \mu_p^{-1}) \nabla f) + k^2(1 - n_p)f & \text{in } D_p \\ \Delta f + k^2 f & = 0 & \text{in } D_p \end{cases}$$

Assumption 5.3.3 implies in particular that  $\text{ITP}(\mu_p, n_p, D_p, 0)$  is well posed. We then obtain that  $f = 0$  in  $D_p$  and  $w_q = 0$  in  $D_p$ . On the other hand, using the fact that  $\mu_p^{-1} = 1$  and  $n_p = 1$  in  $\omega$  we also have  $w_q \in H_0^1(\omega)$  and

$$\Delta w_q + k^2 w_q + \operatorname{div}((\mu^{-1} - 1) \nabla w) + k^2(n - 1)w = \operatorname{div}((1 - \mu^{-1}) \nabla f) + k^2(1 - n)f \quad \text{in } \omega \quad (5.95)$$

The fact that  $w_q \in H_0^2(\omega)$  gives for instance that

$$\int_{\omega} (\Delta w_q + k^2 w_q) \theta = 0 \quad (5.96)$$

for all  $\theta \in \tilde{H}_{\text{inc}}(\omega) := \{\theta \in H^1(\omega), \Delta \theta + k^2 \theta = 0\}$ . Now, taking the  $L^2$  scalar product of (5.95) with  $\theta$  we arrive at

$$\int_{\omega} \operatorname{div}((1 - \mu^{-1})(\nabla w + \nabla f)) \bar{\theta} + \int_{\omega} k^2(1 - n)(f + w) \bar{\theta} = 0 \quad (5.97)$$

for all  $\theta \in H_{inc}(\omega)$ . Recall that the solution  $w$  of (5.47) can be represented as

$$w(x) = \operatorname{div} \int_{D_p \cup \omega} (\mu^{-1} - 1)(\nabla w + \nabla f)(y) \Phi(x; y) \, dy + \int_{D_p \cup \omega} k^2(n-1)(w+f)(y) \Phi(x; y) \, dy \quad x \in \Omega_M. \quad (5.98)$$

Since  $y \mapsto \Phi(x; y) \in H_{inc}(\omega)$  for  $x \notin \omega$  and  $f = 0$  in  $D_p$ , we then obtain from (5.97) that

$$w(x) = \operatorname{div} \int_{D_p} (1 - \mu_p^{-1}) \nabla w(y) \Phi(x; y) \, dy + \int_{D_p} k^2(1 - n_p) w(y) \Phi(x; y) \, dy \quad \text{for } x \notin \omega. \quad (5.99)$$

Let us define  $\tilde{w} \in H^1(\omega)$  by

$$\tilde{w}(x) := \operatorname{div} \int_{D_p} (1 - \mu_p^{-1}) \nabla w(y) \Phi(x; y) \, dy + \int_{D_p} k^2(1 - n_p) w(y) \Phi(x; y) \, dy \quad x \in \omega.$$

Then obviously  $u = w + f$  and  $f$  satisfy

$$\begin{cases} \operatorname{div}(\mu^{-1} \nabla u) + k^2 n u = 0 & \text{in } \omega \\ \Delta f + k^2 f = 0 & \text{in } \omega \\ u - f = \tilde{w} & \text{on } \partial \omega \\ \frac{\partial(u-f)}{\partial \nu} = \frac{\partial \tilde{w}}{\partial \nu} & \text{on } \partial \omega \end{cases} \quad (5.100)$$

Since

$$\Delta \tilde{w} + k^2 \tilde{w} = 0 \text{ in } \omega$$

and  $\text{ITP}(\mu, n, \omega, \tilde{w})$  is well posed (by Assumption 5.3.3 and the fact that  $D_p \cap \omega_p = \emptyset$ ) we get that

$$f = -\tilde{w} \quad \text{and} \quad u = 0 \quad \text{in } \omega.$$

We then conclude that  $w + f = 0$  in  $\omega$ , i.e.  $w = \tilde{w}$  in  $\omega$ . This implies for instance that

$$w(x) = \operatorname{div} \int_{D_p} (1 - \mu_p^{-1}) \nabla w(y) \Phi(x; y) \, dy + \int_{D_p} k^2(1 - n_p) w(y) \Phi(x; y) \, dy \quad \text{for } x \notin \Omega_M \quad (5.101)$$

which means in particular that  $w$  is solution to (5.47) with  $\mu^{-1} = \mu_p^{-1}$ ,  $n = n_p$  and  $f = 0$ . We then infer that  $w = 0$  by uniqueness of the solution to the scattering problem for  $\mu^{-1} = \mu_p^{-1}$  and  $n = n_p$  (third assumption of our Lemma). This proves that  $f = 0$  in  $\omega$  and finishes the arguments for the injectivity of the operator  $G_q^+$ .

Using the injectivity we now prove the denseness of the range of  $G_q^\pm$ . The proof can be done as in Theorem 5.3.6. Let  $g \in \overline{\mathcal{R}(G_q^+)}^\perp$ . Then

$$\langle G_q^+(\varphi), g \rangle_{\ell^2(\mathbb{Z})} = 0, \quad \forall \varphi = (\nabla f, f) \in H_{inc}^q(D).$$

Consider  $\varphi$  of the form  $\varphi = \overline{H_q^+(a)}$  for some  $a \in \ell^2(\mathbb{Z})$ . Since  $G_q^+ = (H_q^+)^* T$ , we then have

$$\langle T(\overline{H_q^+(a)}), H_q^+(g) \rangle_{L^2(D)} = 0, \quad \forall a \in \ell^2(\mathbb{Z}). \quad (5.102)$$

Let  $w(a)$  and  $w(g)$  solution to (5.47) with respect to  $\overline{H_q^+(a)}$  and  $\overline{H_q^+(g)}$ . Let us denote by  $H_{q,1}^\pm := H_1^\pm \mathbf{I}_q$ ,  $H_{q,2}^\pm := H_2^\pm \mathbf{I}_q$ , where  $H_2^\pm = 1$  and  $H_2^\pm$  are defined in the proof of Theorem 5.3.6. From Lemma 5.3.5 we get

$$\begin{aligned} \left( T(\overline{H_q^+(a)}), H_q^+(g) \right)_{L^2(D)} &= \int_D q(\overline{H_{q,1}^+ a} + \nabla w(a)) \cdot \overline{H_{q,1}^+ g} \, dx + k^2 \int_D p(\overline{H_{q,2}^+ a} + w(a)) \overline{H_{q,2}^+ g} \, dx \\ &= \int_D q(\overline{H_{q,1}^+ g} + \nabla w(g)) \cdot \overline{H_{q,1}^+ a} \, dx + k^2 \int_D p(\overline{H_q^+ g} + w(g)) \overline{H_q^+ a} \, dx. \end{aligned}$$

Therefore,

$$\left( T(\overline{H_q^+(g)}), H_q^+(a) \right)_{L^2(D)} = \left( T(\overline{H_q^+(a)}), H_q^+(g) \right)_{L^2(D)}, \quad \forall a \in \ell^2(\mathbb{Z}). \quad (5.103)$$

We deduce from (5.102) that

$$\left( G_q^+(\overline{H_q^+(g)}), a \right)_{\ell^2(\mathbb{Z})} = 0, \quad \forall a \in \ell^2(\mathbb{Z}),$$

which implies  $G_q^+(\overline{H_q^+(g)}) = 0$ . The injectivity of  $G_q^+$  ensure that  $H_q^+ g = 0$  in  $D$  and consequently  $g = 0$  (using Lemma 5.3.12). This proves the denseness of the range of  $G_q^+$ .  $\square$

**Lemma 5.3.14.** *Under the same assumptions as in Lemma 5.3.13 we have that  $\mathbf{I}_q^* \widehat{\Phi}_q^\pm(\cdot; z) \in \mathcal{R}(G_q^\pm)$  if and only if  $z \in D_p \cup \omega_p$ .*

*Proof.* We recall that  $\widehat{\Phi}_q^\pm(\cdot; z)$  is the Rayleigh sequence of  $\Phi_q(\cdot; z)$ .

Let  $z \in D_p$ . We consider  $(u_q, v_q)$  solution of ITP( $\mu, n, D, \Phi_q(\cdot; z)$ ). Let us define

$$w = \begin{cases} u_q - v_q & \text{in } D_p. \\ \Phi_q & \text{in } \Omega_M \setminus D_p. \end{cases}$$

Then  $w \in H_{\#, \text{loc}}^1(\Omega_M)$  and satisfies the scattering problem (5.47) with  $f = v_q$  (remark that  $v_q = -\Phi_q$  in  $\omega$ ). Therefore  $G^\pm(v_q) = \widehat{\Phi}_q^\pm(\cdot; z)$ . Applying  $\mathbf{I}_q^*$  to the two sides of the equality we get  $G_q^\pm(v_q) = \mathbf{I}_q^* \widehat{\Phi}_q^\pm(\cdot; z)$ .

Now let  $z \in \omega$  and consider  $\varphi = (\nabla f, f) \in H_{\text{inc}}(D)$  is such that  $G(\varphi) = \widehat{\Phi}^\pm(\cdot; z)$  as in Lemma. By construction

$$\mathbf{I}_q^* G(M\varphi) = M \mathbf{I}_q^* \widehat{\Phi}^\pm(\cdot; z) = \mathbf{I}_q^* \widehat{\Phi}_q^\pm(\cdot; z) \quad (5.104)$$

However the function  $\varphi \notin H_{\text{inc}}^q(D)$  since the function  $f$  is constructed such that  $(u, f) \in H^2(D) \times H^2(D)$  is a solution of ITP( $\mu, n, D, \Phi(\cdot; z)$ ) and therefore  $f = -\Phi(\cdot; z)$  in  $D_p$  (which is not an  $\alpha_q$ -quasi-periodic function with period  $L$ ). We use the same procedure as in the proof of Lemma 4.6.3 to construct from  $\varphi$  a function  $\tilde{\varphi} \in H_{\text{inc}}^q(D)$  such that  $G_q(\varphi) = G_q(\tilde{\varphi})$ . We first decompose  $f|_{D_p}$  into

$$f = \frac{1}{M} \sum_{q' \in \mathbb{Z}_M} f_{q'}$$

where  $f_{q'}$  is  $\alpha_{q'}$ -quasi-periodic function with period  $L$  as in 4.52. Each of the  $f_{q'}$  is also a solution to  $\Delta f_{q'} + k^2 f_{q'} = 0$  in  $D_p$ . Remark that for our case  $f_{q'} = \Phi_{q'}(\cdot, z)$ . We set  $\tilde{f} = f - \frac{1}{M} f_q$ . Now consider the solution  $\tilde{w} \in H_{\#, \text{loc}}^1(\Omega_M)$  to

$$\operatorname{div}(\mu_p^{-1} \nabla \tilde{w}) + k^2 n_p \tilde{w} = \operatorname{div}((1 - \mu_p^{-1}) \nabla \tilde{f}) + k^2 (1 - n_p) \tilde{f} \quad \text{in } \Omega_M$$

with the Rayleigh radiation condition. Using the  $L^2(\Omega_M)$  orthogonality of  $(1 - \mu_p^{-1})\nabla\tilde{f}$  and  $(1 - n_p)\tilde{f}$  with  $\alpha_q$ -quasi-periodic functions and uniqueness the solution of the scattering problem for we simply get that

$$\tilde{w}_q = 0.$$

If we denote by  $\tilde{a}^\pm$  the Rayleigh sequences associated with  $\tilde{w}$ , then the latter ensures that  $I_q^*(\tilde{a}^\pm) = 0$ . Now observe that we can also write the equation for  $\tilde{w}$  as

$$\begin{aligned} & \operatorname{div}(\mu^{-1}\nabla\tilde{w}) + k^2n\tilde{w} \\ &= \operatorname{div}((1 - \mu_p^{-1})\nabla\tilde{f}) + k^2(1 - n_p)\tilde{f} + \operatorname{div}((\mu_p^{-1} - \mu^{-1})\nabla(-\tilde{w})) + k^2(n_p - n)(-\tilde{w}) \quad \text{in } \Omega_M \end{aligned}$$

and since  $\omega \cap D_p = \emptyset$ , we have  $\Delta\tilde{w} + k^2\tilde{w} = 0$  in  $\omega$ . Therefore the function  $\hat{f}$  defined by  $\hat{f} = \tilde{f}$  in  $D_p$  and  $\hat{f} = -\tilde{w}$  in  $\omega$  satisfies  $\hat{\varphi} := (\nabla\hat{f}, \hat{f}) \in H_{\text{inc}}(D)$  and

$$G^\pm(\hat{\varphi}) = \tilde{a}^\pm.$$

Therefore  $I_q^*G^\pm(\hat{\varphi}) = 0$ ,  $(\varphi - \hat{\varphi}) \in H_{\text{inc}}^q(D)$  and

$$I_q^*G^\pm(\varphi - \hat{\varphi}) = I_q^*G^\pm(\varphi).$$

(Remark that if  $\hat{\varphi} \neq 0$  then  $G^\pm(\varphi - \hat{\varphi}) \neq G^\pm(\varphi)$ ). Applying this procedure to the  $\varphi$  in 5.104 we obtain that,  $M(\varphi - \hat{\varphi}) \in H_{\text{inc}}^q(D)$  and

$$G_q(M(\varphi - \hat{\varphi})) = \mathbf{I}_q^*\hat{\Phi}_q^\pm(\cdot; z).$$

Now consider the case where  $z \in \omega + mL$  with  $m - m_0 \in \mathbb{Z}_M$ . We recall that  $\hat{\Phi}_q^\pm(\cdot; z) = e^{imL \cdot \alpha_q} \hat{\Phi}_q^\pm(\cdot; z - mL)$ . Therefore, if we consider  $\varphi \in H_{\text{inc}}^q(D)$  such that  $G_q(\varphi) = \mathbf{I}_q^*\hat{\Phi}_q^\pm(\cdot; z - mL)$ , which is possible by the previous step since  $z - mL \in \omega$ , then

$$G_q(e^{imL \cdot \alpha_q} \varphi) = \mathbf{I}_q^*(\hat{\Phi}_q^\pm(\cdot; z)).$$

We finally consider the case  $z \notin D_p \cup \omega_p$ . If  $G_q(\varphi) = \mathbf{I}_q^*\hat{\Phi}_q^\pm(\cdot; z - mL)$ , then using the same unique continuation argument as in the proof of Lemma 5.3.13 we get that  $w_q = \Phi_q$  in  $\Omega_M \setminus \{D_p \cup \omega_p\}$  where  $w$  is the solution of (5.47) with  $f = \varphi$  and  $w_q$  is defined by 5.92. This gives a contradiction since  $w_q \in H_{\text{loc}}^1(\Omega_M \setminus \{D_p \cup \omega_p\})$  while  $\Phi_q(\cdot; z) \notin H_{\text{loc}}^1(\Omega_M \setminus \{D_p \cup \omega_p\})$ .  $\square$

### Application to the differential imaging functional

The material above allows us to conclude as in Section 4.6.3 and design a method that would reconstruct the domain  $\omega$ . We here rephrase the main result for this differential imaging functional. The proof is exactly the same as for Theorem 4.6.7.

We recall that in addition to the assumptions on the geometry that was made in the previous section we need to assume that for all  $z \in \omega$  there exists  $m \in \mathbb{Z}^{d-1}$  such that  $z + mL \in \Omega_M \setminus D$ . This means in particular that  $M > 1$  and therefore there is a least one period that does not contain a defect. We only present the results for the operator  $N^+$ . We obtain exactly the same results by changing the exponent  $+$  with the exponent  $-$ . We define

$$N_{q, \#}^\pm := \mathbf{I}_q^* N_\#^\pm \mathbf{I}_q.$$

We then consider for  $\phi$  and  $a$  in  $\ell^2(\mathbb{Z})$  the cost functionals

$$\begin{aligned} J_\alpha^+(\phi, a) &:= \alpha(\mathbf{N}_\#^+ a, a) + \|\mathbf{N}^+ a - \phi\|^2, \\ J_{\alpha,q}^+(\phi, a) &:= \alpha(\mathbf{N}_{q,\#}^+ a, a) + \|\mathbf{N}_q^+ a - \phi\|^2. \end{aligned} \quad (5.105)$$

Let  $a^{\alpha,z}$ ,  $a_q^{\alpha,z}$  and  $\tilde{a}_q^{\alpha,z}$  be in  $\ell(\mathbb{Z})$  verifying

$$\begin{aligned} J_\alpha^+(\widehat{\Phi}^\pm(\cdot; z), a^{\alpha,z}) &\leq \inf_{a \in \ell^2(\mathbb{Z})} J_\alpha^+(\widehat{\Phi}^\pm(\cdot; z), a) + c(\alpha) \\ J_\alpha^+(\widehat{\Phi}_q^\pm(\cdot; z), a_q^{\alpha,z}) &\leq \inf_{a \in \ell^2(\mathbb{Z})} J_\alpha^+(\widehat{\Phi}_q^\pm(\cdot; z), a) + c(\alpha) \\ J_{\alpha,q}^+(\mathbf{I}_q^* \widehat{\Phi}_q^\pm(\cdot; z), \tilde{a}_q^{\alpha,z}) &\leq \inf_{a \in \ell^2(\mathbb{Z})} J_{\alpha,q}^+(\mathbf{I}_q^* \widehat{\Phi}_q^\pm(\cdot; z), a) + c(\alpha) \end{aligned} \quad (5.106)$$

with  $\frac{c(\alpha)}{\alpha} \rightarrow 0$  as  $\alpha \rightarrow 0$ . We then consider the following imaging functional to characterize  $\omega$ :

$$\mathcal{I}_\alpha^+(z) = \left( (\mathbf{N}_\#^+ a^{\alpha,z}, a^{\alpha,z}) \left( 1 + \frac{(\mathbf{N}_\#^+ a^{\alpha,z}, a^{\alpha,z})}{D^+(a_q^{\alpha,z}, \tilde{a}_q^{\alpha,z})} \right) \right)^{-1} \quad (5.107)$$

where for  $a$  and  $b$  in  $\ell^2(\mathbb{Z})$ ,

$$D^+(a, b) := \left( \mathbf{N}_\#^+(a - \mathbf{I}_q b), (a - \mathbf{I}_q b) \right).$$

**Theorem 5.3.15.** *Under the assumptions of Theorems 5.3.10, and of Lemma 5.3.11, we have that*

$$z \in \omega \text{ if and only if } \lim_{\alpha \rightarrow 0} \mathcal{I}_\alpha^+(z) > 0.$$

As mentioned earlier, the proof is exactly the same as the proof of Theorem 4.95. One also can state the version for noisy operators as in Theorem 4.6.8. We finally indicate that Theorems 4.6.5 and 4.6.6 that contain statements on the reconstruction of  $D_p \cup \omega_p$  from the single Floquet-Bloch near field operator still hold (under obvious modifications for the assumptions).





# Surface potential formulation of a scattering problem with sign changing coefficients

## Abstract

This chapter is dedicated to the analysis of the well posedness of the scattering problem for negative materials with contrast  $= -1$ . After introducing the problem and recalling some known results on surface potentials, we reformulate the problem as a surface integral equation problem. We then recall the potential properties for the traces of  $L^2$  solutions and exploit them to prove the Fredholm properties of the problem in dimension 2.

## Contents

<b>6.1</b>	<b>Introduction</b>	<b>141</b>
<b>6.2</b>	<b>Setting of the problem</b>	<b>142</b>
<b>6.3</b>	<b>Reformulation of the problem via surface potentials</b>	<b>143</b>
6.3.1	Surface potentials and some classical results	144
6.3.2	The system of surface integral equations	145
<b>6.4</b>	<b>The Fredholm property of the surface integral operator <math>Z(k_i, k_e)</math></b>	<b>147</b>
6.4.1	The existence and uniqueness of solutions to the scattering problem	150

## 6.1 Introduction

The present short chapter is dedicated to the analysis of the scattering problem by media with negative index. It contains a material that has been developed at the beginning of my PhD work and that is somehow not really directly related with the main subject on scattering from locally perturbed periodic media. It is however linked with the main theme of the Metamath ANR project that funded my PhD work and has been motivated by the analysis of the interior transmission problem in [50].

We specifically consider the scattering problem for case where the contrast between two media  $= -1$ . It is well known that this case is the most difficult case to analyze since for instance the problem is no longer of Fredholm type in the classical  $H^1$  setting of the problem [18, 17, 22, 52]. We refer to [41, 97, 108, 93] for an overview on the mathematical and numerical analysis of some scattering problems with sign changing coefficients. Employing a surface potential method, we show that the problem is in fact well posed in the  $L^2$  sense for the two dimensional setting. We rely on the surface potential mapping properties for densities that are traces of  $L^2$  solutions and

that were analyzed in [50, 49] for the purpose of establishing the Fredholm property of the so-called interior transmission problem. Our result complement the results in [94] where the three dimensional problem is analyzed. We also refer to [93] where complementary results are derived using variational techniques.

Since this is not the core of my PhD work, we kept this chapter as short as possible with only the main key points and proofs. We introduce in Section 6.2 the direct problem and the main assumptions. We reformulate the problem in terms of surface potentials and define  $L^2$  solutions in Section 6.3. We also recall in this section the main properties for surface potentials that are needed to solve the problem. The well-posedness of the problem is established in Section 6.4.

## 6.2 Setting of the problem

Let  $\Omega \subset \mathbb{R}^d$ ,  $d = 2$  or  $d = 3$ , be a bounded open set with a smooth boundary  $\Gamma := \partial\Omega$  and such that the domain  $\mathbb{R}^d \setminus \overline{\Omega}$  is connected.

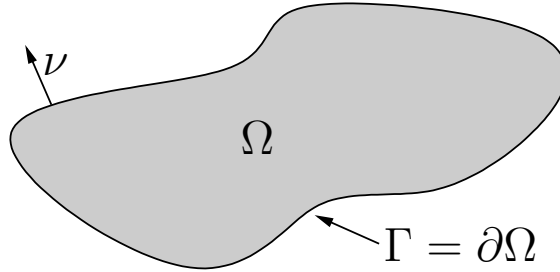


Figure 6.1: Sketch of the geometry

The scattering problem can be formulated as

$$\begin{aligned}
 \operatorname{div}(\mu^{-1}\nabla u) + k^2\varepsilon u &= 0 && \text{in } \Omega, \\
 \Delta v + k^2v &= 0 && \text{in } \mathbb{R}^d \setminus \overline{\Omega}, \\
 u &= v && \text{on } \Gamma, \\
 \mu^{-1}\frac{\partial u}{\partial \nu} &= \frac{\partial v}{\partial \nu} && \text{on } \Gamma.
 \end{aligned} \tag{6.1}$$

We hereafter shall only consider the simple case where  $\varepsilon$  and  $\mu$  are negative constants in  $\Omega$  (the results can be easily extended to the cases where they are constant only in a neighborhood of  $\Gamma$ ). Let us denote by  $k_i := k\sqrt{\varepsilon\mu}$ . The first equation of (6.1) is equivalent to

$$\Delta u + k_i^2 u = 0 \quad \text{in } \Omega. \tag{6.2}$$

We denote by  $u^i$  an incident wave

$$\Delta u^i + k^2 u^i = 0 \quad \text{in } \mathbb{R}^d \tag{6.3}$$

and  $u^s$  a scattered wave such that  $v = u^s + u^i$ . Then,  $u^s$  verifies

$$\Delta u^s + k^2 u^s = 0 \quad \text{in } \mathbb{R}^d \setminus \Omega. \tag{6.4}$$

We require in addition that  $u^s$  satisfies the Sommerfeld radiation condition [46], i.e.

$$\lim_{R \rightarrow \infty} \int_{r=R} \left| \frac{\partial u^s}{\partial r} - ik u^s \right|^2 dx = 0. \tag{6.5}$$

The transmission conditions on the boundary can be rewritten as

$$\begin{aligned} u &= u^s + u^i && \text{on } \Gamma \\ \mu^{-1} \frac{\partial u}{\partial \nu} &= \frac{\partial u^s}{\partial \nu} + \frac{\partial u^i}{\partial \nu} && \text{on } \Gamma. \end{aligned} \quad (6.6)$$

System (6.1) now becomes

$$\begin{aligned} \operatorname{div} (\mu^{-1} \nabla u) + k^2 \varepsilon u &= 0 && \text{in } \Omega, \\ \Delta u^s + k^2 u^s &= 0 && \text{in } \mathbb{R}^d \setminus \overline{\Omega}, \\ u &= u^s + u^i && \text{on } \Gamma, \\ \mu^{-1} \frac{\partial u}{\partial \nu} &= \frac{\partial u^s}{\partial \nu} + \frac{\partial u^i}{\partial \nu} && \text{on } \Gamma. \end{aligned} \quad (6.7)$$

If we use the notation  $f = u^i|_{\Gamma}$ ,  $g := \frac{\partial u^i}{\partial \nu}|_{\Gamma}$ , problem (6.7) has the form

$$\begin{aligned} \operatorname{div} (\mu^{-1} \nabla u) + k^2 \varepsilon u &= 0 && \text{in } \Omega, \\ \Delta u^s + k^2 u^s &= 0 && \text{in } \mathbb{R}^d \setminus \overline{\Omega}, \\ u &= u^s + f && \text{on } \Gamma \\ \mu^{-1} \frac{\partial u}{\partial \nu} &= \frac{\partial u^s}{\partial \nu} + g, && \text{on } \Gamma \end{aligned} \quad (6.8)$$

where we require the Sommerfeld radiation condition at infinity, i.e.  $u^s$  satisfies (6.5). The solvability of equation (6.8) for a given  $(f, g) \in H^s(\Gamma) \times H^{s-1}(\Gamma)$ ,  $s > 0$  is in general studied for the case  $\mu \neq -1$  in a neighborhood of the boundary  $\Gamma$ . In this case, it is well known, see for instance [51], that for any given  $(f, g) \in H^s(\Gamma) \times H^{s-1}(\Gamma)$ , equation (6.8) admits a unique solution and such that  $(u|_{\Gamma}, u^s|_{\Gamma}) \in H^s(\Gamma)^2$ . In the case that  $\mu = -1$ , using explicit calculations the Dirichlet to Neumann symbol at smooth boundaries  $\Gamma$  for the Helmholtz operator, the author obtained in [94] the Fredholm property of the problem (6.8) in space dimension  $d \geq 3$  with the assumption that  $\Omega$  is a strictly convex. For space dimension 2, the well-posedness for this equation is proved if  $\mu = -1$  only at a finite set of points on  $\Gamma$ . In the following, we shall consider the solvability of Problem (6.7) in 2D when  $\mu = -1$  for  $L^2$  solutions: i.e. find  $(u, u^s) \in L^2_{\Delta}(\Omega) \times L^2_{\text{loc}, \Delta}(\mathbb{R}^d \setminus \overline{\Omega})$  solving equation (6.7), where

$$\begin{aligned} L^2_{\Delta}(\Omega) &:= \{w \in L^2(\Omega) \mid \Delta w \in L^2(\Omega)\} \\ L^2_{\text{loc}, \Delta}(\mathbb{R}^d \setminus \overline{\Omega}) &:= \{w \in L^2_{\text{loc}}(\mathbb{R}^d \setminus \overline{\Omega}) \mid \Delta w \in L^2_{\text{loc}}(\mathbb{R}^d \setminus \overline{\Omega})\}. \end{aligned} \quad (6.9)$$

In this case the transmission conditions on  $\Gamma$  has to be understood in the sense of  $H^{-1/2}(\Gamma)$  and  $H^{-3/2}(\Gamma)$  traces respectively. We shall apply the same method proposed in [50]. From now on, we use  $k_e$  to denote the value of  $k$  outside  $\Omega$ .

### 6.3 Reformulation of the problem via surface potentials

Let  $\nu$  be the unit normal to  $\Gamma$  directed into the exterior of  $\Omega$  (see Figure 6.1). For a regular function  $u$  defined in  $\mathbb{R}^d \setminus \Gamma$ , we denote by

$$u^{\pm}(x) := \lim_{h \rightarrow 0} u(x \pm h\nu(x)), \quad \frac{\partial u^{\pm}}{\partial \nu}(x) := \lim_{h \rightarrow 0} \nabla u(x \pm h\nu(x)) \cdot \nu(x) \quad \text{for } x \in \Gamma. \quad (6.10)$$

We use the same notation if the function is not regular but the limits (6.10) exist in the sense of traces on  $\Gamma$ . In this case, we further define the jumps on  $\Gamma$  as

$$[u]_{\Gamma}(x) := u^+(x) - u^-(x), \quad \left[ \frac{\partial u}{\partial \nu} \right]_{\Gamma}(x) := \frac{\partial u^+}{\partial \nu}(x) - \frac{\partial u^-}{\partial \nu}(x) \quad \text{for } x \in \Gamma.$$

### 6.3.1 Surface potentials and some classical results

Let  $\Phi_k$  be the fundamental solution to the Helmholtz associated with wavenumber  $k \in \mathbb{C}$ . We recall that

$$\Phi_k(x, y) = \begin{cases} \frac{e^{ik|x-y|}}{4\pi|x-y|}, & x \neq y, & \text{if } d = 3 \\ \frac{i}{4}H_0^{(1)}(k|x-y|), & & \text{if } d = 2, \end{cases} \quad (6.11)$$

where  $H_0^{(1)}$  denotes the Hankel function of the first kind of order 0. We then define the single and double layer potentials for a regular density  $\varphi$ , respectively by

$$(\text{SL}_k\varphi)(x) := \int_{\Gamma} \Phi_k(x, y)\varphi(y)ds(y), \quad x \in \mathbb{R}^d \setminus \Gamma \quad (6.12)$$

and

$$(\text{DL}_k\varphi)(x) := \int_{\Gamma} \frac{\partial \Phi_k}{\partial \nu(y)}(x, y)\varphi(y)ds(y), \quad x \in \mathbb{R}^d \setminus \Gamma. \quad (6.13)$$

The following theorem summarizes some classical results from potential theory that are presented, for instance, in [88], [90] or [65]. Let  $B_R(0)$  be a ball containing  $\Omega$  and denote by  $\Omega^c := B_R(0) \cap \mathbb{R}^d \setminus \bar{\Omega}$ .

**Theorem 6.3.1.** *The single-layer potential  $\text{SL}_k : H^{-1/2}(\Gamma) \rightarrow H^1(\tilde{\Omega})$  and the double layer potential  $\text{DL}_k : H^{1/2}(\Gamma) \rightarrow H^1(\tilde{\Omega})$ , ( $\tilde{\Omega} = \Omega$  or  $\Omega^c$ ) are bounded and give rise to bounded linear operators*

$$\begin{aligned} S_k : H^{-1/2}(\Gamma) &\rightarrow H^{1/2}(\Gamma), & K_k : H^{1/2}(\Gamma) &\rightarrow H^{1/2}(\Gamma), \\ K'_k : H^{-1/2}(\Gamma) &\rightarrow H^{-1/2}(\Gamma), & T_k : H^{1/2}(\Gamma) &\rightarrow H^{-1/2}(\Gamma), \end{aligned}$$

such that for all  $\varphi \in H^{-1/2}(\Gamma)$  and  $\psi \in H^{1/2}(\Gamma)$ ,

$$\begin{cases} (\text{SL}_k\varphi)^{\pm} = S_k\varphi & \text{and} & (\text{DL}_k\psi)^{\pm} = K_k\psi \pm \frac{1}{2}\psi & \text{in } H^{1/2}(\Gamma), \\ \frac{\partial(\text{SL}_k\varphi)^{\pm}}{\partial \nu} = K'_k\varphi \mp \frac{1}{2}\varphi & \text{and} & \frac{\partial(\text{DL}_k\psi)^{\pm}}{\partial \nu} = T_k\psi & \text{in } H^{-1/2}(\Gamma). \end{cases}$$

We recall that for the regular densities  $\varphi$  and  $\psi$ , the surface potentials  $S_k$ ,  $K_k$ ,  $K'_k$  and  $T_k$  can be expressed as

$$\begin{aligned} (S_k\varphi)(x) &= \int_{\Gamma} \Phi_k(x, y)\varphi(y)ds(y), & (K_k\psi)(x) &= \int_{\Gamma} \frac{\partial \Phi_k}{\partial \nu(y)}(x, y)\psi(y)ds(y), \\ (K'_k\varphi)(x) &= \int_{\Gamma} \frac{\partial \Phi_k}{\partial \nu(x)}(x, y)\varphi(y)ds(y), & (T_k\psi)(x) &= \lim_{\varepsilon \rightarrow 0} \int_{\Gamma, |y-x|>\varepsilon} \frac{\partial^2 \Phi_k}{\partial \nu(y)\partial \nu(x)}(x, y)\psi(y)ds(y) \end{aligned} \quad (6.14)$$

for all  $x \in \Gamma$ . More generally, it is showed in [50] that for any densities  $\varphi \in H^{-3/2}(\Gamma)$  and  $\psi \in H^{-1/2}(\Gamma)$ ,  $\text{SL}_k\varphi$  and  $\text{DL}_k\psi$  are locally  $L^2$  functions that satisfy the Helmholtz equation in the distributional sense in  $\mathbb{R}^d \setminus \Gamma$ . In particular, the operators  $\text{SL}_k : H^{-3/2}(\Gamma) \rightarrow Y$  and  $\text{DL}_k : H^{-1/2}(\Gamma) \rightarrow Y$ , where  $Y = L^2_\Delta(\Omega)$  or  $Y = L^2_{\text{loc},\Delta}(\mathbb{R}^d \setminus \Omega)$ , are continuous. We in fact have a more precise results that generalize the results of Theorem 6.3.1 in the following sense (see [50]).

**Theorem 6.3.2.** *The single-layer potential  $\text{SL}_k : H^{-3/2}(\Gamma) \rightarrow L^2_\Delta(\tilde{\Omega})$  and the double layer potential  $\text{DL}_k : H^{-1/2}(\Gamma) \rightarrow L^2_\Delta(\tilde{\Omega})$ , with  $(\tilde{\Omega} = \Omega \text{ or } \Omega^c)$ , are bounded and give rise to bounded linear operators*

$$\begin{aligned} S_k &: H^{-3/2}(\Gamma) \rightarrow H^{-1/2}(\Gamma), & K_k &: H^{-1/2}(\Gamma) \rightarrow H^{-1/2}(\Gamma), \\ K'_k &: H^{-3/2}(\Gamma) \rightarrow H^{-3/2}(\Gamma), & T_k &: H^{-1/2}(\Gamma) \rightarrow H^{-3/2}(\Gamma), \end{aligned}$$

such that for all  $\varphi \in H^{-3/2}(\Gamma)$  and  $\psi \in H^{-1/2}(\Gamma)$ ,

$$\begin{cases} (\text{SL}_k\varphi)^\pm = S_k\varphi & \text{and} & (\text{DL}_k\psi)^\pm = K_k\psi \pm \frac{1}{2}\psi & \text{in } H^{-1/2}(\Gamma), \\ \frac{\partial(\text{SL}_k\varphi)^\pm}{\partial\nu} = K'_k\varphi \mp \frac{1}{2}\varphi & \text{and} & \frac{\partial(\text{DL}_k\psi)^\pm}{\partial\nu} = T_k\psi & \text{in } H^{-3/2}(\Gamma). \end{cases}$$

We now shall exploit these results to equivalently reformulate the scattering problem as a surface integral equation on  $\Gamma$ .

### 6.3.2 The system of surface integral equations

We first formally establish the system of boundary integral equations that are satisfied by the traces and the normal traces of the solution  $\Gamma$ . The solutions  $u^s$  and  $u$  can be represented in the form

$$u^s(x) = - \int_\Gamma \left( \frac{\partial v}{\partial\nu}(y) \Phi_{k_e}(x, y) - v \frac{\partial \Phi_{k_e}(x, y)}{\partial\nu(y)} \right), \quad \forall x \in \mathbb{R}^d \setminus \Omega. \quad (6.15)$$

and

$$u(x) = \int_\Gamma \left( \frac{\partial u}{\partial\nu}(y) \Phi_{k_i}(x, y) - u \frac{\partial \Phi_{k_i}(x, y)}{\partial\nu(y)} \right), \quad \forall x \in \Omega \quad (6.16)$$

where,  $\nu$  denote the unit normal vector to the boundary  $\Gamma$  directed into the exterior of  $\Omega$ . Let us denote by

$$\varphi := v|_\Gamma = u|_\Gamma \quad \text{and} \quad \psi := \frac{\partial v}{\partial\nu} \Big|_\Gamma = \mu^{-1} \frac{\partial u}{\partial\nu} \Big|_\Gamma \quad (6.17)$$

Equations (6.15)-(6.16) are then equivalent to

$$\begin{aligned} u^s(x) &= -\text{SL}_{k_e}(\psi) + \text{DL}_{k_e}(\varphi), \quad \forall x \in \mathbb{R}^d \setminus \Omega \\ u(x) &= \text{SL}_{k_i}(\mu\psi) - \text{DL}_{k_i}(\varphi), \quad \forall x \in \Omega \end{aligned} \quad (6.18)$$

We then deduce, using the operators  $S_k, K_k, K'_k$  and  $T_k$  defined in Theorem 6.3.1, that

$$\begin{aligned} u|_\Gamma &= \mu S_{k_i}(\psi) - K_{k_i}(\varphi) - \frac{1}{2}\varphi & \text{and} & \quad u^s|_\Gamma &= K_{k_e}(\varphi) - \frac{1}{2}\varphi - S_{k_e}(\psi) \\ \frac{\partial u}{\partial\nu} \Big|_\Gamma &= \mu K'_{k_i}(\psi) - T_{k_i}(\varphi) & \text{and} & \quad \frac{\partial u^s}{\partial\nu} \Big|_\Gamma &= -K'_{k_e}(\psi) + T_{k_e}(\varphi). \end{aligned} \quad (6.19)$$

Combining (6.19) with the transmission conditions (6.7) we get

$$\begin{aligned} (\mu S_{k_i} + S_{k_e})(\psi) - (K_{k_i} + K_{k_e})(\varphi) &= u^i|_{\Gamma}, \\ -(\mu^{-1}T_{k_i} + T_{k_e})(\varphi) + (K'_{k_i} + K'_{k_e})(\psi) &= \frac{\partial u^i}{\partial \nu} \Big|_{\Gamma}. \end{aligned} \quad (6.20)$$

We synthetically write this system of equations as: Find  $(\psi, \varphi) \in H^{-3/2}(\Gamma) \times H^{-1/2}(\Gamma)$  satisfy

$$Z(\mu; k_i, k_e) \begin{bmatrix} \psi \\ \varphi \end{bmatrix} = [U^i], \quad (6.21)$$

where,

$$Z(\mu; k_i, k_e) = \begin{bmatrix} (\mu S_{k_i} + S_{k_e}) & -(K_{k_i} + K_{k_e}) \\ (K'_{k_i} + K'_{k_e}) & -(\mu^{-1}T_{k_i} + T_{k_e}) \end{bmatrix} \quad \text{and} \quad [U^i] = \begin{bmatrix} u^i|_{\Gamma} \\ \frac{\partial u^i}{\partial \nu} \Big|_{\Gamma} \end{bmatrix} \quad (6.22)$$

**Proposition 6.3.3.** *Let  $u^i$  be an incident wave satisfying (6.3). If  $(\psi, \varphi) \in H^{-\frac{3}{2}}(\Gamma) \times H^{-\frac{1}{2}}(\Gamma)$  is solution to (6.21) then  $u := \mu \text{SL}_{k_i}(\psi) - \text{DL}_{k_i}(\varphi)$  and  $v := -\text{SL}_{k_e}(\psi) + \text{DL}_{k_e}(\varphi) + u^i$  are respectively in  $L^2_{\Delta}(\Omega)$  and  $L^2_{\text{loc}, \Delta}(\mathbb{R}^d \setminus \bar{\Omega})$  and satisfy problem (6.1)-(6.5). Conversely, if  $(u, v) \in L^2_{\Delta}(\Omega) \times L^2_{\text{loc}, \Delta}(\mathbb{R}^d \setminus \bar{\Omega})$  satisfy (6.1)-(6.5) then  $(\psi, \varphi)$  defined in (6.17) are in  $H^{-\frac{3}{2}}(\Gamma) \times H^{-\frac{1}{2}}(\Gamma)$  and are solutions to (6.21).*

*Proof.* The establishment of system (6.21) together with Theorem 6.3.2 shows that if  $(u, v) \in L^2_{\Delta}(\Omega) \times L^2_{\text{loc}, \Delta}(\mathbb{R}^d \setminus \bar{\Omega})$  is solution to equation (6.1)-(6.5) then  $(\psi, \varphi) \in H^{-\frac{3}{2}}(\Gamma) \times H^{-\frac{1}{2}}(\Gamma)$  is solution to (6.21). We now prove the opposite implication. Assume that  $(\psi, \varphi) \in H^{-\frac{3}{2}}(\Gamma) \times H^{-\frac{1}{2}}(\Gamma)$  is solution to (6.21). Then, using Theorem 6.3.2,  $u := \mu \text{SL}_{k_i}(\psi) - \text{DL}_{k_i}(\varphi)$  is in  $L^2_{\Delta}(\Omega)$  and satisfies

$$\Delta u + k_i^2 u = 0 \quad \text{in } \Omega,$$

and  $u^s := -\text{SL}_{k_e}(\psi) + \text{DL}_{k_e}(\varphi)$  is in  $L^2_{\text{loc}, \Delta}(\mathbb{R}^d \setminus \bar{\Omega})$  and satisfies

$$\Delta u^s + k_e^2 u^s = 0 \quad \text{in } \mathbb{R}^d \setminus \Omega,$$

together with Sommerfeld radiation condition (since the single and double layer potentials verify Sommerfeld radiation condition, see [46]). Taking the traces and the normal traces on  $\Gamma$  as in Theorem 6.3.2 we end up with

$$\begin{aligned} (\mu S_{k_i} + S_{k_e})(\psi) - (K_{k_i} + K_{k_e})(\varphi) &= u|_{\Gamma} - u^s|_{\Gamma} \\ -(\mu^{-1}T_{k_i} + T_{k_e})(\varphi) + (K'_{k_i} + K'_{k_e})(\psi) &= \mu^{-1} \frac{\partial u}{\partial \nu} \Big|_{\Gamma} - \frac{\partial u^s}{\partial \nu} \Big|_{\Gamma}. \end{aligned} \quad (6.23)$$

Using system (6.20) we arrive at

$$\begin{aligned} u^i|_{\Gamma} &= u|_{\Gamma} - u^s|_{\Gamma} \\ \frac{\partial u^i}{\partial \nu} \Big|_{\Gamma} &= \mu^{-1} \frac{\partial u}{\partial \nu} \Big|_{\Gamma} - \frac{\partial u^s}{\partial \nu} \Big|_{\Gamma}. \end{aligned} \quad (6.24)$$

which finishes the proof.  $\square$

For the case  $\mu \neq -1$  it is well known (see for instance [51]) that (6.21) has a unique solution. We shall consider in the following the case  $\mu = -1$  (and specifically study the problem in space dimension 2). Let us denote the operator  $Z(-1; k_i, k_e)$  by  $Z(k_i, k_e)$ . Then our problem is:  $(\psi, \varphi) \in H^{-\frac{3}{2}}(\Gamma) \times H^{-\frac{1}{2}}(\Gamma)$  and satisfies

$$Z(k_i, k_e) \begin{bmatrix} \psi \\ \varphi \end{bmatrix} = [U^i]. \quad (6.25)$$

Theorem 6.3.3 ensures that solving equation (6.1) in  $L^2_\Delta(\Omega) \times L^2_{\text{loc}, \Delta}(\mathbb{R}^d \setminus \Omega)$  is equivalent to solving equation (6.25) in  $H^{-\frac{3}{2}}(\Gamma) \times H^{-\frac{1}{2}}(\Gamma)$ .

## 6.4 The Fredholm property of the surface integral operator $Z(k_i, k_e)$

Let  $k$  and  $k'$  be two complex valued wave numbers with non-negative imaginary part. For  $k \neq k'$ , we define

$$\begin{aligned} \mathcal{S}\mathcal{L}_{k, k'} &:= \text{SL}_k - \text{SL}_{k'} \\ \mathcal{D}\mathcal{L}_{k, k'} &:= \text{DL}_k - \text{DL}_{k'}. \end{aligned}$$

Then, following [50, Theorem 5.1.2], the operators  $\mathcal{S}\mathcal{L}_{k, k'} : H^{-3/2}(\Gamma) \rightarrow H^2(D)$  and  $\mathcal{D}\mathcal{L}_{k, k'} : H^{-1/2}(\Gamma) \rightarrow H^2(D)$  are continuous for all  $k \neq k'$ . This means in particular that the operators  $S_{i|k_e|} - S_{i|k_i|} : H^{-\frac{3}{2}}(\Gamma) \rightarrow H^{\frac{3}{2}}(\Gamma)$  and  $T_{i|k_e|} - T_{i|k_i|} : H^{-\frac{1}{2}}(\Gamma) \rightarrow H^{\frac{1}{2}}(\Gamma)$  are continuous. Moreover, if we denote by

$$\gamma(k, k') := \frac{k^2 - k'^2}{|k|^2 - |k'|^2}. \quad (6.26)$$

then  $\mathcal{S}\mathcal{L}_{k, k'} + \gamma(k, k')\mathcal{S}\mathcal{L}_{i|k|, i|k'|} : H^{-3/2}(\Gamma) \rightarrow H^3(D)$  and  $\mathcal{D}\mathcal{L}_{k, k'} + \gamma(k, k')\mathcal{D}\mathcal{L}_{i|k|, i|k'|} : H^{-1/2}(\Gamma) \rightarrow H^3(D)$  are continuous for all  $k \neq k'$ ,  $k, k' \in \mathbb{C} \setminus \mathbb{R}^-$ . This means in particular that  $S_{k_e} - S_{k_i} + \gamma(k_e, k_i)(S_{i|k_e|} - S_{i|k_i|}) : H^{-\frac{3}{2}}(\Gamma) \rightarrow H^{\frac{3}{2}}(\Gamma)$  and  $T_{k_e} - T_{k_i} + \gamma(k_e, k_i)(T_{i|k_e|} - T_{i|k_i|}) : H^{-\frac{1}{2}}(\Gamma) \rightarrow H^{\frac{1}{2}}(\Gamma)$  are compact. These results are valid for dimensions 2 and 3. Let us set

$$Z_1(k_i, k_e) = \begin{bmatrix} S_{k_e} - S_{k_i} & 0 \\ 0 & -T_{k_e} + T_{k_i} \end{bmatrix}, \quad Z_2(k_i, k_e) = \begin{bmatrix} 0 & -K_{k_e} - K_{k_i} \\ K'_{k_e} + K'_{k_i} & 0 \end{bmatrix}. \quad (6.27)$$

Then we decompose  $Z(k_i, k_e)$  as

$$Z(k_e, k_i) = -\gamma(k_e, k_i)Z_1(i|k_e|, i|k_i|) + (Z_1(k_e, k_i) + \gamma(k_e, k_i)Z_1(i|k_e|, i|k_i|)) + Z_2(k_e, k_i). \quad (6.28)$$

From previous discussion we first observe that:

**Lemma 6.4.1.** *The operator  $(Z_1(k_e, k_i) + \gamma(k_e, k_i)Z_1(i|k_e|, i|k_i|)) : H^{-\frac{3}{2}}(\Gamma) \times H^{-\frac{1}{2}}(\Gamma) \rightarrow H^{\frac{3}{2}}(\Gamma) \times H^{\frac{1}{2}}(\Gamma)$  is compact.*

We now prove the following:

**Lemma 6.4.2.** *Assume that  $k_e \neq k_i$ , then the operators*

1.  $S_{i|k_e|} - S_{i|k_i|} : H^{-\frac{3}{2}}(\Gamma) \rightarrow H^{\frac{3}{2}}(\Gamma)$ ,
2.  $T_{i|k_e|} - T_{i|k_i|} : H^{-\frac{1}{2}}(\Gamma) \rightarrow H^{\frac{1}{2}}(\Gamma)$



are coercive.

*Proof.* We here just give a sketch of the proof of the first claim, since the proof of the second one is similar. Let  $\psi \in H^{-3/2}(\Gamma)$ . We first consider the problem to find  $u \in H^2(\mathbb{R}^d)$  such that,

$$\begin{cases} (\Delta - k_i^2)(\Delta - k_e^2)u = 0, & \text{in } \mathbb{R}^d \setminus \Gamma \\ [\Delta u]_\Gamma = 0, & \text{on } \Gamma \\ \left[ \frac{\partial(\Delta u)}{\partial \nu} \right] = (k_e^2 - k_i^2)\psi, & \text{on } \Gamma. \end{cases} \quad (6.29)$$

This is equivalent to find  $u \in H^2(\mathbb{R}^d)$  such that

$$\int_{\mathbb{R}^d \setminus \Gamma} (\Delta u - k_i^2 u)(\Delta \bar{v} - k_e^2 \bar{v}) dx = \int_{\Gamma} (k_e^2 - k_i^2)\psi \bar{v} ds(x), \quad \forall v \in H^2(\mathbb{R}^d). \quad (6.30)$$

Since

$$a(u, v) := \int_{\mathbb{R}^d \setminus \Gamma} (\Delta u - k_i^2 u)(\Delta \bar{v} - k_e^2 \bar{v}) dx$$

is continuous on  $H^2(\mathbb{R}^d) \times H^2(\mathbb{R}^d)$  and

$$a(u, u) = \|\Delta u\|_{L^2(\mathbb{R}^d)}^2 + (k_i^2 + k_e^2)\|\nabla u\|_{L^2(\mathbb{R}^d)}^2 + k_i^2 k_e^2 \|u\|_{L^2(\mathbb{R}^d)}^2 \geq C \|u\|_{H^2(\mathbb{R}^d)}^2 \quad (6.31)$$

for some constant  $C > 0$ . Then, using Lax-Milgram theorem, problem (6.30) admits a unique solution  $u \in H^2(\mathbb{R}^d)$ . Moreover,  $\tilde{u} = \mathcal{S}\mathcal{L}_{ik_i, ik_e}\psi = \mathcal{S}\mathcal{L}_{ik_i}\psi - \mathcal{S}\mathcal{L}_{ik_e}\psi$  satisfies the same equations as  $u$  and therefore  $\tilde{u} = u$  and

$$u|_\Gamma = (S_{i|k_e|} - S_{i|k_i|})\psi.$$

Replacing  $v$  by  $u$  in (6.30) we then deduce, using the coercivity of  $a(\cdot, \cdot)$ , that

$$\left| \langle \psi, u|_\Gamma \rangle_{H^{-3/2}(\Gamma) \times H^{3/2}(\Gamma)} \right| = \left| \int_{\Gamma} \psi \bar{u} ds(x) \right| \geq \tilde{C} \|u\|_{H^2(\Omega)}^2, \quad (6.32)$$

with  $\tilde{C} := C/|k_e^2 - k_i^2|$ . Therefore,

$$\left| \langle (S_{i|k_e|} - S_{i|k_i|})\psi, \psi \rangle \right| = \left| \langle \psi, u|_\Gamma \rangle_{H^{-3/2}(\Gamma) \times H^{3/2}(\Gamma)} \right| \geq \tilde{C} \|u\|_{H^2(\Omega)}^2. \quad (6.33)$$

In the other hand, for such  $\tilde{\varphi} \in H^{3/2}(\Gamma)$ , there exists  $\tilde{v} \in H^2(\Omega)$  such that  $\tilde{v}|_\Gamma = \tilde{\varphi}$  and  $\|\tilde{v}\|_{H^2(\Omega)} \leq \theta \|\tilde{\varphi}\|_{H^{3/2}(\Gamma)}$ , where  $\theta$  is independent from  $\tilde{\varphi}$ . We then have that,

$$\begin{aligned} \left| \langle \psi, \tilde{\varphi} \rangle_{H^{-3/2}(\Gamma) \times H^{3/2}(\Gamma)} \right| &= \frac{a(u, \tilde{v})}{k_e^2 - k_i^2} \leq C' \|u\|_{H^2(\Omega)} \|\tilde{v}\|_{H^2(\Omega)} \\ &\leq C' \theta \|u\|_{H^2(\Omega)} \|\tilde{\varphi}\|_{H^{3/2}(\Gamma)}, \quad \forall \tilde{\varphi} \in H^{3/2}(\Gamma). \end{aligned} \quad (6.34)$$

This implies that

$$\|\psi\|_{H^{-3/2}(\Gamma)} \leq C' \theta \|u\|_{H^2(\Omega)} \quad (6.35)$$

We finally deduce from (6.33) and (6.35) the coercivity of  $S_{i|k_e|} - S_{i|k_i|}$ ,

$$\left| \langle (S_{i|k_e|} - S_{i|k_i|})\psi, \psi \rangle \right| \geq C_0 \|\psi\|_{H^{-3/2}(\Gamma)}^2 \quad (6.36)$$

where  $C_0 = \tilde{C}C'^2\theta^2$ . For the coercivity of the operator  $T_{|k_e|} - T_{|k_i|}$ , one follows the same approach considering  $u \in H^2(\mathbb{R}^d)$  such that,

$$\begin{cases} (\Delta - k_i^2)(\Delta - k_e^2)u = 0, & \text{in } \mathbb{R}^d \setminus \Gamma \\ [\Delta u]_\Gamma = (k_e^2 - k_i^2)\varphi, & \text{on } \Gamma \\ \left[ \frac{\partial(\Delta u)}{\partial \nu} \right] = 0, & \text{on } \Gamma. \end{cases} \quad (6.37)$$

for some  $\varphi \in H^{-1/2}(\Gamma)$ . □

An immediate corollary of Lemma 6.4.2 is following result.

**Corollary 6.4.3.** *The operator  $Z_1(|k_i|, |k_e|) : H^{-\frac{3}{2}}(\Gamma) \times H^{-\frac{1}{2}}(\Gamma) \rightarrow H^{\frac{3}{2}}(\Gamma) \times H^{\frac{1}{2}}(\Gamma)$  is invertible.*

The last ingredient for proving the final result is the following compactness result. The proof of this result can be found in [94], but since the proof is simple, we give it her for the reader's convenience.

**Lemma 6.4.4.** *For space dimension  $d = 2$ , the operators  $K_{k_e} + K_{k_i} : H^{-\frac{1}{2}}(\Gamma) \rightarrow H^{\frac{3}{2}}(\Gamma)$  and  $K'_{k_e} + K'_{k_i} : H^{-\frac{3}{2}}(\Gamma) \rightarrow H^{\frac{1}{2}}(\Gamma)$  are compact.*

*Proof.* We shall prove the result for the operator  $K_{k_e} + K_{k_i}$  since the proof for the operator  $K'_{k_e} + K'_{k_i}$  is similar. The Lemma follows from the observation that the operator  $K_k$  can be (locally) represented as an integral operator with a regular  $C^\infty$  kernel if the curve is smooth  $C^\infty$ . Let  $x = (x_1, x_2) \in \Gamma$ . There is a neighborhood  $\mathcal{U}$  of  $x$  such that  $\mathcal{U} \cap \Gamma$  is formulated by a smooth function  $y_2 = \phi(y_1)$ . Let  $\psi \in C_0^\infty(\mathcal{U})$ . We shall show that kernel of the map

$$K_\psi : C^\infty(\Gamma) \rightarrow D'(\Gamma), \quad u \mapsto K_\psi(u) := \psi K_k(\psi u) \quad (6.38)$$

is a smooth function across  $x = y$ . By definition of the operator  $K_\psi$

$$K_\psi(u)(x) = \int_\Gamma \frac{\partial \Phi_k}{\partial \nu_y}(x - y) \psi(x) \psi(y) u(y) dy \quad (6.39)$$

Let us denote

$$\mathcal{K}(x, y) := \left( \frac{\partial \Phi_{k_e}}{\partial \nu_y} + \frac{\partial \Phi_{k_i}}{\partial \nu_y} \right) (x - y) \psi(x) \psi(y)$$

Using the limit form for small argument of the Hankel function (see [1, page 360, 9.1.8]) we have that for mall enough arguments, the Hankel function  $H_0^{(1)}(z)$  is approximated, up to a smooth function, by  $-\frac{1}{2\pi} \ln |z|$ . Hence, this is enough to consider the kernel  $\mathcal{K}(x, y)$  of the map  $u \mapsto K_\psi(u)$  as given by the formula

$$\mathcal{K}(x, y) = -\frac{1}{\pi} \frac{\partial \ln(|x - y|)}{\partial \nu_y} \psi(x) \psi(y). \quad (6.40)$$

Since  $y_2 = \phi(y_1)$ , the normal vector at  $(y_1, y_2)$  is given by  $n_y := (\phi'(y_1), -1)$ , we then have

$$\begin{aligned} \mathcal{K}(x, y) &= -\frac{1}{\pi} \nabla_y \ln(|x - y|) n_y \psi(x) \psi(y) \\ &= -\frac{1}{\pi} \left( \frac{\partial \ln(|x - y|)}{\partial y_1} \phi'(y_1) - \frac{\partial \ln(|x - y|)}{\partial y_2} \right) (x - y) \psi(x) \psi(y) \end{aligned} \quad (6.41)$$

We then arrive that

$$\mathcal{K}(x, y) = \frac{\frac{\phi''(y_1)}{2} + R(x_1, y_1)}{\pi + \pi \left( \frac{\phi(x_1) - \phi(y_1)}{x_1 - y_1} \right)^2} \psi(y_1, \phi(y_1)) \psi(x_1, \phi(x_1)), \quad (6.42)$$

where

$$R(x_1, y_1) = \frac{x_1 - y_1}{2} \int_0^1 (1-t)^2 \phi^{(3)}(y_1 - t(y_1 - x_1)) dt$$

It shows that the kernel  $\mathcal{K}(x, y)$  can be extended (if the surface is smooth) as a smooth function across  $x_1 = y_1$ , which proves the lemma.  $\square$

### 6.4.1 The existence and uniqueness of solutions to the scattering problem

**Theorem 6.4.5.** *Let  $d = 2$ . Assume that  $k_i \neq k_e$  and that  $\text{Im}(\varepsilon) \geq 0$ . Then the operator  $Z(k_i, k_e) : H^{-\frac{3}{2}}(\Gamma) \times H^{-\frac{1}{2}}(\Gamma) \rightarrow H^{\frac{3}{2}}(\Gamma) \times H^{\frac{1}{2}}(\Gamma)$  is invertible and therefore problem (6.7) admits a unique solution in (6.9).*

*Proof.* We recall that the operator  $Z(k_i, k_e)$  can be decomposed as

$$Z(k_e, k_i) = -\gamma(k_e, k_i) Z_1(i|k_e|, i|k_i|) + Z_1(k_e, k_i) + \gamma(k_e, k_i) Z_1(i|k_e|, i|k_i|) + Z_2(k_e, k_i),$$

We first have, using Corollary 6.4.3 that  $-\gamma(k_e, k_i) Z_1(i|k_e|, i|k_i|)$  is invertible. In addition, we deduce from Lemma 6.4.4 that the operator  $Z_2(k_e, k_i)$  is compact. Therefore, combining with Lemma 6.4.1 we obtain that  $Z_1(k_e, k_i) + \gamma(k_e, k_i) Z_1(i|k_e|, i|k_i|) + Z_2(k_e, k_i)$  is compact. This prove that

$$Z(k_e, k_i) = -\gamma(k_e, k_i) Z_1(i|k_e|, i|k_i|) + (Z_1(k_e, k_i) + \gamma(k_e, k_i) Z_1(i|k_e|, i|k_i|)) + Z_2(k_e, k_i)$$

is Fredholm of index 0. In order to conclude the proof, we then only need to prove that the operator  $Z(k_i, k_e)$  is injective. We shall prove that in the case  $u^i = 0$ , equation (6.25) admits a trivial unique solution  $(\psi, \varphi) = (0, 0)$ . Since  $u^i = 0$  then from (6.7),  $u^s$  extended by  $u$  in  $\Omega$  verifies

$$\text{div}(\mu^{-1} \nabla u^s) + k^2 \varepsilon u^s = 0 \quad \text{in } \mathbb{R}^d \quad (6.43)$$

where,  $\mu$  and  $\varepsilon$  are extended by 1 outside  $\Omega$ . Let  $B_R(0)$  be a ball that contains  $\Omega$ . The second Green identity shows that

$$\int_{r=R} \overline{u^s} \frac{\partial u^s}{\partial r} dr - \int_{r=R} u^s \frac{\partial \overline{u^s}}{\partial r} dr = \int_{B_R(0)} \text{div}(\mu^{-1} \nabla u^s) \overline{u^s} - u^s \text{div}(\mu^{-1} \nabla \overline{u^s}) dx.$$

(The integrals on  $|x| = r$  are to be understood in the classical sense sine the scattered field is regular outside  $\Gamma$  by elliptic regularity results) Therefore,

$$\text{Im} \int_{r=R} u^s \frac{\partial \overline{u^s}}{\partial r} dr = - \int_{\Omega} k^2 \text{Im}(\varepsilon) |u^s|^2 dx \leq 0 \quad (6.44)$$

by assumption on  $\varepsilon$ . Combining condition (6.44) with the Sommerfeld radiation condition we deduce that

$$\lim_{R \rightarrow 0} \int_{|x|=R} |u^s(x)|^2 = 0. \quad (6.45)$$

This implies, using the Rellich lemma ([46]) and unique continuation principle we get  $u^s = 0$  in  $\mathbb{R}^d \setminus \Omega$ . Furthermore, from (6.43) we obtain that  $u^s$  also satisfies

$$\Delta u^s + k^2 u^s = 0 \quad \text{in } \mathbb{R}^d.$$

We then get by unique continuation that  $u = u^s = 0$  in  $\Omega$ . Using the definition of  $\psi$  and  $\varphi$  in (6.17) we get that  $\psi = 0$  and  $\varphi = 0$ . This proves the injectivity of  $Z(k_i, k_e)$  and finishes the proof.  $\square$



# Electromagnetic scattering by periodic structures with sign-changing coefficients

## Abstract

This short chapter is extracted from [92], a joint work with Dinh-Liem Nguyen that has been conducted independently from the other parts of this thesis. It is the only place in this manuscript where the full Maxwell problem is considered. We analyze well-posedness of a scattering problem of time-harmonic electromagnetic waves by periodic structures with sign-changing coefficients. Transmission problems for Maxwell's equations with sign-changing coefficients in bounded domains have been studied in [41] using the so-called T-coercivity framework. In this work we generalize such framework for periodic scattering problems relying on an integral equation approach. The periodic scattering problem is formulated by a hypersingular integral equation of Lipmann-Schwinger type. We prove that the integral equation satisfies a Gårding-type estimate, which allows us to establish well-posedness of the problem in the sense of Fredholm.

## Contents

<b>7.1</b>	<b>Introduction</b>	<b>153</b>
<b>7.2</b>	<b>Periodic Electromagnetic Scattering</b>	<b>154</b>
<b>7.3</b>	<b>T-Coercivity Framework</b>	<b>156</b>
<b>7.4</b>	<b>Fredholm Alternative</b>	<b>157</b>

## 7.1 Introduction

Electromagnetic scattering theory in periodic structures is well-known as a topic of great interest in applications, e.g., for the construction and optimization of optical filters, lenses, and beam-splitters in optics (see [96]). There is a large body of applied math literature on forward and inverse scattering problems related to periodic dielectric materials, see for example [5, 91] and references therein. Recently, periodic metamaterials with applications in photonics and optics has received intense research interest in engineering and applied physics community (see for instance [2, 95]). However to the knowledge of the authors there have been only a few works on periodic metamaterials in applied math literature. The paper [84] studies Fredholm property of a scattering problem for periodic metamaterials known as gratings in two dimensions.

We consider in this work an electromagnetic scattering problem for biperiodic gratings consisting of dielectric materials and metamaterials. By biperiodic, we mean that the grating is periodic in the, say,  $x_1$ - and  $x_2$ -direction, while it is bounded in the  $x_3$ -direction. This is modeled by a periodic scattering problem for 3D Maxwell's equations with sign-changing coefficients.

Transmission Maxwell problems with sign-changing coefficients in bounded domains have been recently studied in [19] by the so-called  $T$ -coercivity framework. In this article we generalize such framework for periodic scattering problems relying on an integral equation approach. It is known in [75, 91] that scattering problems for Maxwell's equations can be formulated as a hypersingular integral equation of Lipmann-Schwinger type. We aim to prove a Gårding-type estimate for the integral equation with sign-changing coefficients. The idea is to investigate such estimates for the integral equation in a truncation of the unit cell. This, roughly speaking, enables the use of variational formulation of the integral operators which allows us to incorporate the idea of  $T$ -coercivity. It also turns out that one needs similar assumptions as in [19] when applying the  $T$ -coercivity framework.

The integral equation approach in this work has some advantages in the sense that Fredholm property obtained is valid at Rayleigh frequencies which is typically excluded in variational approach. It further avoids technical complication that one might have when treating boundary terms, with Calderón maps, perturbed by the abstract operator  $T$  in the  $T$ -coercivity framework.

*Notation:* Let  $\mathcal{O}$  be a bounded domain (connected and open) with Lipschitz boundary  $\partial\mathcal{O}$ . We indistinctly denote by  $\langle \cdot, \cdot \rangle$  the inner products of  $L^2(\mathcal{O})$  and  $(L^2(\mathcal{O}))^3$  and by  $\| \cdot \|$  the associated norms. We set  $L_p^\infty(\mathbb{R}^3) = \{v \in L^\infty(\mathbb{R}^3) : v \text{ is } 2\pi\text{-periodic in } x_1 \text{ and } x_2\}$ ,  $H_\alpha(\text{curl}, \mathcal{O})$  is the closure with respect to the norm  $\| \cdot \| + \| \text{curl} \cdot \|$  of space of smooth functions which is  $\alpha$ -quasi-periodic in  $x_1$  and  $x_2$ .

## 7.2 Periodic Electromagnetic Scattering

We consider scattering of time-harmonic electromagnetic waves from a diffraction grating consisting of dielectric materials and metamaterials. The electric field  $E$  and the magnetic field  $H$  are governed by the time-harmonic Maxwell's equations at frequency  $\omega > 0$  in  $\mathbb{R}^3$

$$\text{curl } H + i\omega\varepsilon E = 0, \quad \text{curl } E - i\omega\mu H = 0 \quad \text{in } \mathbb{R}^3, \quad (7.1)$$

where the electric permittivity  $\varepsilon$ , the magnetic permeability  $\mu$  are real-valued functions in  $L_p^\infty(\mathbb{R}^3)$ . We assume that there are positive constants  $\varepsilon_0$  and  $\mu_0$  such that  $\varepsilon = \varepsilon_0$ ,  $\mu = \mu_0$  outside the grating. We define the wave number  $k = \omega(\varepsilon_0\mu_0)^{1/2}$ .

The grating is illuminated by an electromagnetic plane wave with wave vector  $d = (d_1, d_2, d_3) \in \mathbb{R}^3$ ,  $d_3 \neq 0$  such that  $d \cdot d = k^2$ . The polarizations  $p, s \in \mathbb{R}^3$  of the incident wave satisfy  $p \cdot d = 0$  and  $s = 1/(\omega\varepsilon_0)(p \times d)$ . With these definitions, the incident plane waves  $E^i$  and  $H^i$  are given by

$$E^i = se^{id \cdot x}, \quad H^i = pe^{id \cdot x}, \quad x \in \mathbb{R}^3. \quad (7.2)$$

For  $d = (d_1, d_2, d_3) \in \mathbb{R}^3$  defined in (7.2), we set  $\alpha = (\alpha_1, \alpha_2, 0) = (d_1, d_2, 0)$ . Then a function  $u : \mathbb{R}^3 \rightarrow \mathbb{C}^3$  is then called  $\alpha$ -quasi-periodic if, for all  $x = (x_1, x_2, x_3)^\top \in \mathbb{R}^3$ ,  $n = (n_1, n_2, 0)^\top \in \mathbb{Z}^3$ ,

$$u(x_1 + 2\pi n_1, x_2 + 2\pi n_2, x_3) = e^{2\pi i \alpha \cdot n} u(x_1, x_2, x_3).$$

Note that the incident fields  $E^i, H^i$  defined in (7.2) are  $\alpha$ -quasi-periodic functions. The relative material parameters are defined by  $\varepsilon_r = \varepsilon/\varepsilon_0$ ,  $\mu_r = \mu/\mu_0$ . We wish to reformulate (7.1) in terms of the scattered field  $u$ , defined by  $u := H - H^i$ . Since, by construction,  $\text{curl } \text{curl } H^i - k^2 H^i = 0$ , eliminating the electric field  $E$  from (7.1), and subtracting the latter equation implies that

$$\text{curl}(\varepsilon_r^{-1} \text{curl } u) - k^2 \mu_r u = \text{curl}(q \text{curl } H^i) + k^2 p H^i \quad \text{in } \mathbb{R}^3, \quad (7.3)$$

where the contrasts  $q, p$  are defined by

$$q := 1 - \varepsilon_r^{-1}, \quad p := \mu_r - 1.$$

It is required that  $u$  also be  $\alpha$ -quasi-periodic in  $x_1$  and  $x_2$  and that it admits a Rayleigh expansion radiation condition of the form (see, e.g., [5, 91])

$$u(x) = \sum_{n \in \mathbb{Z}^2} c_n^\pm e^{i(\alpha_n \cdot x \pm \beta_n(x_3 \mp h))} \quad \text{for } x_3 \gtrless \pm h, \quad (7.4)$$

where  $c_n^\pm \in \mathbb{C}^3$ ,  $\alpha_n = (\alpha_1 + n_1, \alpha_2 + n_2, 0)$ ,  $\beta_n = \sqrt{k^2 - |\alpha_n|^2}$  with the branch of the square root chosen such that  $\text{Re}(\beta_n) \geq 0$  and  $\text{Im}(\beta_n) \geq 0$ , and  $h > \sup\{|x_3| : (x_1, x_2, x_3)^\top \in \text{supp}(q) \cup \text{supp}(p)\}$ . Now we define

$$\Omega := (-\pi, \pi)^2 \times \mathbb{R}, \quad \bar{D} := [\text{supp}(q) \cup \text{supp}(p)] \cap \Omega.$$

Recall that  $\varepsilon_r, \mu_r$  belong to  $L_p^\infty(\mathbb{R}^3)$ . We further make the following assumption on  $\varepsilon_r$  and  $\mu_r$  and  $D$  for our analysis.

**Assumption 7.2.1.** *We assume that the support  $D$  is a bounded and simply connected domain in  $\mathbb{R}^3$  such that its Lipschitz boundary  $\partial D$  is connected, and that  $\varepsilon_r^{-1}, \mu_r^{-1}$  belong to  $L_p^\infty(\mathbb{R}^3)$ .*

The problem (7.3)–(7.4) can be reduced to one period  $\Omega$  due to its periodicity. We now consider a more general problem as follows: Given  $f, g \in (L^2(D))^3$ , find  $u : \Omega \rightarrow \mathbb{C}^3$  such that

$$\text{curl}(\varepsilon_r^{-1} \text{curl} u) - k^2 \mu_r u = \text{curl} f + k^2 g \quad \text{in } \Omega, \quad (7.5)$$

and the radiation condition (7.4). To study the well-posedness of the latter problem we reformulate it as an integral equation of Lippmann-Schwinger type. This has been done in [75] for scattering problems in bounded inhomogeneous media. The following lines follow from [91].

We denote by  $G_k$  the  $\alpha$ -quasi-periodic Green's function to the Helmholtz equation in  $\mathbb{R}^3$ . From [5] we know that  $G_k(x) = \exp(ik|x|)/(4\pi|x|) + \Psi_k(x)$  for  $x \neq 0$  where  $\Psi_k$  is an analytic function solving the Helmholtz equation  $\Delta \Psi_k + k^2 \Psi_k = 0$  in  $(-2\pi, 2\pi)^2 \times \mathbb{R}$ . Now, for any  $R > 0$ , the truncation  $\Omega_R$  of the unit cell  $\Omega$  is defined by

$$\Omega_R = (-\pi, \pi)^2 \times (-R, R), \quad \Gamma_{\pm R} = (-\pi, \pi)^2 \times \{\pm R\}.$$

We define the volume potential  $V_k$  by

$$(V_k f)(x) = \int_D G_k(x-y) f(y) dy, \quad x \in \Omega, \quad (7.6)$$

for  $f \in L^2(D)$ . The following lemma is the main ingredient for the integral equation formulation which is also necessary for our analysis later on. Its proof can be found in chapter 3 of [91].

**Lemma 7.2.2.** *The volume potential  $V_k$  defined in (7.6) is bounded from  $L^2(D)$  into  $H_\alpha^2(\Omega_R)$  for all  $R > 0$ . The potentials  $A_k = \text{curl} V_k$  and  $B_k = (k^2 + \nabla \text{div}) V_k$  are bounded from  $(L^2(D))^3$  into  $H_\alpha(\text{curl}, \Omega_R)$  for all  $R > 0$ . Further, for  $g \in (L^2(D))^3$ ,  $A_k g$  and  $B_k g$  are the unique solution to*

$$\begin{aligned} \int_\Omega (\text{curl} A_k g \cdot \text{curl} \bar{\psi} - k^2 A_k g \cdot \bar{\psi}) dx &= \int_D g \cdot \text{curl} \bar{\psi} dx, \\ \int_\Omega (\text{curl} B_k g \cdot \text{curl} \bar{\psi} - k^2 B_k g \cdot \bar{\psi}) dx &= k^2 \int_D g \cdot \bar{\psi} dx, \end{aligned}$$

for all  $\psi \in H_\alpha(\text{curl}, \Omega)$  with compact support, and additionally the radiation condition (7.4).

The scattering problem (7.4)–(7.5) is equivalent to the integral equation (see [91] for more details)

$$u - A_k(q \text{curl} u + f) - B_k(pu + g) = 0 \quad \text{in } \Omega_h. \quad (7.7)$$



### 7.3 $T$ -Coercivity Framework

In this section we study the framework of  $T$ -coercivity in quasi-periodic function spaces. For  $\xi \in L_p^\infty(\mathbb{R}^3)$ , we define

$$\begin{aligned}\mathcal{S}_\alpha(\Omega_h) &= \left\{ v \in H_\alpha^1(\Omega_h) : \int_{\partial\Omega_h} v = 0 \right\}, \\ \mathcal{V}_\alpha(\xi, \Omega_h) &= \{ w \in H_\alpha(\text{curl}, \Omega_h) : \langle \xi w, \nabla \psi \rangle = 0 \text{ for all } \psi \in \mathcal{S}_\alpha \}, \\ \mathcal{X}_\alpha(\xi, \Omega_h) &= \{ u \in H_\alpha(\text{curl}, \Omega_h) : \text{div}(\xi u) = 0 \text{ in } \Omega_h, n \times u = 0 \text{ on } \partial\Omega_h \}.\end{aligned}$$

It is well-known that  $\langle \nabla \cdot, \nabla \cdot \rangle$  defines an inner product on  $\mathcal{S}_\alpha(\Omega_h)$  with an equivalent norm given by  $\|u\|_{\mathcal{S}_\alpha(\Omega_h)} = \|\nabla u\|$ . Let  $X$  be  $\mathcal{V}_\alpha(1, \Omega_h)$  or  $\mathcal{X}_\alpha(1, \Omega_h)$ . Proof of the following lemmas can be done similarly as in the free space case, see [3, 19].

**Lemma 7.3.1.** *The embedding of  $X$  in  $(L^2(\Omega_h))^3$  is compact. Further, there exists a positive constant  $C$  such that  $\|w\| \leq C\|\text{curl} w\|$  for all  $w \in X$ . Thus,  $\langle \text{curl} \cdot, \text{curl} \cdot \rangle$  defines an inner product on  $X$  with an equivalent norm given by  $\|w\|_X = \|\text{curl} w\|$ .*

**Lemma 7.3.2.** *A function  $u \in H_\alpha(\text{div}, \Omega_h)$  satisfies*

$$\text{div} u = 0 \quad \text{in } \Omega_h \quad \text{and} \quad \int_{\partial\Omega_h} (n \cdot u) \bar{\phi} \, ds = 0 \quad \text{for all } \phi \in \mathcal{S}_\alpha(\Omega_h) \quad (7.8)$$

*if and only if there exists a vector potential  $\psi$  in  $\mathcal{X}_\alpha(1, \Omega_h)$  such that  $u = \text{curl} \psi$ . Further, this function  $\psi$  is unique.*

As in [19] the following assumption is important for the  $T$ -coercivity framework. We also refer to the cited paper for a detailed investigation on geometric configurations related to this assumption.

**Assumption 7.3.3.** ( $H^{\varepsilon_r}$ ): *There exists an isomorphism  $T^{\varepsilon_r}$  in  $H_0^1(\Omega_h)$  and a positive constant  $C$  such that*

$$|\langle \varepsilon_r \nabla u, \nabla T^{\varepsilon_r} u \rangle| \geq C \|\nabla u\|^2, \quad \text{for all } u \in H_0^1(\Omega_h).$$

( $H^{\mu_r}$ ): *There exists an isomorphism  $T^{\mu_r}$  in  $\mathcal{S}_\alpha(\Omega_h)$  and a positive constant  $C$  such that*

$$|\langle \mu_r \nabla u, \nabla T^{\mu_r} u \rangle| \geq C \|\nabla u\|^2, \quad \text{for all } u \in \mathcal{S}_\alpha(\Omega_h).$$

**Lemma 7.3.4.** *Suppose the assumptions ( $H^{\varepsilon_r}$ ) and ( $H^{\mu_r}$ ) hold true, there exist isomorphisms  $T$  in  $\mathcal{V}_\alpha(\mu_r, \Omega_h)$  and  $\tilde{T}$  in  $\mathcal{X}_\alpha(1, \Omega_h)$  such that*

$$\begin{aligned}\langle \varepsilon_r^{-1} \text{curl} u, \text{curl} T v \rangle &= \langle \varepsilon_r^{-1} \text{curl} T u, \text{curl} v \rangle = \langle \text{curl} u, \text{curl} v \rangle, \quad \text{for all } u, v \in \mathcal{V}_\alpha(\mu_r, \Omega_h) \\ \langle \mu_r^{-1} \text{curl} u, \text{curl} \tilde{T} v \rangle &= \langle \mu_r^{-1} \text{curl} \tilde{T} u, \text{curl} v \rangle = \langle \text{curl} u, \text{curl} v \rangle, \quad \text{for all } u, v \in \mathcal{X}_\alpha(1, \Omega_h).\end{aligned}$$

We refer to [19] for the proof of Lemma 7.3.4.

**Lemma 7.3.5.** *Suppose the assumption ( $H^{\mu_r}$ ) holds true.  $\mathcal{V}_\alpha(\mu_r, \Omega_h)$  is compactly embedded in  $(L^2(\Omega_h))^3$ .*

*Proof.* Assume that  $(u_n)$  is a bounded sequence in  $\mathcal{V}_\alpha(\mu_r, \Omega_h)$ . Since  $\mu_r u_n$  satisfies (7.8), Lemma 7.3.2 implies that there exists  $w_n \in \mathcal{X}_\alpha(1, \Omega_h)$  such that  $u_n = \mu_r^{-1} \text{curl} w_n$ . Recall that  $\mu_r^{-1} \in L^\infty(\Omega_h)$ , it is sufficient now to show that  $(\text{curl} w_n)$  has a subsequence that converges in  $(L^2(\Omega_h))^3$ .

Since  $(u_n)$  is a bounded sequence in  $\mathcal{V}_\alpha(\mu_r, \Omega_h)$ , the equation  $u_n = \mu_r^{-1} \operatorname{curl} w_n$  implies that  $(w_n)$  is also bounded in  $\mathcal{X}_\alpha(1, \Omega_h)$ . The compact embedding  $\mathcal{X}_\alpha(1, \Omega_h) \subset (L^2(\Omega_h))^3$  deduces that  $(w_n)$  has a subsequence, still denoted  $(w_n)$ , that converges in  $(L^2(\Omega_h))^3$ . Therefore,  $(w_n)$  is a Cauchy sequence in  $(L^2(\Omega_h))^3$ . Now, for  $\phi \in \mathcal{X}_\alpha(1, \Omega_h)$ , we have  $\langle \operatorname{curl}(\mu_r^{-1} \operatorname{curl} w_{nm}), \phi \rangle = \langle \mu_r^{-1} \operatorname{curl} w_{nm}, \operatorname{curl} \phi \rangle$ , where  $w_{nm} = w_n - w_m$ . Taking  $\phi = \tilde{T}w_{nm}$ , where  $\tilde{T}$  is the isomorphism from Lemma 7.3.4, we find that

$$\langle \operatorname{curl} u_{nm}, \tilde{T}w_{nm} \rangle = \langle \mu_r^{-1} \operatorname{curl} w_{nm}, \operatorname{curl} \tilde{T}w_{nm} \rangle = \|\operatorname{curl} w_{nm}\|^2. \quad (7.9)$$

Furthermore, we have that  $(\tilde{T}w_n)$  is also a Cauchy sequence in  $(L^2(\Omega_h))^3$  and  $(\operatorname{curl} u_n)$  is bounded in  $(L^2(\Omega_h))^3$ . Therefore, we obtain from (7.9) that  $(\operatorname{curl} w_n)$  is a Cauchy sequence in  $(L^2(\Omega_h))^3$ .  $\square$

To prove the Fredholm property for the integral equation we need the following Hodge decomposition. We notice that for the case that  $\mu_r$  does not change sign this result is classical, see [89].

**Lemma 7.3.6.** *Suppose the assumption  $(H^{\mu_r})$  holds true, we have*

$$H_\alpha(\operatorname{curl}, \Omega_h) = \mathcal{V}_\alpha(\mu_r, \Omega_h) \oplus \nabla \mathcal{S}_\alpha(\Omega_h).$$

Furthermore, for all  $u = u_0 + \nabla p$ ,

$$\|u\|_{H_\alpha(\operatorname{curl}, \Omega_h)}^2 = \|\operatorname{curl} u_0\|^2 + \|\nabla p\|^2.$$

*Proof.* For  $u \in H_\alpha(\operatorname{curl}, \Omega_h)$ , we know from the assumption  $(H^{\mu_r})$  that there exists a unique  $p \in \mathcal{S}_\alpha(\Omega_h)$  such that

$$\int_{\Omega_h} \mu_r \nabla p \cdot \nabla \bar{\varphi} \, dx = \int_{\Omega_h} \mu_r u \cdot \nabla \bar{\varphi} \, dx, \quad \text{for all } \varphi \in \mathcal{S}_\alpha(\Omega_h).$$

This means that  $\int_{\Omega_h} \mu_r (u - \nabla p) \cdot \nabla \bar{\varphi} \, dx = 0$ , for all  $\varphi \in \mathcal{S}_\alpha(\Omega_h)$ , or  $u_0 := u - \nabla p$  belongs to  $\mathcal{V}_\alpha(\mu_r, \Omega_h)$ .

Now it remains to prove that

$$\mathcal{V}_\alpha(\mu_r, \Omega_h) \cap \nabla \mathcal{S}_\alpha(\Omega_h) = \{0\}.$$

Let  $w \in \mathcal{V}_\alpha(\mu_r, \Omega_h) \cap \nabla \mathcal{S}_\alpha(\Omega_h)$ , then there exists  $\psi \in \mathcal{S}_\alpha(\Omega_h)$  such that  $w = \nabla \psi$ , and  $\langle \mu_r w, \nabla \varphi \rangle = 0$  for all  $\varphi \in \mathcal{S}_\alpha(\Omega_h)$ . Choosing  $\varphi = T^{\mu_r} \psi$  implies that

$$\langle \mu_r \nabla \psi, \nabla T^{\mu_r} \psi \rangle = 0.$$

Again, from the assumption  $(H^{\mu_r})$  we have  $0 = |\langle \mu_r \nabla \psi, \nabla T^{\mu_r} \psi \rangle| \geq C \|\nabla \psi\|^2$  which leads to that  $\nabla \psi = 0$  or  $w = 0$ .  $\square$

## 7.4 Fredholm Alternative

**Theorem 7.4.1.** *Suppose that the assumptions  $(H^{\varepsilon_\tau})$  and  $(H^{\mu_\tau})$  hold true. Then there exist an isomorphism  $\mathbb{T}$ , a compact operator  $K$  in  $H_\alpha(\operatorname{curl}, \Omega_h)$  and a positive constant  $C$  such that, for all  $u \in H_\alpha(\operatorname{curl}, \Omega_h)$ ,*

$$\operatorname{Re} \langle u - A_k(q \operatorname{curl} u) - B_k(pu), \mathbb{T}u \rangle_{H_\alpha(\operatorname{curl}, \Omega_h)} \geq C \|u\|_{H_\alpha(\operatorname{curl}, \Omega_h)}^2 + \operatorname{Re} \langle Ku, u \rangle_{H_\alpha(\operatorname{curl}, \Omega_h)}.$$

*Proof.* Let  $u \in H_\alpha(\text{curl}, \Omega_h)$ . Hodge decomposition implies that  $u = u_0 + \nabla p$  where  $u_0 \in \mathcal{V}_\alpha(\mu_r, \Omega_h)$ , and  $p \in \mathcal{S}_\alpha(\Omega_h)$ . Let us consider the operator  $\mathbb{T} : H_\alpha(\text{curl}, \Omega_h) \rightarrow H_\alpha(\text{curl}, \Omega_h)$  defined by

$$u = (u_0 + \nabla p) \mapsto (Tu_0 + \nabla T^{\mu_r} p).$$

where  $T$  and  $T^{\mu_r}$  are defined in Lemma 7.3.4 and Assumption 7.3.3, respectively. It is easy to check that  $\mathbb{T}$  is an isomorphism. Now we define  $w$  by

$$\begin{aligned} w &= A_i(q \text{curl } u) + B_i(pu) \\ &= \text{curl} \int_D G_i(\cdot - y) q(y) \text{curl } u(y) \, dy + (-1 + \nabla \text{div}) \int_D G_i(\cdot - y) p(y) u(y) \, dy. \end{aligned}$$

From Lemma 7.2.2 we have, for all  $v \in H_\alpha(\text{curl}, \Omega_h)$ ,

$$\langle w, v \rangle_{H_\alpha(\text{curl}, \Omega_h)} + \int_{\Gamma_{\pm h}} (n \times \text{curl } w) \cdot (n \times \bar{v}) \times n \, ds = \int_D (q \text{curl } u \cdot \text{curl } \bar{v} \, dx - pu \cdot \bar{v}) \, dx$$

which implies that

$$\langle u - w, v \rangle_{H_\alpha(\text{curl}, \Omega_h)} = \int_{\Omega_h} (\varepsilon_r^{-1} \text{curl } u \cdot \text{curl } \bar{v} + \mu_r u \cdot \bar{v}) \, dx + \int_{\Gamma_{\pm h}} (n \times \text{curl } w) \cdot (n \times \bar{v}) \times n \, ds.$$

Choosing  $v = \mathbb{T}u$  we have

$$\begin{aligned} \langle u - w, \mathbb{T}u \rangle_{H_\alpha(\text{curl}, \Omega_h)} &= \int_{\Omega_h} (\varepsilon_r^{-1} \text{curl } u_0 \cdot \text{curl } \overline{Tu_0} + \mu_r u_0 \cdot \overline{Tu_0} + \mu_r \nabla p \cdot \overline{\nabla T^{\mu_r} p}) \, dx \\ &\quad + \int_{\Gamma_{\pm h}} (n \times \text{curl } w) \cdot (n \times \overline{\mathbb{T}u}) \times n \, ds \geq C \|u\|_{H_\alpha(\text{curl}, \Omega_h)}^2 + \langle K_1 u, u \rangle_{H_\alpha(\text{curl}, \Omega_h)}. \end{aligned} \quad (7.10)$$

Here, due to Lemma 7.3.5 and the smoothness of  $w$  on  $\Gamma_{\pm h}$ ,  $K_1$  defined by

$$\langle K_1 u, u \rangle_{H_\alpha(\text{curl}, \Omega_h)} = \int_{\Omega_h} \mu_r u_0 \cdot \overline{Tu_0} \, dx + \int_{\Gamma_{\pm h}} (n \times \text{curl } w) \cdot (n \times \overline{\mathbb{T}u}) \times n \, ds$$

is a compact operator in  $H_\alpha(\text{curl}, \Omega_h)$ . From (7.10) we have

$$\begin{aligned} \langle u - A_k(q \text{curl } u) - B_k(pu), \mathbb{T}u \rangle_{H_\alpha(\text{curl}, \Omega_h)} &\geq C \|u\|_{H_\alpha(\text{curl}, \Omega_h)}^2 + \langle K_1 u, u \rangle_{H_\alpha(\text{curl}, \Omega_h)} \\ &\quad - \langle (A_k - A_i)(q \text{curl } u) + (B_k - B_i)(pu), \mathbb{T}u \rangle_{H_\alpha(\text{curl}, \Omega_h)}. \end{aligned} \quad (7.11)$$

Recall that the Green function  $G_k(x, y) = \Phi_k(x, y) + \Psi_k(x, y)$  where  $\Phi_k(x, y) = \exp(ik|x - y|)/(4\pi|x - y|)$  and  $\Psi_k(x, y)$  is an analytic function. Hence thanks to the smoothness of  $\Phi_k(x, y) - \Phi_i(x, y)$ , the last term in (7.11) can be written as  $\langle K_2 u, u \rangle_{H_\alpha(\text{curl}, \Omega_h)}$  with a compact operator  $K_2$  in  $H_\alpha(\text{curl}, \Omega_h)$  (see [75, 91]). The latter argument completes the proof.  $\square$

**Corollary 7.4.2.** *For  $f, g \in (L^2(D))^3$ , the problem of finding  $u \in H_\alpha(\text{curl}, \Omega_h)$  such that  $u - A_k(q \text{curl } u) - B_k(pu) = A_k f + B_k g$  in  $H_\alpha(\text{curl}, \Omega_h)$ , satisfies the Fredholm alternative, i.e., uniqueness of solution implies existence of solution.*

*Proof.* The Corollary is directly implied by Theorem 7.4.1 and the use of the  $T$ -coercivity method (see [19]): for  $\mathbb{T}$  being an isomorphism on  $H_\alpha(\text{curl}, \Omega_h)$ , the two following problems  $\mathcal{P}$  and  $\mathcal{P}_\mathcal{T}$  are equivalent:

( $\mathcal{P}$ ) : Seek  $u \in H_\alpha(\text{curl}, \Omega_h)$  such that

$$\langle u - A_k(q \text{curl } u) - B_k(pu), v \rangle_{H_\alpha(\text{curl}, \Omega_h)} = \langle A_k f + B_k g, v \rangle_{H_\alpha(\text{curl}, \Omega_h)}, \quad \forall v \in H_\alpha(\text{curl}, \Omega_h),$$

( $\mathcal{P}_{\mathcal{T}}$ ) : Seek  $u \in H_\alpha(\text{curl}, \Omega_h)$  such that

$$\langle u - A_k(q \text{curl } u) - B_k(pu), \mathbb{T}v \rangle_{H_\alpha(\text{curl}, \Omega_h)} = \langle A_k f + B_k g, \mathbb{T}v \rangle_{H_\alpha(\text{curl}, \Omega_h)}, \quad \forall v \in H_\alpha(\text{curl}, \Omega_h).$$

□



# Conclusion and Perspectives

---

We presented in this manuscript some mathematical and numerical developments for solutions to forward and inverse scattering problems from locally perturbed periodic layers. This problem has connexions with many real life applications like photonics structures, optical fibers, gratings, etc. The main motivation was to provide numerical solutions for non destructive testings of these periodic structures. In this problem, some incident (electromagnetic) wave is used to illuminate the structure and the response of the structure is measured at some distance from the structure. One is interested in identifying the presence of defects and if possibly find the location and the shape of these defects. We considered the problem at a fixed frequency meaning that the wave is harmonic in time with a known pulsation. For most of the thesis, only a simplified scalar problem is considered (modeling for instance some specific polarizations of the electromagnetic wave or some special configuration of acoustic waves).

We first analyzed the forward scattering problem and designed a numerical method that can simulate the scattering of waves from unbounded periodic layers that would contain a defect. We revisited in Chapter 2 the analysis of the scattering problem for unbounded layers and the conditions that would ensure the non existence of guided waves. We showed in particular that if absorption is present in the layer, then one can avoid the strict monotonicity assumption for the real part of the refractive index. This allows in particular considering the problem where the layer is embedded in the free space. The employed technique can easily be extended to Maxwell problem based for instance on [58]. A more challenging perspective would be to find a unified theory that allows to treat the problem in its general setting, including the case where propagative modes would be present (using for instance limiting absorption principles in the manner of the analysis in [54]).

We proposed a spectral numerical method to solve the direct scattering problem. Using the Floquet-Bloch transform, the locally perturbed problem is transformed into coupled quasi-periodic problems. Our method is then based on discretizing first the Floquet-Bloch parameter and then apply to the semi-discrete problems a spectral discretization method for the Lipmann-Schwinger formulation of the problem. The convergence analysis for this methodology is performed only for wave numbers with non vanishing imaginary part (that would model some absorption in the media). Obtaining some weaker convergence results for the non absorbing case can be obtained as in [87] if some regularity and decay properties on the solution is assumed. A broader perspective for this part would be to design a numerical method that can handle the propagative modes. One possibility would be to construct suitable Dirichlet-to-Neumann maps as in [68, 54, 55, 16]. We also quote the method proposed in [23, 15] that handle open waveguides as an alternative numerical approach or the method in [4] based on surface integral equations. We also think that one could extend our method to treat the case with propagative modes if one can adapt the quadrature rule at the Floquet-Bloch parameters where some guided modes would appear.

Our major contribution is on the inverse problem and is related to the design of a sampling method capable of directly reconstruct the defect see Chapters 4-5. The method that we propose is inspired by the Differential LSM introduced in [13, 12]. In the latter, one relies on differential measurements (measurements for the configuration with defects and measurements for the

configuration without defects) to construct an indicator function that characterises the defect geometry. The construction of this indicator function relies on the GLSM formalism and the link with solutions to the over-mentioned ITP. For our problem we exploited the information that the background is periodic with known periodicity length to develop an indicator function that does not use differential measurements. We construct from the measurement operator a so-called single Floquet-Bloch mode measurement operator. This operator plays the role of the background operator in the case of differential measurements. The analysis of sampling methods associated with these single Floquet-Bloch modes operators for the perturbed periodic problem constitutes our main theoretical contribution. The numerical results illustrate how the indicator function is capable of identifying the defect even though the sampling methods fail to capture the (complex) structure of the periodic background. This constitutes the main advantage of our differential imaging functional. Indeed the perspectives for this part are very large since this is the first time such a method has been proposed in the literature. The short term perspective would be to remove the assumption on the geometry of the media where we assumed that the defect does not intersect the background media. To remove this assumption one should analyse a new type of interior transmission problem that couples the interior transmission problem for one Floquet-Bloch mode with scattering problems for the other modes. In our analysis, we directly solve this problem using the simplification on the geometry. It is also desirable to remove the technical assumption that consists into truncating the infinite periodic domain into  $M$  periods with periodicity conditions. As we indicated in our analysis this is mainly justified with the need to rely on some properties for the discrete Floquet-Bloch transform that does not generalize easily to the continuous case. Another perspective in this vein is to consider differential imaging functional using the restricted measurement operators for several Floquet-Bloch modes instead of assuming that we truncate the physical domain. Other perspectives would include studying the case where only propagating plane waves can be used but this seems to be challenging from the theoretical point of view. A more accessible perspective is probably the case where point sources are used instead of plane waves. Numerical results show that our differential imaging method behaves better in this configuration. Indeed a natural perspective for all this would be to consider the 3D Maxwell problem.

# Abstract theoretical foundations of the sampling methods

---

We summarize in this appendix the main theoretical results of the literature that allows us to establish the sampling methods for the locally perturbed periodic domains. We here follow the formulations of the theorems as given in [30] with obvious simplifications in the case of Hilbert spaces which is sufficient for our problem.

## A.1 Main theorem for the $F_{\sharp}$ method

This method is one of two versions of the Factorization method, which was introduced by A.Kirsch in [71] and detailed in [76]. Consider a separable Hilbert space  $X$  and an operator  $N : X \rightarrow X$  such that factorization

$$N = (\mathcal{H})^* T \mathcal{H} \quad (\text{A.1})$$

where  $\mathcal{H} : X \rightarrow Y$  and  $T : Y \rightarrow Y$  are bounded operators and  $Y$  is also a separable Hilbert space. In the following we shall denote by  $(\cdot, \cdot)$  the scalar product and by  $\|\cdot\|$  the associated norm in  $X$  and use the same notation for  $Y$  since there is no risk of confusion.

We first make the following assumption on the operator  $T$ .

**Assumption A.1.1.** *We assume that operator  $T : Y \rightarrow Y$  satisfies*

$$\text{Im}(T\varphi, \varphi) \geq 0 \quad \text{or} \quad \text{Im}(T\varphi, \varphi) \leq 0 \quad (\text{A.2})$$

for all  $\varphi \in \overline{\mathcal{R}(\mathcal{H})}$ ,  $\text{Re } T = T_0 + C$  where  $C$  is compact on  $\overline{\mathcal{R}(\mathcal{H})}$  and

$$(T_0\varphi, \varphi) \geq \alpha \|\varphi\|^2 \quad (\text{A.3})$$

for all  $\varphi \in \overline{\mathcal{R}(\mathcal{H})}$  and some  $\alpha > 0$ . Moreover, we assume that one of the following assumptions holds:

- (i)  $T$  is injective on  $\overline{\mathcal{R}(\mathcal{H})}$ ;
- (ii)  $\text{Im}(T)$  is injective on  $\overline{\mathcal{R}(\mathcal{H})} \cap \ker \text{Re } T$ .

Define the operator  $N_{\sharp}$  as

$$N_{\sharp} := \frac{1}{2}|N + N^*| + \frac{1}{2}|N - N^*|.$$

Then the main theorem for the factorization method can be formulated as:

**Theorem A.1.2.** *Let  $N$  be given by (A.1) and assume that  $H : X \rightarrow Y$  is compact and injective and that  $T$  satisfies Assumption A.1.1. Then*

$$N_{\sharp} = (\mathcal{H})^* T_{\sharp} \mathcal{H} \quad (\text{A.4})$$

where  $T_{\sharp} : Y \rightarrow Y$  is self-adjoint and satisfies the coercivity property on  $\overline{\mathcal{R}(\mathcal{H})}$

$$(T_{\sharp}\varphi, \varphi) \geq \alpha \|\varphi\|^2 \quad \forall \varphi \in \overline{\mathcal{R}(\mathcal{H})}.$$

Moreover, the ranges  $\mathcal{R}((\mathcal{H})^*)$  and  $\mathcal{R}((N_{\sharp})^{1/2})$  coincide.



## A.2 Theoretical foundations of GLSM

This method has been introduced in [14, 12] and we provide here the main theorems that are used in our case. They are mainly extracted from [30, 13]. We consider two separable Hilbert spaces  $X$  and  $Y$  (as in the previous section) and two operators  $N : X \rightarrow X$  and  $B : X \rightarrow X$ , which have the following factorizations

$$N = G \mathcal{H} \quad \text{and} \quad B = (\mathcal{H})^* T_{\sharp} \mathcal{H}. \quad (\text{A.5})$$

where  $\mathcal{H} : X \rightarrow Y$  and  $T : Y \rightarrow Y$  are bounded operators and where

$$G : \overline{\mathcal{R}(\mathcal{H})} \mapsto X$$

is also bounded. We consider first the GLSM for noisy free operators. Let  $\alpha > 0$  be a given parameter and  $\phi \in X$ . We define the functional  $J_{\alpha}(\phi; \cdot) : X \rightarrow \mathbb{R}$  by

$$J_{\alpha}(\phi; a) := \alpha |(Ba, a)| + \|Na - \phi\|^2 \quad \forall a \in X \quad (\text{A.6})$$

and set

$$j_{\alpha}(\phi) := \inf_{g \in X} J_{\alpha}(\phi; g). \quad (\text{A.7})$$

The first main theorem for GLSM is the following:

**Theorem A.2.1.** *Assume that*

- $G$  is compact and  $N$  has dense range.
- $T_{\sharp}$  satisfies the coercivity property

$$|(T_{\sharp}\varphi, \varphi)| > \mu \|\varphi\|^2 \quad \forall \varphi \in \mathcal{R}(\mathcal{H}), \quad (\text{A.8})$$

where  $\mu > 0$  is a constant independent of  $\varphi$ . Let  $c(\alpha)$  such that  $c(\alpha)/\alpha \rightarrow 0$  as  $\alpha \rightarrow 0$  and  $\phi \in X$ , an element  $a_{\alpha} \in X$  such that

$$J_{\alpha}(\phi; a_{\alpha}) \leq j_{\alpha}(\phi) + c(\alpha). \quad (\text{A.9})$$

Then the following holds.

- If  $\phi \in \mathcal{R}(G)$  then  $\limsup_{\alpha \rightarrow 0} |(Ba, a)| < \infty$ .
- If  $\phi \notin \mathcal{R}(G)$  then  $\liminf_{\alpha \rightarrow 0} |(Ba, a)| = \infty$ .

In the previous theorem nothing is said on the convergence or not of  $(Ba, a)$ . The latter is possible if the operator  $T_{\sharp}$  is selfadjoint (meaning that  $B$  is also selfadjoint).

**Theorem A.2.2.** *We assume, in addition to the hypothesis of Theorem A.2.1, that  $N$  is injective and that  $T_{\sharp}$  is selfadjoint. Consider for  $\alpha > 0$  and  $\phi \in X$ ,  $a_{\alpha} \in X$  such that*

$$J_{\alpha}(\phi; a_{\alpha}) \leq j_{\alpha}(\phi) + c(\alpha) \quad (\text{A.10})$$

where  $\frac{c(\alpha)}{\alpha} \rightarrow 0$  as  $\alpha \rightarrow 0$ . Then  $\phi \in \mathcal{R}(G)$  if and only if  $\lim_{\alpha \rightarrow 0} (Ba_{\alpha}, a_{\alpha}) < \infty$ . Moreover, in the case  $\phi = G\varphi$ , the sequence  $\mathcal{H}a_{\alpha}$  strongly converges to  $\varphi$  in  $Y$  as  $\alpha$  goes to zero.

**Remark A.2.3.** In the case  $B = F$ , one can avoid the extra assumption on the operator  $T_{\sharp}$  in Theorem A.2.2 by replacing the cost functional  $J_{\alpha}$  with

$$J_{\alpha}(\phi; g) := \alpha |(\mathbf{N}g, g)| + \alpha^{1-\eta} |(\mathbf{N}g - \phi, g)| + \|\mathbf{N}g - \phi\|^2 \quad \forall g \in X, \quad (\text{A.11})$$

with  $\eta \in ]0, 1]$  being a fixed parameter. This type of functional is more suited for the case of limited aperture data which is not considered in our case.

We now state the RGLSM theorem for the case of noisy data. More precisely, we shall assume that one has access to two noisy operators  $\mathbf{B}^{\delta}$  and  $\mathbf{N}^{\delta}$  such that

$$\|\mathbf{N}^{\delta} - \mathbf{N}\| \leq \delta \quad \text{and} \quad \|\mathbf{B}^{\delta} - \mathbf{B}\| \leq \delta$$

for some  $\delta > 0$ . We also assume that the operators  $\mathbf{B}$ ,  $\mathbf{B}^{\delta}$ ,  $\mathbf{N}^{\delta}$  and  $\mathbf{N}$  are compact. We then consider for  $\alpha > 0$  and  $\phi \in X^*$  the functional  $J_{\alpha}^{\delta}(\phi; \cdot) : X \rightarrow \mathbb{R}$  defined by

$$J_{\alpha}^{\delta}(\phi; a) := \alpha (|\mathbf{B}^{\delta} a, a| + \delta \alpha^{-\eta} \|a\|^2) + \|\mathbf{N}^{\delta} a - \phi\|^2 \quad \forall a \in X, \quad (\text{A.12})$$

where  $\eta < 1$  is a fixed positive parameter.

**Theorem A.2.4.** *Assume that the assumptions of Theorem A.2.2 hold. Let  $a_{\alpha}^{\delta}$  be the minimizer of  $J_{\alpha}^{\delta}(\phi; \cdot)$  (defined by (A.12)) for  $\alpha > 0$ ,  $\delta > 0$  and  $\phi \in X^*$ . Then  $\phi \in \mathcal{R}(G)$  if and only if  $\lim_{\alpha \rightarrow 0} \limsup_{\delta \rightarrow 0} (|\mathbf{B}^{\delta} a_{\alpha}^{\delta}, a_{\alpha}^{\delta}| + \delta \alpha^{-\eta} \|a_{\alpha}^{\delta}\|^2) < \infty$ . Moreover, in the case  $\phi = G\varphi$ ,  $\lim_{\alpha \rightarrow 0} \limsup_{\delta \rightarrow 0} \|\mathcal{H}a_{\alpha} - \varphi\| = 0$ .*

**Remark A.2.5.** If  $\mathbf{B}^{\delta}$  is a positive selfadjoint operator one can directly compute the minimizer  $a_{\alpha}^{\delta}$  of  $J_{\alpha}^{\delta}(\phi; \cdot)$  (defined by (A.12)) for  $\alpha > 0$ ,  $\delta > 0$  and  $\phi \in X^*$  as the solution of

$$(\alpha \mathbf{B}^{\delta} + \alpha \delta I + (\mathbf{N}^{\delta})^* \mathbf{N}^{\delta}) a_{\alpha}^{\delta} = (\mathbf{N}^{\delta})^* \phi. \quad (\text{A.13})$$



# Bibliography

- [1] M. ABRAMOWITZ AND I. STEGUN, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.*, Dover Publications, 1964. (Cited on page 149.)
- [2] A. ALÛ, *First-principles homogenization theory for periodic metamaterials*, Physical Review B, 84 (2011), p. 075135. (Cited on pages 9 and 153.)
- [3] C. AMROUCHE, C. BERNARDI, M. DAUGE, AND V. GIRAULT, *Vector potentials in three-dimensional non-smooth domains*, Math. Meth. Appl. Sci., 21 (1998), pp. 823–864. (Cited on page 156.)
- [4] X. ANTOINE, C. GEUZAINÉ, AND K. RAMDANI, *Computational Methods for Multiple Scattering at High Frequency with Applications to Periodic Structures Calculations.*, in Wave Propagation in Periodic Media, Progress in Computational Physics, Vol. 1, Bentham, 2010, pp. 73–107. (Cited on pages 3, 13 and 161.)
- [5] T. ARENS, *Scattering by biperiodic layered media: The integral equation approach*. Habilitation Thesis, Universität Karlsruhe, 2010. (Cited on pages 2, 9, 12, 153 and 155.)
- [6] T. ARENS, *Uniqueness for Elastic Wave Scattering by Rough Surfaces*, SIAM Journal on Mathematical Analysis, 33 (2001), pp. 461–476. (Cited on pages 5 and 17.)
- [7] T. ARENS AND N. GRINBERG, *A complete factorization method for scattering by periodic structures*, Computing, 75 (2005), pp. 111–132. (Cited on pages 3, 7, 13 and 76.)
- [8] T. ARENS, K. HASELOH, AND S. N. CHANDLER-WILDE, *Solvability and spectral properties of integral equations on the real line: II.  $L^p$ -spaces and applications*, J. Int. Equ. Appl., 15 (2003), pp. 1–35. (Cited on pages 2 and 12.)
- [9] T. ARENS AND T. HOHAGE, *On radiation conditions for rough surface scattering problems*, IMA J. Applied Math., 70 (2005), pp. 839–847. (Cited on pages 5, 17, 18 and 43.)
- [10] T. ARENS AND A. KIRSCH, *The factorization method in inverse scattering from periodic structures*, Inverse Problems, 19 (2003), pp. 1195–1211. (Cited on pages 3, 7, 13, 76 and 87.)
- [11] T. ARENS, A. LECHLEITER, AND D. R. LUKE, *Music for extended scatterers as an instance of the factorization method*, SIAM Journal on Applied Mathematics, 70 (2009), pp. 1283–1304. (Cited on page 87.)
- [12] L. AUDIBERT, *Qualitative methods for heterogeneous media*, theses, Ecole Doctorale Polytechnique, Sept. 2015. (Cited on pages 3, 4, 7, 13, 14, 76, 88, 161 and 164.)
- [13] L. AUDIBERT, A. GIRARD, AND H. HADDAR, *Identifying defects in an unknown background using differential measurements*, Inverse Problems and Imaging, 9 (2015), pp. 625–643. (Cited on pages 3, 4, 8, 13, 14, 76, 95, 161 and 164.)
- [14] L. AUDIBERT AND H. HADDAR, *A generalized formulation of the linear sampling method with exact characterization of targets in terms of farfield measurements*, Inverse Problems, 30 (2014), p. 035011. (Cited on pages 3, 7, 13, 76, 88, 89 and 164.)

- [15] V. BARONIAN, A.-S. BONNET-BEN DHIA, S. FLISS, AND A. TONNOIR, *Iterative methods for scattering problems in isotropic or anisotropic elastic waveguides*, Wave Motion, 64 (2016), pp. 13–33. (Cited on pages 2, 13 and 161.)
- [16] C. BESSE, J. COATLÉVEN, S. FLISS, I. LACROIX-VIOLET, AND K. RAMDANI, *Transparent boundary conditions for locally perturbed infinite hexagonal periodic media*, Commun. Math. Sci., 11 (2013), pp. 907–938. (Cited on pages 2, 13 and 161.)
- [17] A.-S. BONNET-BEN DHIA, L. CHESNEL, AND P. CIARLET, JR., *T-coercivity for scalar interface problems between dielectrics and metamaterials*, ESAIM Math. Model. Numer. Anal., 46 (2012), pp. 1363–1387. (Cited on pages 4, 9, 14 and 141.)
- [18] A.-S. BONNET-BEN DHIA, L. CHESNEL, AND P. CIARLET, JR., *Two-dimensional Maxwell's equations with sign-changing coefficients*, Appl. Numer. Math., 79 (2014), pp. 29–41. (Cited on pages 4, 9, 14 and 141.)
- [19] A.-S. BONNET-BEN DHIA, L. CHESNEL, AND P. CIARLET JR., *T-coercivity for the Maxwell problem with sign-changing coefficients*, Comm. Partial Differential Equations (accepted), (2014). (Cited on pages 9, 10, 154, 156 and 158.)
- [20] A.-S. BONNET-BEN DHIA, G. DAKHIA, C. HAZARD, AND L. CHORFI, *Diffraction by a defect in an open waveguide: A mathematical analysis based on a modal radiation condition*, SIAM Journal on Applied Mathematics, 70 (2009), pp. 677–693. (Cited on pages 2 and 12.)
- [21] A.-S. BONNET-BENDHIA, L. CHESNEL, AND H. HADDAR, *On the use of T-coercivity to study the interior transmission eigenvalue problem.*, C. R. Acad. Sci. Mathematics, 11-12 (2011), pp. 647–651. (Cited on page 126.)
- [22] A.-S. BONNET-BENDHIA, M. DAUGE, AND K. RAMDANI, *Analyse spectrale et singularités d'un problème de transmission non coercif*, C. R. Acad. Sci. Paris Sér. I Math., 328 (1999), pp. 717–720. (Cited on pages 4, 9, 14 and 141.)
- [23] A.-S. BONNET-BENDHIA AND G. LEGENDRE, *An alternative to dirichlet-to-neumann maps for waveguides*, C. R. Acad. Sci. Ser. I, 349 (2011), pp. 1005–1009. (Cited on pages 2, 13 and 161.)
- [24] A.-S. BONNET-BENDHIA AND K. RAMDANI, *Diffraction by an acoustic grating perturbed by a bounded obstacle*, Modeling and Computation of Optics and Electromagnetics, Special Issue, Advances in Computational Mathematics, 16 (2002). (Cited on pages 2 and 12.)
- [25] A.-S. BONNET-BENDHIA AND F. STARLING, *Guided waves by electromagnetic gratings and non-uniqueness examples for the diffraction problem*, Mathematical Methods in the Applied Sciences, 17 (1994), pp. 305–338. (Cited on pages 2, 6, 12, 22 and 42.)
- [26] Y. BOUKARI AND H. HADDAR, *The Factorization method applied to cracks with impedance boundary conditions*, Inverse Problems and Imaging, 7 (2013). (Cited on page 87.)
- [27] L. BOURGEOIS AND S. FLISS, *On the identification of defects in a periodic waveguide from far field data*, Inverse Problems, 30 (2014), p. 095004. (Cited on pages 4, 7, 8, 14, 76 and 87.)
- [28] L. BOURGEOIS AND E. LUNÉVILLE, *On the use of the linear sampling method to identify cracks in elastic waveguides*, Inverse Problems, 29 (2013), p. 025017. (Cited on pages 4, 8, 14 and 76.)

- [29] F. CAKONI AND D. COLTON, *Qualitative Methods in Inverse Scattering Theory. An Introduction.*, Springer, Berlin, 2006. (Cited on pages 3, 7, 13 and 76.)
- [30] F. CAKONI, D. COLTON, AND H. HADDAR, *Inverse Scattering Theory and Transmission Eigenvalues*, vol. 88 of CBMS Series, SIAM publications, 2016. (Cited on pages 3, 7, 13, 76, 81, 88, 163 and 164.)
- [31] F. CAKONI AND H. HADDAR, *On the existence of transmission eigenvalues in an inhomogeneous medium*, *Applicable Analysis*, 88 (2009), pp. 475–493. (Cited on page 81.)
- [32] F. CAKONI AND H. HADDAR, *Transmission eigenvalues*, *Inverse Problems*, 29 (2013), p. 100201. Editorial. (Cited on page 81.)
- [33] S. CHANDLER-WILDE, C. ROSS, AND B. ZHANG, *Scattering by infinite one-dimensional rough surfaces*, *Proceedings of the Royal Society A*, 455 (1999), pp. 3767–3787. (Cited on pages 2 and 12.)
- [34] S. N. CHANDLER-WILDE, *The Impedance Boundary Value Problem for the Helmholtz Equation in a Half-Plane*, *Mathematical Methods in the Applied Sciences*, 20 (1997), pp. 813–840. (Cited on pages 5 and 17.)
- [35] S. N. CHANDLER-WILDE AND J. ELSCHNER, *Variational approach in weighted Sobolev spaces to scattering by unbounded rough surfaces*, *SIAM J. Math. Anal.*, 42 (2010), pp. 2554–2580. (Cited on pages 2, 5, 12 and 18.)
- [36] S. N. CHANDLER-WILDE, E. HEINEMEYER, AND R. POTTHAST, *Acoustic scattering by mildly rough surfaces in three dimensions*, *SIAM J. Appl. Math.*, 66 (2006), pp. 1002–1026. (Cited on pages 2 and 12.)
- [37] S. N. CHANDLER-WILDE AND P. MONK, *Existence, uniqueness, and variational methods for scattering by unbounded rough surfaces*, *SIAM J. Math. Anal.*, 37 (2005), pp. 598–618. (Cited on pages 2, 5, 12, 17, 18, 19 and 25.)
- [38] S. N. CHANDLER-WILDE, P. MONK, AND M. THOMAS, *The mathematics of scattering by unbounded, rough, inhomogeneous layers*, *Journal of Computational and Applied Mathematics*, 204 (2007), pp. 549–559. (Cited on pages 2, 6, 12, 18 and 42.)
- [39] S. N. CHANDLER-WILDE AND B. ZHANG, *A uniqueness result for scattering by infinite dimensional rough surfaces*, *SIAM J. Appl. Math.*, 58 (1998), pp. 1774–1790. (Cited on pages 2 and 12.)
- [40] A. CHARALAMBOPOULUS, A. KIRSCH, K. A. ANAGNOSTOPOULUS, D. GINTIDES, AND K. KIRIAKI, *The factorization method in inverse elastic scattering from penetrable bodies*, *Inverse Problems*, 23 (2007), pp. 27–51. (Cited on page 87.)
- [41] L. CHESNEL, *Investigation of some transmission problems with sign-changing coefficients. Application to metamaterials.*, theses, Ecole Polytechnique X, Oct. 2012. (Cited on pages 4, 5, 9, 14, 15, 141 and 153.)
- [42] J. COATLÉVEN, *Analyse mathématique et numérique de quelques problèmes d’ondes en milieu périodique*, PhD thesis, 10 2011. (Cited on pages 6, 7, 42, 45 and 58.)
- [43] J. COATLÉVEN, *Mathematical and numerical analysis of wave problems in periodic media*, theses, Ecole Polytechnique X, Nov. 2011. (Cited on pages 2 and 12.)

- [44] J. COATLÉVEN, *Helmholtz equation in periodic media with a line defect*, Journal of Computational Physics, 231 (2012). (Cited on pages 6 and 42.)
- [45] D. COLTON AND A. KIRSCH, *A simple method for solving inverse scattering problems in the resonance region*, Inverse Problems, 12 (1996), pp. 383–393. (Cited on pages 3 and 13.)
- [46] D. COLTON AND R. KRESS, *Inverse acoustic and electromagnetic scattering theory*, vol. 93 of Applied Mathematical Sciences, Springer, New York, third ed., 2013. (Cited on pages 3, 7, 13, 34, 46, 76, 142, 146 and 151.)
- [47] D. COLTON, M. PIANA, AND R. POTTHAST, *A simple method using Morozov’s discrepancy principle for solving inverse scattering problems*, Inverse Problems, 13 (1997), pp. 1477–1493. (Cited on page 86.)
- [48] D. L. COLTON AND R. KRESS, *Inverse acoustic and electromagnetic scattering theory*, Springer, 2nd ed., 1998. (Cited on pages 1, 3, 11, 13 and 113.)
- [49] A. COSSONNIÈRE, *Valeurs propres de transmission et leur utilisation dans l’identification d’inclusions à partir de mesures électromagnétiques*, PhD thesis, University of Toulouse, 2011. (Cited on pages 5, 9, 15 and 142.)
- [50] A. COSSONNIÈRE AND H. HADDAR, *Surface integral formulation of the interior transmission problem*, J. Integral Equations Appl., 25 (2013), pp. 341–376. (Cited on pages 5, 9, 15, 141, 142, 143, 145 and 147.)
- [51] M. COSTABEL AND E. STEPHAN, *A direct boundary integral equation method for transmission problems*, J. Math. Anal. Appl., 106 (1985), pp. 367–413. (Cited on pages 143 and 147.)
- [52] M. COSTABEL AND E. STEPHAN, *A direct boundary integral equation method for transmission problems*, J. Math. Anal. Appl., 106 (1985), pp. 367–413. (Cited on pages 4, 9, 14 and 141.)
- [53] J. ELSCHNER AND G. HU, *Inverse scattering of elastic waves by periodic structures: uniqueness under the third or fourth kind boundary conditions*, Methods and Applications of Analysis, 18 (2011), pp. 215–244. (Cited on pages 3, 7, 13 and 76.)
- [54] S. FLISS, *Analyse mathématique et numérique de problèmes de propagation des ondes dans des milieux périodiques infinis localement perturbés*, PhD thesis, Ecole Polytechnique X, 2009. (Cited on pages 2, 6, 13, 42 and 161.)
- [55] S. FLISS AND P. JOLY, *Wave propagation in locally perturbed periodic media (case with absorption): Numerical aspects*, Journal of Computational Physics, 231 (2012), pp. 1244–1271. (Cited on pages 2, 13 and 161.)
- [56] B. GEBAUER AND N. HYVÖNEN, *Factorization method and irregular inclusions in electrical impedance tomography*, Inverse Problems, 23 (2007), pp. 2159–2170. (Cited on page 87.)
- [57] H. HADDAR AND A. LECHLEITER, *Electromagnetic Wave Scattering from Rough Penetrable Layers*, SIAM Journal on Mathematical Analysis, 43 (2011), pp. 2418–2443. (Cited on pages 2 and 12.)
- [58] H. HADDAR AND A. LECHLEITER, *Electromagnetic wave scattering from rough penetrable layers*, SIAM J. Math. Anal., (2011). (Cited on pages 6, 42 and 161.)

- [59] H. HADDAR AND G. MIGLIORATI, *Numerical analysis of the factorization method for EIT with a piecewise constant uncertain background*, *Inverse Problems*, 29 (2013), p. 065009. (Cited on page 87.)
- [60] H. HADDAR AND T.-P. NGUYEN, *A volume integral method for solving scattering problems from locally perturbed infinite periodic layers*, *Applicable Analysis*, (2016), p. 29. (Cited on pages 15 and 41.)
- [61] M. HANKE AND B. SCHAPPEL, *The factorization method for electrical impedance tomography in the half space*, *SIAM J. Appl. Math.*, 68 (2008), pp. 907–924. (Cited on page 87.)
- [62] C. HAZARD, *On the absence of trapped modes in locally perturbed open waveguides*, *IMA Journal of Applied Mathematics*, 80 (2015), pp. 1049–1062. (Cited on pages 2 and 12.)
- [63] T. HOHAGE, *On the numerical solution of a three-dimensional inverse medium scattering problem*, *Inverse Problems*, 17 (2001), pp. 1743–1763. (Cited on pages 6 and 42.)
- [64] ———, *Fast numerical solution of the electromagnetic medium scattering problem and applications to the inverse problem*, *J. Comp. Phys.*, 214 (2006), pp. 224–238. (Cited on pages 6 and 42.)
- [65] G. C. HSIAO AND W. L. WENDLAND, *Boundary integral equations*, Springer, Berlin, 2008. (Cited on page 144.)
- [66] G. HU, Y. LU, AND B. ZHANG, *The factorization method for inverse elastic scattering from periodic structures*, *Inverse Problems*, 29 (2013), p. 115005. (Cited on pages 3, 7, 13 and 76.)
- [67] F. IHLENBURG, *Finite element analysis of acoustic scattering*, Springer, 1998. (Cited on page 21.)
- [68] P. JOLY, J.-R. LI, AND S. FLISS, *Exact boundary conditions for periodic waveguides containing a local perturbation*, *Commun. Comput. Phys.*, 1 (2006), pp. 945–973. (Cited on pages 2, 6, 13, 42 and 161.)
- [69] A. KIRSCH, *Diffraction by periodic structures*, in *Proc. Lapland Conf. on Inverse Problems*, L. Pävarinta and E. Somersalo, eds., Springer, 1993, pp. 87–102. (Cited on pages 33 and 46.)
- [70] ———, *Characterization of the shape of a scattering obstacle using the spectral data of the far field operator*, *Inverse Problems*, 14 (1998), pp. 1489–1512. (Cited on pages 3, 7, 13, 76 and 87.)
- [71] ———, *Factorization of the far field operator for the inhomogeneous medium case and an application in inverse scattering theory*, *Inverse Problems*, 15 (1999), pp. 413–429. (Cited on pages 7, 76, 87 and 163.)
- [72] A. KIRSCH, *The music-algorithm and the factorization method in inverse scattering theory for inhomogeneous media*, *Inverse Problems*, 18 (2002), p. 1025. (Cited on pages 3 and 13.)
- [73] A. KIRSCH, *The factorization method for a class of inverse elliptic problems*, *Math. Nachr.*, 278 (2004), pp. 258–277. (Cited on pages 3 and 13.)
- [74] ———, *The factorization method for Maxwell’s equations*, *Inverse Problems*, 20 (2004), pp. S117–S134. (Cited on pages 3 and 13.)



- [75] ———, *An integral equation approach and the interior transmission problem for Maxwell's equations*, Inverse. Probl. Imag., 1 (2007), pp. 159–180. (Cited on pages 9, 154, 155 and 158.)
- [76] A. KIRSCH AND N. GRINBERG, *The Factorization Method for Inverse Problems*, Oxford Lecture Series in Mathematics and its Applications 36, Oxford University Press, 2008. (Cited on pages 3, 7, 13, 76, 87 and 163.)
- [77] R. KRESS, *A factorization method for an inverse Neumann problem for harmonic vector fields*, Georgian Mathematical Journal, 10 (2003), pp. 549–560. (Cited on page 87.)
- [78] P. KUCHMENT, *Floquet Theory for Partial Differential Equations*, vol. 60 of Operator Theory. Advances and Applications, Birkhäuser, Basel, 1993. (Cited on page 45.)
- [79] A. LECHLEITER, *Factorization Methods for Photonics and Rough Surface Scattering*, PhD thesis, Universität Karlsruhe, Karlsruhe, Germany, 2008. (Cited on page 87.)
- [80] ———, *The Factorization method is independent of transmission eigenvalues*, Inverse Problems and Imaging, 3 (2009), pp. 123–138. (Cited on pages 3, 13 and 87.)
- [81] A. LECHLEITER, N. HYVÖNEN, AND H. HAKULA, *The factorization method applied to the complete electrode model of impedance tomography*, Preprint, 68 (2008), pp. 1097–1121. (Cited on page 87.)
- [82] A. LECHLEITER AND D.-L. NGUYEN, *Factorization Method for Electromagnetic Inverse Scattering from Biperiodic Structures*, SIAM Journal on Imaging Sciences, 6 (2013), pp. 1111–1139. (Cited on pages 3, 4, 7, 8, 13, 14, 76 and 87.)
- [83] ———, *On uniqueness in electromagnetic scattering from biperiodic structures*, ESAIM Math. Model. Numer. Anal., 47 (2013), pp. 1167–1184. (Cited on pages 6, 7 and 42.)
- [84] A. LECHLEITER AND D.-L. NGUYEN, *Volume integral equations for scattering from anisotropic diffraction gratings*, Math. Methods Appl. Sci, 36 (2013), pp. 262–274. (Cited on pages 2, 6, 9, 12, 42 and 153.)
- [85] A. LECHLEITER AND D.-L. NGUYEN, *A trigonometric Galerkin method for volume integral equations arising in TM grating scattering*, Adv. Comput. Math., 40 (2014), pp. 1–25. (Cited on pages 6, 34, 42, 43, 46, 65 and 115.)
- [86] A. LECHLEITER AND S. RITTERBUSCH, *A variational method for wave scattering from penetrable rough layers*, IMA J. Appl. Math., 75 (2010), pp. 366–391. (Cited on pages 2, 5, 6, 12, 17, 18, 19, 22, 25, 27, 42 and 43.)
- [87] A. LECHLEITER AND R. ZHANG, *A convergent numerical scheme for scattering of aperiodic waves from periodic surfaces based on the floquet-bloch transform*, arXiv:1603.07860, (2016). (Cited on pages 2, 12, 69 and 161.)
- [88] W. MCLEAN, *Strongly Elliptic Systems and Boundary Integral Operators*, Cambridge University Press, Cambridge, UK, 2000. (Cited on pages 24 and 144.)
- [89] P. MONK, *Finite Element Methods for Maxwell's Equations*, Oxford Science Publications, Oxford, 2003. (Cited on page 157.)

- [90] J.-C. NÉDÉLEC, *Acoustic and Electromagnetic Equations*, Springer, New York etc, 2001. (Cited on page 144.)
- [91] D. L. NGUYEN, *Spectral Methods for Direct and Inverse Scattering from Periodic Structures*, PhD thesis, 2012. (Cited on pages 2, 3, 7, 9, 12, 13, 76, 98, 114, 115, 116, 153, 154, 155 and 158.)
- [92] D.-L. NGUYEN AND T.-P. NGUYEN, *Electromagnetic scattering by periodic structures with sign-changing coefficients*, *Comptes Rendus Mathématique*, (2015), pp. 893–898. (Cited on pages 16 and 153.)
- [93] H.-M. NGUYEN, *Limiting absorption principle and well-posedness for the Helmholtz equation with sign changing coefficients*, *J. Math. Pures Appl.* (9), 106 (2016), pp. 342–374. (Cited on pages 4, 5, 9, 14, 15, 141 and 142.)
- [94] P. OLA, *Remarks on a transmission problem*, *Journal of Mathematical Analysis and Applications*, 196 (1995). (Cited on pages 5, 9, 15, 142, 143 and 149.)
- [95] J. PENDRY, *Photonics: Metamaterials in the sunshine*, *Nature Materials*, 5 (2006), pp. 599–600. (Cited on pages 9 and 153.)
- [96] R. PETIT, ed., *Electromagnetic theory of gratings*, Springer, 1980. (Cited on pages 9 and 153.)
- [97] K. RAMDANI, *Lignes supraconductrices : analyse mathématique et numérique*, PhD thesis, 1999. (Cited on pages 4, 9, 14 and 141.)
- [98] K. SANDFORT, *The factorization method for inverse scattering from periodic inhomogeneous media*, PhD thesis, Karlsruhe Institut für Technologie, 2010. (Cited on pages 3, 7, 13, 76 and 87.)
- [99] J. SARANEN AND G. VAINIKKO, *Periodic integral and pseudodifferential equations with numerical approximation*, Springer, 2002. (Cited on pages 2, 12 and 51.)
- [100] S. SAUTER AND C. SCHWAB, *Boundary Element Methods*, Springer, 1. ed., 2007. (Cited on page 63.)
- [101] J. SUN AND C. ZHENG, *Reconstruction of obstacles embedded in periodic waveguides*, *Contemporary Mathematics*, 586 (2013), pp. 341–351. (Cited on pages 4, 7, 8, 14 and 76.)
- [102] J. SYLVESTER, *Discreteness of Transmission Eigenvalues via Upper Triangular Compact Operators*, *SIAM Journal on Mathematical Analysis*, 44 (2012), pp. 341–354. (Cited on page 81.)
- [103] G. VAINIKKO, *Fast solvers of the Lippmann-Schwinger equation*, in *Direct and Inverse Problems of Mathematical Physics*, D. Newark, ed., *Int. Soc. Anal. Appl. Comput.* 5, Dordrecht, 2000, Kluwer, p. 423. (Cited on pages 6, 7 and 42.)
- [104] R. WEDER, *Spectral and scattering theory for wave propagation in perturbed stratified media*, vol. 87 of *Applied Mathematical Sciences*, Springer-Verlag, New York, 1991. (Cited on pages 2, 12 and 22.)

- 
- [105] J. YANG, B. ZHANG, AND R. ZHANG, *A sampling method for the inverse transmission problem for periodic media*, *Inverse Problems*, 28 (2012), p. 17pp. (Cited on pages 3, 7, 13 and 76.)
- [106] B. ZHANG AND S. N. CHANDLER-WILDE, *Acoustic Scattering by an Inhomogeneous Layer on a Rigid Plate*, *SIAM Journal on Applied Mathematics*, 58 (1998), pp. 1931–1950. (Cited on pages 5 and 17.)
- [107] —, *Integral equation methods for scattering by infinite rough surfaces*, *Mathematical Methods in the Applied Sciences*, 26 (2003), pp. 463–488. (Cited on pages 5 and 17.)
- [108] C. M. ZWÖLF, *Méthodes variationnelles pour la modélisation des problèmes de transmission d’onde électromagnétique entre diélectrique et méta-matériau.*, PhD thesis, Université Versailles - Saint-Quentin, 2008. (Cited on pages 4, 9, 14 and 141.)

**Titre : Solveurs directes et inverses pour la diffraction par des couches périodiques infinies localement perturbées**

**Mots-clés :** problème inverse, méthodes d'échantillonnage, surfaces périodiques, problème de diffraction.

Nous sommes intéressés par l'analyse de la diffraction directe et inverse des ondes par des couches infinies périodiques localement perturbées à une fréquence fixe. Nous analysons d'abord le problème direct et établissons certaines conditions pour lesquelles il n'existe pas de modes guidés. Nous proposons ensuite une méthode numérique pour résoudre le problème de diffraction basée sur l'utilisation de la transformée de Floquet-Bloch dans les directions de périodicité. Nous discrétisons le problème de manière uniforme dans la variable de Floquet-Bloch et utilisons une méthode spectrale dans la discrétisation spatiale. La discrétisation en espace exploite une reformulation Lipmann-Schwinger et une périodisation du noyau dans la direction perpendiculaire à la périodicité. L'analyse de la convergence est faite dans le cas avec absorption et la validation numérique est réalisée sur des exemples 2D. Pour le problème inverse, nous étendons trois méthodes d'échantillonnage pour la reconstruction de la géométrie du défaut à partir de la connaissance de données multistatiques associées à des ondes incidentes planes. Nous analysons ces méthodes pour le problème semi-discrétisé dans la variable Floquet-Bloch. Nous proposons ensuite une nouvelle méthode d'imagerie capable de visualiser directement la géométrie du défaut sans connaître ni les propriétés physiques du milieu périodique, ni les propriétés physiques du défaut. Cette méthode d'imagerie-différentielle est basée sur l'analyse des méthodes d'échantillonnage pour un seul mode de Floquet-Bloch et la relation avec les solutions de problèmes de transmission intérieurs d'un type nouveau. Nous complétons notre travail par deux résultats sur l'analyse du problème de diffraction pour des matériaux périodiques ayant des indices négatifs: caractère bien posé du problème en 2D dans le cas d'un contraste égal à -1 et caractère Fredholm de la formulation Lipmann-Schwinger du problème dans le cas d'un contraste différent de -1.

**Title: Direct and inverse solvers for scattering problems from locally perturbed infinite periodic layers**

**Keywords:** inverse problem, sampling methods, periodic metamaterial structures, scattering problems.

We are interested by the analysis of scattering and inverse scattering problems for locally perturbed periodic infinite layers at a fixed frequency. This problem has connexions with non destructive testings of periodic structures. We first analyze the forward scattering problem and establish some conditions under which there exist no guided modes. We then propose a numerical method that solves the direct scattering problem based on Floquet-Bloch transform in the periodicity directions of the background. We discretize the problem uniformly in the Floquet-Bloch variable and use a spectral method in the space variable. The discretization in space exploits a volumetric reformulation of the problem in a cell and a periodization of the kernel in the direction orthogonal to the periodicity. The convergence analysis is done for the case with absorption and numerical validating results are conducted in 2D. For the inverse problem we extend the use of three sampling methods to solve the problem of retrieving the defect from the knowledge of multistatic data associated with incident plane waves. We analyze these methods for the semi-discretized problem in the Floquet-Bloch variable. We then propose a new method capable of retrieving directly the defect without knowing neither the background material properties nor the defect properties. The differential-imaging functional that we propose is based on the analysis of sampling methods for a single Floquet-Bloch mode and the relation with solutions to interior transmission problems. The theoretical investigations are corroborated with numerical experiments on synthetic data. We complement our thesis with two results on the analysis of the scattering problem for periodic materials with negative indices: well posedness of the problem in 2D in the case of a contrast equals -1 and Fredholm properties of the volume potential formulation of the problem using the T-coercivity approach in the case of a contrast different from -1.



