Analysis of a multiscale finite element method applied to the design of photovoltaic cells : a multiscale hybrid-mixed method for the Helmholtz equation with quasi-periodic boundary conditions

Zakaria Kassali

UNIVERSITÉ CÔTE D'AZUR

ÉCOLE DOCTORALE
SCIENCES FONDAMENTALES ET APPLIQUÉES

Analyse d'une méthode d'éléments finis multi-échelles appliquée à la conception de cellules photovoltaïques

Une méthode hybride-mixte multi-échelle pour l'équation de Helmholtz avec des conditions aux limites quasi-périodiques

by

Zakaria Kassali

Inria Sophia Antipolis-Méditerranée

Présentée en vue de l'obtention du grade de docteur en sciences mathématiques appliquées d'Université Côte d'Azur

Dirigée par: Stéphane Lanteri
Co-dirigée par: Théophile Chaumont-Frelet
Date de soutenance: 11/01/2023

Devant le jury, composé de :
Didier Auroux, PR, Univ. Côte d'Azur
Hélène Barucq, DR, Inria
Théophile Chaumont-Frelet, CR, Inria
Sonia Fliss, PR Associée, ENSTA
Stéphane Lanteri, DR, Inria
Serge Nicaise, PR, Univ. Polytechnique Hauts-de-France
Frédéric Valentin, DR, LNCC-Brazil

Inria

Analysis of a multiscale finite element method applied to the design of photovoltaic cells

A multiscale hybrid-mixed method for the Helmholtz equation with quasi-periodic boundary conditions

by

Zakaria Kassali

Jury :

Président du jury
Didier Auroux, professeur, Univ. Côte d'Azur
Rapporteurs
Hélène Barucq, Directrice de Recherches, Inria
Serge Nicaise, professeur , Univ. Polytechnique Hauts-de-France
Examinateurs
Sonia Fliss, professeure associée, ENSTA
Frédéric Valentin, directeur de recherche, LNCC-Brazil
Invités
Théophile Chaumont-Frelet, chargé de recherche, Inria
Stéphane Lanteri, directeur de recherche, Inria

Title: Analysis of a multiscale finite element method applied to the design of photovoltaic cells

Abstract: The objective of this thesis is the mathematical and numerical study of wave propagation in periodic and heterogeneous media modeled by the Helmholtz equation with quasi-periodic boundary conditions. In the current context of climate change, photovoltaic solar devices are emerging as an effective tool for a clean energy transition. This circumstance significantly pushes scientific research on the development of these devices. In turn, this background motivates the study of light propagation in these solar cells, which the Helmholtz equation can model with a quasi-periodic boundary condition. This unusual boundary condition represents a particular case of trapping geometries and gives rise to the appearance of some quasi-resonant frequencies. This work presents frequency-explicit stability results in the homogeneous case revealing the effect of these quasi-resonant frequencies on the use of perfectly matched layers (PML) and finite element discretizations. The Fourier expansion available in this case allows our study to go through the analysis of some parameterized one-dimensional Helmholtz problems satisfied by the Fourier modes. We also provide a frequency-explicit analysis for more general physical coefficients where Fourier expansion does not work. Specifically, we consider multilayer media, and our study uses the alternative "Morawetz multiplier" technique to obtain frequency-explicit results, which are of particular interest since they enter into the stability and convergence analysis of finite element discretizations. The second part of this work is devoted to the use of a two-level finite element method named the multiscale hybrid-mixed (MHM) method to solve our model problem. This method arises from a hybridization procedure using coarse mesh, and its multiscale basis functions are locally computed via independent cell problems. We first provide frequency-explicit error estimates, showing that the MHM method is more accurate and stable than the standard finite element method in the presence of quasi-resonant frequencies. Then, having in mind the nanoscale texturation used to ameliorate solar cells efficiently, an MHM multiscale convergence analysis is presented. The obtained error estimates hold uniformly when the characteristic length δ of the texturation goes to zero, which signifies that the MHM method keeps its robustness and capture small-scale heterogeneities using coarse meshes.

Keywords: Wave propagation, periodic structure, Helmholtz equation, quasi-periodic boundary condition, numerical analysis, finite element methods, multiscale methods.

Titre: Analyse d'une méthode d'éléments finis multi-échelles appliquée à la conception de cellules photovoltaïques

Résumé: L'objectif de cette thèse est l'étude mathématique et numérique de la propagation des ondes dans des milieux périodiques et hétérogènes modélisés par l'équation de Helmholtz avec des conditions aux limites quasi-périodiques. Dans le contexte actuel du changement climatique, les dispositifs solaires photovoltaïques apparaissent comme un outil efficace pour une transition énergétique propre. Ces circonstances encouragent considérablement la recherche scientifique sur le développement de ces dispositifs. À son tour, ce cadre motive l'étude de la propagation de la lumière dans ces cellules solaires, que l'équation de Helmholtz peut modéliser avec une condition limite quasi-périodique. Cette condition aux limites inhabituelle représente un cas particulier de géométries "captantes" et donne lieu à l'apparition de certaines fréquences quasi-résonantes. Ce travail présente des résultats de stabilité explicites en fréquence dans le cas homogène révélant l'effet de ces fréquences quasi-résonantes sur l'utilisation de couches parfaitement adaptées (PML) et sur les discrétisations par éléments finis. L'expansion de Fourier disponible dans ce cas permet à notre étude de passer par l'analyse de quelques problèmes de Helmholtz unidimensionnels paramétrés satisfaits par les modes de Fourier. Nous fournissons également une analyse explicite en fréquence pour des coefficients physiques plus généraux pour lesquels l'expansion de Fourier ne fonctionne pas. Plus précisément, nous considérons des milieux multicouches, et notre étude utilise la technique du "multiplicateur de Morawetz" pour obtenir des résultats explicites en fréquence, qui sont d'un intérêt particulier puisqu'ils interviennent dans l'analyse de stabilité et de convergence des discrétisations par éléments finis. La deuxième partie de ce travail est consacrée à l'utilisation d'une méthode d'éléments finis à deux niveaux, appelée la méthode Multiéchelle Hybride-Mixte (MHM), pour résoudre notre problème modèle. Cette méthode est issue d'une procédure d'hybridation utilisant un maillage grossier, et ses fonctions de base multi-échelles sont calculées localement via des problèmes indépendants dans chaque cellule. Nous fournissons d'abord des estimations d'erreurs explicites en fréquence, montrant que la méthode MHM est plus précise et plus stable que la méthode des éléments finis standard en présence de fréquences quasi-résonantes. Ensuite, ayant à l'esprit la texturation à l'échelle nanométrique utilisée pour améliorer l'efficacité des cellules solaires, une analyse de convergence multi-échelle de la méthode MHM est présentée. Les estimations d'erreur obtenues sont uniformes lorsque la longueur caractéristique δ de la texturation tend vers zéro, ce qui signifie que la méthode MHM garde sa robustesse et capture les hétérogénéités à petite échelle en utilisant des mailles grossières.

Mots clés: Propagation des ondes, structure périodique, équation de Helmholtz, condition aux limites quasi-périodique, analyse numérique, méthodes des éléments finis, méthodes multi-échelles.

Contents

List of Figures

Chapter 1

Introduction

Classified as a way of energy displacement, electromagnetic radiation is an extremely present phenomenon in our daily lives. For example, sunlight, microwaves oven, and medical scanners (including X-rays) all represent forms of electromagnetic wave propagation. This wide range of applications has greatly accelerated and pushed scientific research in the field of numerical simulations of electromagnetic wave propagation. Therefore, robust numerical methods are required, especially when the propagation domain has particular properties such as heterogeneity and periodicity.

The main objective of this Ph.D. thesis is to develop, analyze and evaluate a new multi-scale finite element method to simulate the propagation of electromagnetic waves in highly heterogeneous and periodic media. To achieve our goal, it is necessary to first go through mathematical study and analysis of the properties of our model problem. In the time-harmonic context, this propagation is perfectly modeled by the Helmholtz equation. Having in mind the immense influence of frequency and heterogeneities on the Helmholtz solution, we focus on explicit results that allow us to identify these influences.

Photovoltaics (PV). Among the possible applications mentioned in the first paragraph, PV cells for exploiting solar energy have undergone considerable progress over the past decade. Sunlight is considered the largest source of energy received or existing on the earth, and solar cells represent a useful way to harness solar energy to generate electric power. Having in mind the world's general orientation towards renewable energies, researchers and engineers are working intensively to improve the efficiency and reduce the production cost of these solar cells. In the research projects, scientists are trying to overcome weaknesses that may affect the essential properties of a solar cell, namely electrical efficiency, optical absorption, and manufacturing cost. To minimize manufacturing costs, the trend has always been to develop solar devices with an increasingly thin thickness. But at the same time, this reduction in thickness can lead to optical losses and, therefore, a reduction in overall solar device efficiency. Hence, fabrication materials represent the most influencing factor on the aforementioned properties, and scientific research is necessarily oriented towards the optimization of light trapping (in order to maximize the absorption

and minimize the reflection and the cost), which requires numerical simulations of light propagation in these materials. Currently, the solar panel market consists of three generations, crystalline silicon wafer-based devices, thin-film solar cells, and a new generation based on nanostructured materials. For each device of the three mentioned generations, the solar cell is the core building block. In detail, as represented in Figure 1.1, a solar array is a series of modules/panels, which also consists of a periodic arrangement of cells. This periodic construction in cells allows the scientific study to be focused on a single cell and then extended to the whole device.

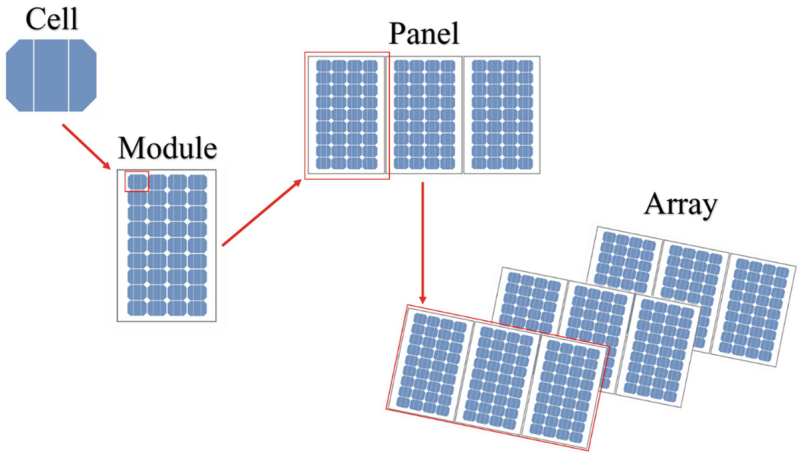


Figure 1.1: Solar cell, module, panel, and array. Source: DOI 10.1007/978-3-030-27824-3

Absorption and light trapping techniques Since their inception, the majority of solar cells have been using semiconductor materials to convert sunlight into electrical energy. The chemical transformation that allows semiconductor materials, when illuminated, to produce an electrical voltage is known as the photovoltaic effect. Unfortunately, due to their physical properties, each one of these materials absorbs only a part of the solar spectrum, and they tend to produce some undesirable reflections of sunlight. In order to overcome the first difficulty and to maximize the absorption, crystalline silicon and thin-film technologies of solar cells are stacking layers of varied materials: each layer absorbs its specific portion of the solar spectrum as in [131]. On the other hand, to minimize reflection losses, the standard approach is to cover these layers with an anti-reflective coating to limit these optical reflection losses. In the last decades, nano-texturing and light trapping techniques (Figure 1.2) have been used to deal with reflection problems. These techniques allow the sunlight to be focused and trapped in the absorber layer and to increase its optical path in the active layer [25, 44, 93, 132, 107]. Compared to the random one, periodic nano-texturation (see Figure 1.3) provides an additional approach to improve solar cells efficiency. For instance, both mono- and bi-periodic nano-texturation have been studied and proved viable options for improving the efficiency of organic solar cells [126, 104, 128, 50]. Similarly, periodic texturation have been designed to reduce the undesired reflection of sunlight for both photonic-crystal solar cells [135] and thin-film cells [97]. We also refer

to the reviews [64, 21] for more detailed discussions of the nano-texturation benefits on different solar devices.

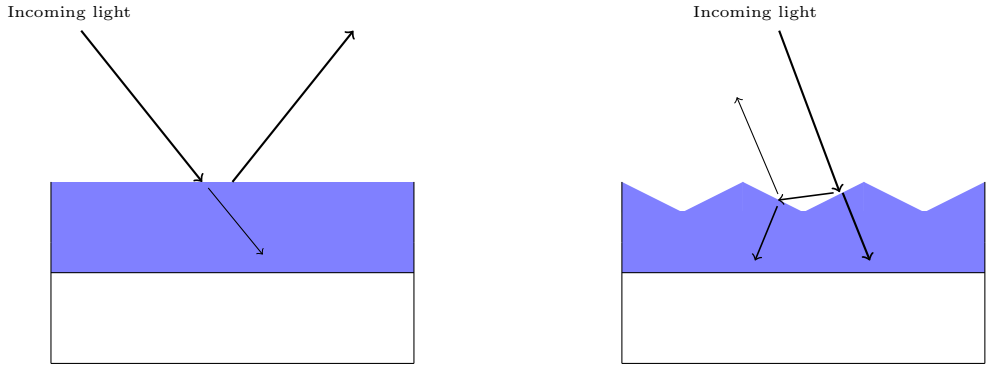


Figure 1.2: Standard solar surface (left) and light trapping surface (right).

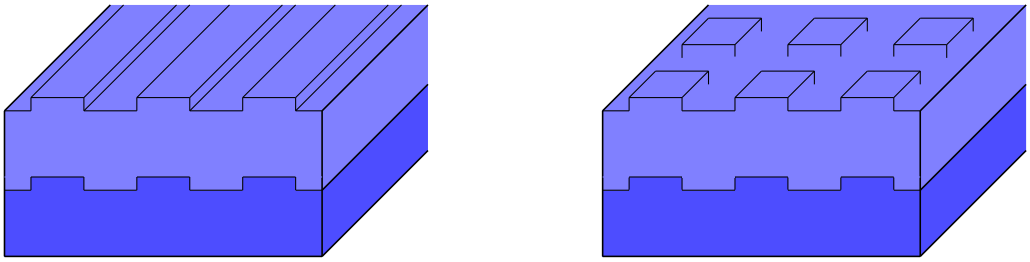
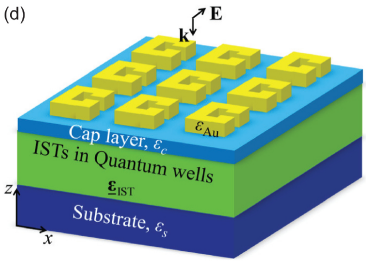


Figure 1.3: mono-periodic texturation (left) and bi-periodic texturation (right).

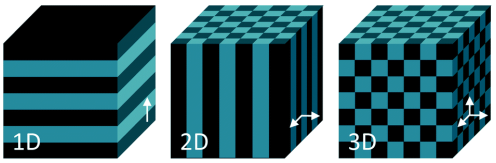
Motivated by the results of this great amount of research targeting solar cell improvement, we are led to study the propagation of the electromagnetic field in these periodic structures of solar cells. In particular, we will study the continuous properties of the problem that models this propagation and use a new multi-scale finite element method to numerically approximate this model and analyze its robustness in the considered circumstances.

Wave propagation in periodic structures has many challenging and vital applications, e.g., metasurfaces [100, 37], photonic crystals [92], electromagnetic band gap structures [91], frequency selective surfaces [105], photovoltaic devices, and many other systems (Figure 1.4). To improve the performance of these devices, wave propagation in periodic structures is receiving considerable attention in the numerical simulation community. A key point of the periodic characterization is the possibility of focusing the study on the unit cell level. Furthermore, the results sought by the numerical simulation in a unit periodic cell can be divided into two components. First, the performance study of a periodic device by providing accurate numerical solutions to mathematical equations (direct problems).

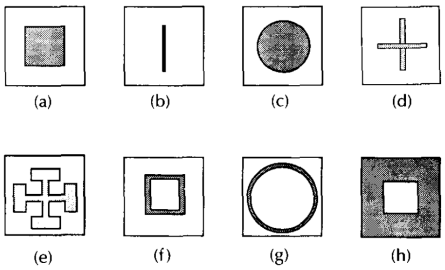
Second, developing new and more efficient periodic configurations through the numerical approximation of optimal design problems (inverse problems). In addition to their interest in the context of solar cell design, all the results found in this work can be interesting for these different applications that use periodical and/or multilayer structures.



(a) A periodic construction of an optical metasurface (source: [37]).



(b) Periodic modulation of one, two- and three-dimensional photonic crystal structures.



(c) Shape examples of unit constructive cells of frequency selective surfaces (source: [105]).

Figure 1.4: Three examples of periodic structures.

Maxwell's equations are the fundamental laws that govern electromagnetic wave propagation [71]. They are formulated in the time domain, which means that their solution depends on both the spatial variable $\boldsymbol{x}$ and the temporal variable t. In this work, we focus on bi-dimensional time-harmonic wave propagation problems (the solution time dependence is $U(\boldsymbol{x}, t) = u(\boldsymbol{x})e^{i\omega t}$). In this case, we switch to the frequency domain, where the unknown (i.e., $u(\boldsymbol{x})$) depends only on the spatial variable. Indeed, Maxwell's system is simplified into a bi-dimensional Helmholtz equation by considering the two fundamental polarizations, the transverse electric (TE) polarization and the transverse magnetic (TM) polarization (see, e.g., [106] Remark 2.1). Once obtained, the Helmholtz equation has two main parameters, the frequency, which depends on the propagating wave, and the permittivity and the permeability, which depend on the propagation medium (see subsection 2.1.2 for more details).

Domain truncation. Bounded propagation media represent a critical factor in the numerical simulation of wave propagation. Therefore, truncation of the unbounded domain appears to be an essential step before applying domain-based numerical methods (such as finite element methods). Certainly, some domains are naturally and physically reduced and easier to truncate than others. For example, in the problems studied in this thesis, which are the mono-periodic texturation case (Figure 1.3, also known as the 1D-grating problem), the domain and the texturation are constant in one of the space directions. In this context, we can naturally switch from a 3D to a 2D simulation setting by taking a cross-section. Also, as mentioned earlier, the natural truncation of a periodic medium by cells is to restrict the study to one cell and consider some periodic properties on the cell edges. However, the truncation of an unbounded propagation domain is much more difficult to achieve. Mathematically, propagation in an unbounded domain is represented by the Sommerfeld radiation condition (s.g. Section 1.1.3 in [87]). Numerically, when using domain-based methods, different truncation approaches have been used to bound the computational domain. Certainly, the ultimate objective of these truncation approaches is to replace the Sommerfeld condition without causing much perturbation to the original solution. The main idea is to create artificial boundaries, without physical signification, to limit the computational domain. Then, "absorbing" or "non-reflective" boundary conditions are imposed on these artificial boundaries. The works of Enguist and Madja [56, 57] represent the first results of these truncation approaches. There, the authors develop perfectly absorbing boundary conditions using a non-local pseudo-differential operator named the "Dirichlet-to-Neumann" (DtN) operator. In general, this operator is not explicitly computable except for special geometries (circular, spherical, cartesian); in these cases, it allows us to treat the exact solution in the domain of interest (s.g. [73, 87]). Unfortunately, since the DtN operator is necessarily non-local, it is not really suited for numerical calculations as its discretization is very expensive. For this reason, several authors have approached the DtN operator by developing improved local boundary conditions (e.g. [56, 45, 67, 68]). On the other hand, introduced by Berenger to truncate the unbounded propagation of electromagnetic waves in [17, 18], the perfectly matched layer (PML) approach quickly became very popular due to its efficiency and ease of implementation [86, 133, 43, 42, 85]. In fact, the physical domain is surrounded by an additional layer, characterized by its length ℓ_{P} and its absorption function ν_{P}, and in which the outgoing waves are absorbed without reflecting waves. Thanks to these simple features, the PML technique has been used in many configurations and has shown its mathematical performance due to the exponential convergence (depending on the PML characteristics ℓ_{P} and ν_{P}) of the PML solution towards the original one [137, 138, 96, 82]. Unfortunately, in the case of scattering by periodic structures in a single direction, the standard PML method (with a Dirichlet condition in the external boundary of the PML layer) can lose its efficiency and accuracy. Indeed, the periodic conditions caused by the periodicity of the medium generate quasi-resonant modes, also known as anomalous modes (or Rayleigh frequencies) [137]. The main characteristic of these quasi-resonant modes is their propagation in the periodic direction and their constant value in the direction truncated by the PML. Thus, the amplitude of the reflected wave by the PML is not small enough, and it can contaminate the solution.

In this work, we focus on analyzing the effect of quasi-resonant modes on the stability of the PML problem and on the error between the PML solution and the original solution. We note that Chen and Wu in [40] developed a FE method with PML based on an a-posteriori error analysis which excludes the case of quasi-resonant modes. And recently, in [134], they proposed another FE-PML method that handles well these quasi-resonant modes. Similarly, in the thesis [136], and in order to have an automatic adaptive PML method, the author uses an a-posteriori criterion based on a PML error that takes into account the quasi-resonant modes. Here, we give an explicit and detailed error analysis of the PML error for our case, but not for the purpose of adapting a finite element method with optimized PML. We focus more on the well-posedness of the PML problem, the effect of quasi-resonant modes on its stability, and the convergence of the PML solution to the exact one using energy Sobolev norms.

Stability analysis. Like a wide range of applications of time-harmonic wave propagation, electromagnetic wave propagation in solar cells is often characterized by high values of the frequency. Coupled with the sensitivity of the numerical methods to the frequency, this feature impacts the choice of the numerical simulation method. More precisely, in the high-frequency regime, the Helmholtz solution oscillates much more and becomes difficult to simulate. Theoretically, when the frequency is large, the negative term depending on the frequency in the Helmholtz operator (i.e., $-k^2u-\Delta u$), becomes larger, which causes a lack of coercivity of the sesquilinear forms that appear naturally by integration by parts of the Helmholtz equation. This lack of coercivity impacts the stability and quasi-optimality of standard FE schemes on coarse meshes (the so-called "pollution effect") [88, 89]. In addition, it is found that the stability of a numerical scheme is necessarily linked to the stability of the continuous problem, which is obtained by controlling the norm of the solution operator with respect to the norm of the data multiplied by a stability constant. Namely, when the quasi-optimality of the FE solution is studied, we usually use the Schatz argument, e.g., [125, 8] (an adaptation of the standard Aubin-Nitsche duality argument used for coercive problems, e.g., [114, 6]), and it is established that the stability constant appears in the quasi-optimality conditions on the mesh (e.g., [103, 101]). Therefore, frequency-explicit stability estimates play a crucial role in the numerical analysis of the numerical schemes. In particular, they allow for a better understanding of the behavior of numerical schemes with respect to the frequency and then optimally choose the discretization parameters (mesh size, polynomial degree, etc.).

Frequency-explicit stability estimates are then a major key to rigorously understanding the performance of numerical methods for solving the Helmholtz equation. Their importance is illustrated by their usage in several works aiming at an explicit convergence analysis in frequency. For example, for convergence analysis studies that use frequency-explicit stability estimate, we cite [32, 33] for finite element methods, [58, 59] for DG methods, [70, 62] for integral-equation methods and [15, 36] for some multiscale finite element methods. Many other papers, motivated by applications in numerical analysis, have sought to prove frequency-explicit stability estimates [16, 30, 29, 106, 124]. However, this multiplicity of works is due to the significant influence of medium characteristics (its regularity,

homogeneous/heterogeneous, the coefficients smoothness and the boundary conditions) on the stability constant and on the technique to prove it. In general, the stability constant is strongly related to the inf-sup condition of Helmholtz problem, which is equivalent to the well-posedness of the problem [7].
Having in mind the strong influence of the properties of the medium on the frequency dependence of the stability constant and the fact that scattering by periodic structures represents a particular case of geometries and boundary conditions that can have a significant impact on the stability constant, we will seek in this work frequency-explicit stability estimates in two different cases, either when the solar cell consists of one material, or when it is made up several finely textured layers to show the effect of this periodic conditions. To do this, in the case of a single layer, we will use the periodic property, and we will proceed with a stability analysis of parameterized one-dimensional Helmholtz problems. In the multi-layer case, we will adapt and extend another proof approach named the "Morawetz multiplier" technique.

Multiscale methods. As mentioned earlier, nanoscale texturations can be extremely beneficial for solar cells to achieve good energy performance. Therefore, numerical simulations are required to optimize the nanotexturation layout. Nonetheless, the simulation of light propagation in a nanostructured material is difficult to realize with precision. This is due to the various spatial scales within the problem, including cell size, wavelength, and nanostructure size (see Figure 1.5).

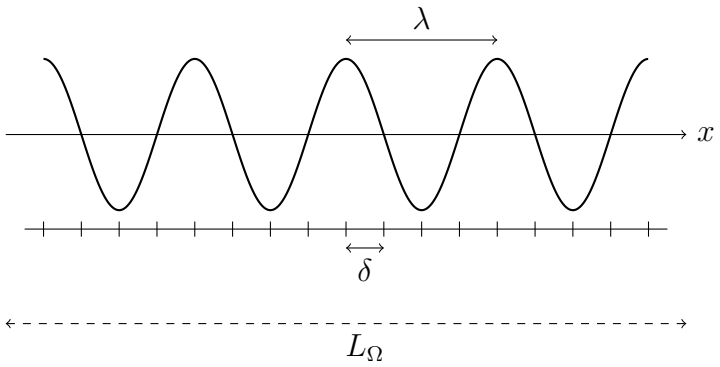


Figure 1.5: Example of the different scales of a one-dimensional propagation problem: the size of the domain L_Ω, the wavelength λ and the size of the heterogeneities δ.

For such multiscale problems, standard simulation methods show limitations. For example, to get a significant solution that captures all the different information from the small-scale heterogeneities, the discretization step must be very small [83], which gives more degrees of freedom and a very large system to solve. In the context of second-order elliptic problems, it has been observed that these limitations are due to polynomial basis functions that are unable to manage small-scale heterogeneities. With this in mind, multiscale methods have been developed where the basis functions are constructed as local

solutions of the PDE in each mesh element. Since the analytical solution of these local PDEs is not available in general, a common essential point between multiscale approaches is going through a second level allowing to approximate these local basis functions. As a result, these basic functions are adapted to the local properties and capture small-scale information in each mesh element. Well-known multiscale methods include, the Generalized multiscale finite element methods (GMsFEM) [84, 55, 54], the Heterogeneous Multiscale Method (HMM) [130, 53], the Localized Orthogonal Decomposition (LOD) [79, 112] and the Multiscale Hybrid Mixed Method [76, 3].

For the highly heterogeneous Helmholtz equation, adaptations of these methods and other strategies have been recently proposed. In [115], the authors adapted the HMM method to capture the heterogeneous Helmholtz solution on coarse meshes, but their analysis is limited to locally periodic media. In the spirit of the LOD method, Peterseim and his collaborators have developed the Petrov-Galerkin stabilization technique to eliminate the pollution effect in the homogeneous case [63, 119], then for a particular class of smooth heterogeneous coefficients assumed constant in each mesh element in [22] and recently for a more general class of piecewise constant coefficients in [120]. During his thesis, Chaumont-Frelet introduced the multiscale medium approximation method (MMAM) based on high-order polynomial basis functions, which gave excellent results for the acoustic Helmholtz equation, provided that the density is constant. Unfortunately, the results of the MMAM are not fully satisfactory for non-constant density media and elastic mediums.

In [36], Chaumont-Frelet and Valentin approximate Helmholtz solution using an adaptation of the MHM method that was initially developed for the Darcy equation in [76, 3] and used later for other problems [77, 75, 78, 95, 14, 69]. In the context of wave propagation, the MHM method was first used by Lanteri and his collaborators to solve time-domain Maxwell's equations in [95], then for the same model problem Gobe gave additional numerical details and validations in his thesis [69]. In [36], an MHM analysis was presented for highly heterogeneous Helmholtz coefficients allowing variation within the mesh elements. There, the MHM method appears as a very efficient multiscale strategy in the case of heterogeneous media using coarse meshes, and also it is computationally efficient since the local problems are elementwise and naturally parallelized. Furthermore, the authors have shown that using polynomial basis functions for the first MHM level, the MHM method can produce exact solutions for some propagation directions. Based on these properties, we believe that the MHM method can achieve very accurate results when simulating propagation in periodic and/or nanotextured solar cells. To clarify, MHM has shown its ability and effectiveness against the two major problems encountered in solar cell simulation. First, it performs very well in highly heterogeneous media, and this has been shown in [116] by a multiscale convergence analysis. Therefore, it can capture the small-scale information represented by the nanotexturation. Second, we know that quasi-resonant frequencies are plane waves traveling in the periodic direction, and the fact that the MHM method can give an exact solution for this direction makes the MHM desirable for these periodic cases.

Contributions and thesis outline The manuscript is structured as follow:

- In chapter 2, we introduce our model of the Helmholtz problem, especially the quasi-periodic boundary condition, which explains how periodic geometries affect Helmholtz solutions. We then turn to the main objective of this chapter, namely the frequency-explicit stability analysis of our model problem for two different cases. First, we consider the case of a homogeneous propagation medium and go through the stability analysis of specific one-dimensional Helmholtz problems satisfied by the Fourier expansion modes. Additionally, we show that our stability bounds are sharp with respect to the frequency and present numerical examples illustrating the impact of these stability results on the stability of finite element discretizations. Second, we employ the "Morawetz multiplier" technique to provide stability estimates for physical coefficients that model the case of finely structured layered media.

- In chapter 3, we use the PML technique to approximate the DtN operator and analyze the well-posedness properties of the resulting "PML problem". We start by establishing error estimates that control the difference between the original solution and the solution of the PML problem and reveal the effect of the presence of quasi-resonant modes on this convergence. We then turn to the frequency-explicit stability analysis of the PML problem in two different cases. On the one hand, we will treat the case of a homogeneous medium for which we can show an optimal stability estimate with the same frequency dependence as for the DtN problem. This result will be proved by following the same approach used for the homogeneous DtN problem. On the other hand, we will provide a general well-posedness result of the PML problem as soon as the corresponding DtN problem is well-posed.

- Chapter 4 is concerned wit the periodic homogenization theory applied to our model problem. Specifically, our study is motivated by the convergence analysis of multiscale numerical methods. We consider the case of finely textured layered media with a periodicity assumption and analyze the problem through the lens of periodic homogenization theory. In particular, using the explicit stability results found in Chapters 2 and 3, we derive frequency-explicit error estimates controlling the difference between the solutions of the homogenized problem and the solution of the oscillating problem.

- Chapter 5 is dedicated to the MHM method for the Helmholtz problem with PML and quasi-periodic boundary conditions. As a two-level method, the MHM method characterizes the solution as a collection of local contributions that are tied together through a global problem. We start by presenting an MHM formulation of our model problem showing the effect of the considered boundary conditions and analyzing the well-posedness of the one- and two-level formulations. Then, we rely on the Fourier expansion and some properties of the MHM method to provide a one-level convergence analysis of the MHM scheme showing its robustness to the presence of quasi-resonances, and we also present numerical examples that illustrate these MHM

performances. Finally, motivated by the nanoscale texturation used to improve the efficiency of solar cells, we consider the case of periodically finely textured layered media. In fact, we use the homogenization error results of Chapter 4 to obtain robust MHM error estimates when the small characteristic length of the texturing goes to zero. We then present numerical examples illustrating our convergence estimates.

- The numerical illustrations were obtained by adapting and modifying two existing codes written in Fortran (for MHM) and C++ (for FEM).

Chapter 2

Stability analysis with DtN boundary conditions

Contents

The goal of this chapter is to study the properties of our model boundary value problem. Specifically, we are interested in stability properties that are explicit with respect to the frequency. We start by rigorously introducing our model problem in Section 2.1. To do so, we specify the functional framework in Subsection 2.1.1 and collect preliminary definitions in Subsections 2.1.3 to 2.1.8. We introduce our model Helmholtz problem in Subsection 2.1.9, and discuss the choice of energy norm in which we derive our stability results in Subsection 2.1.10. Sections 2.2 and 2.3 then contain our stability analysis. Specifically, the case of a homogeneous propagation medium is considered in Section 2.2, whereas Section 2.3 treats the case of finely structured layered media.

2.1 Helmholtz problem in periodic structures

2.1.1 Functional spaces

The precise definition and analysis of weak formulations for Helmholtz problems requires suitable functional spaces that we introduce here.

If $D \subset \mathbb{R}^2$ is an open domain with Lipschitz boundary, $L^2(D)$ is the usual Lebesgue space of complex-valued square integrable functions defined on D. We denote by

$$\|v\|_D^2 := \int_D |v|^2 \qquad \forall v \in L^2(D)$$

the usual norm of $L^2(D)$. $\boldsymbol{L}^2(D) := [L^2(D)]^2$ contains vector valued functions, and we still employ the notation $\|\cdot\|_D$ for its usual norm. The notation $(\cdot,\cdot)_D$ is used for both the inner products of $L^2(D)$ and $\boldsymbol{L}^2(D)$.

If $\rho : D \to \mathbb{R}$ is measurable function such that $0 < \rho_{\min} \leq \rho \leq \rho_{\max} < +\infty$ a.e. in D for two constant $\rho_{\min}, \rho_{\max} \in \mathbb{R}$, then

$$\|v\|_{\rho,D}^2 := \int_D \rho|v|^2 \qquad \forall v \in L^2(D)$$

is a norm on $L^2(D)$ equivalent to the standard one that we shall frequently employ. Similarly, if $\boldsymbol{A} : D \to \mathbb{R}^{2\times 2}$ is a symmetric matrix with two constant $\alpha_{\min}, \alpha_{\max} \in \mathbb{R}$ such that

$$0 < \alpha_{\min} \leq \min_{\substack{\boldsymbol{\xi}\in\mathbb{R}^2 \\ |\boldsymbol{\xi}|=1}} \boldsymbol{A}(\boldsymbol{x})\boldsymbol{\xi}\cdot\boldsymbol{\xi} \qquad \max_{\substack{\boldsymbol{\xi},\boldsymbol{\xi}'\in\mathbb{R}^2 \\ |\boldsymbol{\xi}|=|\boldsymbol{\xi}'|=1}} \boldsymbol{A}(\boldsymbol{x})\boldsymbol{\xi}\cdot\boldsymbol{\xi}' \leq \alpha_{\max} < +\infty$$

for a.e. $\boldsymbol{x}$ in D, then we will often employ the following norm

$$\|\boldsymbol{v}\|_{\boldsymbol{A},D}^2 := \int_D \boldsymbol{A}\boldsymbol{v}\cdot\boldsymbol{v} \qquad \forall \boldsymbol{v} \in \boldsymbol{L}^2(D)$$

on $\boldsymbol{L}^2(D)$.

If $\Gamma \subset \mathbb{R}^2$ is a one-dimensional manifold, then $L^2(\Gamma)$, $(\cdot,\cdot)_\Gamma$ and $\|\cdot\|_\Gamma$ are defined similarly using the surface measure on Γ.

If $v \in L^2(D)$, and $n = 1$ or 2, the symbol $\partial v/\partial \boldsymbol{x}_n$ stands for the weak derivative of v in the sense of distributions. Then, the following Sobolev spaces

$$H^1(D) := \left\{ v \in L^2(D) \mid \frac{\partial v}{\partial \boldsymbol{x}_n} \in L^2(D);\ 1 \leq n \leq 2 \right\},$$

and

$$H^2(D) := \left\{ v \in L^2(D) \mid \frac{\partial^2 v}{\partial \boldsymbol{x}_n \boldsymbol{x}_m} \in L^2(D);\ 1 \leq n, m \leq 2 \right\}$$

will be useful. If $v \in H^1(D)$, we will employ the compact notation

$$\boldsymbol{\nabla} v := \left(\frac{\partial v}{\partial \boldsymbol{x}_1}, \frac{\partial v}{\partial \boldsymbol{x}_2} \right) \in \boldsymbol{L}^2(\Omega)$$

for its the weak gradient.

We respectively refer the reader to chapters 4 and 9 of [20] for more details on the Lebesgue and Sobolev spaces.

2.1.2 From Maxwell to Helmholtz equations

Since light is an electromagnetic wave, its propagation is modeled by Maxwell's equations (see e.g. [4, Chapter 1]). Namely, the electric field $\boldsymbol{E}$ and the magnetic field $\boldsymbol{H}$ are linked through the equations

$$\begin{aligned} \boldsymbol{\nabla} \times \boldsymbol{E} &= -\mu \frac{\partial \boldsymbol{H}}{\partial t}, \\ \boldsymbol{\nabla} \times \boldsymbol{H} &= \varepsilon \frac{\partial \boldsymbol{E}}{\partial t}. \end{aligned} \tag{2.1.1}$$

As can be seen from (2.1.1), $\boldsymbol{E}$ and $\boldsymbol{H}$ are in general time-dependent space-varying vector fields. In this work, we focus on the time-harmonic framework, where we assume a sinusoidal time-behaviour. Specifically, we assume that

$$\boldsymbol{E}(\boldsymbol{x}, t) = \mathrm{Re}(\boldsymbol{E}(\boldsymbol{x}) e^{-ikt}) \qquad \text{and} \qquad \boldsymbol{H}(\boldsymbol{x}, t) = \mathrm{Re}(\boldsymbol{H}(\boldsymbol{x}) e^{-ikt}). \tag{2.1.2}$$

for a fixed and known frequency k. Substituting (2.1.2) in (2.1.1), electromagnetic fields satisfy the following frequency-domain Maxwell equations:

$$\begin{aligned} \boldsymbol{\nabla} \times \boldsymbol{E} &= ik\mu \boldsymbol{H}, \\ \boldsymbol{\nabla} \times \boldsymbol{H} &= -ik\varepsilon \boldsymbol{E}, \end{aligned} \tag{2.1.3}$$

where the unknown fields now only depend on the space variable $\boldsymbol{x}$.

The equations in (2.1.3) are still complicated to handle mathematically, as they involve vector-valued unknowns and curl operators, for which the functional framework is very involved [4, Chapters 2 and 3]. In this work, we focus on a two-dimensional setting for which,

fortunately, we can reformulate (2.1.3) with a scalar unknown only. Indeed, assuming that $\boldsymbol{E}$ and $\boldsymbol{H}$ are independent of $\boldsymbol{x}_3$, (2.1.3) decouples into two different two-dimensional problems corresponding to distinct polarization. (i) In the transverse magnetic (TM) mode, the magnetic field $\boldsymbol{H}$ is aligned with the $\boldsymbol{x}_3$-axis. Then, $\boldsymbol{H} = (0, 0, H_3)$, where $H_3 = H_3(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is independent of $\boldsymbol{x}_3$ and satisfies the following Helmholtz problem

$$-k^2\mu H_3 - \boldsymbol{\nabla} \cdot \left(\frac{1}{\varepsilon}\boldsymbol{\nabla} H_3\right) = 0. \tag{2.1.4}$$

On the other hand (ii), In the transverse electric (TE) mode, the electric field takes the form $\boldsymbol{E} = (0, 0, E_3(\boldsymbol{x}_1, \boldsymbol{x}_2))$, with E_3 solution to

$$-k^2\varepsilon E_3 - \boldsymbol{\nabla} \cdot \left(\frac{1}{\mu}\boldsymbol{\nabla} E_3\right) = 0. \tag{2.1.5}$$

We can reformulate both (2.1.4) and (2.1.5) under the same setting, namely, the following Helmholtz equation

$$-k^2\kappa u - \boldsymbol{\nabla} \cdot \left(\frac{1}{\rho}\boldsymbol{\nabla} u\right) = 0,$$

where κ, ρ and u depend on the polarization.

2.1.3 Plane waves and right-hand sides

Plane waves are a special case of waves whose physical characteristics are constant in one spatial direction. As a result, a plane wave can be characterized by its frequency and traveling direction. The general expression of a plane wave is thus

$$\xi_{\boldsymbol{d}}(\boldsymbol{x}) := e^{ik\boldsymbol{d}\cdot\boldsymbol{x}},$$

where k is the frequency, $\boldsymbol{d} := (d_1, d_2)$ is a unit vector representing the direction of propagation and $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2)$ represents the space coordinates.

In the case of solar cell simulations, the incoming waves actually originate from the Sun. Since relative to the characteristic wavelength and size of the device, the Sun is located extremely far from the Earth, the incoming light can be accurately modeled by a plane wave, which we will do hereafter. This is depicted in Figure 2.1.

The physical phenomenon we want to model thus corresponds to an incoming plane wave, which can be injected through a boundary of the domain as a surfacic right-hand side. However, as we will describe later on, we will carry out convergence and stability analysis of numerical schemes using duality arguments that require volumic right-hand sides. With this in mind, we will consider as a model problem a Helmholtz equation with a volumic right-hand side. We will see that it allows covering both duality arguments for the analysis of numerical schemes, and the injection of incoming plane waves for the actual simulations.

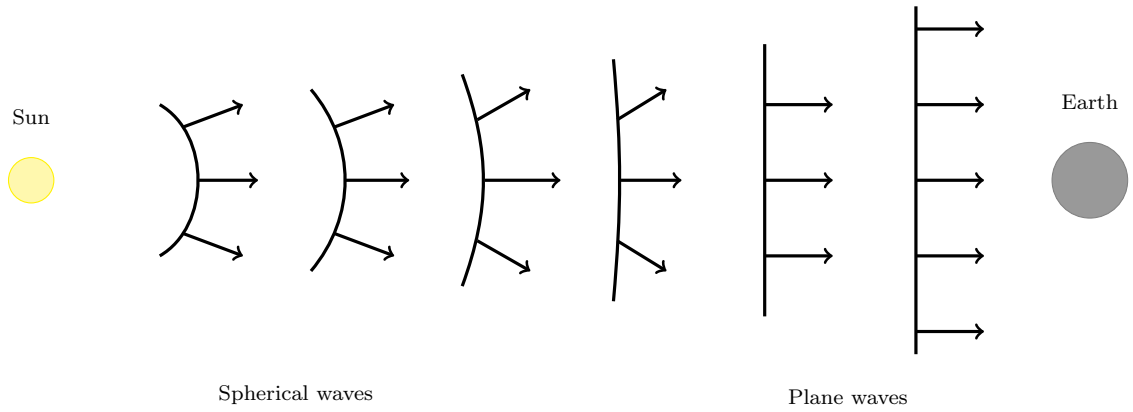


Figure 2.1: Sunlight: from spherical to plane waves.

2.1.4 Description of the propagation medium

A mono-periodic structure as the one shown in Figure 1.3, can be described without loss of generality by assuming that its physical characteristics are constant in the $\boldsymbol{x}_3$ direction, and periodic in $\boldsymbol{x}_1$ direction, with a period equal to a ℓ_1. In this context, taking a cross-section, we obtain a two-dimensional medium posed in the $\boldsymbol{x}_1 - \boldsymbol{x}_2$ plane, and the ℓ_1-periodicity allows us to restrict our study to a single cell of width ℓ_1.

Mathematically, we are led to the setting illustrated in Figure 2.3, which is known in the literature as a "1D grating problem". It is related to the two-dimensional scattering of waves by periodic structures. Since the 80's, many works have been interested in this problem in the applied mathematical community, see e.g. [38, 113, 47, 9, 13].

Let $\Omega = (0, \ell_1) \times (0, \ell_2)$ be a two-dimensional rectangular domain. Its boundary is divided into three parts:

- $\Gamma_\sharp$ corresponds to the vertical sides of the rectangle. Namely $\Gamma_\sharp := {\Gamma_\sharp}^+ \cup {\Gamma_\sharp}^-$, ${\Gamma_\sharp}^- = \{0\} \times (0, \ell_2)$ ${\Gamma_\sharp}^+ = \{\ell_1\} \times (0, \ell_2)$. On $\Gamma_\sharp$, we impose quasi-periodic boundary conditions (see subsection 2.1.5 for more details). To simplify further discussions, we introduce the notations $v_+ = v|_{{\Gamma_\sharp}^+}$ and $v_- = v|_{{\Gamma_\sharp}^-}$ for all $v \in H^1(\Omega)$.

- $\Gamma_{\mathrm{D}} := (0, \ell_1) \times \{0\}$ is the bottom of the rectangle. We assume that no light is transmitted through this interface, which we model with a Dirichlet condition.

- $\Gamma_{\mathrm{A}} := (0, \ell_2) \times \{\ell_2\}$ is the top of rectangle. This interface is not physical, and corresponds to an artificial boundary used to close the computational domain. On Γ_{A}, we prescribe a "transparent" boundary condition to account for a semi-infinite propagation medium. This is detailed in subsection 2.1.8 below.

The materials contained in Ω are physically characterized by their dielectric permittivity $\varepsilon\colon \Omega \to \mathbb{R}$ and magnetic permeability $\boldsymbol{\mu}\colon \Omega \to \mathbb{R}^{2\times 2}$. In fact, the inverse of the permeability $\boldsymbol{A} := \boldsymbol{\mu}^{-1}$ directly enters the equation.

In this chapter, we will focus on two different situations. First, we will consider the one layer case, where we assume that $\varepsilon = 1$ and $\boldsymbol{A} = \boldsymbol{I}$ in Ω. This correspond to a homogeneous medium. Second, we will consider the multi-layers case, where the coefficients ε and $\boldsymbol{A}$ are respectively required to increase and decrease when $\boldsymbol{x}_2$ increases. This is detailed in Assumption 2.1.1 (e.g., the illustration in Figure 2.2).

Assumption 2.1.1. *The matrix $\boldsymbol{A}$ is diagonal, i.e.*

$$\boldsymbol{A} := \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

for two measurable scalar functions $A_1, A_2 : \Omega \to \mathbb{R}$, and $\varepsilon : \Omega \to \mathbb{R}$ is a measurable function. There exist constants $\varepsilon_{\min}, \varepsilon_{\max} \in \mathbb{R}$ and $A_{\min}, A_{\max} \in \mathbb{R}$ such that $0 < \varepsilon_{\min} \leq \varepsilon \leq \varepsilon_{\max} < +\infty$, $0 < A_{\min} \leq A_1 \leq A_{\max} < +\infty$ and $0 < A_{\min} \leq A_2 \leq A_{\max} < +\infty$ a.e. in Ω. There exists a neighborhood of Γ_{A} in which $\varepsilon \equiv 1$ and $\boldsymbol{A} \equiv \boldsymbol{I}$.

In addition, we assume that Ω is partitioned into N subdomains $\{\Omega_j\}_{j=1}^N$ in such way that:

(i) For $1 \leq j \leq N-1$ let $\Gamma_j := \overline{\Omega_j} \cap \overline{\Omega_{j+1}}$, and $\boldsymbol{n}^j := \boldsymbol{n}_{\Omega_j}$. We assume that

$$\boldsymbol{n}_2^j \geq 0.$$

and that $\partial\Omega_j \setminus \Gamma_\sharp \subset \Gamma_{j-1} \cup \Gamma_j$ for $1 \leq j \leq N$, where we wrote $\Gamma_0 := \overline{\Gamma_{\mathrm{D}}}$ and $\Gamma_N := \overline{\Gamma_{\mathrm{A}}}$ for the sake of simplicity.

(ii) The coefficients are smooth in each subdomains, meaning that $\varepsilon|_{\Omega_j}, A_1|_{\Omega_j}, A_2|_{\Omega_j} \in C^{1,1}(\overline{\Omega_j})$, for $1 \leq j \leq N$.

(iii) For $1 \leq j \leq N$, we assume that the periodic extension $\Omega_j^\sharp$ of Ω_j is of class C^2 in the sense of [20, section 9.6]. Roughly speaking, it means that the periodic extension $\Gamma_j^\sharp$ of the interface of each interface $\Gamma_j^\sharp$ can be locally expressed as the graph of a function of class C^2.

(iv) ε is increasing with $\boldsymbol{x}_2$ whereas $\boldsymbol{A}$ is decreasing. Specifically

$$\frac{\partial \varepsilon|_{\Omega_j}}{\partial \boldsymbol{x}_2} \geq 0 \quad \frac{\partial A_1|_{\Omega_j}}{\partial \boldsymbol{x}_2} \leq 0 \quad \frac{\partial A_2|_{\Omega_j}}{\partial \boldsymbol{x}_2} \leq 0$$

for $1 \leq j \leq N$, and

$$[\![\varepsilon]\!]_{\Gamma_j} \geq 0 \quad [\![A_1]\!]_{\Gamma_j} \leq 0 \quad [\![A_2]\!]_{\Gamma_j} \leq 0$$

for $1 \leq j \leq N-1$, where we employed the standard notation

$$[\![v]\!]_{\Gamma_j} = \left(v|_{\Omega_{j+1}}\right)|_{\Gamma_j} - \left(v|_{\Omega_j}\right)|_{\Gamma_j}$$

for the jump of a function $v \in C^{1,1}(\Omega_j) \cup C^{1,1}(\Omega_{j+1})$ through an interface Γ_j.

Remark 2.1.2 (Minimal wavespeed)**.** *An important consequence of (iv) in Assumption (2.1.1) is that*

$$\varepsilon_{\max} = 1 \quad \text{and} \quad A_{\min} = 1. \tag{2.1.6}$$

Physically, it implies in particular that the minimal wave speed is 1.

The assumption that of Γ_{A} *we have* $\varepsilon \equiv 1$ *and* $\boldsymbol{A} \equiv \boldsymbol{I}$ *in a neighborhood is made for convenience, and a similar analysis could be done by only assuming that* ε *is constant and that* $\boldsymbol{A}$ *is constant and isotropic in a neighborhood of* Γ_{A}.

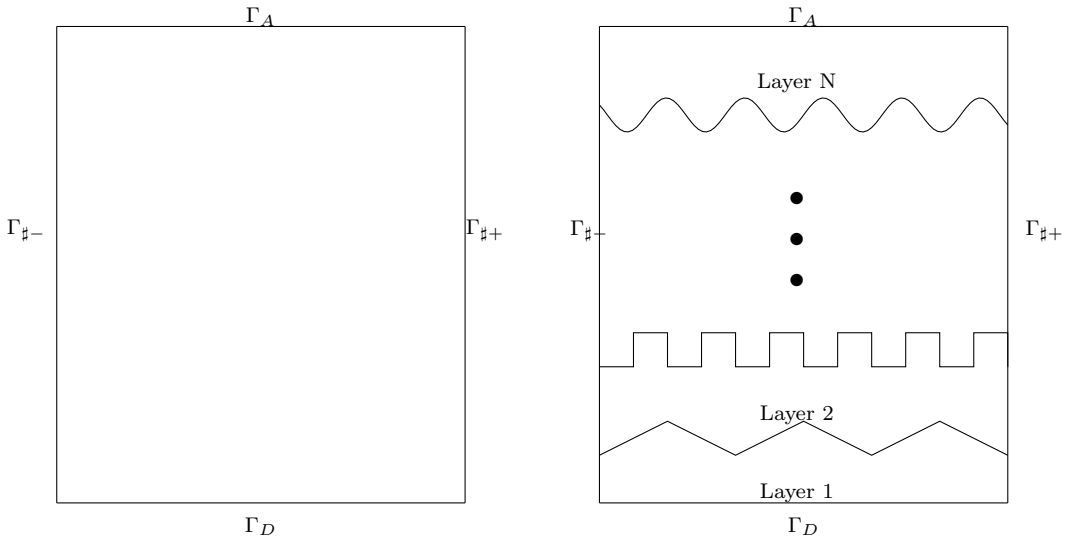


Figure 2.2: One layer case (left) and multi-layers case (right).

2.1.5 Quasi-periodicity

As explained above and depicted in Figure 2.3, we consider a two dimensional structure periodic in the $\boldsymbol{x}_1$ direction, with period ℓ_1, and an incoming plane wave $\xi_\theta(\boldsymbol{x}) := e^{ik\boldsymbol{d}\cdot\boldsymbol{x}}$, where k is the frequency and $\boldsymbol{d} := (\sin\theta, -\cos\theta)$ is a unit vector representing the direction of propagation with θ being the angle of incidence. When $\theta := 0$, we have $\boldsymbol{d} = (0, -1)$ and the incident wave approaches vertically. This is know as "normal incidence". The limiting values of θ are $\pm\pi/2$ corresponding to horizontal incoming waves. The cases where $\theta \neq 0$ are known as "oblique incidences".

Straightforward computations show that

$$\xi_\theta(\boldsymbol{x}_1 + \ell_1, \boldsymbol{x}_2) = e^{i\alpha\ell_1}\xi_\theta(\boldsymbol{x}_1, \boldsymbol{x}_2), \qquad \alpha := k\sin\theta.$$

and we say that the incident wave satisfies a quasi-periodicity condition. In this case, the field u generated by the incident wave is assumed to satisfy the same quasi-periodicity condition than the incident wave and its values can always been deduced from the interval $(0, \ell_1)$. As a result, we will simply say that $u : (0, \ell_1) \times \mathbb{R}_+ \to \mathbb{C}$ is quasi-periodic if

$$u(\ell_1, \boldsymbol{x}_2) = e^{i\alpha\ell_1}u(0, \boldsymbol{x}_2),$$

and we are going to write

$$u_+ - e^{i\alpha\ell_1} u_- = 0 \quad \text{on} \quad \Gamma_\sharp. \tag{2.1.7}$$

This definition also applies with slight modifications to univariate functions $u : (0, \ell_1) \to \mathbb{C}$.

The following Sobolev space, incorporating the quasi-periodic boundary conditions, will be useful

$$H^1_\sharp(\Omega) := \left\{ v \in H^1(\Omega, \mathbb{C}) \mid v_{|\Gamma_D} = 0 \text{ and } v_+ = e^{i\alpha\ell_1} v_- \right\}.$$

We denote by $H^{1/2}_\sharp(\Gamma_{\mathrm{A}})$ the image of $H^1_\sharp(\Omega)$ through the trace operator on Γ_{A}.

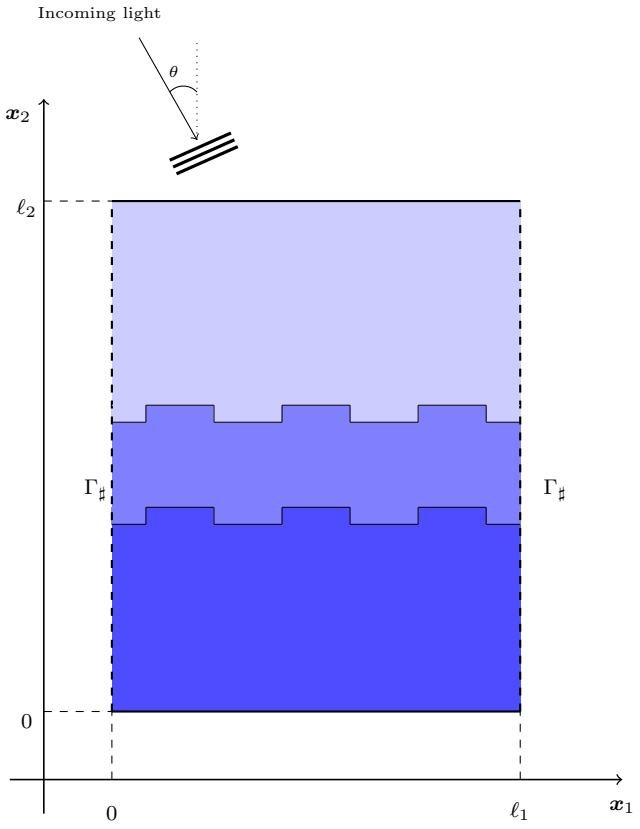


Figure 2.3: A cross section of a mono-periodic texturation (presented in Figure 1.3).

2.1.6 Fourier expansion

Expansion into Fourier series is a standard tool that will often be useful throughout this manuscript. In particular, we will employ it to reduce the study of a two-dimensional PDE to a system of uncoupled one-dimensional PDEs. Standard Fourier expansion applies to periodic functions. However, we show here that the technique also works on quasi-periodic functions, up to slight adjustments.

Let $I_1, I_2 \subset \mathbb{R}$ be two open bounded interval, $D := I_1 \times I_2$, and consider a quasi-periodic function $u \in H^1(D)$. It is easily seen that the function $v(\boldsymbol{x}) := u(\boldsymbol{x}) e^{-i\alpha \boldsymbol{x}_1}$ is ℓ_1-periodic

in the $\boldsymbol{x}_1$ direction. Applying Fourier expansion to v, we get the expansion

$$u(\boldsymbol{x}) = \sum_{n\in\mathbb{Z}} \widehat{u}_n(\boldsymbol{x}_2) e^{i(\alpha+\alpha_n)\boldsymbol{x}_1}. \tag{2.1.8}$$

where the convergence takes place in $H^1(\Omega)$, and

$$\alpha_n := \frac{2n\pi}{\ell_1}, \qquad \widehat{u}_n(\boldsymbol{x}_2) := \frac{1}{\ell_1}\int_0^{\ell_1} u(\boldsymbol{x}_1,\boldsymbol{x}_2) e^{-i(\alpha+\alpha_n)\boldsymbol{x}_1} d\boldsymbol{x}_1,$$

for a.e. $\boldsymbol{x}_2 \in I_2$. We also have the Parseval identity

$$\|u\|_D^2 = \ell_1 \sum_{n\in\mathbb{Z}} \|\widehat{u}_n\|_{I_2}^2. \tag{2.1.9}$$

Now if $u \in H^2(\Omega)$, then the function $f := -k^2 u - \Delta u \in L^2(\Omega)$ is quasi-periodic, and each mode satisfies

$$-\left[k^2 - (\alpha+\alpha_n)^2\right]\widehat{u}_n - \frac{d^2\widehat{u}_n}{d\boldsymbol{x}_2^2} = \widehat{f}_n.$$

To simplify the above equation, we define n_c as the smallest integer such that $(\alpha+\alpha_n)^2 - k^2 \le 0$ for $-n_c \le n \le n_c$. We then define $\beta_n := \sqrt{|k^2-(\alpha+\alpha_n)^2|}$, and

$$k_n := \begin{cases} \beta_n & \text{if } |n| \le n_c \\ i\beta_n & \text{otherwise.} \end{cases} \tag{2.1.10}$$

Notice that k_n is either a non-negative real number and $k_n = |k_n|$, or a pure imaginary number with positive imaginary part and $k_n = i|k_n|$. The notation

$$k_\star := \min_{n\in\mathbb{Z}} |k_n| \tag{2.1.11}$$

will be useful in many places throughout the manuscript. We also notice that due to our definition of k_n, we have

$$-k_n^2\widehat{u}_n - \widehat{u}_n'' = \widehat{f}_n. \tag{2.1.12}$$

We can equip $H_\sharp^{1/2}(\Gamma_A)$ with the norm

$$\|v\|^2_{H_\sharp^{1/2}(\Gamma_A)} := \ell_1 \sum_{n\in\mathbb{Z}} (1+n)|\widehat{v}_n|^2 \quad \forall v \in H_\sharp^{1/2}(\Gamma_A). \tag{2.1.13}$$

2.1.7 Quasi-resonant modes

An important aspect of the mathematical and numerical studies of 1D grating problems is the presence of so-called quasi-resonant frequencies (also called anomalous modes or Rayleigh frequencies) [137, 136]. Indeed, the presence of these quasi-resonances substantially affects the properties of the solution. In fact, because their presence makes the analysis more complex, many works in the literature on 1D gratings exclude these anomalous modes from their study. Among them, we may cite, for instance [10, 11, 12, 40, 46, 47].

Here, we pay particular attention to the presence of anomalous modes and to their impact on various properties of the solution. This is especially important as anomalous modes also strongly impact numerical schemes.

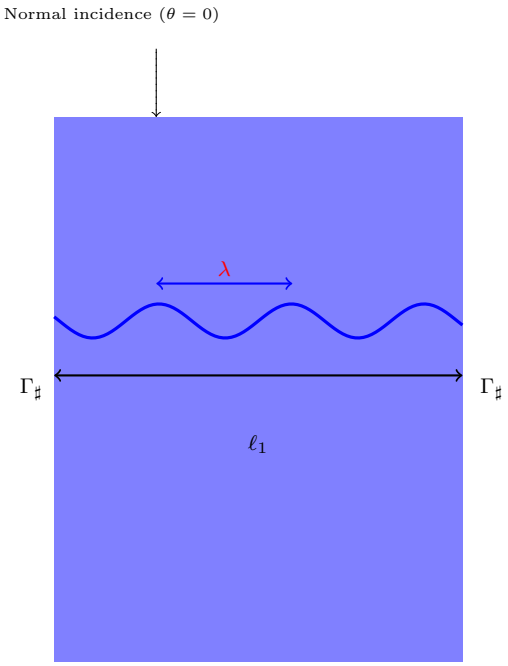


Figure 2.4: Quasi-resonant mode for a normal incidence in homogeneous cases.

A quasi-resonant mode $\widehat{u}_n$ corresponds to a value of n for which $k_n = 0$. In our setting, it can appear if there exists a value of $n \in \mathbb{N}$ such that

$$(1 - \sin\theta)\ell_1 k = 2\pi n.$$

In the particular case of normal incidence when $\theta = 0$, this expression simplifies to

$$\ell_1 k = 2\pi n,$$

meaning that the periodicity length ℓ_1 is a multiple the wavelength $\lambda = 2\pi/k$. Notice that for "most" frequencies k, there is no quasi-resonance. However, we shall see that quasi-periodic boundary conditions impact the stability properties of the Helmholtz problem for all frequencies.

2.1.8 Dirichlet-to-Neumann map

As previously mentioned, we will eventually need bounded computational domains in view of finite element discretizations. The Dirichlet-to-Neumann (DtN) operator is a convenient way to represent unbounded problems in a bounded domain with suitable boundary conditions. The truncation of the domain is not only crucial for FEM discretizations, but it

is also very convenient for the abstract analysis of the properties of the solution. In the context of wave propagation, such boundary conditions, based on the DtN operator are often refered to as transparent boundary conditions.

In our setting, the DtN operator can be explicitly constructed through Fourier series [11]. Consider a quasi-periodic function $u \in H^2_{\mathrm{loc}}((0,\ell_1)\times(\ell_2,+\infty))$ such that $-k^2u-\Delta u = 0$. We then know from the previous section that we have the decomposition

$$u(\boldsymbol{x}) = \sum_{n\in\mathbb{N}} \widehat{u}_n(\boldsymbol{x}_2) e^{i(\alpha+\alpha_n)\boldsymbol{x}_1}$$

with each mode $\widehat{u}_n$ satisfying the one-dimensional equation

$$-k_n^2\widehat{u}_n - \widehat{u}_n'' = 0.$$

It is then easily seen that for each $n \in \mathbb{N}$,

$$\widehat{u}_n(\boldsymbol{x}_2) = c_+ e^{ik_n\boldsymbol{x}_2} + c_- e^{-ik_n\boldsymbol{x}_2}.$$

for two complex numbers $c_\pm \in \mathbb{C}$. Then, there are two distinct scenarios. If $|n| \leq n_{\mathrm{c}}$, then $k_n = |k_n|$. The first term corresponds to a wave going up, while the second term corresponds to a wave going down. In the second case where $|n| > n_{\mathrm{c}}$, we have $ik_n = -|k_n|$ and the second term is blowing up, while the first term corresponds to an evanescent wave. In both cases, only the first term is physically releveant, and the condition

$$\widehat{u}_n'(\ell_2) = ik_n\widehat{u}_n(\ell_2) \tag{2.1.14}$$

will ensure that $c_- = 0$. This leads to the formal definition of the DtN operator

$$\mathscr{R}v := i\sum_{n\in\mathbb{N}} k_n\widehat{v}_n e^{i(\alpha+\alpha_n)\boldsymbol{x}_1}$$

for any "suitable" function $v : (0,\ell_1)\times(0,\ell_2) \to \mathbb{C}$. Mathematically, we can summarize the key properties of the DtN operator as follows.

Theorem 2.1.3 (DtN operator)**.** *We have $\mathscr{R} : H^{1/2}_\sharp(\Gamma_{\mathrm{A}}) \to \left(H^{1/2}_\sharp(\Gamma_{\mathrm{A}})\right)'$, with the definition*

$$\mathscr{R}v := i\sum_{n\in\mathbb{Z}} k_n\widehat{v}_n e^{i(\alpha+\alpha_n)\boldsymbol{x}_1} \qquad \forall v \in H^{1/2}_\sharp(\Gamma_{\mathrm{A}}). \tag{2.1.15}$$

In addition, we have

$$\frac{1}{\ell_1}\langle\mathscr{R}v, v\rangle_{\Gamma_{\mathrm{A}}} = i\sum_{|n|\leq n_{\mathrm{c}}} |k_n||\widehat{v}_n|^2 - \sum_{|n|\geq n_{\mathrm{c}}} |k_n||\widehat{v}_n|^2 \quad \forall v \in H^{1/2}_\sharp(\Gamma_{\mathrm{A}}) \tag{2.1.16}$$

and

$$\frac{1}{\ell_1}\|\mathscr{R}v\|^2_{\Gamma_{\mathrm{A}}} = \sum_{|n|\leq n_{\mathrm{c}}} |k_n|^2|\widehat{v}_n|^2 + \sum_{|n|\geq n_{\mathrm{c}}} |k_n|^2|\widehat{v}_n|^2 \quad \forall v \in H^1_\sharp(\Gamma_{\mathrm{A}}). \tag{2.1.17}$$

Proof. Let us consider two smooth functions $v, \phi \in C^\infty(\overline{\Gamma_{\rm A}})$. We have

$$(\mathscr{R}v, \phi)_{\Gamma_{\rm A}} = i \sum_{n\in\mathbb{Z}} k_n \widehat{v}_n (e^{i(\alpha+\alpha_n)\boldsymbol{x}_1}, \phi)_{\Gamma_{\rm A}} = i\ell_1 \sum_{n\in\mathbb{Z}} k_n \widehat{v}_n \overline{\widehat{\phi}_n},$$

so that by Holdër inequality

$$\begin{aligned}|(\mathscr{R}v, \phi)_{\Gamma_{\rm A}}| \leq \ell_1 \sum_{n\in\mathbb{Z}} |k_n||\widehat{v}_n||\widehat{\phi}_n| \leq \left(\sup_{n\in\mathbb{Z}} \frac{|k_n|}{1+n}\right) \ell_1 \sum_{n\in\mathbb{Z}} (1+|n|)|\widehat{v}_n||\widehat{\phi}_n| \leq \\ \left(\sup_{n\in\mathbb{Z}} \frac{|k_n|}{1+n}\right) \left(\ell_1 \sum_{n\in\mathbb{Z}} (1+|n|)|\widehat{v}_n|^2\right)^{1/2} \left(\ell_1 \sum_{n\in\mathbb{Z}} (1+|n|)|\widehat{\phi}_n|^2\right)^{1/2} \\ = \left(\sup_{n\in\mathbb{Z}} \frac{|k_n|}{1+|n|}\right) \|v\|_{H^{1/2}_\sharp(\Gamma_{\rm A})} \|\phi\|_{H^{1/2}_\sharp(\Gamma_{\rm A})},\end{aligned}$$

and by density of $C^\infty(\overline{\Gamma_{\rm A}})$ in $H^{1/2}_\sharp(\Gamma_{\rm A})$, we obtain that

$$\|\mathscr{R}v\|_{(H^{1/2}_\sharp(\Gamma_{\rm A}))'} \leq \left(\sup_{n\in\mathbb{Z}} \frac{|k_n|}{1+|n|}\right) \|v\|_{H^{1/2}_\sharp(\Gamma_{\rm A})}.$$

Then, a careful inspection of the definition $k_n :=$ given at (2.1.10) reveals that

$$|k_n| = \sqrt{\left|k^2 - \left(\alpha + \frac{2n\pi}{\ell_1}\right)^2\right|},$$

so that

$$\lim_{|n|\to+\infty} \frac{|k_n|}{1+|n|} = \frac{2\pi}{\ell_1}$$

and

$$\|\mathscr{R}v\|_{(H^{1/2}_\sharp(\Gamma_{\rm A}))'} \leq C\|v\|_{H^{1/2}_\sharp(\Gamma_{\rm A})},$$

for all smooth v, showing that $\mathscr{R}$ indeed maps $H^{1/2}_\sharp(\Gamma_{\rm A})$ into its dual using again the density of smooth functions.

Then, (2.1.16) and (2.1.17) follows from the following simple computations:

$$\begin{aligned}\frac{1}{\ell_1}\langle \mathscr{R}v, v\rangle_{\Gamma_{\rm A}} = i \sum_{n\in\mathbb{N}} k_n \widehat{v}_n \langle \frac{1}{\ell_1} e^{i(\alpha+\alpha_n)}, v\rangle_{\Gamma_{\rm A}} = i \sum_{n\in\mathbb{N}} k_n |\widehat{v}_n|^2 \\ = i \sum_{|n|\leq n_{\rm c}} |k_n||\widehat{v}_n|^2 - \sum_{|n|\geq n_{\rm c}} |k_n||\widehat{v}_n|^2 \quad \forall v \in H^{1/2}_\sharp(\Gamma_{\rm A}),\end{aligned}$$

and

$$\frac{1}{\ell_1}\|\mathscr{R}v\|^2_{\Gamma_{\rm A}} = -\sum_{n\in\mathbb{N}} k_n^2 |\widehat{v}_n|^2 = -\sum_{|n|\leq n_{\rm c}} |k_n|^2|\widehat{v}_n|^2 + \sum_{|n|\geq n_{\rm c}} |k_n|^2|\widehat{v}_n|^2 \quad \forall v \in H^1_\sharp(\Gamma_{\rm A}).$$

□

2.1.9 The model Helmholtz problem

We are now ready to state our model boundary value problem modeling the propagation of the electromagnetic field in the domain of interest. Let $\Omega := (0,\ell_1)\times(0,\ell_2)$ for $0 < \ell_1,\ell_2 < +\infty$. The electric field, denoted here by $u : \Omega \to \mathbb{C}$, is solution of the following Helmholtz problem

$$\left\{\begin{array}{rcll} -k^2\varepsilon u - \boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla}u) &=& \varepsilon f & \text{in } \Omega,\\ \boldsymbol{A}\boldsymbol{\nabla}u\cdot\boldsymbol{n} - \mathscr{R}u &=& 0 & \text{on } \Gamma_{\mathrm{A}},\\ u &=& 0 & \text{on } \Gamma_{\mathrm{D}},\\ u_+ - e^{i\alpha\ell_1}u_- &=& 0 & \text{on } \Gamma_\sharp. \end{array}\right. \tag{2.1.18}$$

where ε and $\boldsymbol{A}$ describe the permittivity and (inverse of the) permeability of the medium and $f \in L^2(\Omega)$ represents the electromagnetic source.

The variational (or weak) formulation of (2.1.18) consists in looking for $u \in H^1_\sharp(\Omega)$ such that

$$b(u,v) = (\varepsilon f, v)_\Omega \quad \forall v \in H^1_\sharp(\Omega), \tag{2.1.19}$$

where

$$b(u,v) := -k^2(\varepsilon u, v)_\Omega - \langle \mathscr{R}u, v\rangle_{\Gamma_{\mathrm{A}}} + (\boldsymbol{A}\boldsymbol{\nabla}u, \boldsymbol{\nabla}v)_\Omega.$$

2.1.10 Inf-sup condition and energy norms

Hereafter, we are interested in well-posedness in the sense of Hadamar, which means that for each right-hand side f, there exists a unique u solution to (2.1.19), and that u continuously depends on f. Since we are considering a linear problem, it is well-known that well-posedness is equivalent to an inf-sup condition [20, Theorems 2.20 and 2.21]. Specifically, Problem (2.1.19) is well-posed if and only if

$$\inf_{\substack{u\in H^1_\sharp(\Omega)\\ \|u\|_\star=1}} \sup_{\substack{v\in H^1_\sharp(\Omega)\\ \|v\|_\star=1}} \mathrm{Re}\, b(u,v) > 0,$$

where $\|\cdot\|_\star$ is any Hilbertian norm on $H^1_\star(\Omega)$.

For a fixed frequency, the choice of the $\|\cdot\|_\star$ norm is not extremely important. However, we will be interested in the dependence of the inf-sup constant on the frequency. In this case, it turns out that the "correct" norm has to be k-weighted, see e.g. [101]. We will thus consider the following "energy" norm

$$|\!|\!| v |\!|\!|^2_{k,\Omega} := k^2\|v\|^2_{\varepsilon,\Omega} + \|\boldsymbol{\nabla}v\|^2_{\boldsymbol{A},\Omega} \qquad \forall v \in H^1(\Omega), \tag{2.1.20}$$

and introduce the inf-sup constant $\mathscr{C}_{\mathrm{is}}$,

$$\frac{1}{\mathscr{C}_{\mathrm{is}}} := \inf_{\substack{u\in H^1_\sharp(\Omega)\\ |\!|\!|u|\!|\!|_{k,\Omega}=1}} \sup_{\substack{v\in H^1_\sharp(\Omega)\\ |\!|\!|v|\!|\!|_{k,\Omega}=1}} \mathrm{Re}\, b(u,v) > 0, \tag{2.1.21}$$

that satisfies $0 < \mathscr{C}_{\mathrm{is}} < +\infty$ whenever Problem (2.1.19) is well-posed.

The inf-sup condition is actually equivalent to the existence of a unique solution $u \in H^1_\sharp(\Omega)$ for every generalized right-hand side $f \in (H^1_\star(\Omega))'$ with the estimate

$$\|\!|u|\!\|_{k,\Omega} \leq \mathscr{C}_{\rm is} \sup_{\substack{v \in H^1_\sharp(\Omega) \\ \|\!|v|\!\|_{k,\Omega}=1}} \operatorname{Re} \langle f, v \rangle_\Omega \,.$$

In the next Lemma, we follow [102] to show that in the particular case of Helmholtz problems, establishing a stability estimate for $L^2(\Omega)$ right-hand sides is equivalent to derive an inf-sup condition.

Lemma 2.1.4. *Assume that for every $f \in L^2(\Omega)$, there exists a unique $u \in H^1_\sharp(\Omega)$ solution to (2.1.19) and that*

$$k\|u\|_{\varepsilon,\Omega} \leq \frac{\mathscr{C}_{\rm st}}{k}\|f\|_{\varepsilon,\Omega}. \tag{2.1.22}$$

Then, we have

$$\|\!|u|\!\|_{k,\Omega} \leq \frac{\mathscr{C}_{\rm st,e}}{k}\|f\|_{\varepsilon,\Omega}, \qquad \mathscr{C}_{\rm st,e} := 1 + 2\mathscr{C}_{\rm st} \tag{2.1.23}$$

and (2.1.21) holds true with

$$\mathscr{C}_{\rm is} \leq 1 + 2\mathscr{C}_{\rm st,e} = 3 + 4\mathscr{C}_{\rm st}. \tag{2.1.24}$$

Conversely, if (2.1.21) holds true, then (2.1.22) and (2.1.23) hold with the constants

$$\mathscr{C}_{\rm st} \leq \mathscr{C}_{\rm st,e} \leq \mathscr{C}_{\rm is}. \tag{2.1.25}$$

Proof. We start by establishing (2.1.23). To do so, we fix $f \in L^2(\Omega, \mathbb{C})$ and assume that $u \in H^1_\sharp(\Omega)$ solves (2.1.19) with the estimate in (2.1.22). We start by observing that due to (2.1.16), we have

$$\operatorname{Re}\{-\langle \mathscr{R} u, u\rangle_{\Gamma_{\rm A}}\} \geq 0,$$

so that

$$\operatorname{Re} b(u,u) \geq \|\boldsymbol{\nabla} u\|^2_{\boldsymbol{A},\Omega} - k^2\|u\|^2_{\varepsilon,\Omega}. \tag{2.1.26}$$

Hence,

$$\begin{aligned}
\|\!|u|\!\|^2_{k,\Omega} \leq \operatorname{Re} b(u,u) + 2k^2\|u\|^2_{\varepsilon,\Omega} &= \operatorname{Re}(\varepsilon f, u)_\Omega + 2k^2\|u\|^2_{\varepsilon,\Omega} \\
&\leq \left(\frac{1}{k}\|f\|_{\varepsilon,\Omega} + 2k\|u\|_{\varepsilon,\Omega}\right) k\|u\|_{\varepsilon,\Omega} \\
&\leq \left(\frac{1}{k}\|f\|_{\varepsilon,\Omega} + 2k\|u\|_{\varepsilon,\Omega}\right) \|\!|u|\!\|_{k,\Omega}\,,
\end{aligned}$$

and it follows from the assumption (2.1.22) that

$$\|\!|u|\!\|_{k,\Omega} \leq \frac{1}{k}\|f\|_{\varepsilon,\Omega} + 2k\|u\|_{\varepsilon,\Omega} \leq \frac{1 + 2\mathscr{C}_{\rm st}}{k}\|f\|_{\varepsilon,\Omega},$$

thus showing (2.1.23).

We then establish (2.1.24). Let $u \in H^1_\sharp(\Omega)$ be arbitary, and define ξ as the only element of $H^1_\sharp(\Omega)$ such that

$$b(w,\xi) = 2k^2(\varepsilon w, u)_\Omega$$

for all $w \in H^1_\sharp(\Omega)$. Letting $v = u + \xi \in H^1_\sharp(\Omega)$, we have

$$\operatorname{Re} b(u,v) = \operatorname{Re} b(u,u) + \operatorname{Re} b(u,\xi) = |\!|\!| u |\!|\!|^2_{k,\Omega}\,.$$

On the other hand, recalling (2.1.23)

$$|\!|\!| \xi |\!|\!|_{k,\Omega} \le 2\mathscr{C}_{\mathrm{st,e}} k \|u\|_{\varepsilon,\Omega} \le 2\mathscr{C}_{\mathrm{st,e}} |\!|\!| u |\!|\!|_{k,\Omega}$$

by assumption, and therefore $|\!|\!| v |\!|\!|_{k,\Omega} \le (1 + 2\mathscr{C}_{\mathrm{st,e}}) |\!|\!| u |\!|\!|_{k,\Omega}$, and

$$\operatorname{Re} b(u,v) = |\!|\!| u |\!|\!|^2_{k,\Omega} \ge \frac{1}{1 + 2\mathscr{C}_{\mathrm{st,e}}} |\!|\!| u |\!|\!| \, |\!|\!| v |\!|\!|\,,$$

which shows (2.1.24).

We finally turn to the converse estimate, and assume that the inf-sup condition (2.1.21) holds true and that $u \in H^1_\sharp(\Omega)$ solves (2.1.19). Then, we have

$$\begin{aligned} |\!|\!| u |\!|\!|_{k,\Omega} \le \mathscr{C}_{\mathrm{is}} \sup_{\substack{v \in H^1_\star(\Omega) \\ |\!|\!| v |\!|\!|_{k,\Omega} = 1}} \operatorname{Re} b(u,v) &\le \mathscr{C}_{\mathrm{is}} \sup_{\substack{v \in H^1_\star(\Omega) \\ |\!|\!| v |\!|\!|_{k,\Omega} = 1}} \operatorname{Re} (\varepsilon f, v)_\Omega \\ &\le \mathscr{C}_{\mathrm{is}} \sup_{\substack{v \in H^1_\star(\Omega) \\ |\!|\!| v |\!|\!|_{k,\Omega} = 1}} \left(\frac{1}{k} \|f\|_{\varepsilon,\Omega} |\!|\!| v |\!|\!|_{k,\Omega} \right) \le \frac{\mathscr{C}_{\mathrm{is}}}{k} \|f\|_{\varepsilon,\Omega}, \end{aligned}$$

which shows that (2.1.23) holds with $\mathscr{C}_{\mathrm{st,e}} \le \mathscr{C}_{\mathrm{is}}$. The estimate in (2.1.22) then follows since $k\|u\|_{\varepsilon,\Omega} \le |\!|\!| u |\!|\!|_{k,\Omega}$. □

2.2 Frequency-explicit stability estimates in the one-layer case

In this section, our goal is to derive stability estimates of the form

$$k\|u\|_{\varepsilon,\Omega} \le \frac{\mathscr{C}_{\mathrm{st}}}{k} \|f\|_{\varepsilon,\Omega}, \tag{2.2.1}$$

where $f \in L^2(\Omega)$ is an arbitrary right-had side, $u \in H^1_\sharp(\Omega)$ is the associated solution to (2.1.19), and $\mathscr{C}_{\mathrm{st}}$ is a "stability" constant independent of f. We are especially interested in "frequency-explicit" stability estimates where the dependency of $\mathscr{C}_{\mathrm{st}}$ on k is explicitly tracked. Besides their intrinsic value, frequency-explicit stability estimates are of particular

interest, as they immediately enters the stability and convergence analysis of finite element discretizations [33, 94, 101].

We shall first mention that in the case of coercive elliptic problems, stability estimates easily follows from Lax-Milgram theory [20, Corollary 5.8]. Unfortunately, this theory is not applicable to Helmholtz problems except for small frequencies. In fact, the sesquilinear form $b(\cdot,\cdot)$ in (2.1.19) is not coercive in general, but satisfies the weaker property

$$\operatorname{Re} b(v,v) \geq |\!|\!| v |\!|\!|_{k,\Omega}^2 - 2k^2 \|v\|_{\varepsilon,\Omega}^2 \qquad \forall v \in H^1_\sharp(\Omega) \tag{2.2.2}$$

called the "Gårding inequality". It shows that the sesquilinear form fails to be coercive, up to a "compact perturbation". Indeed, the negative term in (2.2.2) is an $L^2(\Omega)$ norm, and the injection $H^1(\Omega) \subset L^2(\Omega)$ is compact, due to the Rellich-Kondrachow theorem [20, Theorem 9.16]. As a result, Fedholm alternative applies [20, Theorem 6.6], and existence and stability hold if one can show uniqueness. Uniqueness itself can often be showed by invoking the unique continuation principle [1]. Unfortunately, although this method is very powerful to show the well-posedness (2.1.19) in general settings, it provides no information about the stability constant $\mathscr{C}_{\rm st}$ a part from the fact that it is finite.

For Helmholtz problems, the behavior of $\mathscr{C}_{\rm st}$ may be linked with the geometry of the domain Ω. The simplest setting to explain this link is the case of scattering by a bounded (as opposed to periodic) obstacle with a Dirichlet boundary. In this case direct connections between the trajectories of the rays reflected by the obstacle and the stability constant can be drawn, and the rigorous way to do so is through semi-classical analysis, see e.g. [139, Section 5.3]. In particular, there is an important difference between non-trapping where all rays escape the domain after a finite amount of time, and trapping geometries where some rays are trapped on a periodic orbit, see Figure 2.5.

For non-trapping geometries, the stability constant behaves as in a homogeneous medium, and it can be shown that

$$\mathscr{C}_{\rm st} \sim k\ell. \tag{2.2.3}$$

Since on the other hand, it is shown in [31] that $\mathscr{C}_{\rm st} \gtrsim k\ell$, this situation is the best possible. Trapping geometries can be further broken down into three categories named hyperbolic, parabolic and elliptic trapping as depicted on Figure 2.6.

In hyperbolic trapping, it can shown that

$$\mathscr{C}_{\rm st} \sim \ln(1+k\ell)k\ell, \tag{2.2.4}$$

so that the stability loss is very mild [24, 90]. This is due to the fact only a single ray is trapped. In the parabolic case, an infinite number of rays are trapped, but their trajectories are not stable, polynomial estimates of the form

$$\mathscr{C}_{\rm st} \sim (k\ell)^3, \tag{2.2.5}$$

can be shown in this case [29]. Finally, in the elliptic case [23, 26], it is shown that the best possible estimates are of the form

$$\mathscr{C}_{\rm st} \sim e^{\alpha(k\ell)}, \qquad \alpha > 0. \tag{2.2.6}$$

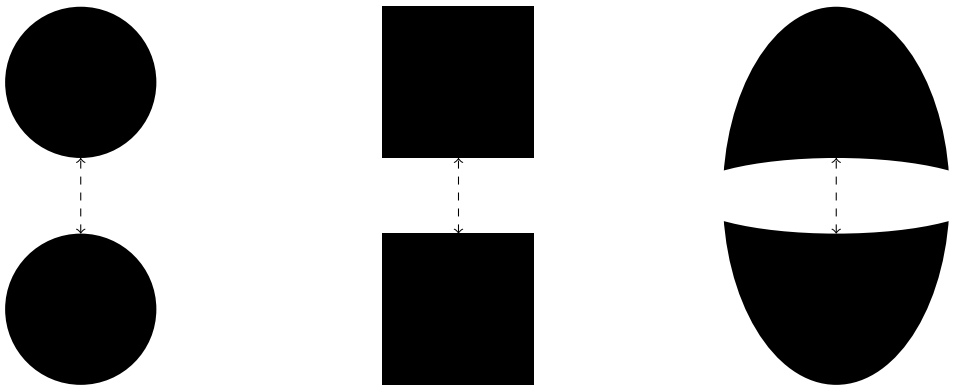
We will see that the quasi-periodic conditions act very similarly to the parallel boundaries of the two squares in Figure 2.6b. In fact, the Helmholtz problems we consider in this manuscript undergo a parabolic trapping, and we will obtain polynomial stability bounds similar to (2.2.5).

In this section, we focus on the one layer case where $\varepsilon \equiv 1$ and $\boldsymbol{A} \equiv \boldsymbol{I}$. We employ the Fourier expansion techniques of subsection 2.1.6, and then study the resulting family of one-dimensional homogeneous Helmholtz problems with frequency k_n. For one-dimensional homogeneous Helmholtz problems with a real frequency κ, several approaches are available to derive explicit stability estimates. In particular, one can employ the explicit representation of the solution using the Green's function, this approach is pursued for instance in [49, 88]. Another possibility used in [8, 30, 99] is to multiply the PDE by $x\overline{u}'$ and perform integration by parts. We will follow this approach here.

The funtion $x\overline{u}'$ is a particular case of a Morawetz (or Rellich) multiplier. Multiplying the PDE by a multiplier is a technique that also work in multiple space dimensions and some types of heterogeneous media. We will discuss this aspect in more depth in section 2.3.

(a) All rays escape (b) Some rays are trapped

Figure 2.5: Examples of trapping and non-trapping situations.



(a) Hyperbolic (b) Parabolic (c) Elliptic

Figure 2.6: Different kinds of trapping

2.2.1 Analysis of the one dimensional Helmholtz problem

As described in subsection 2.1.6, after expansion in Fourier series, we are left with a parameterized one-dimensional Helmholtz problem that we analyze in this section. Hence, going through the stability analysis of these one-dimensional problems is a key step in obtaining a stability estimate for our 2D Helmholtz problem. This desired 2D-stability result is then obtained by using the relation between the norms of the unknown u and the norms of its Fourier modes $\widehat{u}_n$. Using the definition of $k_n \in \mathbb{C}$ given in (2.1.10), if we substitute the Fourier expansion (2.1.8) in the Helmholtz problem (2.1.18), we find then that Fourier modes $\widehat{u}_n$ satisfy

$$\left\{\begin{array}{rcll} -k_n^2\widehat{u}_n - \widehat{u}_n'' & = & \widehat{f}_n & \text{in } (0,\ell_2) \\ \widehat{u}_n(0) & = & 0 & \\ \widehat{u}_n'(\ell_2) - ik_n u(\ell_2) & = & 0. & \end{array}\right. \tag{2.2.7}$$

In this subsection, $L^2(0,\ell_2)$ stands for the set of complex-valued functions that are square integrable on $(0,\ell_2)$, and $\|\cdot\|$ and $(\cdot,\cdot)$ denote the standard norm and inner product of $L^2(0,\ell_2)$. The following Sobolev space

$$W := \left\{w \in H^1(0,\ell_2) \mid w(0) = 0\right\}$$

will also be useful.

Formally, assuming that $\widehat{f}_n \in L^2(0,\ell_2)$ the weak formulation of (2.2.7) then consists in finding $\widehat{u}_n \in W$ such that

$$b_n(\widehat{u}_n,\widehat{v}_n) = (\widehat{f}_n,\widehat{v}_n) \quad \forall \widehat{v}_n \in W, \tag{2.2.8}$$

where

$$b_n(\widehat{u}_n,\widehat{v}_n) := -k_n^2(\widehat{u}_n,\widehat{v}_n) - ik_n u(\ell_2)\overline{\widehat{v}_n(\ell_2)} + (\widehat{u}_n',\widehat{v}_n').$$

We note that when $k_n = i\beta_n$, we have

$$b_n(\widehat{u}_n, \widehat{v}_n) = \beta_n^2(\widehat{u}_n, \widehat{v}_n) + \beta_n \widehat{u}_n(\ell_2)\overline{\widehat{v}_n(\ell_2)} + (\widehat{u}_n', \widehat{v}_n'). \tag{2.2.9}$$

We will prove now a useful Poincaré inequality

Lemma 2.2.1 (Poincaré inequality)**.** *We have*

$$\|\widehat{v}_n\| \leq 2\ell_2 \|\widehat{v}_n'\| \tag{2.2.10}$$

for all $\widehat{v}_n \in W$.

Proof. Let $\widehat{v}_n \in W$. Since $(|v|^2)' = 2\,\mathrm{Re}\, v\overline{v}'$, and since $\widehat{v}_n(0) = 0$, we have

$$|\widehat{v}_n(x)|^2 = 2\,\mathrm{Re}\int_0^x \widehat{v}_n\overline{\widehat{v}_n}' \leq 2\int_0^x |\widehat{v}_n||\widehat{v}_n'| \leq 2\int_0^{\ell_2} |\widehat{v}_n||\widehat{v}_n'| \leq 2\|\widehat{v}_n\|\|v'\|,$$

and the result follows after integrating both sides on $(0, \ell_2)$. □

Depending on the nature of the wave number k_n, the proof of a stability estimate for $\widehat{u}_n$ will be done in three main steps: imaginary wave numbers, small real wave numbers, and large real wave numbers.

Lemma 2.2.2 (Imaginary wave numbers)**.** *Assume that* $k_n = i\beta_n$. *Then there exists a unique* $\widehat{u}_n$ *solution to* (2.2.8). *In addition, the estimate*

$$\|\widehat{u}_n\| \leq \min\left(4, (|k_n|\ell_2)^{-2}\right)\ell_2^2\|\widehat{f}_n\| \tag{2.2.11}$$

holds true. In addition, we have

$$\|\widehat{u}_n\| \leq \min\left(4, 2(|k_n|\ell_2)^{-1}\right)\ell_2^2\|\widehat{f}_n\|. \tag{2.2.12}$$

Proof. It is clear from (2.2.9) that we have

$$\beta_n^2\|\widehat{u}_n\|^2 + \beta_n|\widehat{u}_n(\ell_2)|^2 + \|\widehat{u}_n'\|^2 = \mathrm{Re}(\widehat{f}_n, \widehat{u}_n) \leq \|\widehat{f}_n\|\|\widehat{u}_n\|. \tag{2.2.13}$$

Hence, we immediatly see that $\beta_n^2\|\widehat{u}_n\| \leq \|\widehat{f}_n\|$, and since $|k_n| = \beta_n$,

$$\|\widehat{u}_n\| \leq (|k_n|\ell_2)^{-2}\ell_2^2\|\widehat{f}_n\|. \tag{2.2.14}$$

On the other hand, using Poincaré inequality (2.2.10) in the right-hand side of (2.2.13) reveals that $\|\widehat{u}_n'\| \leq 2\ell_2\|\widehat{f}_n\|$, and using (2.2.10) again shows that

$$\|\widehat{u}_n\| \leq 4\ell_2^2\|\widehat{f}_n\|. \tag{2.2.15}$$

Estimate (2.2.11) then follows from (2.2.14) and (2.2.15).

To establish (2.2.12), we observe that the "second branch" of the of the minimum in (2.2.11) is only achived if $(|k_n|\ell_2)^{-1} \leq 2$. As a result, we have $(|k_n|\ell_2)^{-2} \leq 2(|k_n|\ell_2)^{-1}$ in this regime. □

We now consider the case where k is real positive, but "small".

Lemma 2.2.3 (Small real wavenumbers)**.** *Assume that $0 \leq k_n \ell_2 < 1/2$. Then there exists a unique $\widehat{u}_n$ solution to (2.2.7), and we have*

$$\|\widehat{u}_n\| \leq \frac{4\ell_2^2}{1 - 4k_n^2\ell_2^2}\|\widehat{f}_n\|. \tag{2.2.16}$$

In particular, if $0 \leq k_n\ell_2 \leq 1/4$, then

$$\|\widehat{u}_n\| \leq 6\ell_2^2\|\widehat{f}_n\|. \tag{2.2.17}$$

Proof. Poincaré inequality (2.2.10) imply that

$$\left(1 - 4|k_n|^2\ell_2^2\right)\|\widehat{u}_n'\|^2 \leq -k_n^2\|\widehat{u}_n\|^2 + \|\widehat{u}_n'\|^2 = \operatorname{Re} b_n(\widehat{u}_n, \widehat{u}_n),$$

so that b is coercive, and (2.2.8) admits a unique solution. We then have

$$(1 - 4k_n^2\ell_2^2)\|\widehat{u}_n'\|^2 = \operatorname{Re}(\widehat{f}_n, \widehat{u}_n) \leq \|\widehat{f}_n\|\|\widehat{u}_n\| \leq 2\ell_2\|\widehat{f}_n\|\|\widehat{u}_n'\|,$$

so that

$$\|\widehat{u}_n'\| \leq \frac{2\ell_2}{1 - 4k_n^2\ell_2^2}\|\widehat{f}_n\|,$$

and (2.2.16) follows from employing (2.2.10) again. When $k_n\ell_2 \leq 1/4$, estimate (2.2.17) follows from (2.2.16) since

$$\frac{4}{1 - 4k_n^2\ell_2^2} \leq \frac{4}{1 - 1/4} = \frac{16}{3} \leq 6.$$

□

The last scenario we need to cover is the case of a "large" real wave number.

Lemma 2.2.4 (Large real wave numbers)**.** *Assume that $k_n > 0$. Then we have*

$$\|\widehat{u}_n\| \leq \frac{3}{k_n\ell_2}\ell_2^2\|\widehat{f}_n\|. \tag{2.2.18}$$

Proof. Our analysis hinges on the so-called "Morawetz identity"

$$k_n^2\|w\|^2 + \|w'\|^2 = k_n^2\ell_2|w(\ell_2)|^2 + \ell_2|w'(\ell_2)|^2 + 2\operatorname{Re}\int_0^{\ell_2}(-k_n^2 w - w'')x\overline{w}' \tag{2.2.19}$$

valid for $w \in H^2(0, \ell_2)$.
Picking the test function $\widehat{v}_n = \widehat{u}_n$ in (2.2.8) and taking the imaginary part yields

$$-k_n|\widehat{u}_n(\ell_2)|^2 = \operatorname{Im}(\widehat{f}_n, \widehat{u}_n).$$

On the other hand, since $\widehat{f}_n \in L^2(0,\ell_2)$, then if a solution $\widehat{u}_n$ to (2.2.7) exists, then it is $H^2(0,\ell_2)$. In addition, Morawetz identity (2.2.19) shows that

$$k_n^2\|\widehat{u}_n\|^2 + \|\widehat{u}_n'\|^2 = 2\,\mathrm{Re}(\widehat{f}_n, x\widehat{u}_n') + 2k_n^2\ell_2|\widehat{u}_n(\ell_2)|^2 = 2\,\mathrm{Re}(\widehat{f}_n, xu') - 2k_n\ell_2\,\mathrm{Im}(\widehat{f}_n, \widehat{u}_n)$$

As a result, we have

$$k_n^2\|\widehat{u}_n\|^2 + \|\widehat{u}_n'\|^2 \le 2\ell_2\|\widehat{f}_n\|\|\widehat{u}_n'\| + 2k_n\ell_2\|\widehat{f}_n\|\|\widehat{u}_n\| \le \ell_2^2\|\widehat{f}_n\|^2 + \|\widehat{u}_n'\|^2 + 2\ell_2^2\|\widehat{f}_n\|^2 + \frac{k_n^2}{2}\|\widehat{u}_n\|^2,$$

and as a result

$$\frac{k_n^2}{2}\|\widehat{u}_n\|^2 \le 3\ell_2^2\|\widehat{f}_n\|^2,$$

and (2.2.18) follows since

$$\sqrt{6} \le 3.$$

□

The main stability result for one-dimensional problems is obtained by combining the lemmas covering imaginary, small real, and large real wave numbers.

Theorem 2.2.5 (Estimate for the 1D case)**.** *For all k_n, there exists a unique $\widehat{u}_n$ solution to* (2.2.7) *and we have*

$$\|\widehat{u}_n\| \le 12\min(1,(|k_n|\ell_2)^{-1})\ell_2^2\|\widehat{f}_n\|. \tag{2.2.20}$$

Proof. We focus first on the real wave numbers case where $k_n \ge 0$. Assuming that $k_n\ell_2 \ge 1$, estimate (2.2.18) imply

$$\|\widehat{u}_n\| \le 3(k_n\ell_2)^{-1}\ell_2^2\|\widehat{f}_n\| \le 3\min(1,(|k_n|\ell_2)^{-1})\ell_2^2\|\widehat{f}_n\| \le 12\min(1,(|k_n|\ell_2)^{-1})\ell_2^2\|\widehat{f}_n\|.$$

Next, if $k_n\ell_2 \le 1$, we distinguish two scenarios. When $k_n\ell_2 \le 1/4$, estimate (2.2.17) yields

$$\|\widehat{u}_n\| \le 6\ell_2^2\|\widehat{f}_n\| \le 12\ell_2^2\|\widehat{f}_n\| \le 12\min(1,(|k_n|\ell_2)^{-1})\ell_2^2\|\widehat{f}_n\|.$$

On the other hand, for $1/4 \le k_n\ell_2 \le 1$, we have $1 \le (k_n\ell_2)^{-1} \le 4$, and it follows from (2.2.18) that

$$\|\widehat{u}_n\| \le 3(k_n\ell_2)^{-1}\ell_2^2\|\widehat{f}_n\| \le 12\ell_2^2\|\widehat{f}_n\| \le 12\min(1,(|k_n|\ell_2)^{-1})\ell_2^2\|\widehat{f}_n\|.$$

This shows (2.2.20) for the case of real wave numbers.

We now consider an imaginary wave number $k_n = i\beta_n$. If $|k_n|\ell_2 \ge 1$, estimate (2.2.12) implies that

$$\|\widehat{u}_n\| \le 2(|k_n|\ell_2)^{-1}\ell_2^2\|\widehat{f}_n\| \le 2\min(1,(|k_n|\ell_2)^{-1})\ell_2^2\|\widehat{f}_n\| \le 12\min(1,(|k_n|\ell_2)^{-1})\ell_2^2\|\widehat{f}_n\|.$$

When $|k_n|\ell_2 \le 1$, we distinguish two scenarios. Assuming $|k_n|\ell_2 \le 1/2$, estimate (2.2.12) yields

$$\|\widehat{u}_n\| \le 4\ell_2^2\|\widehat{f}_n\| \le 12\ell_2^2\|\widehat{f}_n\| \le 12\min(1,(|k_n|\ell_2)^{-1})\ell_2^2\|\widehat{f}_n\|.$$

For the remaining cases where $1/2 \le |k_n|\ell_2 \le 1$, estimate (2.2.12) shows that

$$\|\widehat{u}_n\| \le 2(|k_n|\ell_2)^{-1}\ell_2^2\|\widehat{f}_n\| \le 4\ell_2^2\|\widehat{f}_n\| \le 12\min(1,(|k_n|\ell_2)^{-1})\ell_2^2\|\widehat{f}_n\|,$$

which concludes the proof. □

2.2.2 Frequency-explicit stability estimates

In this subsection, a frequency-explicit stability estimate is obtained by combining the Fourier expansion introduced in subsection 2.1.6 with the one-dimensional stability estimate (2.2.20). We first provide a preliminary result linking $k_\star$ to k.

Lemma 2.2.6 (One dimensional wave numbers)**.** *We have*

$$k_\star \ell_1 = \min_{n\in\mathbb{Z}} |k_n|\ell_1 \geq \Lambda_{k,\boldsymbol{\ell},\theta}\sqrt{k\ell_1} \tag{2.2.21}$$

where $\Lambda_{k,\boldsymbol{\ell},\theta} \in [0,1]$ *is defined by*

$$\Lambda_{k,\boldsymbol{\ell},\theta} := \sqrt{2\pi \min\left(\left\{\frac{k\ell_1(1-\sin\theta)}{2\pi}\right\}, \left\{\frac{k\ell_1(1+\sin\theta)}{2\pi}\right\}\right)},$$

with $\{x\} = \min_{n\in\mathbb{N}} |x-n|$ *for* $x \in \mathbb{R}$.

Proof. For all $n \in \mathbb{Z}$, we have

$$|k_n|^2 = |\beta_n|^2 = |k^2 - (\alpha+\alpha_n)^2| = |(k-(\alpha+\alpha_n))(k+(\alpha+\alpha_n))|.$$

We then write

$$|k_n|^2 = |k-(\alpha+\alpha_n)||k+(\alpha+\alpha_n)|$$

Then, we first treat the case where $(\alpha+\alpha_n) \geq 0$. In this case, we have

$$|k_n|^2 \geq k|k-(\alpha+\alpha_n)| = k\left|k - k\sin\theta - \frac{2n\pi}{\ell_1}\right| = 2\pi\frac{k}{\ell_1}\left|\frac{(k\ell_1)(1-\sin\theta)}{2\pi} - n\right|.$$

It follows that

$$|k_n|^2\ell_1^2 \geq 2\pi k\ell_1 \left\{\frac{k\ell_1(1-\sin\theta)}{2\pi}\right\}. \tag{2.2.22}$$

In the other case where $(\alpha+\alpha_n) < 0$, we have

$$|k_n|^2 \geq k|k+(\alpha+\alpha_n)| = k\left|k + k\sin\theta + \frac{2n\pi}{\ell_1}\right| = 2\pi\frac{k\ell_1}{\ell_1}\left|\frac{(k\ell_1)(1+\sin\theta)}{2\pi} + n\right|$$

thus

$$|k_n|^2\ell_1^2 \geq 2\pi k\ell_1 \left\{\frac{k\ell_1(1+\sin\theta)}{2\pi}\right\}. \tag{2.2.23}$$

Then, (2.2.21) follows from (2.2.22), (2.2.23) and the definition of $\Lambda_{k,\boldsymbol{\ell},\theta}$. □

We are now in a position to establish the main frequency-explicit stability result in the following theorem.

Theorem 2.2.7 (Stability in homogeneous media)**.** *Assume that* $\varepsilon \equiv 1$ *and* $\boldsymbol{A} \equiv \boldsymbol{I}$*. Then estimate* (2.1.22) *holds with*

$$\mathscr{C}_{\rm st} \leq 12 \min\left(1, (k_\star \ell_2)^{-1}\right) (k\ell_2)^2. \tag{2.2.24}$$

In particular, we have

$$\mathscr{C}_{\rm st} \leq 12 \min\left(1, \frac{\ell_1}{\ell_2} \frac{1}{\Lambda_{k,\boldsymbol{\ell},\theta}\sqrt{k\ell_1}}\right) (k\ell_2)^2. \tag{2.2.25}$$

Proof. Using Fourier expansion, we have

$$\|u\|_\Omega^2 = \ell_1 \sum_{n\in\mathbb{Z}} \|\widehat{u}_n\|^2.$$

The one dimensional result (2.2.20) leads, to

$$\|u\|_\Omega^2 \leq 12^2 \ell_1 \sum_{n\in\mathbb{Z}} \min(1, (|k_n|\ell_2)^{-2}) \ell_2^4 \|\widehat{f}_n\|^2.$$

Thus, the result follows by applying (2.2.21)

$$\|u\|_\Omega^2 \leq 12^2 \min\left(1, \frac{\ell_1}{\ell_2^2 k \Lambda_{k,\boldsymbol{\ell},\theta}^2}\right) \ell_2^4 \|f\|_\Omega^2.$$

□

The stability bound (2.2.24) is fully explicit (explicit for all problem parameters: the frequency k, the angle of incidence θ, and the medium parameters ℓ_1 and ℓ_2) ,which is a very useful result because of the significant role played by the stability constant in the analysis of numerical discretizations.

The effect of quasi-resonant modes on the continuous problem stability is neatly identified by the analysis of the bound (2.2.24). For example, if the solution is constructed from a single Fourier mode $\widehat{u}_n$ with a wave number k_n close to zero ($k_n \approx 0$), this implies that $\Lambda_{k,\boldsymbol{\ell},\theta} \approx 0$ and that the stability constant $\mathscr{C}_{\rm st}$ will be equivalent to $(k\ell_2)^2$. Furthermore, the best stability case (in terms of frequency dependence) with quasi-periodic boundary conditions is given when all modes are far from quasi-resonances. In this case, we have $\Lambda_{k,\boldsymbol{\ell},\theta} \approx 1$, and the stability constant is equivalent to $\mathscr{C}_{\rm st} \approx (k\ell_2)^{3/2}$. In contrast, the standard stability estimate for star-shaped domains surounded by a absorbing or DtN condition (thus, without quasi-periodic boundary conditions) is $\mathscr{C}_{\rm st} \approx k\ell$, where ℓ is the diameter of the domain. This is highlighted in Corollary 2.2.8 below.

Corollary 2.2.8 (Simplified stability estimates)**.** *Assuming that* $\varepsilon \equiv 1$ *and* $\boldsymbol{A} \equiv \boldsymbol{I}$*, we have*

$$\mathscr{C}_{\rm st} \leq 12(k\ell_2)^2.$$

Under the additional assumption that $\Lambda_{k,\boldsymbol{\ell},\theta} \geq \delta > 0$*, we have the improved estimate*

$$\mathscr{C}_{\rm st} \leq 12\sqrt{\frac{\ell_1}{\ell_2}\frac{1}{\delta}}(k\ell_2)^{3/2}.$$

2.2.3 Sharpness of the stability bounds

In this subsection, we show that our stability bounds are sharp with respect to the frequency k. Specifically, we want to show that in the worst case scenario, the stability constants scales as $(k\ell_2)^2$, but also highlight that even when we consider a favorable situation avoiding quasi-resonances, the scaling is still $(k\ell_2)^{3/2}$, i.e., half an order higher than for star-shaped domains without quasi-periodicity. These results are proved by constructing two infinite sequences of wave numbers and right-hand sides for which the norm of the solution grows exactly as dictated by our stability bounds.

Theorem 2.2.9 (Sharpness of the stability bounds)**.** *For all $k \in \mathbb{R}_+$ there exists a right-hand side $f \in L^2(\Omega)$ such that the associated solution $u \in H^1_\sharp(\Omega)$ to (2.1.19) satisfies*

$$\|u\|_\Omega \geq \frac{1}{2\sqrt{105}} \frac{1}{1 + k_\star \ell_2} \ell_2^2 \|f\|_\Omega. \tag{2.2.26}$$

In particular, the stability constant $\mathscr{C}_{\mathrm{st}}$ in (2.1.22) satisfies

$$\mathscr{C}_{\mathrm{st}} \geq \frac{1}{2\sqrt{105}} \frac{1}{1 + k_\star \ell_2} (k\ell_2)^2 \geq \frac{1}{4\sqrt{105}} \min(1, (k_\star \ell_2)^{-1})(k\ell_2)^2. \tag{2.2.27}$$

Remark 2.2.10 (Optimality of the stability estimates)**.** *By combining (2.2.24) and (2.2.27), we have*

$$\frac{1}{4\sqrt{105}} \min(1, (k_\star \ell_2)^{-1})(k\ell_2)^2 \leq \mathscr{C}_{\mathrm{st}} \leq 12 \min(1, (k_\star \ell_2)^{-1})(k\ell_2)^2,$$

for any stability constant $\mathscr{C}_{\mathrm{st}}$ in (2.2.24). This means in particular that our frequency dependence is optimal. In addition, our leading constant is sharp up to factor at most

$$12 \times 4\sqrt{105} = 48\sqrt{105} \leq 492.$$

Proof. Consider $k \in \mathbb{R}_+$ and $n \in \mathbb{Z}$. We start by introducing the cutoff

$$\chi(\boldsymbol{x}_2) := \ell_2^{7/2} \boldsymbol{x}_2 (\boldsymbol{x}_2 - \ell_2)^2 \qquad \forall \boldsymbol{x}_2 \in (0, \ell_2).$$

Tedious, but straightfoward, computations reveal that

$$\|\chi\| = \frac{1}{\sqrt{105}}, \quad \|\chi'\| = \sqrt{\frac{2}{15}} \frac{1}{\ell_2}, \quad \|\chi''\| = \frac{2}{\ell_2^2}.$$

We also can easily check that $\chi(0) = \chi(\ell_2) = \chi'(\ell_2) = 0$.

Due to the boundary conditions satisfied by χ the function $w_n(\boldsymbol{x}_2) := \chi(\boldsymbol{x}_2) e^{ik_n \boldsymbol{x}_2}$ satisfies

$$\begin{cases} -k_n^2 w_n - w_n'' &= \phi_n \quad \text{in } (0, \ell_2) \\ \widehat{w}_n(0) &= 0 \\ \widehat{w}_n'(\ell_2) - ik_n w_n(\ell_2) &= 0, \end{cases}$$

with $\phi_n(\boldsymbol{x}_2) := -\left(\chi(\boldsymbol{x}_2)'' + 2ik_n \chi'(\boldsymbol{x}_2)\right) e^{ik_n \boldsymbol{x}_2}$.

Recalling the discussion on Fourier expansion of quasi-periodic functions in Section 2.1.6, we see that introducing

$$u(\boldsymbol{x}_1,\boldsymbol{x}_2) := w_n(\boldsymbol{x}_2)e^{i(\alpha+\alpha_n)\boldsymbol{x}_1} \qquad f(\boldsymbol{x}_1,\boldsymbol{x}_2) := \phi_n(\boldsymbol{x}_2)e^{i(\alpha+\alpha_n)\boldsymbol{x}_1},$$

we have

$$\left\{\begin{array}{rcll} -k^2u-\Delta u &=& f & \text{in } \Omega,\\ \boldsymbol{\nabla} u\cdot\boldsymbol{n}-ik_nu &=& 0 & \text{on } \Gamma_{\mathrm{A}},\\ u &=& 0 & \text{on } \Gamma_{\mathrm{D}},\\ u_+-e^{i\alpha\ell_1}u_- &=& 0 & \text{on } \Gamma_\sharp. \end{array}\right.$$

The remaining of the proof consists in respectively bounding the norms of u and f from below and above. For the solution, we readily compute

$$\ell_1^{-1/2}\|u\|_\Omega = \|\chi\| = \frac{1}{\sqrt{105}}$$

For the right-hand side, we have

$$\ell_1^{-1/2}\|f\|_\Omega = \|\chi''+2ik_n\chi'\| \le \|\chi''\|+2|k_n|\|\chi'\| = \frac{2}{\ell_2^2}+2|k_n|\left(\sqrt{\frac{2}{15}}\frac{1}{\ell_2}\right) \le \frac{2}{\ell_2^2}(1+|k_n|\ell_2),$$

and therefore

$$1 \ge \frac{\ell_2^2}{2}\frac{1}{1+|k_n|\ell_2}\ell_1^{-1/2}\|f\|_\Omega$$

and it follows that

$$\sqrt{105}\|u\|_\Omega = 1 \ge \frac{\ell_2^2}{2}\frac{1}{1+|k_n|\ell_2}\|f\|_\Omega,$$

and

$$\|u\|_\Omega \ge \frac{1}{2\sqrt{105}}\frac{1}{1+|k_n|\ell_2}\ell_2^2\|f\|_\Omega.$$

Then, (2.2.26) follows by selecting $n\in\mathbb{Z}$ such that $|k_n| = k_\star$.

To establish (2.2.27), we observe that the first inequality is a direct consequence of the definition of $\mathscr{C}_{\mathrm{st}}$ in (2.1.22) and (2.2.26). On the other hand, the second inequality follows if we can show that

$$\frac{1}{1+k_\star\ell_2} \ge \frac{1}{2}\min\left(1,(k_\star\ell_2)^{-1}\right).$$

We establish this by distinguishing to cases. First, when $k_\star\ell_2\le 1$, we have

$$\frac{2}{1+k_\star\ell_2} \ge 1 \ge \min(1,(k_\star\ell_2)^{-1}).$$

Second, if $k_\star\ell_2\ge 1$, we have $k_\star\ell_2/2\ge 1/2$, so that

$$k_\star\ell_2 \ge \frac{1+k_\star\ell_2}{2},$$

and

$$\frac{2}{1+k_\star\ell_2} \geq \frac{1}{k_\star\ell_2} \geq \min(1,(k_\star\ell_2)^{-1}).$$

□

Corollary 2.2.11 (Frequency scalings with and without quasi-resonances)**.** *Considering the sequence of wavenumber $k^{(j)} := \alpha+\alpha_j$, $j \in \mathbb{N}$, there exists a sequence of right-hand side $(f^{(j)})_{j\in\mathbb{N}} \subset L^2(\Omega)$ such that*

$$k^{(j)}\|u^{(j)}\|_\Omega \geq \frac{1}{21}(k^{(j)}\ell_2)^2\frac{1}{k^{(j)}}\|f^{(j)}\|_\Omega. \tag{2.2.28}$$

Considering the sequence of wavenumber $k^{(j)} := \left(1+\sqrt{(\alpha+\alpha_j)^2\ell_2^2+1}\right)/(2\ell_2)$, $j \in \mathbb{N}$, there exists a sequence of right-hand side $(f^{(j)})_{j\in\mathbb{N}} \subset L^2(\Omega)$ such that

$$k^{(j)}\|u^{(j)}\|_\Omega \geq \frac{1}{42}(k^{(j)}\ell_2)^{3/2}\frac{1}{k^{(j)}}\|f^{(j)}\|_\Omega. \tag{2.2.29}$$

Proof. Recalling (2.2.26), the estimate in (2.2.28) simply follows from the fact that for the selected sequence, $k_n^{(j)} = 0$, and that $2\sqrt{105} \leq 21$. To establish (2.2.29), we observe that for the second sequence of wave numbers, $(k_n^{(j)})^2 = k^{(j)}/\ell_2$, so that

$$1+k_\star^{(j)}\ell_2 = (k^{(j)}\ell_2)^{1/2}+1 \leq (k^{(j)}\ell_2)^{1/2}+(k^{(j)}\ell_2)^{1/2} \leq 2(k^{(j)}\ell_2)^{1/2}.$$

□

2.2.4 Numerical illustrations

Here, we illustrate our stability results by highlighting how they impact the stability of finite element discretizations. To this end, we consider a first-order finite element discretization of (2.1.19) where the DtN operator is approximated by a perfectly matched layer (with length $\ell_{\mathrm{P}} := 1$ and damping factor $\gamma := 5$). We do not present in detail here the finite element method nor the perfectly matched layer treatment, as this topics will be addressed in detail in Chapters 3 and 5.

We fix the domain $\Omega := (0,1)^2$ (i.e. $\ell_1 = \ell_2 = 1$), and consider the coefficients $\varepsilon \equiv 1$, $\boldsymbol{A} \equiv \boldsymbol{I}$. We select the right-hand side $f \in L^2(\Omega)$ so that

$$u(\boldsymbol{x}) := \chi(\boldsymbol{x})e^{ik\boldsymbol{d}^{\mathrm{in}}\cdot\boldsymbol{x}} + e^{ik\boldsymbol{d}^{\mathrm{out}}\cdot\boldsymbol{x}}, \quad \forall \boldsymbol{x} \in \Omega$$

with $\boldsymbol{d}^{\mathrm{in}}\cdot\boldsymbol{d}^{\mathrm{in}} = \boldsymbol{d}^{\mathrm{out}}\cdot\boldsymbol{d}^{\mathrm{out}} = 1$, $\boldsymbol{d}_1^{\mathrm{in}} = \boldsymbol{d}_1^{\mathrm{out}} = \alpha + m\pi$ for some $m \in N$, $\boldsymbol{d}_2^{\mathrm{in}} \leq 0$ and $\boldsymbol{d}_2^{\mathrm{out}} = -\boldsymbol{d}_2^{\mathrm{in}}$, and the cutoff function

$$\chi(\boldsymbol{x}) := \begin{cases} 1 & \text{if} \quad 0 \leq \boldsymbol{x}_2 \leq \frac{1}{2}, \\ 16\left(\boldsymbol{x}_2-\frac{3}{4}\right)^2(8\boldsymbol{x}_2-3) & \text{if} \quad \frac{1}{2} \leq \boldsymbol{x}_2 \leq \frac{3}{4}, \\ 0 & \text{if} \quad \frac{3}{4} \leq \boldsymbol{x}_2 \leq 1. \end{cases} \tag{2.2.30}$$

The cutoff function actually enables us to consider a plane wave source as volumic data (noted at the end of subsection 2.1.3). Figures 2.7 and 2.8 depict the function f and the associated solution u for the frequency $k = 15\pi$ and the incidence angle $\theta = 20°$.

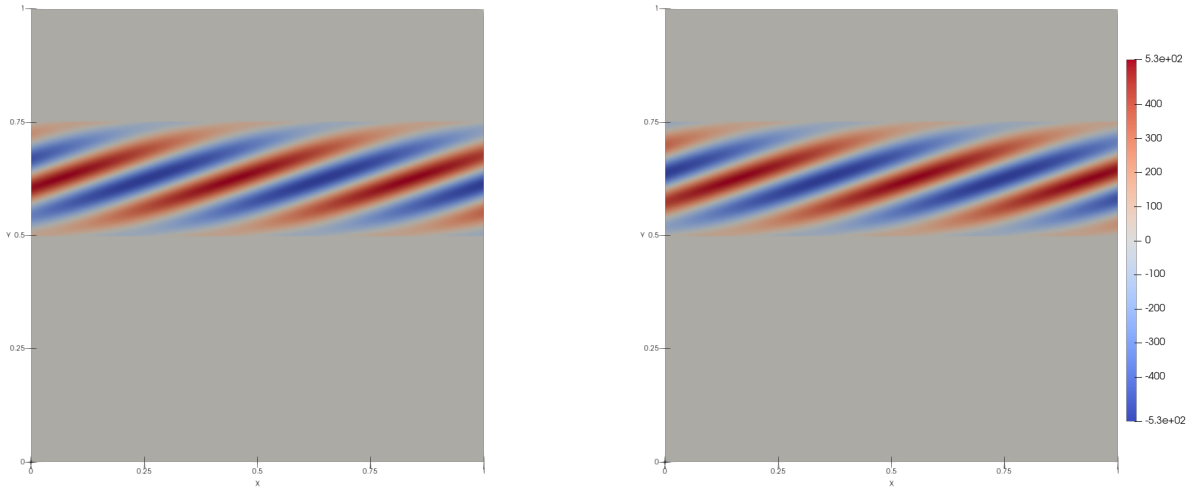


Figure 2.7: Right-hand side f for $k = 15\pi$, $\theta = 20°$ and $m = 0$. Real part (left) and imaginary part (right).

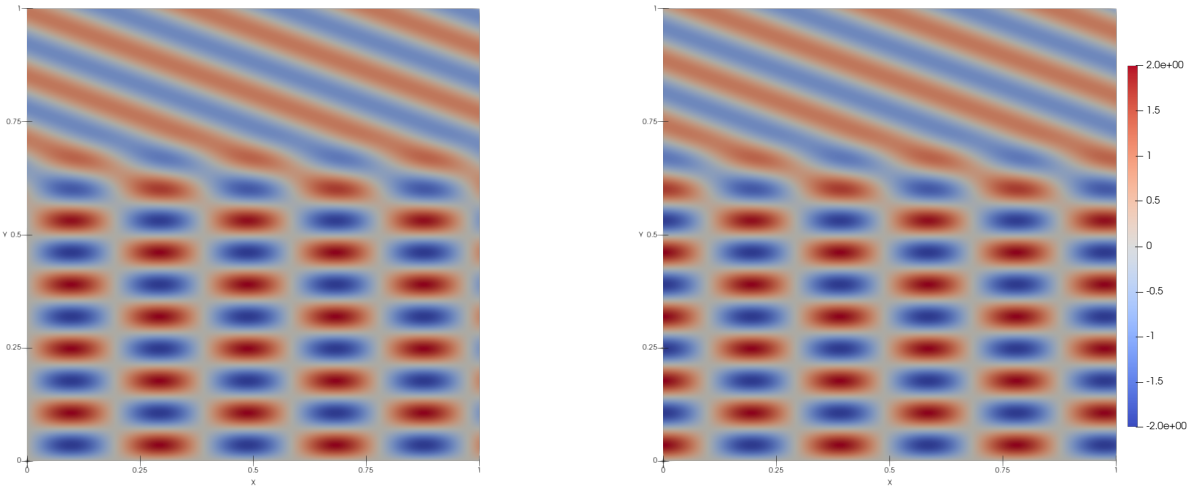


Figure 2.8: Solution u for $k = 15\pi$, $\theta = 20°$ and $m = 0$. Real part (left) and imaginary part (right).

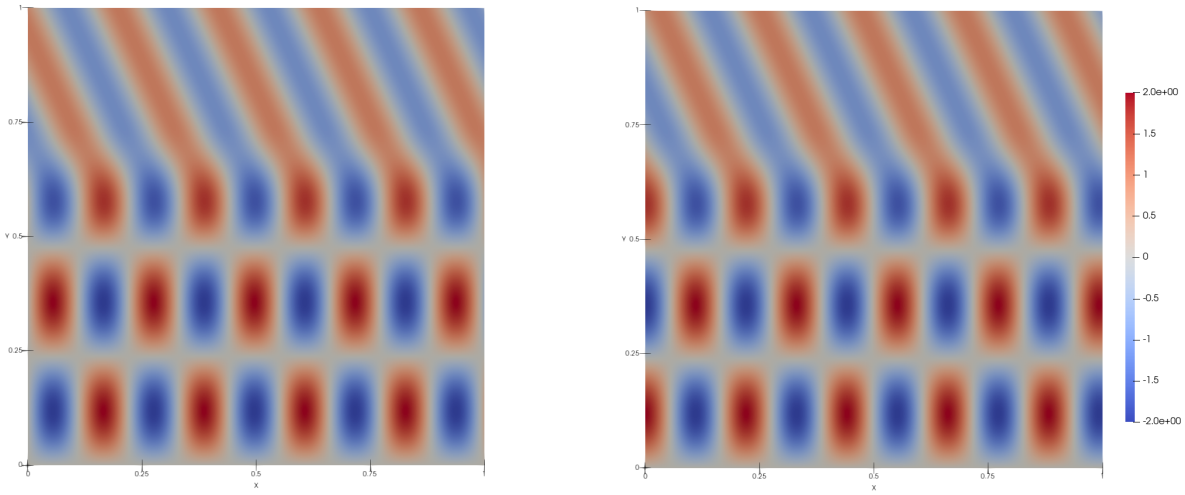


Figure 2.9: Solution u for $k = 10\pi$, $\theta = 45°$ and $m = 1$. Real part (left) and imaginary part (right).

Our goal here is to illustrate how the stability of the finite element method is affected by the angle of incidence θ. For this purpose, we fix a frequency k a mode $m \in \mathbb{N}$, and a mesh, and plot the relative interpolation error and finite element error versus the angle θ on Figures 2.10 and 2.11.

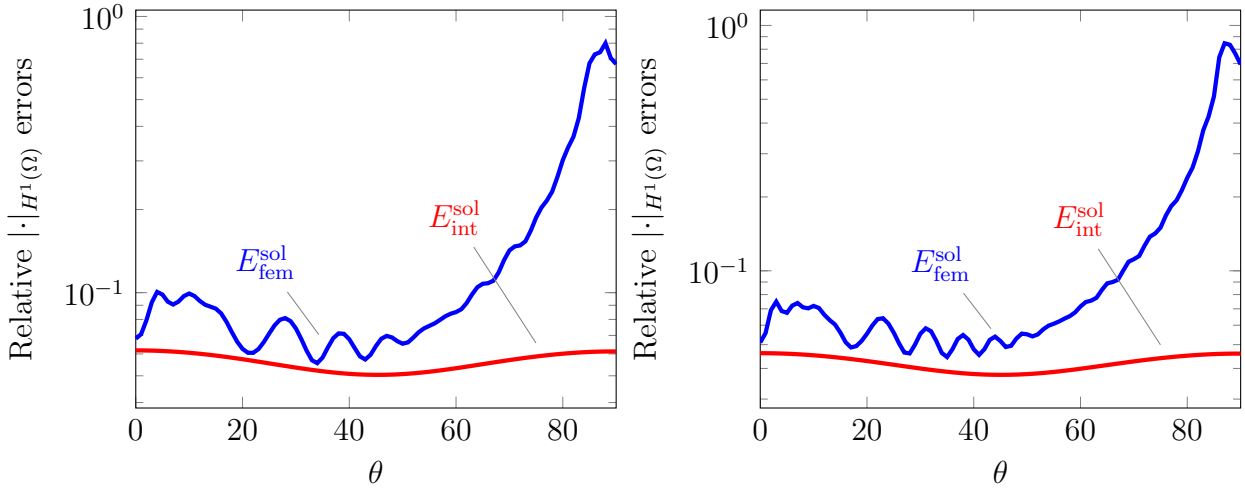


Figure 2.10: Interpolation and finite element errors for $m = 0$ with $k = 10\pi$ (left) and $k = 15\pi$ (right).

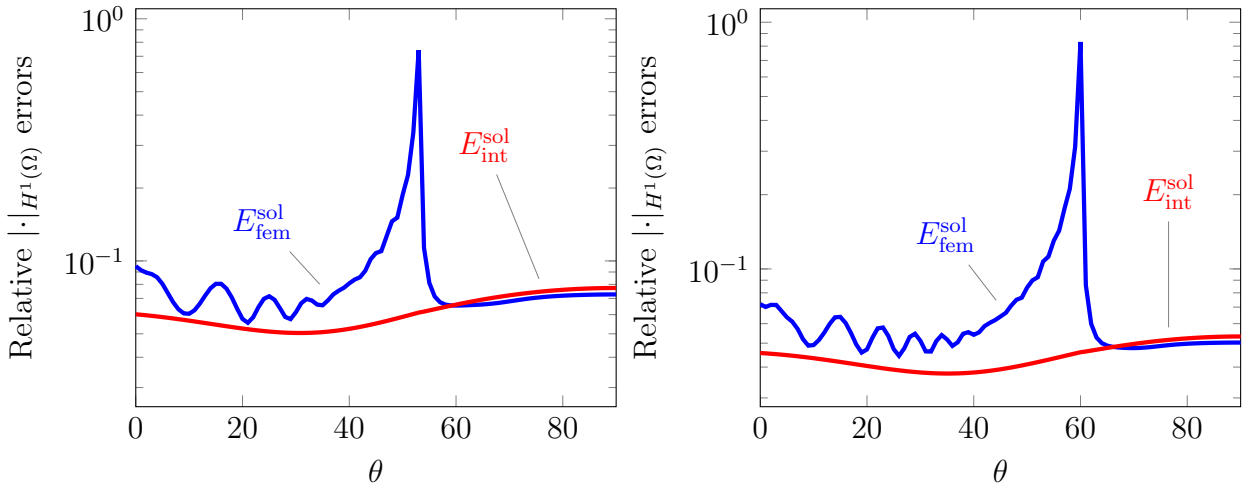


Figure 2.11: Interpolation and finite element errors for $m = 1$ with $k = 10\pi$ (left) and $k = 15\pi$ (right).

Figures 2.10 and 2.11 show that quasi-resonant modes increases the "pollution effect" that manifests as a gap between the finite element error and the interpolation error. More precisely, the phenomenon is particularly notable for $\theta = 90°$ and $k = 10\pi$ and $k = 15\pi$ when $m = 0$ and $\theta = 53.13°$, and $\theta = 60.07°$ for $k = 10\pi$ and $k = 15\pi$, respectively when $m = 1$. On the other hand, the link between these numerical results and the theoretical results of stability is quite clear: the increased gaps correspond to values of θ leading to $k_\star = 0$, worsening the stability constant from $\mathscr{C}_{\mathrm{st}} \approx (k\ell_2)^{3/2}$ to $\mathscr{C}_{\mathrm{st}} \approx (k\ell_2)^2$.

2.3 Frequency-explicit stability estimates: multi-layer case

In this section, we allow for more generality for the coefficients ε and $\boldsymbol{A}$. Specifically, we will consider multi-layered media covered by Assumption 2.1.1. In such a setting, the Fourier expansion techniques do not work, because the coefficients are allowed to depend on the $\boldsymbol{x}_1$ variable. As a result, we will rely on the alternative technique of "Morawetz multiplier" that consists in multiplying the PDE by a well chosen test function and performing integration by parts.

Morawetz multiplier are named after Cathleen S. Morawetz due to her seminal works linked with the stability of Helmholtz problems [108, 109, 111, 110]. They are also sometimes called "Rellich multipliers" due to his earlier work on eigenvalue problems [122]. This key idea was subsequently used in a plethora of works, some of which we discuss hereafter.

At its core, the original idea consists in multiplying the Helmholtz PDE by $\boldsymbol{x} \cdot \boldsymbol{\nabla}\overline{u}$, where u is the solution. Formally, simple integration by parts (see e.g. [15, Equations (4)

and (5)]) show that

$$2\,\mathrm{Re}(f, \boldsymbol{x}\cdot\boldsymbol{\nabla}u)_\Omega = 2k^2\|u\|_\Omega^2 - k^2\int_{\partial\Omega}|u|^2\boldsymbol{x}\cdot\boldsymbol{n} + \int_{\partial\Omega}|\boldsymbol{\nabla}u|^2\boldsymbol{x}\cdot\boldsymbol{n}.$$

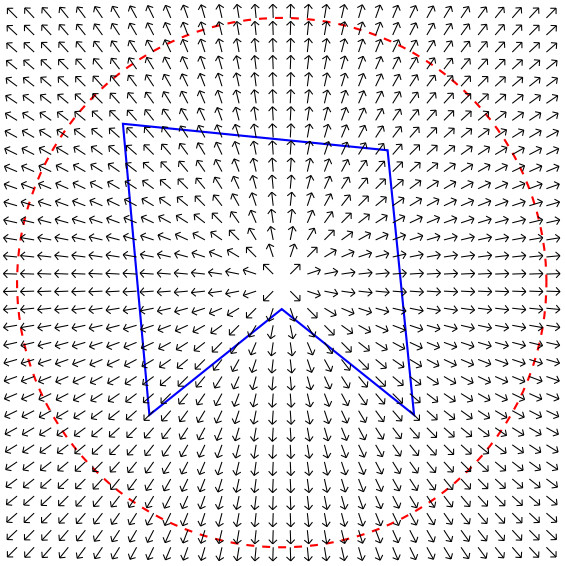
The boundary conditions together with geometrical requirements can then be used to ensure that the boundary terms have proper signs, leading to stability estimates [80, 127]. In particular, when consider a scattering problem by star-shaped Dirichlet obstacle, this technique provides the optimal non-trapping stability estimate

$$\mathscr{C}_{\mathrm{st}} \sim k\ell.$$

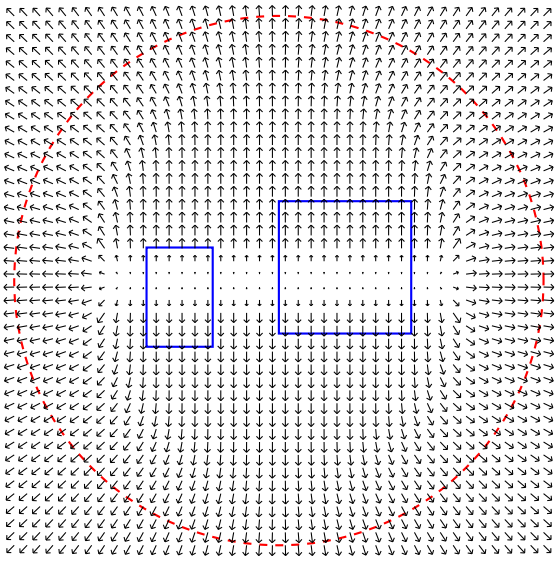
The approach can also be used for non-constant coefficients, leading to the same estimate under radial monotonicity conditions for ε and $\boldsymbol{A}$, see [15, 28, 106, 118].

Although being powerful, the original Morawetz multiplier is not suited to study geometries featuring parabolic trapping. Consider for simplicity a Dirichlet boundary, the field $\boldsymbol{x}$ multiplying $\boldsymbol{\nabla}u$ must enter the domain through the Dirichlet boundary and leave the domain through the DtN boundary. This is for instance the case in Figure 2.12a, where the obstacle is indeed star-shaped. However, for a geometry similar to the one we consider, we can see in Figure 2.12c that the field $\boldsymbol{x}$ would flow outside the domain through the Dirichlet boundary, leading to "wrong" signs in the boundary terms. It turns out that this difficulty may be circumvented by replacing the field $\boldsymbol{x}$ by $(0, \boldsymbol{x}_2)$, as can be seen on Figure 2.12d. We call the function $(0, \boldsymbol{x}_2)\cdot\boldsymbol{\nabla}u = \boldsymbol{x}_2\partial u/\partial\boldsymbol{x}_2$ a "directional" multiplier. To the best of our knowledge, this idea was first introduced in [27] to study scattering by rough surfaces. A blend of the directional multiplier and the original multiplier was then also used in [29] to study scattering by bounded obstacles with parabolically trapped rays, as described in Figure 2.12b. The same multiplier is also used in [35] to study scattering by finely layered obstacles.

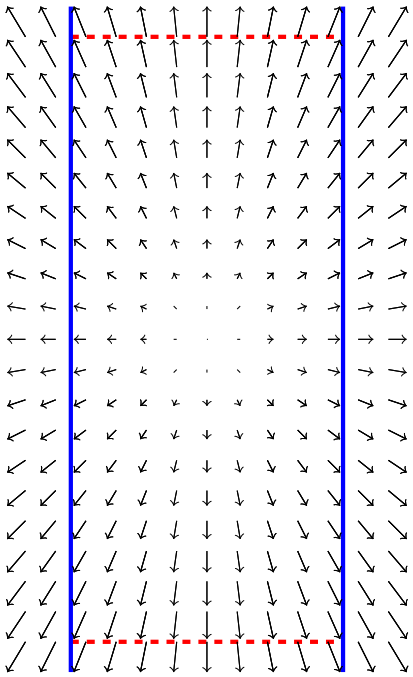
In this section, we will see that the quasi-periodic conditions act similarly as Dirichlet boundary conditions would, as far as Morawetz multipliers are concerned. As a result, as sketched on Figure 2.13, we cannot use the standard Morawetz multiplier for scattering by periodic structures, but the directional multiplier has a favorable behavior along the quasi-periodic boundary conditions.



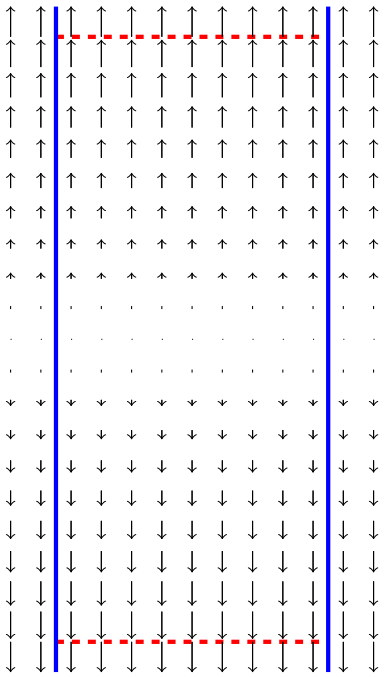
(a) Standard Morawetz multiplier $\boldsymbol{x}$



(b) Directional multiplier $(\chi \boldsymbol{x}_1, \boldsymbol{x}_2)$

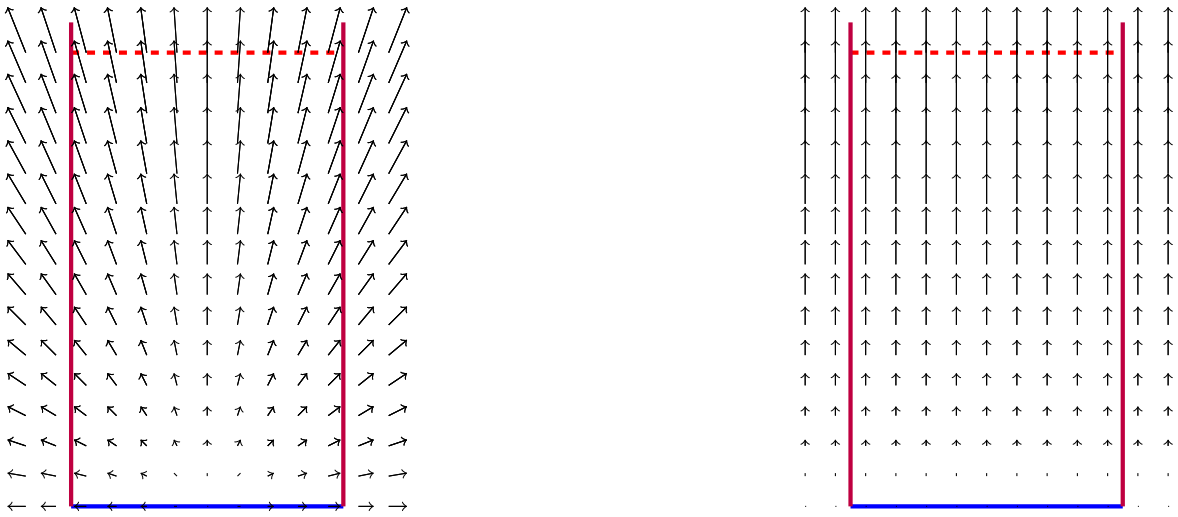


(c) Standard Morawetz multiplier $\boldsymbol{x}$



(d) Directional multiplier $(0, \boldsymbol{x}_2)$

Figure 2.12: Morawetz multiplier fields



(a) The vector field $\boldsymbol{x}$. (b) The vector field $(0, \boldsymbol{x}_2)$.

Figure 2.13: Multipliers in periodic structures

2.3.1 A Morawetz identity for quasi-periodic boundary conditions

Morawetz-type (also known as Rellich-type) identities are obtained after multiplying Helmholtz equation by a Morawetz multiplier and applying the divergence theorem. The purpose of this subsection is to derive a Morawetz identity, which will be used afterwards to bound the energy norm of the solution. Under the conditions of Assumption 2.1.1, we multiply the first equation in (2.1.18) by a directional Morawetz multiplier in order to obtain the identity (2.3.1), which will help get an explicit stability estimate for this case.

Recall from Assumption 2.1.1 that for each $1 \leq j \leq N$, Ω_j corresponds to one of the layers where the coefficients ε and $\boldsymbol{A}$ are smooth. The following Sobolev space of functions which are H^2-regular in each layer will be useful in this subsection

$$H^2(\mathscr{P}) := \left\{ v \in L^2(\Omega) \mid v|_{\Omega_j} \in H^2(\Omega_j);\ 1 \leq j \leq N \right\}.$$

Lemma 2.3.1 (Morawetz identity)**.** *For all $u \in H^1_\sharp(\Omega) \cap H^2(\mathscr{P})$ solution to (2.1.18), we*

have

$$2\int_\Omega A_2\left|\frac{\partial u}{\partial \boldsymbol{x}_2}\right|^2+k^2\sum_{j=1}^{N-1}\int_{\Gamma_j}[\![\varepsilon]\!]|u|^2\boldsymbol{x}_2\boldsymbol{n}_2^j-\sum_{k=1}^{N-1}\int_{\Gamma_j}\left\{[\![A_1]\!]\left|\frac{\partial u}{\partial \boldsymbol{x}_1}\right|^2+[\![A_2]\!]\left|\frac{\partial u}{\partial \boldsymbol{x}_2}\right|^2\right\}\boldsymbol{x}_2\boldsymbol{n}_2^j$$
$$+k^2\int_\Omega\frac{\partial\varepsilon}{\partial \boldsymbol{x}_2}\boldsymbol{x}_2|u|^2-\int_\Omega\left\{\frac{\partial A_1}{\partial \boldsymbol{x}_2}\boldsymbol{x}_2\left|\frac{\partial u}{\partial \boldsymbol{x}_1}\right|^2+\frac{\partial A_1}{\partial \boldsymbol{x}_2}\boldsymbol{x}_2\left|\frac{\partial u}{\partial \boldsymbol{x}_1}\right|^2\right\}-\mathrm{Re}\left\{\int_{\Gamma_A}\mathscr{R}u\,\bar{u}\right\}$$
$$=2\,\mathrm{Re}\int_\Omega\varepsilon f\left(\boldsymbol{x}_2\frac{\partial\bar{u}}{\partial \boldsymbol{x}_2}+\frac{1}{2}\bar{u}\right)+\ell_2\int_{\Gamma_\mathrm{A}}\left\{k^2\varepsilon|u|^2+\left|\frac{\partial u}{\partial \boldsymbol{x}_2}\right|^2-\left|\frac{\partial u}{\partial \boldsymbol{x}_1}\right|^2\right\}. \quad (2.3.1)$$

In particular, we have

$$2\int_\Omega A_2\left|\frac{\partial u}{\partial \boldsymbol{x}_2}\right|^2\le 2\,\mathrm{Re}\int_\Omega\varepsilon f\left(\boldsymbol{x}_2\frac{\partial\bar{u}}{\partial \boldsymbol{x}_2}+\frac{1}{2}\bar{u}\right)+\ell_2\int_{\Gamma_\mathrm{A}}\left\{k^2\varepsilon|u|^2+\left|\frac{\partial u}{\partial \boldsymbol{x}_2}\right|^2-\left|\frac{\partial u}{\partial \boldsymbol{x}_1}\right|^2\right\}. \quad (2.3.2)$$

Proof. We multiply the first equation in (2.1.18) by the (conjugate of the) directional Morawetz multiplier

$$\boldsymbol{x}_2\frac{\partial u}{\partial \boldsymbol{x}_2}+\frac{1}{2}u,$$

we obtain

$$2\,\mathrm{Re}\int_\Omega f\left(\boldsymbol{x}_2\frac{\partial\bar{u}}{\partial \boldsymbol{x}_2}+\frac{1}{2}\bar{u}\right)=2\,\mathrm{Re}\int_\Omega\left(-k^2\varepsilon u-\boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla}u)\right)\left(\boldsymbol{x}_2\frac{\partial\bar{u}}{\partial \boldsymbol{x}_2}+\frac{1}{2}\bar{u}\right). \quad (2.3.3)$$

We will develop the terms in the right hand side of (2.3.3). Let us first recall that the identity

$$2\,\mathrm{Re}\,\phi\frac{\partial\bar{\phi}}{\partial \boldsymbol{x}_j}=\frac{\partial}{\partial \boldsymbol{x}_j}|\phi|^2 \quad (2.3.4)$$

holds for any sufficiently smooth function ϕ and $j=1$ or 2. In particular, we have

$$2\,\mathrm{Re}\,\varepsilon u\frac{\partial\bar{u}}{\partial \boldsymbol{x}_2}=\varepsilon\frac{\partial}{\partial \boldsymbol{x}_j}|u|^2$$

Then, recalling from Assumption 2.1.1 that $\boldsymbol{n}^j=(\boldsymbol{n}_1^j,\boldsymbol{n}_2^j)$ the unit upward normal to

Γ_j, we have

$$\begin{aligned}
2\,\mathrm{Re}\int_\Omega -k^2\varepsilon u\boldsymbol{x}_2\frac{\partial\bar{u}}{\partial\boldsymbol{x}_2} &= -k^2\sum_{j=1}^{N} 2\,\mathrm{Re}\int_{\Omega_j}\varepsilon\boldsymbol{x}_2 2\,\mathrm{Re}\left\{u\frac{\partial\bar{u}}{\partial\boldsymbol{x}_2}\right\}\\
&= -k^2\sum_{j=1}^{N}\int_{\Omega_j}\varepsilon\boldsymbol{x}_2\frac{\partial}{\partial\boldsymbol{x}_2}|u|^2\\
&= -k^2\sum_{j=1}^{N}\left\{-\int_{\Omega_j}\frac{\partial}{\partial\boldsymbol{x}_2}(\varepsilon\boldsymbol{x}_2)|u|^2+\int_{\Gamma_j}\varepsilon\boldsymbol{x}_2|u|^2\boldsymbol{n}_2^j-\int_{\Gamma_{j-1}}\varepsilon\boldsymbol{x}_2|u|^2\boldsymbol{n}_2^{j-1}\right\}\\
&= k^2\int_\Omega\varepsilon|u|^2+k^2\int_\Omega\frac{\partial\varepsilon}{\partial\boldsymbol{x}_2}\boldsymbol{x}_2|u|^2+k^2\sum_{j=1}^{N-1}\int_{\Gamma_j}[\![\varepsilon]\!]_j\boldsymbol{x}_2|u|^2\boldsymbol{n}_2^j-k^2\ell_2\varepsilon_N\int_{\Gamma_A}|u|^2.
\end{aligned}$$

Moreover, we have

$$2\,\mathrm{Re}\int_\Omega -k^2\varepsilon u\left(\frac{1}{2}\bar{u}\right)=-k^2\int_\Omega\varepsilon|u|^2.$$

On the other hand, using the formula

$$\int_\Omega -\boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla}u)\bar{v}=\int_\Omega\boldsymbol{A}\boldsymbol{\nabla}u\cdot\boldsymbol{\nabla}\bar{v}-\int_{\Gamma_A}\mathscr{R}u\;\bar{v}\qquad\forall v\in H^1_\sharp(\Omega),$$

we have

$$\begin{aligned}
2\,\mathrm{Re}\int_\Omega -\boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla}u)\left(\frac{1}{2}\bar{u}\right) &= \int_\Omega|\boldsymbol{\nabla}u|^2-\mathrm{Re}\left\{\int_{\Gamma_A}\mathscr{R}u\;\bar{u}\right\}\\
&= \int_\Omega A_1\left|\frac{\partial u}{\partial\boldsymbol{x}_1}\right|^2+\int_\Omega A_2\left|\frac{\partial u}{\partial\boldsymbol{x}_2}\right|^2-\mathrm{Re}\left\{\int_{\Gamma_A}\mathscr{R}u\;\bar{u}\right\}
\end{aligned}$$

and

$$\begin{aligned}
2\,\mathrm{Re}\int_\Omega -\boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla}u)\left(\boldsymbol{x}_2\frac{\partial\bar{u}}{\partial\boldsymbol{x}_2}\right) &= 2\,\mathrm{Re}\left\{\int_\Omega\boldsymbol{A}\boldsymbol{\nabla}u\cdot\boldsymbol{\nabla}\left(\boldsymbol{x}_2\frac{\partial\bar{u}}{\partial\boldsymbol{x}_2}\right)-\int_{\Gamma_A}\mathscr{R}u\left(\boldsymbol{x}_2\frac{\partial\bar{u}}{\partial\boldsymbol{x}_2}\right)\right\}\\
&= 2\,\mathrm{Re}\int_\Omega A_1\boldsymbol{x}_2\frac{\partial u}{\partial\boldsymbol{x}_1}\frac{\partial^2\bar{u}}{\partial\boldsymbol{x}_2\partial\boldsymbol{x}_1}+2\int_\Omega A_2\left|\frac{\partial u}{\partial\boldsymbol{x}_2}\right|^2\\
&\quad+2\,\mathrm{Re}\int_\Omega A_2\boldsymbol{x}_2\frac{\partial u}{\partial\boldsymbol{x}_2}\frac{\partial}{\partial\boldsymbol{x}_2}\frac{\partial\bar{u}}{\partial\boldsymbol{x}_2}-2\ell_2\int_{\Gamma_A}\left|\frac{\partial u}{\partial\boldsymbol{x}_2}\right|^2 \qquad (2.3.5)
\end{aligned}$$

developing the first expression of (2.3.5) and using again (2.3.4), we obtain

$$
\begin{aligned}
2\,\mathrm{Re}\int_\Omega A_1\boldsymbol{x}_2\frac{\partial u}{\partial \boldsymbol{x}_1}\frac{\partial^2\bar u}{\partial \boldsymbol{x}_2\partial \boldsymbol{x}_1} &= \sum_{j=1}^N\int_{\Omega_j} A_1\boldsymbol{x}_2\frac{\partial}{\partial \boldsymbol{x}_2}\left|\frac{\partial u}{\partial \boldsymbol{x}_1}\right|^2\\
&= \sum_{j=1}^N\left\{-\int_{\Gamma_{j-1}} A_1\boldsymbol{x}_2\left|\frac{\partial u}{\partial \boldsymbol{x}_1}\right|^2\boldsymbol{n}_2^{j-1}+\int_{\Gamma_j} A_1\boldsymbol{x}_2\left|\frac{\partial u}{\partial \boldsymbol{x}_1}\right|^2\boldsymbol{n}_2^{j}-\int_{\Omega_j}\frac{\partial}{\partial \boldsymbol{x}_2}(A_1\boldsymbol{x}_2)\left|\frac{\partial u}{\partial \boldsymbol{x}_1}\right|^2\right\}\\
&= -\int_\Omega A_1\left|\frac{\partial u}{\partial \boldsymbol{x}_1}\right|^2-\int_\Omega\frac{\partial A_1}{\partial \boldsymbol{x}_2}\boldsymbol{x}_2\left|\frac{\partial u}{\partial \boldsymbol{x}_1}\right|^2-\sum_{j=1}^{N-1}\int_{\Gamma_j}[\![A_1]\!]_j\left|\frac{\partial u}{\partial \boldsymbol{x}_1}\right|^2\boldsymbol{x}_2\boldsymbol{n}_2^j+\ell_2\int_{\Gamma_A}\left|\frac{\partial u}{\partial \boldsymbol{x}_1}\right|^2
\end{aligned}
$$

developing the second expression

$$
\begin{aligned}
2\,\mathrm{Re}\int_\Omega A_2\boldsymbol{x}_2\frac{\partial u}{\partial \boldsymbol{x}_2}\frac{\partial}{\partial \boldsymbol{x}_2}\frac{\partial\bar u}{\partial \boldsymbol{x}_2} &= \sum_{j=1}^N\int_{\Omega_j} A_2\boldsymbol{x}_2\frac{\partial}{\partial \boldsymbol{x}_2}\left|\frac{\partial u}{\partial \boldsymbol{x}_2}\right|^2\\
&= \sum_{j=1}^N\left\{-\int_{\Gamma_{j-1}} A_2\boldsymbol{x}_2\left|\frac{\partial u}{\partial \boldsymbol{x}_2}\right|^2\boldsymbol{n}_2^{j}+\int_{\Gamma_j} A_2\boldsymbol{x}_2\left|\frac{\partial u}{\partial \boldsymbol{x}_2}\right|^2\boldsymbol{n}_2^{j}-\int_{\Omega_j}\frac{\partial}{\partial \boldsymbol{x}_2}(A_2\boldsymbol{x}_2)\left|\frac{\partial \bar u}{\partial \boldsymbol{x}_2}\right|^2\right\}\\
&= -\int_\Omega A_2\left|\frac{\partial u}{\partial \boldsymbol{x}_2}\right|^2-\int_\Omega\frac{\partial A_2}{\partial \boldsymbol{x}_2}\boldsymbol{x}_2\left|\frac{\partial u}{\partial \boldsymbol{x}_2}\right|^2-\int_{\Gamma_j}\sum_{j=1}^{N-1}[\![A_2]\!]_j\left|\frac{\partial u}{\partial \boldsymbol{x}_2}\right|^2\boldsymbol{x}_2\boldsymbol{n}_2^j+\ell_2\int_{\Gamma_A}\left|\frac{\partial u}{\partial \boldsymbol{x}_2}\right|^2.
\end{aligned}
$$

and, we obtain (2.3.1) by combining the above results. Furthermore, we establish estimate (2.3.2) since Assumption (2.1.1) and the DtN properties (2.1.16) assert that all therms on the left side of (2.3.1) are positive. □

2.3.2 Frequency-explicit stability estimates

The Morawetz identity of Lemma 2.3.1 applies under the assumption that the solution is piecewise smooth. We start by showing that this is indeed the case, due to the regularity of the interfaces Γ_j.

Lemma 2.3.2 (A priori smoothness)**.** *Assume that ε and $\boldsymbol{A}$ satisfies Assumption 2.1.1. Then, if $f\in L^2(\Omega)$ and $u\in H^1_\sharp(\Omega)$ solves (2.1.19), we have $u\in H^2(\mathscr{P})$.*

Proof. The proof is standard [20, Section 9.6], and we shall only sketch it. First, we extend u by quasi-periodicity in the $\boldsymbol{x}_1$ direction to a function $\widetilde{u}$ defined over $\mathbb{R}\times(0,\ell_2)$. This function now satisfies

$$-\varepsilon\widetilde{u}-\boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla}\widetilde{u})=\widetilde{f}$$

where $\widetilde{f}$ is the quasi-periodic extension of f and ε and $\boldsymbol{A}$ are the periodic extension of ε and $\boldsymbol{A}$. We then consider a family of smooth functions $\{\chi_j\}_j$ such that

$$\sum_j\chi_j=1\text{ in }\Omega$$

and the support of each χ_j only contains one interface Γ_j. We then look at the function $u_j := \chi_j \widetilde{u}$ which solves

$$-\varepsilon\widetilde{u} - \boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla}\widetilde{u}) = \widetilde{f} - \boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla}\chi_j)\,\widetilde{u} - 2\boldsymbol{\nabla}\chi_j\cdot\boldsymbol{A}\boldsymbol{\nabla}\widetilde{u}$$

in $\mathbb{R}^2$ but with only one interface or boundary condition. We can then find a coordinate change that flatten this interface or boundary, and apply the method of tangential differential quotient to show that $u_j \in H^2(\mathbb{R}^2 \setminus \Gamma_j)$, which in turns proves that $u \in H^2(\mathscr{P})$. □

Based on Lemma 2.3.2, the Morawetz identity in (2.3.1) and the DtN map properties in (2.1.16) and (2.1.17), we are now able to derive a frequency-explicit stability estimate for the textured multilayer case, corresponding to coefficients satisfying Assumption 2.1.1.

Theorem 2.3.3 (Stability in layered media)**.** *Assume that ε and $\boldsymbol{A}$ satisfy Assumption 2.1.1. Then, (2.1.19) is well-posed, and we have*

$$\mathscr{C}_{\mathrm{st}} \leq 4\,(1+k\ell_2)\,(k\ell_2)^2. \tag{2.3.6}$$

Proof. We fix $f \in L^2(\Omega)$ and that $u \in H^1_\sharp(\Omega)$ is a solution to (2.1.19). Then, by Lemma 2.3.2, we know that $u \in H^2(\mathscr{P})$, so that the Morawetz identity of Lemma 2.3.1 may be applied.

We start by using the Fourier expansion presented in (2.1.8) to write that

$$\int_{\Gamma_{\mathrm{A}}} k^2|u|^2 = k^2 \int_{\Gamma_{\mathrm{A}}} \sum_{n\in\mathbb{Z}} |\widehat{u}_n|^2 = k^2\ell_1 \sum_{n\in\mathbb{Z}} |\widehat{u}_n(\ell_2)|^2,$$

$$\int_{\Gamma_{\mathrm{A}}} \left|\frac{\partial u}{\partial \boldsymbol{x}_1}\right|^2 = \int_{\Gamma_{\mathrm{A}}} \sum_{n\in\mathbb{Z}} (\alpha+\alpha_n)^2|\widehat{u}_n|^2 = \ell_1 \sum_{n\in\mathbb{Z}} (\alpha+\alpha_n)^2|\widehat{u}_n(\ell_2)|^2,$$

and

$$\int_{\Gamma_{\mathrm{A}}} \left|\frac{\partial u}{\partial \boldsymbol{x}_2}\right|^2 = \int_{\Gamma_{\mathrm{A}}} \sum_{n\in\mathbb{Z}} |\widehat{u}_n'|^2 = \ell_1 \sum_{n\in\mathbb{Z}} |\widehat{u}_n'(\ell_2)|^2 = \ell_1 \sum_{n\in\mathbb{Z}} |k_n|^2|\widehat{u}_n(\ell_2)|^2.$$

Recalling that $\varepsilon = 1$ and $\boldsymbol{A} = \boldsymbol{I}$ in a neighborhood of Γ_{A}, we have

$$\begin{aligned}\ell_2 \int_{\Gamma_{\mathrm{A}}} \left\{ k^2|u|^2 + \left|\frac{\partial u}{\partial \boldsymbol{x}_2}\right|^2 - \left|\frac{\partial u}{\partial \boldsymbol{x}_1}\right|^2 \right\} &= \ell_2\ell_1 \sum_{n\in\mathbb{Z}} \left\{ \left(k^2\varepsilon - (\alpha+\alpha_n)^2\right) + |k_n|^2 \right\} |\widehat{u}_n(\ell_2)|^2 \\ &= \ell_2\ell_1 \sum_{n\in\mathbb{Z}} \left\{ k_n^2 + |k_n|^2 \right\} |\widehat{u}_n(\ell_2)|^2,\end{aligned}$$

and therefore

$$\ell_2 \int_{\Gamma_{\mathrm{A}}} \left\{ k^2\varepsilon|u|^2 + \left|\frac{\partial u}{\partial \boldsymbol{x}_2}\right|^2 - \left|\frac{\partial u}{\partial \boldsymbol{x}_1}\right|^2 \right\} = 2\ell_2\ell_1 \sum_{|n|\leq n_{\mathrm{c}}} |k_n|^2|\widehat{u}_n(\ell_2)|^2. \tag{2.3.7}$$

Next, we pick $v = u$ as a test function in (2.1.19) and taking the imaginary part yields

$$\ell_1 \sum_{|n|\le n_c} k_n \left|\widehat{u}_n(\ell_2)\right|^2 = \mathrm{Im}\langle \mathscr{R} u, u\rangle_{\Gamma_A} = -\,\mathrm{Im}(\varepsilon f, u)_\Omega. \tag{2.3.8}$$

Therefore, since $0 \le k_n \le k$ for $n \le n_c$, we have

$$\begin{aligned} 2\ell_2\ell_1 \sum_{|n|\le n_c} |k_n|^2 |\widehat{u}_n(\ell_2)|^2 &\le 2k\ell_2\ell_1 \sum_{|n|\le n_c} k_n |\widehat{u}_n(\ell_2)|^2 \\ &\le 2k\ell_2 \|f\|_{\varepsilon,\Omega}\|u\|_{\varepsilon,\Omega}. \end{aligned}$$

Now, using (2.3.2) , we have

$$\begin{aligned} \left\| \frac{\partial u}{\partial \boldsymbol{x}_2} \right\|^2_{A_2,\Omega} &\le \frac{1}{2}\left\{ 2\ell_2 \|f\|_{\varepsilon,\Omega} \left\| \frac{\partial u}{\partial \boldsymbol{x}_2} \right\|_{\varepsilon,\Omega} + \|f\|_{\varepsilon,\Omega}\|u\|_{\varepsilon,\Omega} + 2k\ell_2 \|f\|_{\varepsilon,\Omega}\|u\|_{\varepsilon,\Omega} \right\} \\ &\le \left\{ \ell_2 \sqrt{\frac{\varepsilon_{\max}}{A_{\min}}} \|f\|_{\varepsilon,\Omega} \left\| \frac{\partial u}{\partial \boldsymbol{x}_2} \right\|_{A_2,\Omega} + \left(\frac{1}{2} + k\ell_2 \right) \|f\|_{\varepsilon,\Omega}\|u\|_{\varepsilon,\Omega} \right\}. \end{aligned}$$

On the other hand, since $u \in H^1(\Omega)$ and $u = 0$ on Γ_D, the Poincaré inequality (2.2.10) give

$$\begin{aligned} \|u\|_{\varepsilon,\Omega} \le \sqrt{\varepsilon_{\max}}\|u\|_\Omega &\le 2\sqrt{\varepsilon_{\max}}\ell_2 \left\| \frac{\partial u}{\partial \boldsymbol{x}_2} \right\|_{0,\Omega} \\ &\le 2\sqrt{\frac{\varepsilon_{\max}}{A_{\min}}}\ell_2 \left\| \frac{\partial u}{\partial \boldsymbol{x}_2} \right\|_{A_2,\Omega} = 2\ell_2 \left\| \frac{\partial u}{\partial \boldsymbol{x}_2} \right\|_{A_2,\Omega}, \end{aligned} \tag{2.3.9}$$

where we used the fact that $\varepsilon_{\max} = 1$ and $A_{\min} = 1$. As a result

$$\left\| \frac{\partial u}{\partial \boldsymbol{x}_2} \right\|^2_{A_2,\Omega} \le 2\left(1 + k\ell_2\right)\ell_2 \|f\|_{\varepsilon,\Omega} \left\| \frac{\partial u}{\partial \boldsymbol{x}_2} \right\|_{A_2,\Omega},$$

and

$$\left\| \frac{\partial u}{\partial \boldsymbol{x}_2} \right\|_{A_2,\Omega} \le 2\left(1 + k\ell_2\right)\ell_2 \|f\|_{\varepsilon,\Omega}.$$

Finally, re-using (2.3.9)

$$\|u\|_{\varepsilon,\Omega} \le 4\left(1 + k\ell_2\right)\ell_2^2 \|f\|_{\varepsilon,\Omega},$$

so that

$$k\|u\|_{\varepsilon,\Omega} \le 4\left(1 + k\ell_2\right)(k\ell_2)\ell_2 \|f\|_{\varepsilon,\Omega},$$

which implies (2.3.6) under the assumption that u does exist. Now, (2.3.6) implies that if $f = 0$, we must have $u = 0$, so that the solution, if it exists, must be unique. As a result, Fredholm alternative ensures the existence and uniqueness of the solution, completing the proof. □

The one-layer case where $\varepsilon \equiv 1$ and $\boldsymbol{A} = \boldsymbol{I}$ is covered by Assumption 2.1.1 as a particular case. Then, our bounds simplifies as

$$\mathscr{C}_{\mathrm{st}} \leq 4(1 + k\ell_2)(k\ell_2)^2 \approx (k\ell_2)^3,$$

for large frequencies. It is interesting to note that the estimate obtained in the previous subsection, namely

$$\mathscr{C}_{\mathrm{st}} \leq 12 \min\left(1, k_\star \ell_2\right)^{-1}) \, (k\ell_2)^2 \leq 12(k\ell_2)^2,$$

is much sharper (and is even improved away from quasi-resonances). This is the reason why Fourier expansion techniques are interesting, even though they are restricted to specific situations. The stability bound derived in this subsection, on the other hand, has the advantage to apply to a wider range of situations. In particular, the $\boldsymbol{x}_1$ variations of the coefficients are not restricted at all by Assumption 2.1.1. We will see in Chapter 5 that it allows to perform periodic homogenization for layers with highly oscillating interfaces.

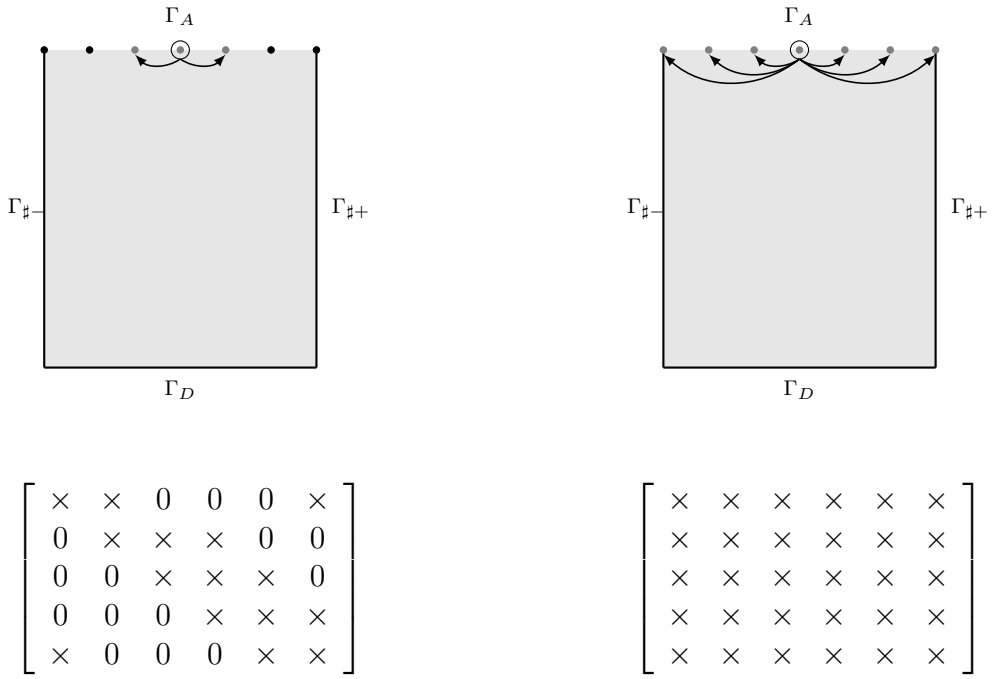
Chapter 3

Error and stability analysis of perfectly matched layers

Contents

Many simulation tools, including finite difference, finite element and Trefftz methods, operate on a bounded computational domain. Wave propagation problems are, on the other hand, often set in unbounded media, so that a "domain truncation" strategy must be applied before discretizing the problem with the aforementioned techniques. The DtN operator we have introduced in the previous chapter is not really suited for this purpose, as it is non-local and leads to a dense block in the discretization matrix (see Figure 3.1).

Several strategies have therefore been developed to approximate the DtN operator by a local, or at least, computationally efficient, boundary condition. We may cite, among many, infinite element methods [5], low- and high-order absorbing boundary conditions [66, 45], FEM-BEM coupling [81] and perfectly matched layers [17, 18, 31]. Here, we will focus on the latter.



$$\begin{bmatrix} \times & \times & 0 & 0 & 0 & \times \\ 0 & \times & \times & \times & 0 & 0 \\ 0 & 0 & \times & \times & \times & 0 \\ 0 & 0 & 0 & \times & \times & \times \\ \times & 0 & 0 & 0 & \times & \times \end{bmatrix} \qquad \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \end{bmatrix}$$

(a) local boundary condition (b) non-local boundary condition

Figure 3.1: Communication and matrix patterns associated with boundary conditions

Perfectly matched layers (PML) were introduced in the seminal works of Berenger [17, 18]. Strictly speaking, a PML is not a boundary condition. Instead, the computational domain is extended by adding an absorbing layer that will damp outgoing waves (see Figure 3.2a). Mathematically, these absorbing layers are represented by material coefficients that are anisotropic and complex-valued. This coefficients are carefully chosen so that the transmission condition from the physical medium to the absorbing layer does not produce unwanted reflections. This may also be interpreted as a complex change of coordinates for the solution in the absorbing layer [52, Chapter 4.5]. In practice, although the

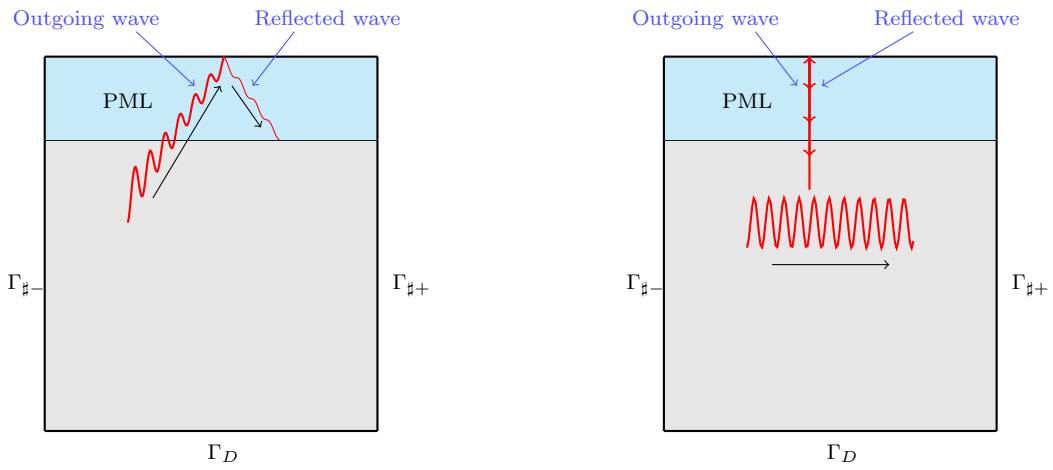
computational domain is slightly enlarged, the computational cost remains relatively minimal compared a non-local boundary condition. In addition, due to the limited accuracy of lowest order ABCs and the difficulty of implementing high order ABCs (containing high-order derivatives), PML techniques have become very popular for the numerical simulation of waves in unbounded domains [43, 42, 85, 86, 133].

When considering scattering by bounded obstacle (i.e. without quasi-periodic boundary conditions), it is shown in [45, 82, 96, 61] that the solution of the PML problem converges exponentially fast to the exact solution when the thickness of the PML and/or the damping coefficient increase. However, we will see in this chapter that the situation is different for scattering by periodic structures. An intuitive way to understand why PMLs behave differently in periodic structures is illustrated in Figure 3.2b. Indeed, quasi-resonant modes are allowed to travel orthogonaly to the layer, so that they are only weakly damped. We will provide quantitative estimates illustrating this phenomenon.

In this chapter, we employ a PML to approximate the DtN operator. Our objective is to analyze the properties of the resulting "PML problem". In particular, we focus on two key sets of results. First (i), we show that the solution to the PML problem converges toward the solution of the original problem if the PML parameters are suitably chosen. Second (ii), we analyze the inf-sup stability of the PML problem. Notice that this is a subtle issue, and in particular, the convergence mentioned at (i) does not imply (ii). Indeed, convergence analysis only applies to (physically meaningful) right-hand sides contained in the original domain, whereas inf-sup stability also requires the anlaysis of right-hand sides contained in the absorbing layer. Although such right-hand sides are non-physical, there are paramount in the stability and convergence study of numerical methods. This analysis is actually complicated, and we will consider two separate cases. On the one hand (iia), we will treat the case of a homogeneous medium with quasi-periodic boundary conditions, for which we provide optimal stability results. On the other hand (iib), we will provide a general inf-sup condition for the PML problem under the general assumption that the original problem is also inf-sub stable. In this case, we believe that our result is sub-optimal, as the PML inf-sup constant is deteriorated compared to the original one.

3.1 The PML Helmholtz problem in periodic structures

In this section, we rigorously introduce the "PML problem" set in an enlarged domain $\widetilde{\Omega}$. We also show that we can equivalently reformulate the PML onto the original domain Ω by introducing a PML boundary operator $\mathscr{R}_{\mathrm{P}}$ that may be viewed as a perturbation of the exact DtN operator $\mathscr{R}$, and we provide an explicit expression for the perturbed operator $\mathscr{R}_{\mathrm{P}}$.



(a) Non quasi-resonant mode (b) Quasi-resonant mode

Figure 3.2: Absorption in a PML

3.1.1 The PML Helmholtz problem and its variational formulation

Our absorbing layer is characterized by a depth $\ell_{\mathrm{P}} > 0$ and two dimensionless constants $\gamma_{\mathrm{r}}, \gamma_{\mathrm{i}} > 0$. The notations

$$\chi_{\mathrm{P}}(\boldsymbol{x}) := \mathbf{1}_{x_2>\ell_2}, \qquad \chi_{\mathrm{I}} := 1 - \chi_{\mathrm{P}}, \qquad \nu := (\gamma_{\mathrm{r}} + i\gamma_{\mathrm{i}})\,\chi_{\mathrm{P}} + \chi_{\mathrm{I}},$$

and

$$\gamma_\star = \min(\gamma_{\mathrm{r}}, \gamma_{\mathrm{i}}), \qquad \gamma^\star = \sup(\gamma_{\mathrm{r}}, \gamma_{\mathrm{i}})$$

will be useful. For the sake of simplicity, we will assume throughout this work that $\gamma_{\mathrm{r}}, \gamma_{\mathrm{i}} \geq 1$. In particular, setting $\nu_{\mathrm{P}} = \gamma_{\mathrm{r}} + i\gamma_{\mathrm{i}}$, we have

$$\sqrt{2} \leq |\nu_{\mathrm{P}}| \leq \sqrt{2}\gamma^\star. \tag{3.1.1}$$

Notice that in essence, our results remain valid if $\gamma_{\mathrm{r}}, \gamma_{\mathrm{i}} < 1$, but the constants may blow up as $\gamma_{\mathrm{r}} \to 0$ or $\gamma_{\mathrm{i}} \to 0$.

The PML problem is set on the enlarged domain $\widetilde{\Omega} := (0, \ell_1) \times (0, \ell_2 + \ell_{\mathrm{P}})$. $\widetilde{\Omega}$ is composed of two parts, the physical domain $\Omega := (0, \ell_1) \times (0, \ell_2)$ and the absorbing layer $\Omega_{\mathrm{P}} := (0, \ell_1) \times (\ell_2, \ell_{\mathrm{P}})$. Its boundary is partitioned as $\partial\widetilde{\Omega} = \Gamma_{\mathrm{P}} \cup \Gamma_{\mathrm{D}} \cup \widetilde{\Gamma}_\sharp$, where

$$\begin{aligned}
\Gamma_{\mathrm{P}} &:= (0, \ell_1) \times \{\ell_2 + \ell_{\mathrm{P}}\} \\
\widetilde{\Gamma}_\sharp, &:= \widetilde{\Gamma}_{\sharp+} \cup \widetilde{\Gamma}_{\sharp-}, \\
\widetilde{\Gamma}_\sharp^- &:= \{0\} \times (0, \ell_2 + \ell_{\mathrm{P}}), \\
\widetilde{\Gamma}_\sharp^+ &:= \{\ell_1\} \times (0, \ell_2 + \ell_{\mathrm{P}}).
\end{aligned} \tag{3.1.2}$$

The geometrical setting is illustrated on Figure 3.3.

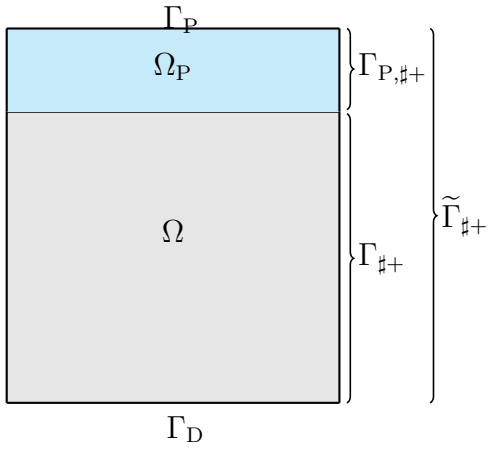


Figure 3.3: Geometrical setting of the PML problem

The PML problem is obtained by formally replacing the $\partial/\partial x_2$ operator by $\nu^{-1}\partial/\partial x_2$. Since ν does not depend on x_1, we can then multiply the resulting equation by ν, leading to the following: Find $\widetilde{u}\colon \widetilde{\Omega} \to \mathbb{C}$ such that

$$\left\{\begin{array}{rcll} -k^2\nu\varepsilon\widetilde{u} - \dfrac{\partial}{\partial x_1}\left(\nu A_1 \dfrac{\partial \widetilde{u}}{\partial x_2}\right) - \dfrac{\partial}{\partial x_2}\left(\nu^{-1} A_2 \dfrac{\partial \widetilde{u}}{\partial x_2}\right) & = & \varepsilon\widetilde{f} & \text{in } \widetilde{\Omega} \\ \widetilde{u} & = & 0 & \text{on } \Gamma_{\mathrm{P}} \\ \widetilde{u} & = & 0 & \text{on } \Gamma_{\mathrm{D}} \\ \widetilde{u}_+ - e^{i\alpha\ell_1}\widetilde{u}_- & = & 0 & \text{on } \widetilde{\Gamma}_\sharp. \end{array}\right. \tag{3.1.3}$$

Introducing the Sobolev space

$$H^1_\sharp(\widetilde{\Omega}) := \left\{ v \in H^1(\widetilde{\Omega}, \mathbb{C}) \mid v_{|\Gamma_{\mathrm{D}}} = v_{|\Gamma_{\mathrm{P}}} = 0 \ \text{ and } \ v_{|\widetilde{\Gamma}_{\sharp+}} = e^{i\alpha\ell_1} v_{|\widetilde{\Gamma}_{\sharp-}} \right\},$$

and assuming that $\widetilde{f} \in L^2(\widetilde{\Omega})$, we can recast the above PML problem as follows: Find $\widetilde{u} \in H^1_\sharp(\widetilde{\Omega})$ such that

$$\widetilde{b}(\widetilde{u}, \widetilde{v}) = (\varepsilon\widetilde{f}, \widetilde{v})_{\widetilde{\Omega}} \quad \forall \widetilde{v} \in H^1_\sharp(\widetilde{\Omega}), \tag{3.1.4}$$

where

$$\widetilde{b}(\widetilde{u}, \widetilde{v}) = -k^2(\nu\varepsilon\widetilde{u}, \widetilde{v})_{\widetilde{\Omega}} + (\nu A_1 \partial_1 \widetilde{u}, \partial_1 \widetilde{v})_{\widetilde{\Omega}} + (\nu^{-1} A_2 \partial_2 \widetilde{u}, \partial_2 \widetilde{v})_{\widetilde{\Omega}}.$$

Remark 3.1.1 (Damping function)**.** *Remark that the PML damping function ν is chosen as piecewise constant, which simplifies the numerical implementation compared to other damping functions. A constant damping function has already been used in [134]. Moreover, different functions have been used for 1D-periodic geometry, including power functions in [40, 39] and non-integrable functions in [19, 123].*

Remark 3.1.2 (Damping coefficient)**.** *The original PML condition takes* $\gamma_{\mathrm{r}} = 1$ *in the PML region. In contrast, we consider a variable* γ_{r} ($\gamma_{\mathrm{r}} \neq 1$) *in order to attenuate both propagating and evanescent waves in this region (see numerical examples 3.2.3 for validation).*

Remark 3.1.3 (Dirichlet boundary condition)**.** *We chose a Dirichlet boundary condition on the external PML boundary* Γ_{P}*. As in [134, 40, 39], one could use other boundary conditions (e.g., Neumann or Robin condition) and the upcoming analysis would be largely similar.*

3.1.2 PML as a DtN approximation

The goal of this subsection is to understand the PML as a perturbation of DtN operator. To do so, we will focus on the case where the right-hand side $\widetilde{f} \in L^2(\widetilde{\Omega})$ of (3.1.4) is supported in Ω, and we will rewrite an equivalent definition $\widetilde{u}|_\Omega$ through a PDE problem only set in Ω. This problem, although set in Ω as the original Helmholtz problem (2.1.19), will have a different boundary condition on Γ_{A}. Actually, we will see that this boundary condition is in fact a DtN operator, but associated with the with the absorbing layer Ω_{P} instead of the semi-infinite strip $(0, \ell_1) \times (\ell_2, +\infty)$.

We first note that any solution $\widetilde{u}$ to (3.1.3) satisfies the original Helmholtz PDE

$$-k^2\varepsilon\widetilde{u} - \boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla}\widetilde{u}) = \widetilde{f},$$

in the physical domain, since $\nu = 1$ in Ω. Therefore, $\widetilde{u}$ must satisfy the original problem in Ω, but with a different boundary condition. To make this boundary condition explicit, we rewrite the PML problem as: Find $u^0 : \Omega \to \mathbb{C}$ and $u^{\mathrm{P}} : \Omega_{\mathrm{P}} \to \mathbb{C}$ such that

$$\left\{\begin{array}{rcll} -k^2\varepsilon u^0 - \boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla}u_0) &=& \varepsilon\widetilde{f} & \text{in } \Omega, \\ u^0 &=& 0 & \text{on } \Gamma_{\mathrm{D}}, \\ {u^0}_+ - e^{i\alpha\ell_1}{u^0}_- &=& 0 & \text{on } \Gamma_\sharp, \\ \boldsymbol{\nabla}u^0\cdot\boldsymbol{n} &=& \nu^{-1}\boldsymbol{\nabla}u^{\mathrm{P}}\cdot\boldsymbol{n} & \text{on } \Gamma_{\mathrm{A}}, \end{array}\right. \tag{3.1.5}$$

and

$$\left\{\begin{array}{rcll} -k^2\nu_{\mathrm{P}}^2 u^{\mathrm{P}} - \nu_{\mathrm{P}}^2\dfrac{\partial^2 u^{\mathrm{P}}}{\partial x_1^2} - \dfrac{\partial^2 u^{\mathrm{P}}}{\partial x_2^2}, &=& \nu_{\mathrm{P}}\widetilde{f} & \text{in } \Omega_{\mathrm{P}} \\ u^{\mathrm{P}} &=& 0 & \text{on } \Gamma_{\mathrm{P}}, \\ u^{\mathrm{P}}_+ - e^{i\alpha\ell_1}u^{\mathrm{P}}_- &=& 0 & \text{on } \Gamma_{\mathrm{P},\sharp}, \\ u^{\mathrm{P}} &=& u^0 & \text{on } \Gamma_{\mathrm{A}}. \end{array}\right. \tag{3.1.6}$$

Notice that the two boundary conditions on Γ_{A} in (3.1.5) and (3.1.6), imply that if set $\widetilde{u} := u^0\chi_{\mathrm{I}} + u^{\mathrm{P}}\chi_{\mathrm{P}}$, then $[\![\widetilde{u}]\!] = 0$ and $[\![\nu^{-1}\partial_2\widetilde{u}]\!] = 0$ on Γ_{A}, so that $\widetilde{u}$ is indeed a solution to (3.1.3).

Assuming that $\widetilde{f}|_\Omega = f$ and $\widetilde{f}|_{\Omega_\mathrm{P}} = 0$ and that Problem (3.1.6) admits a unique solution for any $u^0 \in H^{1/2}_\sharp(\Gamma_\mathrm{A})$, we may define a bounded linear operator $\mathscr{R}_\mathrm{P} : H^{1/2}_\sharp(\Gamma_\mathrm{A}) \to \left(H^{1/2}_\sharp(\Gamma_\mathrm{A})\right)'$ as

$$\mathscr{R}_\mathrm{P} u^0 := \nu^{-1} \boldsymbol{\nabla} u_\mathrm{P} \cdot \boldsymbol{n}, \tag{3.1.7}$$

and (3.1.5) then reads

$$\left\{ \begin{array}{rcll} -k^2 \varepsilon u - \boldsymbol{\nabla} \cdot (\boldsymbol{A} \boldsymbol{\nabla} u) & = & \varepsilon f & \text{in } \Omega, \\ u & = & 0 & \text{on } \Gamma_\mathrm{D}, \\ u_+ - e^{i\alpha\ell_1} u_- & = & 0 & \text{on } \Gamma_\sharp, \\ \boldsymbol{A} \boldsymbol{\nabla} u \cdot \boldsymbol{n} - \mathscr{R}_\mathrm{P} u & = & 0 & \text{on } \Gamma_\mathrm{A}. \end{array} \right. \tag{3.1.8}$$

We thus see that the PML problem is similar to the original one (2.1.18): it amounts to replacing $\mathscr{R}$ by $\mathscr{R}_\mathrm{P}$. Therefore, the error between the operators $\mathscr{R}$ and $\mathscr{R}_\mathrm{P}$ will reflect the convergence of the PML solution in the domain of interest toward the original solution. The sesquilinear form

$$b_\mathrm{P}(\phi, v) := -k^2 (\varepsilon \phi, v)_\Omega - \langle \mathscr{R}_\mathrm{P} \phi, v \rangle_{\Gamma_\mathrm{A}} + (\boldsymbol{A} \boldsymbol{\nabla} \phi, \boldsymbol{\nabla} v)_\Omega \quad \forall \phi, v \in H^1_\sharp(\Omega) \tag{3.1.9}$$

associated with the weak form of (3.1.8) will be useful later. The definition of $\mathscr{R}_\mathrm{P}$ shows that whenever $\widetilde{f} \in L^2(\Omega_\mathrm{P})$ is supported in Ω, if $\widetilde{u} \in H^1_\sharp(\widetilde{\Omega})$ solves (3.1.4), $\widetilde{u}|_\Omega$ can be equivalently defined as the unique element of $H^1_\sharp(\Omega)$ such that

$$b_\mathrm{P}(\widetilde{u}|_\Omega, v) = (\widetilde{f}, v)_\Omega \quad \forall v \in H^1_\sharp(\Omega).$$

The next lemma shows that $\mathscr{R}_\mathrm{P}$ is indeed well-defined, and that an explicit expression is available. Here, we use the fact that we have chosen ν constant inside the PML layer.

Lemma 3.1.4. *For all $u^0 \in H^{1/2}_\sharp(\Gamma_\mathrm{A})$, Problem (3.1.6) admits a unique solution $u^\mathrm{P} \in H^1_\sharp(\Omega_\mathrm{P})$ with $\widetilde{f} = 0$. In addition, we have*

$$\mathscr{R}_\mathrm{P} U = i \sum_{n \in \mathbb{Z}} \frac{1 + e^{2i\nu k_n \ell_\mathrm{P}}}{1 - e^{2i\nu k_n \ell_\mathrm{P}}} k_n \widehat{U}_n e^{i(\alpha + \alpha_n) x_1} \qquad \forall U \in H^{1/2}_\sharp(\Gamma_\mathrm{A}). \tag{3.1.10}$$

Remark 3.1.5. *In the presence of quasi-resonance, there exists values of $n \in \mathbb{Z}$ for which $k_n \ell_\mathrm{P} = 0$, so that the expression*

$$\frac{1 + e^{2i\nu k_n \ell_\mathrm{P}}}{1 - e^{2i\nu k_n \ell_\mathrm{P}}} k_n$$

is not well-defined. In this case, it should be interpreted as

$$\frac{1}{\ell_\mathrm{P}} \frac{1 + e^{2i\nu 0}}{1 - e^{2i\nu 0}} 0 := -\frac{1}{\nu \ell_\mathrm{P}} = \lim_{k\ell_\mathrm{P} \to 0} \frac{1}{\ell_\mathrm{P}} \frac{1 + e^{2i\nu k\ell_\mathrm{P}}}{1 - e^{2i\nu k\ell_\mathrm{P}}} k\ell_\mathrm{P}.$$

Proof. For the sake of simplicity, we write $U \in H_\sharp^{1/2}(\Gamma_{\mathrm{A}})$ for the boundary data and $u \in H_\sharp^1(\Omega_{\mathrm{P}})$ the solution. We have

$$-k^2\nu^2 u - \nu^2 \frac{\partial^2 u}{\partial x_1^2} - \frac{\partial^2 u}{\partial x_2^2} = 0$$

in Ω_{P} together with the boundary conditions $u|_{\Gamma_{\mathrm{P}}} = 0$, $u|_{\Gamma_{\mathrm{A}}} = U$, and quasi-periodicity. Using Fourier expansion, we then obtain that

$$-k_n^2\nu^2\widehat{u}_n - \widehat{u}_n'' = 0 \tag{3.1.11}$$

with the additional conditions that $\widehat{u}_n(\ell_2) = \widehat{U}_n$ and $\widehat{u}_n(\ell_2 + \ell_{\mathrm{P}}) = 0$. We may first write that

$$\widehat{u}_n(x_2) = c_- e^{-ik_n\nu(x_2-\ell_2-\ell_P)} + c_+ e^{ik_n\nu(x_2-\ell_2-\ell_P)}.$$

Then

$$0 = \widehat{u}_n(\ell_2 + \ell_{\mathrm{P}}) = c_- + c_+ = 0,$$

so that

$$\widehat{u}_n(x_2) = c(e^{ik_n\nu(x_2-\ell_2-\ell_{\mathrm{P}})} - e^{-ik_n\nu(x_2-\ell_2-\ell_{\mathrm{P}})}).$$

and then

$$\widehat{U}_n = \widehat{u}_n(\ell_2) = c(e^{-ik_n\nu\ell_{\mathrm{P}}} - e^{ik_n\nu\ell_{\mathrm{P}}}),$$

thus

$$\widehat{u}_n(x_2) = \widehat{U}_n \frac{e^{ik_n\nu(x_2-\ell_2-\ell_{\mathrm{P}})} - e^{-ik_n\nu(x_2-\ell_2-\ell_{\mathrm{P}})}}{e^{-ik_n\nu\ell_{\mathrm{P}}} - e^{ik_n\nu\ell_{\mathrm{P}}}},$$

and

$$\widehat{u}_n'(x_2) = ik_n\nu\widehat{U}_n \frac{e^{ik_n\nu(x_2-\ell_2-\ell_{\mathrm{P}})} + e^{-ik_n\nu(x_2-\ell_2-\ell_{\mathrm{P}})}}{e^{-ik_n\nu\ell_{\mathrm{P}}} - e^{ik_n\nu\ell_{\mathrm{P}}}},$$

as well as

$$\nu^{-1}\widehat{u}_n'(\ell_2) = ik_n\widehat{U}_n \frac{e^{-ik_n\nu\ell_{\mathrm{P}}} + e^{ik_n\nu\ell_{\mathrm{P}}}}{e^{-ik_n\nu\ell_{\mathrm{P}}} - e^{ik_n\nu\ell_{\mathrm{P}}}} = ik_n \frac{1 + e^{2ik_n\nu\ell_{\mathrm{P}}}}{1 - e^{2ik_n\nu\ell_{\mathrm{P}}}}\widehat{U}_n.$$

As a result, if $k_n \neq 0$ for all $n \in \mathbb{Z}$, we have

$$\mathscr{R}_{\mathrm{P}} u = \sum_{n\in\mathbb{Z}} \frac{1 + e^{2ik_n\nu\ell_{\mathrm{P}}}}{1 - e^{2ik_n\nu\ell_{\mathrm{P}}}} ik_n\widehat{u}_n.$$

Now, if $k_n = 0$ for some $n \in \mathbb{Z}$, instead of (3.1.11) we have

$$-\widehat{u}_n'' = 0,$$

with the boundary conditions $\widehat{u}_n(\ell_2) = \widehat{U}_n$ and $\widehat{u}_n(\ell_2 + \ell_{\mathrm{P}}) = 0$. Then

$$\widehat{u}_n(x_2) = \frac{\widehat{U}_n}{\ell_{\mathrm{P}}} (\ell_{\mathrm{P}} + \ell_2 - x_2),$$

as well as

$$\widehat{u}_n'(x_2) = \frac{-\widehat{U}_n}{\ell_{\mathrm{P}}}.$$

As a result, we get

$$\nu^{-1}\widehat{u}_n'(\ell_2) = \frac{-1}{\nu\ell_{\mathrm{P}}}\widehat{U}_n.$$

Since

$$\lim_{k_n\to 0} = \frac{1+e^{2ik_n\nu\ell_{\mathrm{P}}}}{1-e^{2ik_n\nu\ell_{\mathrm{P}}}}ik_n = -\frac{1}{\nu\ell_P},$$

we will maintain the notation

$$\mathscr{R}_{\mathrm{P}}u = \sum_{n\in\mathbb{Z}} \frac{1+e^{2ik_n\nu\ell_{\mathrm{P}}}}{1-e^{2ik_n\nu\ell_{\mathrm{P}}}}ik_n\widehat{u}_n,$$

where the multiplicative coefficients are to be understood as explained in Remark 3.1.5. □

3.2 Error estimates

The goal of this section is to establish error estimates that control (i) the difference between the DtN operators $\mathscr{R}$ and $\mathscr{R}_{\mathrm{P}}$, and (ii) the difference between the original solution u and the solution to the PML problem $\widetilde{u}$. These estimates are certainly of interest in view of error control of finite element discretization. As a result, these questions have already been partly adressed in the literature. For instance, we may cite works on adaptive finite element schemes [40, 39, 134]. One limitation of these works is that they explicitly exclude quasi-resonances from their study. Another interesting work is the PhD [136] where the difference $(\mathscr{R}-\mathscr{R}_{\mathrm{P}})$ is controlled, but the corresponding estimate on $(u-\widetilde{u})|_\Omega$ is not rigorously established.

3.2.1 Error estimates for the DtN approximation

In this subsection, we measure how well $\mathscr{R}_{\mathrm{P}}$ approximates $\mathscr{R}$. Specifically, we will provide an upper bound for $(\mathscr{R}-\mathscr{R}_{\mathrm{P}})$ in a suitable operator norm. This is done thanks to expressions of $\mathscr{R}$ and $\mathscr{R}_{\mathrm{P}}$ that we explicitly obtained in terms of Fourier modes earlier.

Theorem 3.2.1 (Error estimate for the DtNs)**.** *We have $\mathscr{R}-\mathscr{R}_{\mathrm{P}} : L^2(\Gamma_{\mathrm{A}}) \to L^2(\Gamma_{\mathrm{A}})$. In addition, the estimate*

$$\sup_{\substack{U\in H^1_\sharp(\Gamma_{\mathrm{A}})\\ \|U\|_{\Gamma_{\mathrm{A}}}=1}} \|(\mathscr{R}-\mathscr{R}_{\mathrm{P}})U\|_{\Gamma_{\mathrm{A}}} \le \varepsilon_{\min} k\mathscr{E}_{\mathrm{P}} \tag{3.2.1}$$

holds true with

$$\mathscr{E}_{\mathrm{P}} := \frac{1}{\varepsilon_{\min}}\frac{1}{\gamma_\star}\frac{1}{k\ell_{\mathrm{P}}}\left(1+\gamma_\star\Lambda_{k,\boldsymbol{\ell},\theta}\sqrt{k\ell_1}\frac{\ell_{\mathrm{P}}}{\ell_1}\right)\exp\left(-\gamma_\star\Lambda_{k,\boldsymbol{\ell},\theta}\sqrt{k\ell_1}\frac{\ell_{\mathrm{P}}}{\ell_1}\right).$$

Proof. Recalling Definition (2.1.15) of $\mathscr{R}$ and characterization (3.1.10) of $\mathscr{R}_{\mathrm{P}}$, we have

$$(\mathscr{R}-\mathscr{R}_{\mathrm{P}})U = i\sum_{n\in\mathbb{Z}}\left(1-\frac{1+e^{i2\nu k_n\ell_{\mathrm{P}}}}{1-e^{i2\nu k_n\ell_{\mathrm{P}}}}\right)k_n\widehat{U}_n e^{i(\alpha+\alpha_n)x_1} = -\frac{1}{\nu\ell_{\mathrm{P}}}\sum_{n\in\mathbb{Z}}\gamma_n\widehat{U}_n e^{i(\alpha+\alpha_n)x_1},$$

with

$$\gamma_n := (i2\nu k_n\ell_{\mathrm{P}})\frac{e^{i2\nu k_n\ell_{\mathrm{P}}}}{1-e^{i2\nu k_n\ell_{\mathrm{P}}}}.$$

Next, recalling that $\nu_{\mathrm{P}} = \gamma_{\mathrm{r}} + i\gamma_{\mathrm{i}}$, we have

$$e^{i2\nu k_n\ell_{\mathrm{P}}} = e^{i2k_n\gamma_{\mathrm{r}}\ell_{\mathrm{P}}}e^{-2k_n\gamma_{\mathrm{i}}\ell_{\mathrm{P}}}.$$

Then, either $k_n = \beta_n$ and $|e^{i2\nu k_n\ell_{\mathrm{P}}}| = e^{-2|k_n|\gamma_{\mathrm{i}}\ell_{\mathrm{P}}}$ or $k_n = i\beta_n$ and $|e^{i2\nu k_n\ell_{\mathrm{P}}}| = e^{-2|k_n|\gamma_{\mathrm{r}}\ell_{\mathrm{P}}}$, leading to

$$|e^{i2\nu k_n\ell_{\mathrm{P}}}| \le e^{-2\gamma_\star|k_n|\ell_{\mathrm{P}}}. \tag{3.2.2}$$

Combining (3.2.2) with the reverse triangle inequality $|1-e^{i2\nu k_n\ell_{\mathrm{P}}}| \ge 1-|e^{i2\nu k_n\ell_{\mathrm{P}}}|$, we arrive at

$$|\gamma_n| \le 2|\nu||k_n|\ell_{\mathrm{P}}\frac{e^{-2\gamma_\star|k_n|\ell_{\mathrm{P}}}}{1-e^{-2\gamma_\star|k_n|\ell_{\mathrm{P}}}} = \frac{|\nu|}{\gamma_\star}h(2\gamma_\star|k_n|\ell_{\mathrm{P}}),$$

where $h(x) := xe^{-x}/(1-e^{-x})$. Then, elemental analysis shows that

$$h(x) \le g(x) := (1+x)e^{-x},$$

where g is a non-increasing function of $x \ge 0$. It follows that

$$|\gamma_n| \le \frac{|\nu|}{\gamma_\star}g(\gamma_\star|k_n|\ell_{\mathrm{P}}) \le \frac{|\nu|}{\gamma_\star}g(\gamma_\star\min_{n\in\mathbb{Z}}|k_n|\ell_{\mathrm{P}}),$$

and (3.2.1) follows from (2.2.21). □

Remark 3.2.2 (Converge rates)**.** *Far from quasi-resonances, the convergence is exponential of the PML approximation exponential. Suppose for example that* $2k_\star\gamma_\star\ell_{\mathrm{P}} > \ln 2$*, then*

$$\mathscr{E}_{\mathrm{P}} \le \frac{4}{\varepsilon_{\min}}\frac{k_\star}{k}e^{-2\gamma_\star|k_\star|\ell_{\mathrm{P}}}. \tag{3.2.3}$$

Thus, exponential convergence is ensured when the thickness ℓ_{P} *and/or the damping coefficients* $\gamma_\star$ *of the PML layer go to infinity.*

In contrast, we only maintain a linear convergence for modes close to quasi-resonances. In fact, if $k_\star = 0$*, we simply have*

$$\mathscr{E}_{\mathrm{P}} \le \frac{1}{k\varepsilon_{\min}}\frac{1}{\gamma_\star\ell_{\mathrm{P}}}, \tag{3.2.4}$$

and the convergence becomes linear depending on PML thickness ℓ_{P} *and its damping coefficients* $\gamma_\star$*.*

Equally important observation is that the appearance of $\gamma_\star$ in the bound (3.2.1) is due to inequality (3.2.2). Moreover, because of equation (3.2.1), if k_n is associated with an outgoing wave (i.e. non-negative real number $k_n = |k_n|$), then $\gamma_\star$ can be replaced by γ_{i}. Consequently, the variation of the damping coefficient γ_{i} will affect the absorption of the outgoing propagative modes. On the other hand, if k_n is associated with an evanescent wave (pure imaginary number with positive imaginary part and $k_n = i|k_n|$), then $\gamma_\star$ can be replaced by γ_{r}. Therefore, as indicated in remark 3.1.1, the variation of the coefficient γ_{r} will affect the absorption of evanescent modes.

Remark 3.2.3 (Truncation with a Neumann boundary condition)**.** *Let us we rapidly discuss the case of a Neumann boundary condition instead of the Dirichlet boundary condition to truncate the PML, and show that no substantial gain in efficiency should be expected.*

Considering a homogeneous Neumann boundary condition on Γ_{P}, we get then the PML map $\mathscr{R}_{\mathrm{P}}^{\mathrm{N}} : H_\sharp^{1/2}(\Gamma_{\mathrm{A}}) \to \left(H_\sharp^{1/2}(\Gamma_{\mathrm{A}})\right)'$ as

$$\mathscr{R}_{\mathrm{P}}^{\mathrm{N}} U = i \sum_{n\in\mathbb{N}} \tau_n \widehat{U}_n e^{i(\alpha+\alpha_n)x_1},$$

where

$$\tau_n = \frac{1 - e^{2i\nu k_n \ell_{\mathrm{P}}}}{1 + e^{2i\nu k_n \ell_{\mathrm{P}}}} k_n.$$

Thus,

$$\lim_{k_n \to 0} \tau_n = 0,$$

and this result allows the operator to be efficient and to treat exactly the quasi-resistant modes where $k_n = 0$. However, this is not the case for the modes near the quasi-resonance. In fact, we have

$$(\mathscr{R} - \mathscr{R}_{\mathrm{P}}^{\mathrm{N}})U = i \sum_{n\in\mathbb{N}} \left(1 - \frac{1 - e^{i2\nu k_n \ell_{\mathrm{P}}}}{1 + e^{i2\nu k_n \ell_{\mathrm{P}}}}\right) k_n \widehat{U}_n e^{i(\alpha+\alpha_n)x_1} = \frac{1}{\nu \ell_{\mathrm{P}}} \sum_{n\in\mathbb{N}} \delta_n \widehat{U}_n e^{i(\alpha+\alpha_n)x_1},$$

with

$$\delta_n := (i2k_n\ell_{\mathrm{P}}) \frac{e^{i2\nu k_n \ell_{\mathrm{P}}}}{1 + e^{i\nu k_n \ell_{\mathrm{P}}}}.$$

Considering $\nu = 1 + i\gamma_{\mathrm{i}}$, similar calculations to proof of Theorem 3.2.1 show that

$$|\delta_n| \leq \frac{1}{\ell_{\mathrm{P}}} g(2|k_n|\ell_{\mathrm{P}}),$$

where

$$g(t) := \frac{te^{-t}}{|1 + e^{-i\gamma_{\mathrm{i}} t} e^{-t}|} \qquad t \geq 0.$$

In addition, elemental analysis shows that if $t \to 0$ and $x \to \pi/\gamma_{\mathrm{i}}$ then $g(t) \to 1$. Then, the exponential convergence of the PML solution turns linear in this case. To help illustrate

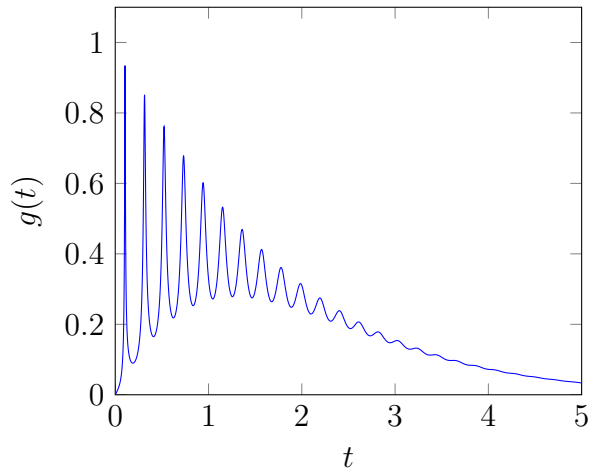


Figure 3.4: Function g for $\gamma_{\mathrm{i}} = 30$.

the variation of the PML error in this case of the Neumann boundary condition, Figure 3.4 depicts the function g for $\gamma_{\mathrm{i}} = 30$.

The graph of g shows that it vanishes at zero, which explains the exact treatment of quasi-resonant modes ($k_n = 0$). But, also shows that for small values of k_n, the PML solution keeps only the linear convergence ($g(t) \approx 1$ for some $t \approx 0$).

3.2.2 Error estimates for the PML solution

Having establish an error error estimate controlling $\mathscr{R} - \mathscr{R}_{\mathrm{P}}$, we now focus on controlling the difference $(u - \widetilde{u})|_\Omega$ between the Helmholtz solution associated with the operator $\mathscr{R}$ and $\mathscr{R}_{\mathrm{P}}$. We start by establishing a useful trace inequality.

Lemma 3.2.4 (Trace inequality)**.** *For all $v \in H^1(\Omega)$, the following trace inequality holds true:*

$$\varepsilon_{\min} k \|v\|^2_{\Gamma_{\mathrm{A}}} \leq |\!|\!| v |\!|\!|^2_{k,\Omega} \,. \tag{3.2.5}$$

Proof. We simply need to show (3.2.5) for smooth functions. The general results will then follow by density. Let us thus consider $v \in C^\infty(\overline{\Omega})$ with $v|_{\Gamma_D} = 0$. Let $\boldsymbol{x}_1 \in (0, \ell_1)$. Since $v(\boldsymbol{x}_1, 0) = 0$, we have

$$k\|v(\boldsymbol{x}_1, \ell_2)\|^2 = 2k \,\mathrm{Re} \int_0^{\ell_2} v(\boldsymbol{x}_1, \boldsymbol{x}_2) \frac{\partial \overline{v}}{\partial \boldsymbol{x}_2}(\boldsymbol{x}_1, \boldsymbol{x}_2) d\boldsymbol{x}_2,$$

and integrating over $\boldsymbol{x}_1$ shows that

$$k\|v\|^2_{\Gamma_{\mathrm{A}}} = 2k \,\mathrm{Re} \int_\Omega v \frac{\partial \overline{v}}{\partial \boldsymbol{x}_2} \leq 2k \|v\|_\Omega \left\| \frac{\partial v}{\partial \boldsymbol{x}_2} \right\|_\Omega \leq 2(k\varepsilon_{\min}^{-1/2} \|v\|_{\varepsilon,\Omega}) \|\boldsymbol{\nabla} v\|_{\boldsymbol{A},\Omega},$$

and (3.2.5) follows from Young's inequality. □

The next step of our analysis is to establish an inf-sup condition for b_{P} assuming that $\mathscr{E}_{\mathrm{P}}$ is sufficiently small.

Lemma 3.2.5 (Inf-sup condition for b_{P})**.** *Assume that $\mathscr{C}_{\mathrm{is}}\mathscr{E}_{\mathrm{P}} < 1$. Then, the inf-sup constant*

$$\frac{1}{\mathscr{C}_{\mathrm{is,P}}} := \inf_{\substack{\phi\in H^1_\sharp(\Omega)\\ \|\phi\|_{k,\Omega}=1}} \sup_{\substack{v\in H^1_\sharp(\Omega)\\ \|v\|_{k,\Omega}=1}} \mathrm{Re}\, b_{\mathrm{P}}(\phi, v) \tag{3.2.6}$$

is finite and we have

$$\mathscr{C}_{\mathrm{is,P}} \leq \frac{\mathscr{C}_{\mathrm{is}}}{1-\mathscr{C}_{\mathrm{is}}\mathscr{E}_{\mathrm{P}}}. \tag{3.2.7}$$

Proof. Let $\phi \in H^1_\sharp(\Omega)$ with $\|\phi\|_{k,\Omega} = 1$. Since $H^1_\sharp(\Omega)$ is Hilbert space, we may replace the supremum in (2.1.21) by maximum, and as a result, there exists $v^\star \in H^1_\sharp(\Omega)$ with $\|v^\star\|_{k,\Omega} = 1$ such that

$$\mathrm{Re}\, b(\phi, v^\star) \geq \frac{1}{\mathscr{C}_{\mathrm{is}}}.$$

Then, we have

$$\mathrm{Re}\, b_{\mathrm{P}}(\phi, v^\star) = \mathrm{Re}\, b(\phi, v^\star) - \mathrm{Re}\langle(\mathscr{R}_{\mathrm{P}} - \mathscr{R})\,\phi, v^\star\rangle_{\Gamma_{\mathrm{A}}} \geq \frac{1}{\mathscr{C}_{\mathrm{is}}} - k\mathscr{E}_{\mathrm{P}}\varepsilon_{\min}\|\phi\|_{\Gamma_{\mathrm{A}}}\|v^\star\|_{\Gamma_{\mathrm{A}}}.$$

Recalling (3.2.5), we have $\varepsilon_{\min}k\|u\|_{\Gamma_{\mathrm{A}}}\|v^\star\|_{\Gamma_{\mathrm{A}}} \leq \|u\|_{k,\Omega}\|v^\star\|_{k,\Omega} = 1$, and as a result

$$\mathrm{Re}\, b_{\mathrm{P}}(\phi, v^\star) \geq \frac{1}{\mathscr{C}_{\mathrm{is}}} - \mathscr{E}_{\mathrm{P}} = \frac{1-\mathscr{C}_{\mathrm{is}}\mathscr{E}_{\mathrm{P}}}{\mathscr{C}_{\mathrm{is}}}.$$

Estimate (3.2.7) then follows since ϕ was arbitrary. □

We are finally ready to deliver the key result of this section.

Theorem 3.2.6 (PML error estimate)**.** *If u and $\widetilde{u}$ respectively denote the solutions to the original and PML Helmholtz problems (2.1.19) and (3.1.4) with right-hand side $f \in L^2(\Omega)$, the error estimates*

$$\|u-\widetilde{u}\|_{k,\Omega} \leq \frac{\mathscr{C}_{\mathrm{is}}\mathscr{E}_{\mathrm{P}}}{1-\mathscr{C}_{\mathrm{is}}\mathscr{E}_{\mathrm{P}}}\|u\|_{k,\Omega}\,, \tag{3.2.8}$$

and

$$k\,\|u-\widetilde{u}\|_{k,\Omega} \leq \frac{\mathscr{C}_{\mathrm{is}}\mathscr{E}_{\mathrm{P}}}{1-\mathscr{C}_{\mathrm{is}}\mathscr{E}_{\mathrm{P}}}\mathscr{C}_{\mathrm{is}}\|f\|_{\varepsilon,\Omega}, \tag{3.2.9}$$

hold true.

Proof. For all $v \in H^1_\sharp(\Omega)$, we have

$$b(u, v) = (f, v) = b_{\mathrm{P}}(\widetilde{u}, v),$$

and it follows in particular that

$$b_{\mathrm{P}}(u-\widetilde{u},v)=b_{\mathrm{P}}(u,v)-b(u,v)=\langle(\mathscr{R}-\mathscr{R}_{\mathrm{P}})u,v\rangle.$$

Employing (3.2.1) and (3.2.5), we have

$$|b_{\mathrm{P}}(u-\widetilde{u},v)|\leq\|(\mathscr{R}-\mathscr{R}_{\mathrm{P}})u\|_{\Gamma_{\mathrm{A}}}\|v\|_{\Gamma_{\mathrm{A}}}\leq\varepsilon_{\min}k\mathscr{E}_{\mathrm{P}}\|u\|_{\Gamma_{\mathrm{A}}}\|v\|_{\Gamma_{\mathrm{A}}}\leq\mathscr{E}_{\mathrm{P}}\,|\!|\!|u|\!|\!|_{k,\Omega}\,|\!|\!|v|\!|\!|_{k,\Omega}\,,$$

Then, (3.2.6) implies that there exists $v^\star\in H^1_\sharp(\Omega)$ with $|\!|\!|v|\!|\!|_{k,\Omega}=1$ such that

$$|\!|\!|u-\widetilde{u}|\!|\!|_{k,\Omega}\leq\mathscr{C}_{\mathrm{is,P}}\,\mathrm{Re}\,b_{\mathrm{P}}(u-\widetilde{u},v^\star)\leq\mathscr{C}_{\mathrm{is,P}}\mathscr{E}_{\mathrm{P}}\,|\!|\!|u|\!|\!|_{k,\Omega}\,|\!|\!|v^\star|\!|\!|_{k,\Omega}=\mathscr{C}_{\mathrm{is,P}}\mathscr{E}_{\mathrm{P}}\,|\!|\!|u|\!|\!|_{k,\Omega}\,,$$

which, together with (3.2.6), proves (3.2.8). Then, (3.2.9) simply follows from (2.1.25). □

3.2.3 Numerical examples

In this subsection, we present numerical experiments illustrating our theoretical convergence results on the PML approach. We employ the same setting than in subsection 2.2.4. Namely, we fix $\ell_1=\ell_2=1$ so that $\Omega=(0,1)^2$, and we select the source term $f\in L^2(\Omega)$ such that the analytical solution $u\in H^1_\sharp(\Omega)$ reads

$$u(\boldsymbol{x}):=\chi(\boldsymbol{x})e^{ik\boldsymbol{d}^{\mathrm{in}}\cdot\boldsymbol{x}}+e^{ik\boldsymbol{d}^{\mathrm{out}}\cdot\boldsymbol{x}},\quad\forall\boldsymbol{x}\in\Omega,$$

with $\boldsymbol{d}^{\mathrm{in}}\cdot\boldsymbol{d}^{\mathrm{in}}=\boldsymbol{d}^{\mathrm{out}}\cdot\boldsymbol{d}^{\mathrm{out}}=1$, $\boldsymbol{d}^{\mathrm{in}}_1=\boldsymbol{d}^{\mathrm{out}}_1=\alpha+m\pi$ for some $m\in N$, $\boldsymbol{d}^{\mathrm{in}}_2\leq 0$ and $\boldsymbol{d}^{\mathrm{out}}_2=-\boldsymbol{d}^{\mathrm{in}}_2$. The cutoff function $\chi\in C^{1,1}(\Omega)$ is defined by

$$\chi(\boldsymbol{x}):=\begin{cases}1 & \text{if}\quad 0\leq\boldsymbol{x}_2\leq\frac{1}{2},\\ 16\left(\boldsymbol{x}_2-\frac{3}{4}\right)^2(8\boldsymbol{x}_2-3) & \text{if}\quad\frac{1}{2}\leq\boldsymbol{x}_2\leq\frac{3}{4},\\ 0 & \text{if}\quad\frac{3}{4}\leq\boldsymbol{x}_2\leq 1.\end{cases}$$

We employ a first-order finite element method to compute an approximate solution $\widetilde{u}_H$ of $\widetilde{u}$.

Similarly to what we have done in subsection 2.2.4, we can select θ in such a way that u is composed of single Fourier mode $\widehat{u}_n$, meaning that it is chosen such that

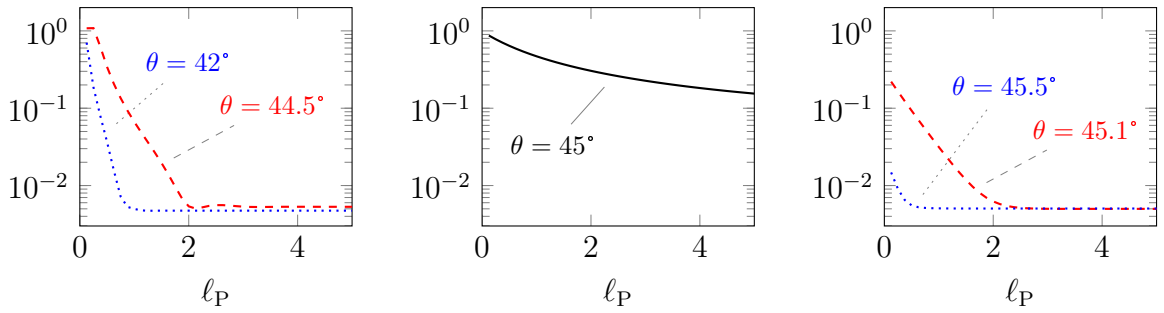
$$k_n^2=k^2-(k\sin(\theta)+2m\pi)^2.$$

We fix $m=1$ and $k=6.8284\pi$, leading to a quasi-resonance for $\theta=45^o$. Figure 3.5 reports the convergence of the PML $\widetilde{u}_H$ towards the original solution u for different values of θ and parameters ℓ_{P}, γ_{i} and γ_{r}. In all cases, we separate the case where (i) $\theta<45^o$ so that $k_n^2>0$, (ii) $\theta=45^o$ so that $k_n^2=0$ and (iii) $\theta>45^o$ so that $k_n^2<0$. In these plots, the y-axis is in log-scale, so that straight lines correspond to exponential convergence.

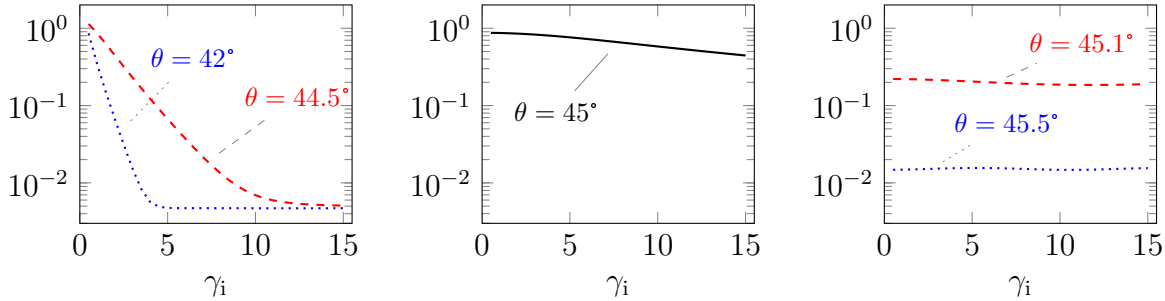
In Figure 3.5a, we observe an exponential convergence rate as ℓ_{P} is increased whenever $k_n^2\neq 0$. Indeed, the curves are straight lines, until they reach a plateau where the finite element error dominates. The slope of this straight lines decreases as k_n^2 approaches zero.

When $k_n^2 = 0$ a linear convergence rate is observed. These results are in agreement with our analysis.

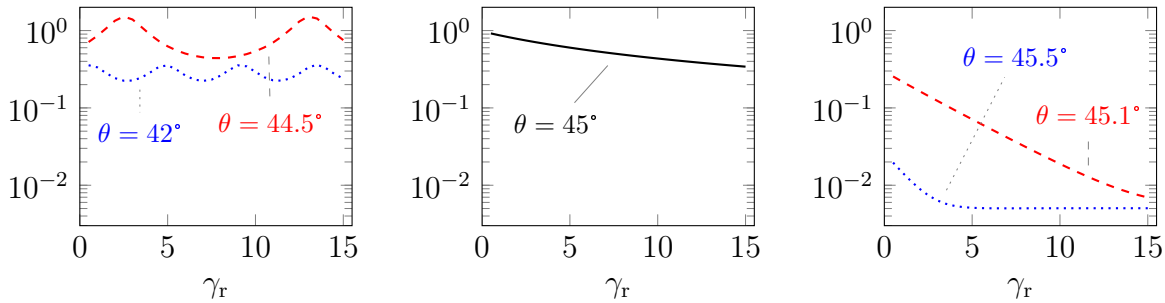
Figures 3.5b and 3.5c also nicely illustrate our analysis. We see a linear convergence rate as γ_i and γ_r as increased for $k_n^2 = 0$. For γ_i, we see exponential convergence if $k_n^2 > 0$ and no convergence for $k_n^2 < 0$. The converse happens for γ_r, where there is no convergence for $k_n^2 > 0$ and exponential convergence when $k_n^2 < 0$.



(a) $\gamma_\mathrm{r} = 1$ and $\gamma_\mathrm{i} = 0.6$.



(b) $\ell_\mathrm{P} := 1/8$ and $\gamma_\mathrm{r} := 1$.



(c) $\ell_\mathrm{P} := 1/8$ and $\gamma_\mathrm{i} := 1$.

Figure 3.5: Convergence of $\|\boldsymbol{\nabla}(u - \widetilde{u}_H)\|_\Omega / \|\boldsymbol{\nabla} u\|_\Omega$ for different PML parameters and incident angles.

3.3 Stability of the PML Helmholtz problem

In the last section, we have shown that the PML problem is well-posed for right-hand sides $f \in L^2(\Omega)$ supported in the original domain, but vanishing in the absorbing layer Ω_{P}. Although this setting does describe the situations that are of "physical interest", duality proofs for the stability and convergence of finite element methods requires the stability of the problem for general right-hand sides $f \in L^2(\widetilde{\Omega})$. More generally, in this section, we address the inf-sup stability of the sesquilinear form $\widetilde{b}$ corresponding to the PML problem set on $\widetilde{\Omega}$.

To the best of our knowledge, the first stability estimate for the PML Helmholtz problem was given in [74] for a one-dimensional homogeneous media. This setting allows the use of a Green function that the authors use to show $\widetilde{\mathscr{C}_{\mathrm{st}}} \sim (k\ell)$, i.e. the same estimate than for the one-dimensional Helmholte problem without the PML. Then, for the two-dimensional homogeneous problem without periodicity assumptions, full-space problem and under certain conditions on the frequency and on the PML coefficients, the estimate $\widetilde{\mathscr{C}_{\mathrm{st}}} \sim (k\ell)^{3/2}$ was proved in [41]. Such estimate half-an-order worst that the corresponding estimates with the exact DtN operator. In fact, it is suboptimal, and it was letter improved to $\widetilde{\mathscr{C}_{\mathrm{st}}} \sim (k\ell)$ for two- and three-dimensional homogeneous media in [98]. The same estimate was then extending to scattering problem by star-shaped obstacles in [31]. Finally, it has been recently be shown that for scattering by a bounded obstacle or heterogeneity, the stability constants of the original and PML problems exhibit the same dependence on $k\ell$, i.e., $\widetilde{\mathscr{C}_{\mathrm{st}}} \sim \mathscr{C}_{\mathrm{st}}$ [61].

Here, we first focus on the homogeneous case, for which we can show an optimal stability estimate with the same frequency dependence as for the DtN problem. This estimate will be proved by following the same approach as the one used for the homogeneous DtN problem. Namely, by analyzing the one-dimensional Helmholtz problems satisfied by the Fourier modes $\widehat{u}_n$ of the solution $\widetilde{u}$. In addition, we will present additional results that will be used in the next chapter, such as the estimation of the L^∞ norms of the Fourier modes derivatives.

In the second subsection, we will show that the PML problem (3.1.3) is well-posed as soon as the corresponding DtN problem is inf-sup stable (this includes in particular heterogeneous media). Unfortunately, we believe that the stability constant $\widetilde{\mathscr{C}_{\mathrm{st}}}$ we obtain for the PML problem is suboptimal, as it increases faster with k than the stability constant $\mathscr{C}_{\mathrm{st}}$ of the corresponding DtN problem. Specifically, $\widetilde{\mathscr{C}_{\mathrm{st}}} \sim (k\ell)^{3/2}\mathscr{C}_{\mathrm{st}}$.

3.3.1 The one-layer case

We start by considering the one-layer case where $\varepsilon \equiv 1$ and $\boldsymbol{A} \equiv \boldsymbol{I}$ in Ω. Similar to section 2.2, we rely on Fourier decomposition and analyze the resulting one-dimensional problem for the Fourier modes.

3.3.1.1 A one dimensional problem

Here, given, $\ell, \ell_{\mathrm{P}} > 0$, $\kappa \in \mathbb{C}$ and $f : (0, \ell+\ell_{\mathrm{P}}) \to \mathbb{C}$, we study the following one-dimensional PML Helmholtz problem with: Find $u : (0, \ell + \ell_{\mathrm{P}}) \to \mathbb{C}$ such that

$$\left\{ \begin{array}{rcll} -\nu\kappa^2 u - (\nu^{-1}u')' & = & f & \text{in } (0, \ell + \ell_{\mathrm{P}}) \\ u(0) & = & 0, & \\ u(\ell + \ell_{\mathrm{P}}) & = & 0, & \end{array} \right. \tag{3.3.1}$$

where $\nu := 1 + i\gamma\chi_{\mathrm{P}}$, and $\chi_{\mathrm{P}}(x) = 1$ if $x \in (\ell, \ell + \ell_{\mathrm{P}})$ and zero otherwise. For the sake of simplicity, we also introduce $\chi_{\mathrm{I}} = 1 - \chi_{\mathrm{P}}$. Notice that for the case of simplicity, we have set $\gamma_{\mathrm{r}} = 1$ and $\gamma_{\mathrm{i}} = \gamma$.

Throughout this subsection, we will employ the notations $\|\cdot\|$ and $(\cdot, \cdot)$ for the natural norm and inner of $L^2(0, \ell + \ell_{\mathrm{P}})$. The additional notations

$$\|v\|_{\mathrm{I}} := \|\chi_{\mathrm{I}} v\| \qquad \|v\|_{\mathrm{P}} := \|\chi_{\mathrm{P}} v\| \qquad \forall v \in L^2(0, \ell + \ell_{\mathrm{P}})$$

and the following Sobolev space

$$W := \left\{ v \in H^1(0, \ell + \ell_{\mathrm{P}}) \mid v(0) = v(\ell + \ell_{\mathrm{P}}) = 0 \right\}.$$

will also be useful.

Assuming that $f \in L^2(0, \ell + \ell_{\mathrm{P}})$, the weak form of (3.3.1) is to find $u \in W$ such that

$$b(u, v) = (f, v) \quad \forall v \in W, \tag{3.3.2}$$

where

$$b(u, v) := -\kappa^2(\nu u, v) + (\nu^{-1}u', v').$$

For later use, we record that

$$|\nu|^2 = 1 + \gamma^2\chi_{\mathrm{P}} \quad \nu^{-1} = \chi_{\mathrm{I}} + \frac{1 - i\gamma}{1 + \gamma^2}\chi_{\mathrm{P}}, \quad \operatorname{Im}\nu = \gamma\chi_{\mathrm{P}}, \quad \operatorname{Im}\nu^{-1} = -\frac{\gamma}{1 + \gamma^2}\chi_{\mathrm{P}}. \tag{3.3.3}$$

Since κ will eventually correspond to a Fourier wave number k_n, we will focus here on the case where $\kappa \geq 0$ or $\kappa = i\beta$, with $\beta \geq 0$. In short, we will write $\kappa \in \mathbb{R}_+ \cup i\mathbb{R}_+$. As in section 2.2, our analysis is divided into three branches where we separately treat imaginary wavenumbers, small real wavenumbers and large real wavenumbers.

To simplify the proofs, for $\zeta > 0$, we introduce the compact notation $\mu_\zeta := 1 + \zeta$. These simple properties will be useful afterwards

$$1 \leq \mu_\zeta \qquad \zeta \leq \mu_\zeta \qquad 1 + \zeta^2 \leq \mu_\zeta^2. \tag{3.3.4}$$

We start by recording a set of inequalities that are simple, but useful.

Lemma 3.3.1 (Poincaré inequalities)**.** *The Poincaré inequalities*

$$\|v\|_{\mathrm{I}} \leq 2\ell\|v'\|_{\mathrm{I}} \qquad \|w\| \leq 2\mu_{\ell_{\mathrm{P}}/\ell}\ell\|w'\| \tag{3.3.5}$$

hold true for all $v \in H^1(0, \ell)$ with $v(0) = 0$ and $w \in H^1(0, \ell + \ell_{\mathrm{P}})$ with $w(0) = 0$.

3.3.1.2 Imaginary wave numbers

We start by addressing imaginary wave numbers, that are easily dealt with since the sesquilinear form b in (3.3.2) is coercive in this case.

Lemma 3.3.2 (Imaginary wave numbers)**.** *Assume that $\kappa \in i\mathbb{R}_+$. Then, for all $f \in L^2(0, \ell + \ell_{\mathrm{P}})$, there exists a unique $u \in W$ solution to (3.3.1). In addition, the estimates*

$$\|u\| \leq \mu_\gamma^2 \mu_{\ell_{\mathrm{P}}/\ell}^2 \min\left(4, (|\kappa|\ell)^{-2}\right) \ell^2 \|f\|, \tag{3.3.6}$$

and

$$\|u\| \leq \mu_\gamma^2 \mu_{\ell_{\mathrm{P}}/\ell}^2 \min\left(4, 2(|\kappa|\ell)^{-1}\right) \ell^2 \|f\|, \tag{3.3.7}$$

hold true.

Proof. We select the test function $v = u$ in (3.3.2) and take the real part. Recalling that $|\nu| \geq 1$, we have

$$\beta^2 \|u\|^2 + |\nu|^{-2} \|u'\|^2 \leq \beta^2 \|u\|^2 + \|u'\|_{\mathrm{I}}^2 + \frac{1}{|\nu|^2} \|u'\|_{\mathrm{P}}^2 = \mathrm{Re}(f, u) \leq \|f\| \|u\|. \tag{3.3.8}$$

Hence, we immediately see that $\beta^2 \|u\| \leq \|f\|$, and since $|\kappa| = \beta$,

$$\|u\| \leq (|\kappa|\ell)^{-2} \ell^2 \|f\|. \tag{3.3.9}$$

On the other hand, using Poincaré inequality (3.3.5), we have

$$\|u\|^2 \leq 4\mu_{\ell_{\mathrm{P}}/\ell}^2 \ell^2 \|u'\|^2 \leq 4\mu_{\ell_{\mathrm{P}}/\ell}^2 |\nu|^2 \ell^2 \|f\| \|u\|,$$

and recalling that $|\nu|^2 = 1 + \gamma^2 \leq \mu_\gamma^2$, we have

$$\|u\| \leq 4\mu_\gamma^2 \mu_{\ell_{\mathrm{P}}/\ell}^2 \ell^2 \|f\|. \tag{3.3.10}$$

Estimate (3.3.6) then follows from (3.3.9) and (3.3.10) since $\mu_\gamma^2 \mu_{\ell_{\mathrm{P}}/\ell}^2 \geq 1$.

To establish (3.3.7), we observe that for $(|\kappa|\ell)^{-1} \leq 2$

$$\min\left(4, (|\kappa|\ell)^{-2}\right) = (|\kappa|\ell)^{-2} \leq 2(|\kappa|\ell)^{-1}.$$

□

3.3.1.3 Small real wave numbers

We then consider small real wave numbers. Again, the stability proof is easy as the sesquilinear form b is coercive due to Poincaré inequality.

Lemma 3.3.3 (Small real wavenumbers)**.** *Assume that*

$$\kappa\ell \leq \frac{1}{\sqrt{8}} \frac{1}{\mu_\gamma \mu_{\ell_{\mathrm{P}}/\ell}}.$$

Then, for all $f \in L^2(0, \ell + \ell_{\mathrm{P}})$, there exists a unique $u \in W$ solution to (3.3.2) and we have

$$\|u\| \leq 8\mu_\gamma^2 \mu_{\ell_{\mathrm{P}}/\ell}^2 \ell^2 \|f\|. \tag{3.3.11}$$

Proof. First, we have

$$\begin{aligned}\operatorname{Re} b(u,u) &= -\kappa^2\|u\|^2 + \|u'\|_{\mathrm{I}}^2 + |\nu|^{-2}\|u'\|_{\mathrm{P}}^2 \\ &\geq \mu_\gamma^{-2}\|u'\|^2 - \kappa^2\|u\|^2 \\ &\geq \left(\mu_\gamma^{-2} - 4\kappa^2\mu_{\ell_{\mathrm{P}}/\ell}\ell^2\right)\|u'\|^2 \\ &= \mu_\gamma^{-2}\left(1 - 4\mu_\gamma^2\mu_{\ell_{\mathrm{P}}/\ell}^2(\kappa\ell)^2\right)\|u'\|^2.\end{aligned}$$

Assuming that $4\mu_\gamma^2\mu_{\ell_{\mathrm{P}}/\ell}^2(\kappa\ell)^2 \leq 1/2$, we thus have

$$\frac{\mu_\gamma^{-2}}{2}\|u'\|^2 \leq \operatorname{Re} b(u,u) \leq \|f\|\|u\|,$$

and it follows that

$$\|u\|^2 \leq 4\mu_{\ell_{\mathrm{P}}/\ell}^2\ell^2\|u'\|^2 \leq 8\mu_\gamma^2\mu_{\ell_{\mathrm{P}}/\ell}^2\ell^2\|f\|\|u\|,$$

and the result follows. □

3.3.1.4 Large real wave numbers

We finally consider large real wave numbers $\kappa \in \mathbb{R}_+$, which is the more subtle case. Similar to the proof with a Robin boundary condition $u'(\ell) = i\kappa u(\ell)$, our analysis will rely on a Morawetz identity. The key difference is the treatment of the dissipative terms. Indeed, considering a Robin boundary condition, the identity

$$\operatorname{Im} b(u,u) = \kappa|u(\ell)|^2 = |u'(\ell)|^2$$

immediately provides control on $u(\ell)$ and $u'(\ell)$ on the boundary. In the PML approach, the situation is more complicated, and the next two Lemmas are dedicated to obtain a similar bound. This is done by first controlling volumic terms in the absorbing layer in Lemma 3.3.4 and then control $u(\ell)$ and $u'(\ell)$ with a trace inequality from the absorbing layer.

Lemma 3.3.4 (Volume estimate in the PML)**.** *Let $\kappa \in \mathbb{R}_+$ and $f \in L^2(0, \ell + \ell_{\mathrm{P}})$, and assume that $u \in W$ solves (3.3.2). Then, we have*

$$\kappa^2\|u\|_{\mathrm{P}}^2 + \|u'\|_{\mathrm{P}}^2 \leq \frac{\mu_\gamma^3}{\gamma^2}\left(2\|f\|_{\mathrm{I}}\|u\|_{\mathrm{I}} + \frac{1}{\kappa^2}\|f\|_{\mathrm{P}}^2.\right) \tag{3.3.12}$$

Proof. First, we pick the test function $v = u$ in (3.3.2). Recalling (3.3.3) and taking the imaginary part, we have

$$\kappa^2\gamma\|u\|_{\mathrm{P}}^2 + \frac{\gamma}{1+\gamma^2}\|u'\|_{\mathrm{P}}^2 = -\operatorname{Im} b(u,u) = -\operatorname{Im}(f,u).$$

On the other hand, the following estimate holds true

$$|(f,u)| \leq \|f\|_{\mathrm{I}}\|u\|_{\mathrm{I}} + \|f\|_{\mathrm{P}}\|u\|_{\mathrm{P}} \leq \|f\|_{\mathrm{I}}\|u\|_{\mathrm{I}} + \frac{1}{2\gamma\kappa^2}\|f\|_{\mathrm{P}}^2 + \frac{\gamma\kappa^2}{2}\|u\|_{\mathrm{P}}^2,$$

leading to

$$\frac{\kappa^2\gamma}{2}\|u\|_{\mathrm{P}}^2+\frac{\gamma}{1+\gamma^2}\|u'\|_{\mathrm{P}}^2\le\|f\|_{\mathrm{I}}\|u\|_{\mathrm{I}}+\frac{1}{2\gamma\kappa^2}\|f\|_{\mathrm{P}}^2.$$

We then obtain (3.3.12) by observing that

$$\frac{1}{2}\frac{\gamma}{\mu_\gamma^2}\left(\kappa^2\|u\|_{\mathrm{P}}^2+\|u'\|_{\mathrm{P}}^2\right)\le\frac{\kappa^2\gamma}{2}\|u\|_{\mathrm{P}}^2+\frac{\gamma}{1+\gamma^2}\|u'\|_{\mathrm{P}}^2$$

and

$$\|f\|_{\mathrm{I}}\|u\|_{\mathrm{I}}+\frac{1}{2\gamma\kappa^2}\|f\|_{\mathrm{P}}^2\le\frac{1}{2}\frac{\mu_\gamma}{\gamma}\left(2\|f\|_{\mathrm{I}}\|u\|_{\mathrm{I}}+\frac{1}{\kappa^2}\|f\|_{\mathrm{P}}^2\right).$$

□

Lemma 3.3.5 (Trace estimate)**.** *We have*

$$\kappa^2\ell|u(\ell)|^2+\ell|u'(\ell)|^2\le 9\frac{\mu_{\ell/\ell_{\mathrm{P}}}^2\mu_\gamma^{10}}{\gamma^4}(1+(\kappa\ell)^{-1})^2\ell^2\|f\|^2+\frac{\kappa^2}{2}\|u\|_{\mathrm{I}}^2. \tag{3.3.13}$$

Proof. We start by presenting the following multiplicative trace inequality

$$\ell|w(\ell)|^2\le\mu_{\ell/\ell_{\mathrm{P}}}\left(\|w\|_{\mathrm{P}}^2+2\ell\|w\|_{\mathrm{P}}\|w'\|_{\mathrm{P}}\right), \tag{3.3.14}$$

valid for all $w\in H^1(\ell,\ell+\ell_{\mathrm{P}})$. To establish it we write that

$$\begin{aligned}\ell_{\mathrm{P}}|w(\ell)|^2&=\left[(x-\ell-\ell_{\mathrm{P}})|w(x)|^2\right]_\ell^{\ell+\ell_{\mathrm{P}}}\\&=\int_\ell^{\ell+\ell_{\mathrm{P}}}|w(x)|^2dx+2\,\mathrm{Re}\int_\ell^{\ell+\ell_{\mathrm{P}}}(x-\ell-\ell_{\mathrm{P}})w(x)\overline{w'(x)}dx\\&\le\|w\|_{\mathrm{P}}+2\ell_{\mathrm{P}}\|w\|_{\mathrm{P}}\|w'\|_{\mathrm{P}},\end{aligned}$$

and multiply both sides by ℓ/ℓ_{P} to obtain

$$\ell|w(\ell)|^2\le\frac{\ell}{\ell_{\mathrm{P}}}\|w\|_{\mathrm{P}}^2+2\ell\|w\|_{\mathrm{P}}\|w'\|_{\mathrm{P}}.$$

Estimate (3.3.14) now follows follows since $\ell/\ell_{\mathrm{P}}\le\mu_{\ell/\ell_{\mathrm{P}}}$ and $1\le\mu_{\ell/\ell_{\mathrm{P}}}$.

Since $f\in L^2(\ell,\ell+\ell_{\mathrm{P}})$, by elliptic regularity, $u\in H^2(\ell,\ell+\ell_{\mathrm{P}})$, and applying (3.3.14) to both u and u', we have

$$\kappa^2\ell|u(\ell)|^2+\ell|u'(\ell)|^2\le\mu_{\ell/\ell_{\mathrm{P}}}\left(\kappa^2\|u\|_{\mathrm{P}}^2+2\ell\kappa^2\|u\|_{\mathrm{P}}\|u'\|_{\mathrm{P}}+\|u'\|_{\mathrm{P}}^2+2\ell\|u'\|_{\mathrm{P}}\|u''\|_{\mathrm{P}}\right). \tag{3.3.15}$$

Next, we want to remove the second derivative in the right-hand side of (3.3.15). To do so, we observe that $-u''=\nu f+\nu^2\kappa^2u$ in $(\ell,\ell+\ell_{\mathrm{P}})$, so that

$$\|u''\|_{\mathrm{P}}\le\mu_\gamma\|f\|_{\mathrm{P}}+\mu_\gamma^2\kappa^2\|u\|_{\mathrm{P}}^2,$$

and

$$\begin{aligned}
&\kappa^2\ell|u(\ell)|^2+\ell|u'(\ell)|^2\\
&\leq \mu_{\ell/\ell_{\mathrm{P}}}\left(\kappa^2\|u\|_{\mathrm{P}}^2+\|u'\|_{\mathrm{P}}^2+2\ell\kappa^2\|u\|_{\mathrm{P}}\|u'\|_{\mathrm{P}}+2\ell\mu_\gamma\|u'\|_{\mathrm{P}}\|f\|_{\mathrm{P}}+2\ell\mu_\gamma^2\kappa^2\|u'\|_{\mathrm{P}}\|u\|_{\mathrm{P}}\right)\\
&\leq \mu_{\ell/\ell_{\mathrm{P}}}\mu_\gamma^2\left(\kappa^2\|u\|_{\mathrm{P}}^2+\|u'\|_{\mathrm{P}}^2+4\ell\kappa^2\|u\|_{\mathrm{P}}\|u'\|_{\mathrm{P}}+2\ell\|u'\|_{\mathrm{P}}\|f\|_{\mathrm{P}}\right).
\end{aligned}$$

By employing the Young's inequalities

$$4\ell\kappa^2\|u\|_{\mathrm{P}}\|u'\|_{\mathrm{P}}\leq 2\kappa\ell\left(\kappa^2\|u\|_{\mathrm{P}}^2+\|u'\|_{\mathrm{P}}^2\right),\qquad 2\ell\|u'\|_{\mathrm{P}}\|f\|_{\mathrm{P}}\leq \ell^2\|f\|_{\mathrm{P}}^2+\|u'\|_{\mathrm{P}}^2,$$

we can further simplify the right-hand side to

$$\kappa^2\ell|u(\ell)|^2+\ell|u'(\ell)|^2\leq 2\mu_{\ell/\ell_{\mathrm{P}}}\mu_\gamma^2\left((1+\kappa\ell)\left(\kappa^2\|u\|_{\mathrm{P}}^2+\|u'\|_{\mathrm{P}}^2\right)+\ell^2\|f\|_{\mathrm{P}}^2\right). \tag{3.3.16}$$

The final step of the proof consists in using the volume estimate (3.3.12) in the absorbing layer from Lemma 3.3.4 in the right-hand side of (3.3.16). We proceed as follows:

$$\begin{aligned}
(1+\kappa\ell)\left(\kappa^2\|u\|_{\mathrm{P}}^2+\|u'\|_{\mathrm{P}}^2\right)&\leq \frac{\mu_\gamma^3}{\gamma^2}(1+\kappa\ell)\left(2\|f\|_{\mathrm{I}}\|u\|_{\mathrm{I}}+\frac{1}{\kappa^2}\|f\|_{\mathrm{P}}^2\right)\\
&\leq 2\frac{\mu_\gamma^3}{\gamma^2}(1+\kappa\ell)\|f\|_{\mathrm{I}}\|u\|_{\mathrm{I}}+\frac{\mu_\gamma^3}{\gamma^2}\frac{1+\kappa\ell}{(\kappa\ell)^2}\ell^2\|f\|_{\mathrm{P}}^2
\end{aligned}$$

leading to

$$\begin{aligned}
\kappa^2\ell|u(\ell)|^2+\ell|u'(\ell)|^2&\leq 2\mu_{\ell/\ell_{\mathrm{P}}}\mu_\gamma^2\left(2\frac{\mu_\gamma^3}{\gamma^2}(1+\kappa\ell)\|f\|_{\mathrm{I}}\|u\|_{\mathrm{I}}+\frac{\mu_\gamma^3}{\gamma^2}\frac{1+\kappa\ell+(\kappa\ell)^2}{(\kappa\ell)^2}\ell^2\|f\|_{\mathrm{P}}^2\right)\\
&\leq 4\frac{\mu_{\ell/\ell_{\mathrm{P}}}\mu_\gamma^5}{\gamma^2}\left((1+\kappa\ell)\|f\|_{\mathrm{I}}\|u\|_{\mathrm{I}}+\frac{1+\kappa\ell+(\kappa\ell)^2}{(\kappa\ell)^2}\ell^2\|f\|_{\mathrm{P}}^2\right).
\end{aligned}$$

Finally, we algebraically simplify the right-hand side to make it easier to read and manipulate. On the one hand, we have

$$\frac{1+\kappa\ell+(\kappa\ell)^2}{(\kappa\ell)^2}\leq\frac{(1+(\kappa\ell))^2}{(\kappa\ell)^2}=(1+(\kappa\ell)^{-1})^2,$$

and on the other hand, we have

$$\begin{aligned}
4\frac{\mu_{\ell/\ell_{\mathrm{P}}}\mu_\gamma^5}{\gamma^2}(1+\kappa\ell)\|f\|_{\mathrm{I}}\|u\|_{\mathrm{I}}&\leq\frac{16}{2\kappa^2}\frac{\mu_{\ell/\ell_{\mathrm{P}}}^2\mu_\gamma^{10}}{\gamma^4}(1+\kappa\ell)^2\|f\|_{\mathrm{I}}^2+\frac{\kappa^2}{2}\|u\|_{\mathrm{I}}^2\\
&=8\frac{\mu_{\ell/\ell_{\mathrm{P}}}^2\mu_\gamma^{10}}{\gamma^4}(1+(\kappa\ell)^{-1})^2\ell^2\|f\|_{\mathrm{I}}^2+\frac{\kappa^2}{2}\|u\|_{\mathrm{I}}^2.
\end{aligned}$$

which leads to

$$\kappa^2\ell|u(\ell)|^2+\ell|u'(\ell)|^2\leq 9\frac{\mu_{\ell/\ell_{\mathrm{P}}}^2\mu_\gamma^{10}}{\gamma^4}(1+(\kappa\ell)^{-1})^2\ell^2\|f\|^2+\frac{\kappa^2}{2}\|u\|_{\mathrm{I}}^2.$$

□

With the boundary estimate (3.3.13) at our disposal, we are finally ready to state our stability results for large real wave numbers using a Morawetz multplier technique.

Lemma 3.3.6 (Large real wave numbers)**.** *Let $\kappa \in \mathbb{R}_+$. For all $f \in L^2(0, \ell + \ell_{\mathrm{P}})$ there exists a unique $u \in W$ solution to (3.3.2) and we have*

$$\|u\| \leq \sqrt{30}\frac{\mu_{\ell/\ell_{\mathrm{P}}}\mu_\gamma^5}{\gamma^2}(1 + (\kappa\ell)^{-1})\frac{1}{\kappa\ell}\ell^2\|f\|. \tag{3.3.17}$$

Proof. We first observe that the following Morawetz identity

$$\kappa^2\|w\|_{\mathrm{I}}^2 + \|w'\|_{\mathrm{I}}^2 = \kappa^2\ell|w(\ell)|^2 + \ell|w'(\ell)|^2 + 2\,\mathrm{Re}\int_0^\ell(-\kappa^2 w - w'')x\overline{w}', \tag{3.3.18}$$

may be easily obtained for $w \in H^2(0, \ell)$ by integration by parts.

We then assume that $u \in W$ solves (3.3.2), and apply (3.3.18) to u, showing that

$$\kappa^2\|u\|_{\mathrm{I}}^2 + \|u'\|_{\mathrm{I}}^2 \leq \kappa^2\ell|u(\ell)|^2 + \ell|u'(\ell)|^2 + 2\ell\|f\|_{\mathrm{I}}\|u'\|_{\mathrm{I}}.$$

Incorporating (3.3.13) and

$$2\ell\|f\|_{\mathrm{I}}\|u'\|_{\mathrm{I}}^2 \leq \ell^2\|f\|_{\mathrm{I}}^2 + \|u'\|_{\mathrm{I}}^2,$$

gives

$$\frac{\kappa^2}{2}\|u\|_{\mathrm{I}}^2 \leq 9\frac{\mu_{\ell/\ell_{\mathrm{P}}}^2\mu_\gamma^{10}}{\gamma^4}(1 + (\kappa\ell)^{-1})^2\ell^2\|f\|^2 + \ell^2\|f\|_{\mathrm{I}}^2,$$

leading to

$$\kappa^2\|u\|_{\mathrm{I}}^2 \leq 20\frac{\mu_{\ell/\ell_{\mathrm{P}}}^2\mu_\gamma^{10}}{\gamma^4}(1 + (\kappa\ell)^{-1})^2\ell^2\|f\|^2,$$

and

$$\|u\|_{\mathrm{I}} \leq 2\sqrt{5}\frac{\mu_{\ell/\ell_{\mathrm{P}}}\mu_\gamma^5}{\gamma^2}(1 + (\kappa\ell)^{-1})\frac{1}{\kappa\ell}\ell^2\|f\|. \tag{3.3.19}$$

At this point, we have control over the "physical" domain $(0, \ell)$. We now need to extend this control into the absorbing layer $(\ell, \ell + \ell_{\mathrm{P}})$. This can be done by recalling the estimate (3.3.12) for $\|u\|_{\mathrm{P}}$ in terms of $\|f\|_{\mathrm{P}}$ and $\|u\|_{\mathrm{I}}$ in Lemma 3.3.4. Specifically, plugging (3.3.12) into (3.3.19), we have

$$\kappa^2\|u\|_{\mathrm{P}}^2 \leq \frac{\mu_\gamma^3}{\gamma^2}\left(4\sqrt{5}\frac{\mu_{\ell/\ell_{\mathrm{P}}}\mu_\gamma^5}{\gamma^2}(1 + (\kappa\ell)^{-1})\frac{1}{\kappa\ell}\ell^2\|f\|^2 + \frac{1}{\kappa^2}\|f\|_{\mathrm{P}}^2\right),$$

and the conclusion follows by algebrically simplifying the right-hand side. Indeed, we have

$$\frac{1}{\kappa^2} \leq \frac{1 + \kappa^2\ell^2}{\kappa^2} \leq (1 + (\kappa\ell)^{-1})^2\ell^2,$$

and

$$\frac{1}{\kappa\ell} \leq \frac{1+\kappa\ell}{\kappa\ell} = 1 + (\kappa\ell)^{-1},$$

leading to

$$\begin{aligned}\kappa^2\|u\|_{\mathrm{P}}^2 &\leq \frac{\mu_\gamma^3}{\gamma^2} 10 \frac{\mu_{\ell/\ell_{\mathrm{P}}}\mu_\gamma^5}{\gamma^2}(1+(\kappa\ell)^{-1})^2\ell^2\|f\|^2 \\ &\leq 10\frac{\mu_{\ell/\ell_{\mathrm{P}}}^2\mu_\gamma^{10}}{\gamma^4}(1+(\kappa\ell)^{-1})^2\ell^2\|f\|^2. \end{aligned} \tag{3.3.20}$$

Then, (3.3.17) follows by adding (3.3.19) and (3.3.20).

So far, we have only established (3.3.17) under the assumption that u solves (3.3.2) in the first place. However, due to Fredholm alternative, a priori stability implies existence, which shows the existence and uniqueness of u. □

3.3.1.5 Stability estimates for the one-dimensional PML problem

We can finally summarize our findings for the one-dimensional PML problem by regrouping the three different branches.

Theorem 3.3.7 (One-dimensional PML problem)**.** *For all $\kappa \in \mathbb{R}_+ \cup i\mathbb{R}_+$, we have*

$$\|u\| \leq 61\frac{(1+\gamma)^7}{\gamma^2}\left(1+\frac{\ell}{\ell_{\mathrm{P}}}\right)\left(1+\frac{\ell_{\mathrm{P}}}{\ell}\right)^2 \min\left(1,(|\kappa|\ell)^{-1}\right)\ell^2\|f\|. \tag{3.3.21}$$

Proof. We first focus on real wave numbers $\kappa \geq 0$. To do, let us set

$$\Theta := \frac{1}{\sqrt{8}}\frac{1}{\mu_\gamma\mu_{\ell_{\mathrm{P}}/\ell}} < 1.$$

and observe that for $|\kappa|\ell \geq \Theta$, we have

$$(1+(\kappa\ell)^{-1})\frac{1}{\kappa\ell} = \sqrt{8}\mu_\gamma\mu_{\ell_{\mathrm{P}}/\ell}(1+\sqrt{8}\mu_\gamma\mu_{\ell_{\mathrm{P}}/\ell}) \leq 11\mu_\gamma^2\mu_{\ell_{\mathrm{P}}/\ell}^2, \tag{3.3.22}$$

whereas if $|\kappa|\ell \leq \Theta$, the estimate

$$(1+(\kappa\ell)^{-1})\frac{1}{\kappa\ell} \leq \frac{2\sqrt{8}\mu_\gamma\mu_{\ell_{\mathrm{P}}/\ell}}{\kappa\ell} \tag{3.3.23}$$

holds true. Assuming that $\Theta \leq \kappa\ell \leq 1$, estimate (3.3.17) together with (3.3.22) imply that

$$\begin{aligned}\|u\| \leq \sqrt{30}\frac{\mu_{\ell/\ell_{\mathrm{P}}}\mu_\gamma^5}{\gamma^2}(1+(\kappa\ell)^{-1})\frac{1}{\kappa\ell}\ell^2\|f\| &\leq 11\sqrt{30}\frac{\mu_{\ell/\ell_{\mathrm{P}}}\mu_{\ell_{\mathrm{P}}/\ell}^2\mu_\gamma^7}{\gamma^2}\ell^2\|f\| \\ &\leq 61\frac{\mu_{\ell/\ell_{\mathrm{P}}}\mu_{\ell_{\mathrm{P}}/\ell}^2\mu_\gamma^7}{\gamma^2}\ell^2\|f\| \leq 61\frac{\mu_{\ell/\ell_{\mathrm{P}}}\mu_{\ell_{\mathrm{P}}/\ell}^2\mu_\gamma^7}{\gamma^2}\min\left(1,(|\kappa|\ell)^{-1}\right)\ell^2\|f\|.\end{aligned}$$

Otherwise, if $\kappa\ell \le \Theta$, estimate (3.3.11) ensures that

$$\|u\| \le 8\mu_\gamma^2\mu_{\ell_{\mathrm{P}}/\ell}^2\ell^2\|f\| \le 61\frac{\mu_{\ell/\ell_{\mathrm{P}}}\mu_{\ell_{\mathrm{P}}/\ell}^2\mu_\gamma^7}{\gamma^2}\ell^2\|f\| \le 61\frac{\mu_{\ell/\ell_{\mathrm{P}}}\mu_{\ell_{\mathrm{P}}/\ell}^2\mu_\gamma^7}{\gamma^2}\min\left(1,(|\kappa|\ell)^{-1}\right)\ell^2\|f\|.$$

Assuming now that $\kappa\ell \ge 1$, (3.3.17) and (3.3.23) yield

$$\begin{aligned}\|u\| &\le \sqrt{30}\frac{\mu_{\ell/\ell_{\mathrm{P}}}\mu_\gamma^5}{\gamma^2}(1+(\kappa\ell)^{-1})\frac{1}{\kappa\ell}\ell^2\|f\| \le 2\sqrt{224}\frac{\mu_{\ell/\ell_{\mathrm{P}}}\mu_{\ell_{\mathrm{P}}/\ell}\mu_\gamma^6}{\gamma^2}(\kappa\ell)^{-1}\ell^2\|f\| \\ &\le 61\frac{\mu_{\ell/\ell_{\mathrm{P}}}\mu_{\ell_{\mathrm{P}}/\ell}^2\mu_\gamma^7}{\gamma^2}(\kappa\ell)^{-1}\ell^2\|f\| \le 61\frac{\mu_{\ell/\ell_{\mathrm{P}}}\mu_{\ell_{\mathrm{P}}/\ell}^2\mu_\gamma^7}{\gamma^2}\min\left(1,(|\kappa|\ell)^{-1}\right)\ell^2\|f\|.\end{aligned}$$

This shows (3.3.7) for the case of real wave numbers.

On the other hand, for imaginary wave numbers, (3.3.7) gives

$$\begin{aligned}\|u\| &\le \mu_\gamma^2\mu_{\ell_{\mathrm{P}}/\ell}^2\min\left(4,2(|\kappa|\ell)^{-1}\right)\ell^2\|f\| \le 4\mu_\gamma^2\mu_{\ell_{\mathrm{P}}/\ell}^2\min\left(1,(|\kappa|\ell)^{-1}\right)\ell^2\|f\| \\ &\le 61\frac{\mu_{\ell/\ell_{\mathrm{P}}}\mu_{\ell_{\mathrm{P}}/\ell}^2\mu_\gamma^7}{\gamma^2}\min\left(1,(|\kappa|\ell)^{-1}\right)\ell^2\|f\|,\end{aligned}$$

which concludes the proof. □

3.3.1.6 Stability estimates for the PML problem in the one-layer case

We are now ready to conclude the study of the one-layer case. As mentioned, using the quasi-periodic boundary conditions, this will be done using the Fourier expansion technique presented in Subsection 2.1.6. For the reader's convenience, we recall that since $\widetilde{u} \in H^1_\sharp(\widetilde{\Omega})$ we may express u as

$$\widetilde{u}(\boldsymbol{x}_1,\boldsymbol{x}_2) = \sum_{n\in\mathbb{Z}} \widehat{u}_n(\boldsymbol{x}_2)e^{i(\alpha+\alpha_n)\boldsymbol{x}_1},$$

with $k_n \in \mathbb{R}_+ \cup i\mathbb{R}_+$ defined in (2.1.10) as $k_n^2 := k^2 - (\alpha+\alpha_n)^2$. Substituting this Fourier expansion in the PML Helmholtz problem (3.1.3), we find that for each $n \in \mathbb{Z}$, the Fourier mode $\widehat{u}_n$ satisfies the following one-dimensional PML problem

$$\left\{\begin{array}{rcll} -\nu k_n{}^2\widehat{u}_n - (\nu^{-1}\widehat{u}_n')' &=& \widehat{f}_n & \text{in } (0,\ell_2+\ell_{\mathrm{P}}),\\ \widehat{u}_n(0) &=& 0, & \\ \widehat{u}_n'(\ell_2+\ell_{\mathrm{P}}) &=& 0, & \end{array}\right.$$

which is exactly (3.3.1) studied the above with $\kappa = k_n$.

Theorem 3.3.8 (Stability of PML problem in homogeneous media)**.** *Assume that $\varepsilon \equiv 1$ and $\boldsymbol{A} \equiv \boldsymbol{I}$. Then the PML problem (3.1.3) is well-posed, and we have*

$$\widetilde{\mathscr{C}_{\mathrm{st}}} \le 61\frac{(1+\gamma)^7}{\gamma^2}\left(1+\frac{\ell_2}{\ell_{\mathrm{P}}}\right)\left(1+\frac{\ell_{\mathrm{P}}}{\ell_2}\right)^2\min\left(1,\frac{1}{k_\star\ell_2}\right)(k\ell_2)^2. \tag{3.3.24}$$

In other words, for all $f \in L^2(\widetilde{\Omega})$, the solution $\widetilde{u} \in H^1_\sharp(\widetilde{\Omega})$ satisfies

$$k\|\widetilde{u}\|_{\widetilde{\Omega}} \leq 61\frac{(1+\gamma)^7}{\gamma^2}\left(1+\frac{\ell_2}{\ell_{\mathrm{P}}}\right)\left(1+\frac{\ell_{\mathrm{P}}}{\ell_2}\right)^2 \min\left(1, \frac{1}{k_\star \ell_2}\right)(k\ell_2)^2 \frac{1}{k}\|\widetilde{f}\|_{\widetilde{\Omega}}. \tag{3.3.25}$$

Proof. Using Fourier expansion technique introduced in section 2.1.6, we have

$$\|\widetilde{u}\|^2_{\widetilde{\Omega}} = \ell_1 \sum_{n\in\mathbb{Z}} \|\widehat{u}_n\|^2,$$

where for each $n \in \mathbb{Z}$, $\widehat{u}_n$ is the only element of $H^1_0(0, \ell+\ell_{\mathrm{P}})$ such that

$$-k_n^2(\nu\widehat{u}_n, \widehat{v}) + (\nu^{-1}\widehat{u}_n', \widehat{v}') = (\widehat{f}_n, \widehat{v}) \quad \forall \widehat{v} \in H^1_0(0, \ell+\ell_{\mathrm{P}}).$$

The one-dimensional stability result (3.3.21) from Theorem 3.3.7 then shows that

$$\begin{aligned}
\|u\|^2_{\widetilde{\Omega}} &\leq \ell_1 \sum_{n\in\mathbb{Z}} \left\{61\frac{(1+\gamma)^7}{\gamma^2}\left(1+\frac{\ell_2}{\ell_{\mathrm{P}}}\right)\left(1+\frac{\ell_{\mathrm{P}}}{\ell_2}\right)^2 \min\left(1, \frac{1}{|k_n|\ell_2}\right)\ell_2^2\right\}^2 \|\widehat{f}_n\|^2 \\
&\leq \left\{61\frac{(1+\gamma)^7}{\gamma^2}\left(1+\frac{\ell_2}{\ell_{\mathrm{P}}}\right)\left(1+\frac{\ell_{\mathrm{P}}}{\ell_2}\right)^2 \min\left(1, \frac{1}{k_\star\ell_2}\right)\ell_2^2\right\}^2 \ell_1 \sum_{n\in\mathbb{Z}} \|\widehat{f}_n\|^2 \\
&= \left\{61\frac{(1+\gamma)^7}{\gamma^2}\left(1+\frac{\ell_2}{\ell_{\mathrm{P}}}\right)\left(1+\frac{\ell_{\mathrm{P}}}{\ell_2}\right)^2 \min\left(1, \frac{1}{k_\star\ell_2}\right)\ell_2^2\right\}^2 \|f\|^2_{\widetilde{\Omega}}.
\end{aligned}$$

□

Interestingly, the stability constant in (3.3.24) is equivalent to the one found for the homogeneous DtN problem (2.2.24), up to a multiplicative constant depending on the PML parameters. In particular, both constants exhibit the same behaviour with respect to the frequency.

3.3.1.7 Improved estimates for the vertical derivative

In section 5.3, we will provide sharp stability properties of the MHM method in homogeneous medium with quasi-periodic boundary conditions. The analysis is subtle, and requires finer stability estimates that we derive here. For the sake of simplicity, we do not track the dependence of the constant on the PML parameters here, and instead, we let C denote a generic constant that may depend on γ and ℓ_2/ℓ_P, but that is independent of k.

Our proofs again rely on Fourier expansion and the associated one-dimensional PML problem (3.3.2). Throughout this subsection, we thus fix a right-hand side $f \in L^2(0, \ell+\ell_{\mathrm{P}})$ and denote by u the (unique) associated solution.

In this subsection, if $v : (0, \ell+\ell_{\mathrm{P}}) \to \mathbb{C}$ is a measurable function, we employ the notation

$$\|v\|_\infty := \inf\{0 \leq M \leq +\infty \mid \lambda(\{|v(x)| \leq M\}) = 0\},$$

for the usual $L^\infty(0, \ell+\ell_{\mathrm{P}})$ norm, where λ denotes the Lebesgue measure.

We start by analyzing imaginary wave numbers.

Lemma 3.3.9 (Imaginary wave numbers)**.** *Assume that* $\kappa \in i\mathbb{R}_+$*, we have*

$$\|u'\| \leq C\ell\|f\|, \tag{3.3.26}$$

and

$$\|u'\|_\infty \leq C\sqrt{\ell}\|f\|. \tag{3.3.27}$$

Proof. Recalling (3.3.8) and using Poincaré inequality (3.3.5), we have

$$\|u'\| \leq 2|\nu|^2(\ell_{\mathrm{P}} + \ell)\|f\| \leq 2\mu_\gamma^2 \mu_{\ell_{\mathrm{P}}/\ell}\ell|f\| = \ell C\|f\|. \tag{3.3.28}$$

Furthermore, since $-u'' = f - |\kappa|^2 u$ in $(0, \ell)$ and $-u'' = \nu f - \nu^2|\kappa|^2 u$ in $(\ell, \ell + \ell_{\mathrm{P}})$, we have

$$\|u''\| \leq |\nu|\|f\| + |\nu|^2|\kappa|^2\|u\| = C\left(\|f\| + |\kappa|^2\|u\|\right) \leq C\|f\|.$$

On the other hand, we have

$$\|u'\|_{L^\infty}^2 \leq \frac{1}{\ell_{\mathrm{P}} + \ell}\|u'\|^2 + 2\|u'\|\|u''\| = C\left(\ell^{-1}\|u'\|^2 + \ell\|u''\|^2\right),$$

and we conclude with the bounds on $\|u'\|$ and $\|u''\|$ obtained above. □

The next step is to consider small real wave numbers.

Lemma 3.3.10 (Small real wave numbers)**.** *Assume that* $0 \leq \kappa \leq 2/\ell$*, we have*

$$\|u'\| \leq C\ell\|f\|, \tag{3.3.29}$$

and

$$\|u'\|_\infty \leq C\sqrt{\ell}\|f\|. \tag{3.3.30}$$

Proof. Suppose that $0 \leq \kappa \leq 2\ell^{-1}$, then

$$\begin{aligned}|\nu|^{-2}\|u'\|^2 = \operatorname{Re} b(u,u) + \kappa^2\|u\| &\leq \|f\|\|u\| + \kappa^2\|u\|^2 \\ &\leq \|f\|\|u\| + \ell^{-2}\|u\|^2 \leq C\left(\ell^2\|f\|^2 + \ell^{-2}\|u\|^2\right),\end{aligned}$$

and recalling from (3.3.11) that

$$\|u\| \leq C\ell^2\|f\|,$$

we obtain

$$\|u'\| \leq C\ell\|f\|.$$

Since $-u'' = f + \kappa^2 u$ in $(0, \ell)$ and $-\nu^{-1}u'' = f + \nu\kappa^2 u$ in $(\ell, \ell + \ell_{\mathrm{P}})$, we conclude that

$$\|u''\| \leq C\|f\|,$$

from the smallness assumption on κ. Then, (3.3.30) follows from the multiplicative trace inequality

$$\|u'\|_\infty^2 \leq C\left(\ell^{-1}\|u'\|^2 + \ell\|u''\|^2\right),$$

we already employed in the proof of Lemma 3.3.9. □

The last key argument is the case of large real wave numbers.

Lemma 3.3.11 (Large real wave numbers)**.** *Assume that $\kappa \geq 2/\ell$, we have*

$$\|u'\| \leq C\ell|f\|, \tag{3.3.31}$$

and

$$\|u'\|_\infty \leq C\sqrt{\ell}\|f\|. \tag{3.3.32}$$

Proof. Since $-\kappa^2 u - u'' = f$ in $(0,\ell)$, with $\kappa > 0$, we can write that

$$u(x) = c_- e^{-i\kappa x} + c_+ e^{i\kappa x} + \frac{i}{2\kappa}\int_0^\ell f(\xi)e^{i\kappa|x-\xi|}d\xi. \quad \forall x \in (0,\ell), \tag{3.3.33}$$

for two constants $c_\pm \in \mathbb{C}$. Since here $(|\kappa|\ell)^{-1} \geq C$ by assumption, the $L^2(0,\ell+\ell_{\mathrm{P}})$ stability bound (3.3.21) we derived in Theorem 3.3.21 simplifies into

$$\|u\| \leq C\frac{\ell}{\kappa}\|f\|. \tag{3.3.34}$$

Besides, the Hölder inequality gives that

$$\left|\frac{i}{2\kappa}\int_0^\ell f(\xi)e^{i\kappa|x-\xi|}d\xi\right| \leq \frac{1}{2\kappa}\int_0^\ell |f(\xi)|d\xi \leq \frac{\sqrt{\ell}}{2\kappa}\|f\|.$$

As a result, we can write that

$$\begin{aligned}
\left\|c_- e^{-i\kappa x} + c_+ e^{i\kappa x}\right\|_I &\leq \|u\| + \left\|\frac{1}{2\kappa}\int_0^\ell f(\xi)e^{i\kappa|x-\xi|}d\xi\right\|_I \\
&\leq \|u\| + \sqrt{\ell}\left|\frac{1}{2\kappa}\int_0^\ell f(\xi)e^{i\kappa|x-\xi|}d\xi\right| \\
&\leq C\frac{\ell}{\kappa}\|f\|.
\end{aligned}$$

But on the other, we also have

$$\begin{aligned}
\left\|c_- e^{-i\kappa x} + c_+ e^{i\kappa x}\right\|^2_{(0,\ell)} &= \int_I |c_-|^2 + |c_+|^2 + 2\,\mathrm{Re}\, c_+\bar{c}_- e^{2i\kappa x} \\
&= (|c_-|^2 + |c_+|^2)\ell + 2\,\mathrm{Re}\, c_+\bar{c}_- \frac{1}{2i\kappa}\left[e^{2i\kappa x}\right]_0^\ell \\
&\geq (|c_-|^2 + |c_+|^2)\ell - \frac{2}{\kappa}|c_+||c_-| \\
&\geq (|c_-|^2 + |c_+|^2)\left(\ell - \frac{1}{\kappa}\right),
\end{aligned}$$

and since $\kappa > 2/\ell$, we obtain that

$$|c_-|^2 + |c_+|^2 \leq C\frac{2\ell}{\kappa^2}\|f\|^2.$$

Recalling expression (3.3.33), we obtain (3.3.32) in the physical domain by observing that

$$\begin{aligned}\|\chi_{\mathrm{I}} u'\|_\infty &\leq \kappa|c_-| + \kappa|c_+| + \frac{1}{2}\int_I |f(\xi)|d\xi \\ &\leq 2\kappa\sqrt{(|c_-|^2 + |c_+|^2)} + \frac{\sqrt{\ell}}{2}\|f\| \\ &\leq C\sqrt{\ell}\|f\|.\end{aligned}$$

Then, holders inequality gives

$$\|u'\|_I^2 \leq \ell\|\chi_{\mathrm{I}} u\|_\infty^2,$$

which is (3.3.31) restricted to the physical domain..

We still have to control the norms of u' in the absorbing layer. To do so, we will again employ the inequality

$$\|\chi_{\mathrm{P}} u'\|_\infty^2 \leq C\left(\ell^{-1}\|u'\|_P^2 + \|u'\|_P\|u''\|_P\right). \tag{3.3.35}$$

Recalling the volume estimate in the PML (3.3.12)

$$\kappa^2\|u\|_{\mathrm{P}}^2 + \|u'\|_{\mathrm{P}}^2 \leq C\left(\|f\|_{\mathrm{I}}\|u\|_{\mathrm{I}} + \frac{1}{\kappa^2}\|f\|_{\mathrm{P}}^2.\right)$$

we derived in Lemma 3.3.4, and using (3.3.34)

$$\kappa^2\|u\|_{\mathrm{P}}^2 + \|u'\|_{\mathrm{P}}^2 \leq C\left(\frac{\ell}{\kappa}\|f\|_{\mathrm{I}}^2\frac{1}{\kappa^2}\|f\|_{\mathrm{P}}^2\right) \leq C\frac{1}{\kappa}\left(\ell + \kappa^{-1}\right)\|f\|^2 \leq C\frac{\ell}{\kappa}\|f\|^2,$$

where we employ the assumption that $\kappa^{-1} \leq C\ell$ in the last inequality. Using again the assumption on κ, we have $\ell/\kappa \leq \ell^2$, so that we have in particular

$$\|u'\|_{\mathrm{P}} \leq C\sqrt{\frac{\ell}{\kappa}}\|f\| \leq C\ell\|f\|, \tag{3.3.36}$$

from which (3.3.31) follows. We also see that

$$\kappa^2\|u\|_{\mathrm{P}} \leq C\sqrt{\kappa\ell}\|f\|.$$

Hence, since $-u'' = \nu f + \nu^2\kappa^2 u$ in $(\ell, \ell + \ell_{\mathrm{P}})$, we have

$$\|u''\|_{\mathrm{P}} \leq C\left(\|f\|_{\mathrm{P}} + \kappa^2\|u\|_{\mathrm{P}}\right) \leq C(1 + \sqrt{\kappa}\ell)\|f\| \leq C\sqrt{\kappa\ell}\|f\|,$$

due to the largness assumption on κ. Thus, using (3.3.35) and the first inequality in (3.3.36), we see that

$$\|\chi_{\mathrm{P}} u'\|_\infty^2 \leq C\left(\frac{1}{\kappa} + \ell\right)\|f\|^2 \leq C\ell\|f\|^2,$$

due again to the assumption that $\kappa \geq 2/\ell$. This concludes the proof. □

We can now provide our improved estimates for the one-dimensional PML problem. The proof simply follows by treating the imaginary, small real, and large real wave numbers by the above lemmas.

Theorem 3.3.12 (Improved estimates for the one-dimensional PML problem)**.** *For all $\kappa \in \mathbb{R}_+ \cup i\mathbb{R}_+$ and $f \in L^2(0, \ell + \ell_{\rm P})$, the solution $u \in H^1_0(0, \ell + \ell_{\rm P})$ satisfies*

$$\|u'\| \leq C\ell\|f\|, \qquad \|u'\|_\infty \leq C\sqrt{\ell}\|f\|. \tag{3.3.37}$$

3.3.2 From DtN to PML stability estimates

In the two previous sections, we analyzed the stability of the PML problem in the one-layer case "from scratch" without relying on the estimates we previously derived for the original problem involving the DtN operator. Here, we follow a different path, and assuming that the DtN problem is well-posed, our goal will be to show that the associated PML problem is also well-posed (for a "absorbing enough" layer), and to link the stability constant of the PML problem to the one of the DtN problem.

3.3.2.1 Auxiliary PML problem

This subsection focuses on the auxiliary problem in the absorbing layer associated with the definition of the DtN PML operator $\mathscr{R}_{\rm P}$. Specifically, given $f^{\rm P} : \Omega_{\rm P} \to \mathbb{C}$, we consider the problem of finding $u^{\rm P} : \Omega_{\rm P} \to \mathbb{C}$ such that

$$\left\{\begin{array}{rcll} -k^2\nu u^{\rm P} - \boldsymbol{\nabla}\cdot(\boldsymbol{D}\boldsymbol{\nabla}u^{\rm P}) &=& f^{\rm P} & \text{in } \Omega_{\rm P}, \\ u^{\rm P}{}_+ - e^{i\alpha\ell_1}u^{\rm P}{}_- &=& 0 & \text{on } \Gamma_\sharp, \\ u^{\rm P} &=& 0 & \text{on } \Gamma_{\rm P}, \\ u^{\rm P} &=& 0 & \text{on } \Gamma_{\rm A}. \end{array}\right.$$

Assuming that $f^{\rm P} \in L^2(\Omega_{\rm P})$, the weak form consists in finding $u^{\rm P} \in H^1_{\sharp,0}(\Omega_{\rm P})$ such that

$$\widetilde{b}^{\rm P}(u^{\rm P}, v^{\rm P}) = (f^{\rm P}, v^{\rm P})_{\Omega_{\rm P}} \quad \forall v^{\rm P} \in H^1_{\sharp,0}(\Omega_{\rm P}), \tag{3.3.38}$$

with

$$H^1_{\sharp,0}(\Omega_{\rm P}) := \left\{ v \in H^1(\Omega_{\rm P}, \mathbb{C}) \mid v_{|\Gamma_{\rm A}} = v_{|\Gamma_{\rm P}} = 0 \ \text{ and } \ v_+ = e^{i\alpha\ell_1}v_-, \right\}$$

and

$$\widetilde{b}^{\rm P}(u^{\rm P}, v^{\rm P}) := -k^2(\nu u^{\rm P}, v^{\rm P})_{\Omega_{\rm P}} + \left(\nu\frac{\partial u^{\rm P}}{\partial \boldsymbol{x}_1}, \frac{\partial v^{\rm P}}{\partial \boldsymbol{x}_1}\right)_{\Omega_{\rm P}} + \left(\nu^{-1}\frac{\partial u^{\rm P}}{\partial \boldsymbol{x}_2}, \frac{\partial v^{\rm P}}{\partial \boldsymbol{x}_2}\right)_{\Omega_{\rm P}}.$$

As usual, we will treat this problem by Fourier expansion in the $\boldsymbol{x}_1$ variable. Clearly, the sesquilinear form associated with each Fourier mode is given by

$$\widehat{b}_n(\widehat{u}, \widehat{v}) = -k_n^2\nu_{\rm P}(\widehat{u}, \widehat{v}) + \nu_{\rm P}^{-1}(\widehat{u}', \widehat{v}') \qquad \forall \widehat{u}, \widehat{v} \in H^1_0(\ell_2, \ell_2 + \ell_{\rm P}),$$

for each $n \in \mathbb{Z}$, where $(\cdot,\cdot)$ denotes the inner product of $L^2(\ell_2, \ell_2 + \ell_{\rm P})$ and $||\cdot||$ will denote its usual norm.

We start by establishing the coercivity of $\widehat{b}_n$ for each $n \in \mathbb{Z}$. The following Poincaré inequality will be useful:

$$\|\widehat{u}\| \le \ell_{\mathrm{P}} \|\widehat{u}'\| \quad \forall \widehat{u} \in H^1_0(\ell_2, \ell_2 + \ell_{\mathrm{P}}), \tag{3.3.39}$$

where $\|\cdot\|$ denotes the usual norm of $L^2(\ell_2, \ell_2 + \ell_{\mathrm{P}})$.

Lemma 3.3.13 (Coercivity)**.** *For all $n \in \mathbb{Z}$ and for all $\widehat{u} \in H^1_0(\ell_2, \ell_2 + \ell_{\mathrm{P}})$, there exists $\theta_n \in \mathbb{C}$ with $|\theta_n| = 1$ such that*

$$\operatorname{Re} \widehat{b}_n(\widehat{u}, \theta_n \widehat{u}) \ge |k_n|^2 \|\widehat{u}\|^2 + |\nu_{\mathrm{P}}|^{-2} \|\widehat{u}'\|^2. \tag{3.3.40}$$

In particular, we have

$$\ell_{\mathrm{P}}^2 \operatorname{Re} \widehat{b}_n(\widehat{u}, \theta_n \widehat{u}) \ge (|\nu_{\mathrm{P}}|^{-2} + |k_n|^2 \ell_{\mathrm{P}}^2) \|\widehat{u}\|^2. \tag{3.3.41}$$

Proof. On the one hand, if $k_n^2 \le 0$, we have

$$\begin{aligned} \operatorname{Re} \widehat{b}_n(\widehat{u}, \widehat{u}) &= -k_n^2 \gamma_{\mathrm{r}} \|\widehat{u}\|^2 + |\nu_{\mathrm{P}}|^{-2} \gamma_{\mathrm{r}} \|\widehat{u}'\|^2 \\ &= \gamma_{\mathrm{r}} \left(|k_n|^2 \|\widehat{u}\|^2 + |\nu_{\mathrm{P}}|^{-2} \|\widehat{u}'\|^2 \right) \\ &\ge |k_n|^2 \|\widehat{u}\|^2 + |\nu_{\mathrm{P}}|^{-2} \|\widehat{u}'\|^2, \end{aligned}$$

since $\gamma_{\mathrm{r}} \ge 1$. This shows (3.3.40) for $k_n^2 \le 0$ with $\theta_n = 1$.

On the other hand, if $k_n^2 > 0$, we have

$$\widehat{b}_n(\widehat{u}, -i\widehat{u}) = -k_n^2 i \nu_{\mathrm{P}} \|\widehat{u}\|^2 + i \nu_{\mathrm{P}}^{-1} \|\widehat{u}'\|^2 = -k_n^2 i \nu_{\mathrm{P}} \|\widehat{u}\|^2 + \frac{i \overline{\nu_{\mathrm{P}}}}{|\nu_{\mathrm{P}}|^2} \|\widehat{u}'\|^2,$$

so that

$$\operatorname{Re} \widehat{b}_n(\widehat{u}, -i\widehat{u}) = \gamma_{\mathrm{i}} \left(k_n^2 \|\widehat{u}\| + \frac{1}{|\nu_{\mathrm{P}}|^2} \|\widehat{u}'\|^2 \right) \ge |k_n|^2 \|\widehat{u}\| + \frac{1}{|\nu_{\mathrm{P}}|^2} \|\widehat{u}'\|^2$$

since $k_n^2 > 0$ and $\gamma_{\mathrm{i}} \ge 1$, which is (3.3.40) for $k_n^2 \le 0$ with $\theta_n = -i$.

The estimate in (3.3.41) then simply follows by multiplying (3.3.40) by ℓ_{P}^2 and applying Poincaré inequality (3.3.39). □

We establish a special boundary estimate that will be useful later on.

Lemma 3.3.14 (Sharp boundary estimate)**.** *For all $\widehat{f}_n \in L^2(\ell_2, \ell_2 + \ell_{\mathrm{P}})$, there exists a unique $\widehat{u}_n \in H^1_0(\ell_2, \ell_2 + \ell_{\mathrm{P}})$ such that*

$$\widehat{b}_n(\widehat{u}_n, \widehat{v}_n) = (\widehat{f}_n, \widehat{v}_n) \qquad \widehat{v}_n \in H^1_0(\ell_2, \ell_2 + \ell_{\mathrm{P}}).$$

In addition $\widehat{u}_n' \in L^\infty(\ell_2, \ell_2 + \ell_{\mathrm{P}})$, and we have

$$|\widehat{u}_n'(\ell_2)| \le 5 |\nu|^2 \frac{1}{1 + (|k_n| \ell_{\mathrm{P}})^{1/2}} \ell_{\mathrm{P}}^{1/2} \|\widehat{f}_n\|. \tag{3.3.42}$$

Proof. Recalling (3.3.40) and (3.3.41). Since,

$$\operatorname{Re}\widehat{b}_n(\widehat{u}_n,\theta_n\widehat{u}_n)=\operatorname{Re}(\widehat{f}_n,\theta_n\widehat{u}_n)\le\|\widehat{f}_n\|\|\widehat{u}_n\|,$$

for all $\theta_n\in\mathbb{C}$ with $|\theta_n|=1$, we have

$$|\nu_{\mathrm{P}}|^{-2}\|\widehat{u}_n'\|^2\le\|\widehat{f}_n\|\|\widehat{u}_n\|, \tag{3.3.43}$$

and

$$(|\nu_{\mathrm{P}}|^{-2}+(|k_n|\ell_{\mathrm{P}})^2)\|\widehat{u}_n\|\le\ell_{\mathrm{P}}^2\|\widehat{f}_n\|. \tag{3.3.44}$$

Plugging (3.3.44) into (3.3.43), we deduce that

$$\|\widehat{u}_n'\|^2\le\frac{|\nu_{\mathrm{P}}|^4}{1+(|k_n|\ell_{\mathrm{P}})^2}\ell_{\mathrm{P}}^2\|\widehat{f}_n\|^2\le|\nu_{\mathrm{P}}|^4\ell_{\mathrm{P}}^2\|\widehat{f}_n\|^2$$

so that

$$\|\widehat{u}_n'\|\le\sqrt{2}|\nu_{\mathrm{P}}|^2\frac{1}{1+|k_n|\ell_{\mathrm{P}}}\ell_{\mathrm{P}}\|\widehat{f}_n\| \tag{3.3.45}$$

and

$$\|\widehat{u}_n'\|^2\le2|\nu_{\mathrm{P}}|^4\frac{1}{1+|k_n|\ell_{\mathrm{P}}}\ell_{\mathrm{P}}^2\|\widehat{f}_n\|^2. \tag{3.3.46}$$

We see from (3.3.44) that $|k_n|^2\|\widehat{u}_n\|\le\|\widehat{f}_n\|$, and since $-\nu_{\mathrm{P}}^{-1}\widehat{u}_n''=\widehat{f}_n+\nu_{\mathrm{P}}k_n^2\widehat{u}_n$, we have

$$\|\widehat{u}_n''\|\le|\nu_{\mathrm{P}}|\|\widehat{f}_n\|+|\nu_{\mathrm{P}}|^2|k_n|^2\|\widehat{u}_n\|\le|\nu_{\mathrm{P}}|\|\widehat{f}_n\|+|\nu_{\mathrm{P}}|^2\|\widehat{f}_n\|\le2|\nu_{\mathrm{P}}|^2\|\widehat{f}_n\|. \tag{3.3.47}$$

We then notice that

$$\begin{aligned}\ell_{\mathrm{P}}|\widehat{u}_n'(\ell)|^2&=[(x-\ell-\ell_{\mathrm{P}})|\widehat{u}_n'(x)|^2]_\ell^{\ell+\ell_{\mathrm{P}}}\\&=\int_\ell^{\ell+\ell_{\mathrm{P}}}|\widehat{u}_n'(x)|^2dx+2\operatorname{Re}\int_\ell^{\ell+\ell_{\mathrm{P}}}(x-\ell-\ell_{\mathrm{P}})\widehat{u}_n'(x)\overline{\widehat{u}_n''(x)}dx\\&\le\|\widehat{u}_n'\|^2+2\ell_{\mathrm{P}}\|\widehat{u}_n'\|\|\widehat{u}_n''\|,\end{aligned} \tag{3.3.48}$$

and inserting (3.3.45), (3.3.46) and (3.3.47) into (3.3.48), it follows that

$$\begin{aligned}\ell_{\mathrm{P}}|\widehat{u}_n'(\ell)|^2&\le\left(2|\nu_{\mathrm{P}}|^4\frac{\ell_{\mathrm{P}}^2}{1+|k_n|\ell_{\mathrm{P}}}+2\ell_{\mathrm{P}}\left(\sqrt{2}\frac{|\nu_{\mathrm{P}}|^2\ell_{\mathrm{P}}}{1+|k_n|\ell_{\mathrm{P}}}\right)\left(2|\nu_{\mathrm{P}}|^2\right)\right)\|\widehat{f}_n\|^2\\&=6\sqrt{2}|\nu_{\mathrm{P}}|^4\frac{1}{1+|k_n|\ell_{\mathrm{P}}}\ell_{\mathrm{P}}^2\|\widehat{f}_n\|^2\\&\le12\sqrt{2}|\nu_{\mathrm{P}}|^4\frac{1}{(1+(|k_n|\ell_{\mathrm{P}})^{1/2})^2}\ell_{\mathrm{P}}^2\|\widehat{f}_n\|^2\\&\le25|\nu_{\mathrm{P}}|^4\frac{1}{(1+(|k_n|\ell_{\mathrm{P}})^{1/2})^2}\ell_{\mathrm{P}}^2\|\widehat{f}_n\|^2.\end{aligned}$$

□

Relying on Fourier expansion, we easily deduce the inf-sup stability of the sesquilinear form $\widetilde{b}^{\mathrm{P}}$ from the coercivity of each $\widehat{b}_n$.

Lemma 3.3.15 (Inf-sup stability)**.** *For all $u^{\mathrm{P}} \in H^1_{\sharp,0}(\Omega_{\mathrm{P}})$, there exists $v^{\mathrm{P}} \in H^1_{\sharp,0}(\Omega_{\mathrm{P}})$ with $\left|\!\left|\!\left|v^{\mathrm{P}}\right|\!\right|\!\right|_{k,\Omega_{\mathrm{P}}} = \left|\!\left|\!\left|u^{\mathrm{P}}\right|\!\right|\!\right|_{k,\Omega_{\mathrm{P}}}$ such that*

$$|\nu|^2 \ell_{\mathrm{P}}^2 \mathrm{Re}\widetilde{b}^{\mathrm{P}}(u^{\mathrm{P}}, v^{\mathrm{P}}) \geq (1 + (k_\star \ell_{\mathrm{P}})^2) \|u^{\mathrm{P}}\|^2_{\varepsilon,\Omega_{\mathrm{P}}}. \tag{3.3.49}$$

In addtion, we have

$$\sup_{\substack{v^{\mathrm{P}} \in H^1_{\sharp,0}(\Omega_{\mathrm{P}}) \\ \left|\!\left|\!\left|v^{\mathrm{P}}\right|\!\right|\!\right|_{k,\Omega_{\mathrm{P}}}=1}} \mathrm{Re}\,\widetilde{b}^{\mathrm{P}}(u^{\mathrm{P}}, v^{\mathrm{P}}) \geq \frac{1}{2|\nu|^4} \frac{1}{1 + (k\ell_{\mathrm{P}})^2} \left|\!\left|\!\left|u^{\mathrm{P}}\right|\!\right|\!\right|_{k,\Omega_{\mathrm{P}}}. \tag{3.3.50}$$

Proof. Consider $u^{\mathrm{P}} \in H^1_{\sharp,0}(\Omega_{\mathrm{P}})$ and set

$$v^{\mathrm{P}} := \sum_{n\in\mathbb{Z}} \theta_n \widehat{u}_n e^{i(\alpha+\alpha_n)x_1}.$$

It is clear that $\left|\!\left|\!\left|v^{\mathrm{P}}\right|\!\right|\!\right|_{k,\Omega_{\mathrm{P}}} = \left|\!\left|\!\left|u^{\mathrm{P}}\right|\!\right|\!\right|_{k,\Omega_{\mathrm{P}}}$, and we have

$$\ell_{\mathrm{P}}^2 \mathrm{Re}\widetilde{b}^{\mathrm{P}}(u^{\mathrm{P}}, v^{\mathrm{P}}) = \ell_{\mathrm{P}}^2 \ell_1 \sum_{n\in\mathbb{Z}} \mathrm{Re}\widehat{b}_n(\widehat{u}_n, \widehat{v}_n) = \ell_{\mathrm{P}}^2 \ell_1 \sum_{n\in\mathbb{Z}} \mathrm{Re}\widehat{b}_n(\widehat{u}_n, \theta_n \widehat{u}_n) \geq \ell_1 \sum_{n\in\mathbb{Z}} (|\nu_{\mathrm{P}}|^{-2} + (|k_n|\ell_{\mathrm{P}})^2) \|\widehat{u}_n\|^2$$

Since $|\nu_{\mathrm{P}}| \geq 1$ and $|k_n| \geq k_\star$, (3.3.49) follows.

We can further write that

$$\begin{aligned}
\mathrm{Re}\,\widetilde{b}^{\mathrm{P}}(u^{\mathrm{P}}, u^{\mathrm{P}}) &= -k^2\,\mathrm{Re}\,\nu \|u^{\mathrm{P}}\|^2_{\varepsilon,\Omega_{\mathrm{P}}} + \mathrm{Re}\,\nu \left\|\frac{\partial u^{\mathrm{P}}}{\partial \boldsymbol{x}_1}\right\|^2_{A_1,\Omega_{\mathrm{P}}} + \mathrm{Re}\,\nu^{-1} \left\|\frac{\partial u^{\mathrm{P}}}{\partial \boldsymbol{x}_2}\right\|^2_{A_2,\Omega_{\mathrm{P}}} \\
&= -k^2\gamma_{\mathrm{r}} \|u^{\mathrm{P}}\|^2_{\varepsilon,\Omega_{\mathrm{P}}} + \gamma_{\mathrm{r}} \left\|\frac{\partial u^{\mathrm{P}}}{\partial \boldsymbol{x}_1}\right\|^2_{A_1,\Omega_{\mathrm{P}}} + \frac{\gamma_{\mathrm{r}}}{|\nu|^2} \left\|\frac{\partial u^{\mathrm{P}}}{\partial \boldsymbol{x}_2}\right\|_{A_2,\Omega_{\mathrm{P}}} \\
&\geq -k^2 \|u^{\mathrm{P}}\|^2_{\varepsilon,\Omega_{\mathrm{P}}} + \left\|\frac{\partial u^{\mathrm{P}}}{\partial \boldsymbol{x}_1}\right\|^2_{A_1,\Omega_{\mathrm{P}}} + \frac{1}{|\nu|^2} \left\|\frac{\partial u^{\mathrm{P}}}{\partial \boldsymbol{x}_2}\right\|_{A_2,\Omega_{\mathrm{P}}},
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Re}\,\widetilde{b}^{\mathrm{P}}(u^{\mathrm{P}}, u^{\mathrm{P}} + 2|\nu|^2(k\ell_{\mathrm{P}})^2 v^{\mathrm{P}}) &\geq k^2 \|u^{\mathrm{P}}\|^2_{\varepsilon,\Omega_{\mathrm{P}}} + \left\|\frac{\partial u^{\mathrm{P}}}{\partial \boldsymbol{x}_1}\right\|^2_{A_1,\Omega_{\mathrm{P}}} + \frac{1}{|\nu|^2} \left\|\frac{\partial u^{\mathrm{P}}}{\partial \boldsymbol{x}_2}\right\|_{A_2,\Omega_{\mathrm{P}}}, \\
&\geq |\nu|^{-2} \left|\!\left|\!\left|u^{\mathrm{P}}\right|\!\right|\!\right|^2_{k,\Omega_{\mathrm{P}}}.
\end{aligned}$$

□

Lemma 3.3.16 (Auxiliary PML operator). *For all $f^{\mathrm{P}} \in L^2(\Omega_{\mathrm{P}})$, there exists a unique solution $\mathscr{L}_{\mathrm{P}} f^{\mathrm{P}} := u^{\mathrm{P}}$ solution to (3.3.38), and the estimates*

$$\left|\!\left|\!\left| \mathscr{L}_{\mathrm{P}} f^{\mathrm{P}} \right|\!\right|\!\right|_{k,\Omega_{\mathrm{P}}} \leq 2|\nu_{\mathrm{P}}|^3 \left(\frac{1}{k\ell_{\mathrm{P}}} + \frac{k\ell_{\mathrm{P}}}{1+(k_\star \ell_{\mathrm{P}})^2} \right) \ell_{\mathrm{P}} \|f^{\mathrm{P}}\|_{\varepsilon,\Omega_{\mathrm{P}}}, \tag{3.3.51a}$$

$$\|\partial_2(\mathscr{L}_{\mathrm{P}} f^{\mathrm{P}})\|_{\Gamma_{\mathrm{A}}} \leq \frac{5|\nu_{\mathrm{P}}|^2}{1+(k_\star \ell_{\mathrm{P}})^{1/2}} \ell_{\mathrm{P}}^{1/2} \|f\|^2_{\varepsilon,\Omega_{\mathrm{P}}}, \tag{3.3.51b}$$

hold true.

Proof. We fix $f^{\mathrm{P}} \in L^2(\Omega_{\mathrm{P}})$ and let $u^{\mathrm{P}} := \mathscr{L}_{\mathrm{P}} f^{\mathrm{P}}$. We have

$$(1+(k_\star \ell_{\mathrm{P}})^2)\|u^{\mathrm{P}}\|^2_{\varepsilon,\Omega_{\mathrm{P}}} \leq |\nu|^2 \ell_{\mathrm{P}}^2 \operatorname{Re} \widetilde{b}^{\mathrm{P}}(u^{\mathrm{P}}, v_\star^{\mathrm{P}}) = |\nu|^2 \ell_{\mathrm{P}}^2 (\varepsilon f^{\mathrm{P}}, v_\star^{\mathrm{P}})_{\Omega_{\mathrm{P}}} \leq |\nu|^2 \ell_{\mathrm{P}}^2 \|f^{\mathrm{P}}\|_{\varepsilon,\Omega_{\mathrm{P}}} \|u^{\mathrm{P}}\|_{\varepsilon,\Omega_{\mathrm{P}}},$$

hence

$$k\|u^{\mathrm{P}}\|_{\varepsilon,\Omega_{\mathrm{P}}} \leq |\nu_{\mathrm{P}}|^2 \frac{k\ell_{\mathrm{P}}}{1+(k_\star \ell_{\mathrm{P}})^2} \ell_{\mathrm{P}} \|f^{\mathrm{P}}\|_{\varepsilon,\Omega_{\mathrm{P}}}.$$

On the other hand, we have

$$\begin{aligned} -k^2\|u^{\mathrm{P}}\|^2_{\varepsilon,\Omega_{\mathrm{P}}} + \|\partial_1 u^{\mathrm{P}}\|^2_{A_1,\Omega_{\mathrm{P}}} + |\nu_{\mathrm{P}}|^{-2}\|\partial_2 u^{\mathrm{P}}\|^2_{A_2,\Omega_{\mathrm{P}}} &= \frac{1}{\gamma_{\mathrm{r}}} \operatorname{Re}\big(-k^2 \nu_{\mathrm{P}} \|u^{\mathrm{P}}\|^2_{\varepsilon,\Omega_{\mathrm{P}}} + \nu_{\mathrm{P}} \|\partial_1 u^{\mathrm{P}}\|^2_{A_1,\Omega_{\mathrm{P}}} \\ &\quad + \nu_{\mathrm{P}}^{-1} \|\partial_2 u^{\mathrm{P}}\|^2_{A_2,\Omega_{\mathrm{P}}} \big) \\ &= \frac{1}{\gamma_{\mathrm{r}}} \operatorname{Re}(\varepsilon f^{\mathrm{P}}, u^{\mathrm{P}})_{\Omega_{\mathrm{P}}}. \end{aligned}$$

Then, since

$$\operatorname{Re}(\varepsilon f^{\mathrm{P}}, u^{\mathrm{P}}) \leq 2k^2 \|u^{\mathrm{P}}\|^2_{\varepsilon,\Omega_{\mathrm{P}}} + \frac{1}{8k^2}\|f^{\mathrm{P}}\|^2_{\varepsilon,\Omega_{\mathrm{P}}} \leq 2k^2 \|u^{\mathrm{P}}\|^2_{\varepsilon,\Omega_{\mathrm{P}}} + \frac{1}{k^2}\|f^{\mathrm{P}}\|^2_{\varepsilon,\Omega_{\mathrm{P}}},$$

we have

$$\begin{aligned} |u^{\mathrm{P}}|^2_{\boldsymbol{A},\Omega_{\mathrm{P}}} &\leq |\nu_{\mathrm{P}}|^2 \|\partial_1 u^{\mathrm{P}}\|_{A_1,\Omega_{\mathrm{P}}} + \|\partial_2 u^{\mathrm{P}}\|_{A_2,\Omega_{\mathrm{P}}} \\ &= \frac{|\nu_{\mathrm{P}}|^2}{\gamma_{\mathrm{r}}} \operatorname{Re}(\varepsilon f^{\mathrm{P}}, u^{\mathrm{P}})_{\Omega_{\mathrm{P}}} + k^2 |\nu_{\mathrm{P}}|^2 \|u^{\mathrm{P}}\|^2_{\varepsilon,\Omega_{\mathrm{P}}} \\ &\leq \frac{|\nu_{\mathrm{P}}|^2}{k^2} \|f^{\mathrm{P}}\|^2_{\varepsilon,\Omega_{\mathrm{P}}} + 3k^2 |\nu_{\mathrm{P}}|^2 \|u^{\mathrm{P}}\|^2_{\varepsilon,\Omega_{\mathrm{P}}}, \end{aligned}$$

and

$$\begin{aligned} \left|\!\left|\!\left| u^{\mathrm{P}} \right|\!\right|\!\right|^2_{k,\Omega_{\mathrm{P}}} &\leq 2\frac{|\nu_{\mathrm{P}}|^2}{k^2} \|f^{\mathrm{P}}\|^2_{\varepsilon,\Omega_{\mathrm{P}}} + 4k^2 |\nu_{\mathrm{P}}|^2 \|u^{\mathrm{P}}\|^2_{\varepsilon,\Omega_{\mathrm{P}}}, \\ &\leq 4|\nu_{\mathrm{P}}|^6 \left(\frac{1}{(k\ell_{\mathrm{P}})^2} + \left(\frac{k\ell_{\mathrm{P}}}{1+(k_\star \ell_{\mathrm{P}})^2} \right)^2 \right) \ell_{\mathrm{P}}^2 \|f^{\mathrm{P}}\|^2_{\varepsilon,\Omega_{\mathrm{P}}}, \end{aligned}$$

and (3.3.51a) follows since $\sqrt{a^2+b^2} \le a+b$.

We then show (3.3.51b). Recalling the quasi-periodicity condition satisfied by u^{P}, we get the Fourier expansion

$$u^{\mathrm{P}} = \sum_{n\in\mathbb{Z}} \widehat{u}_n e^{i(\alpha+\alpha_n)x_1},$$

where for each $n \in \mathbb{Z}$, $\widehat{u}_n$ satisfies

$$\widehat{b}_n(\widehat{u}_n, \widehat{v}) = (\widehat{f}_n, \widehat{v}) \quad \forall \widehat{v} \in H^1_0(\ell_2, \ell_2+\ell_{\mathrm{P}}).$$

Using (3.3.42), we then have

$$\begin{aligned}
\|\partial_2(u^{\mathrm{P}})\|^2_{\Gamma_{\mathrm{A}}} &= \ell_1 \sum_{n\in\mathbb{Z}} \|\widehat{u}_n'(\ell_2)\|^2 \\
&\le 25|\nu_{\mathrm{P}}|^4 \ell_{\mathrm{P}} \ell_1 \sum_{n\in\mathbb{Z}} \left(\frac{1}{1+(|k_n|\ell_{\mathrm{P}})^{1/2}}\right)^2 \|\widehat{f}_n\|^2 \\
&\le 25|\nu_{\mathrm{P}}|^4 \ell_{\mathrm{P}} \left(\frac{1}{1+(k_\star \ell_{\mathrm{P}})^{1/2}}\right)^2 \|f\|^2_{\varepsilon,\Omega_{\mathrm{P}}}.
\end{aligned}$$

□

3.3.2.2 Stability of the PML problem

We now present a key lemma showing how we can relate the stability of the PML problem to that of the original problem.

Lemma 3.3.17. *Let $\widetilde{f} \in L^2(\widetilde{\Omega})$ and assume that $\widetilde{u} \in H^1_\sharp(\widetilde{\Omega})$ satisfies*

$$\widetilde{b}(\widetilde{u}, \widetilde{v}) = (\varepsilon \widetilde{f}, \widetilde{v})_{\widetilde{\Omega}} \quad \forall \widetilde{v} \in H^1_\sharp(\widetilde{\Omega}). \tag{3.3.52}$$

Then, we have

$$b_{\mathrm{P}}(\widetilde{u}|_\Omega, v) = (\varepsilon \widetilde{f}|_\Omega, v)_\Omega + \nu^{-1}\langle \partial_2(\mathscr{L}_{\mathrm{P}}(\widetilde{f}|_{\Omega_{\mathrm{P}}}), v\rangle_{\Gamma_{\mathrm{A}}}, \quad \forall v \in H^1_\sharp(\Omega). \tag{3.3.53}$$

Proof. Let $\widetilde{f} \in L^2(\widetilde{\Omega})$ and let $\widetilde{u} \in H^1_\sharp(\widetilde{\Omega})$ satisfy (3.3.52). For the sake of simplicity, let us write

$$u = \widetilde{u}|_\Omega, \qquad u^{\mathrm{P}} := \widetilde{u}|_{\Omega_{\mathrm{P}}}.$$

Since $\widetilde{u} \in H^1(\Omega)$, its trace is the same on both sides of Γ_{A}, and we have

$$\left\{\begin{array}{rcll}
-k^2\nu_{\mathrm{P}} u^{\mathrm{P}} - \frac{\partial}{\partial \boldsymbol{x}_1}\left(\nu_{\mathrm{P}} \frac{\partial u^{\mathrm{P}}}{\partial \boldsymbol{x}_1}\right) - \frac{\partial}{\partial \boldsymbol{x}_2}\left(\nu_{\mathrm{P}}^{-1} \frac{\partial u^{\mathrm{P}}}{\partial \boldsymbol{x}_2}\right) &=& f^{\mathrm{P}} & \text{in } \Omega_{\mathrm{P}} \\
u^{\mathrm{P}} &=& u|_{\Gamma_{\mathrm{A}}} & \text{on } \Gamma_{\mathrm{A}} \\
u^{\mathrm{P}}_+ - e^{i\alpha\ell_1} u^{\mathrm{P}}_- &=& 0 & \text{on } \Gamma_\sharp \\
u^{\mathrm{P}} &=& 0 & \text{on } \Gamma_{\mathrm{P}},
\end{array}\right.$$

and in particular

$$\nu_{\mathrm{P}}^{-1}(\partial_2 u^{\mathrm{P}})|_{\Gamma_{\mathrm{A}}} = \nu_{\mathrm{P}}^{-1}\partial_2(\mathscr{L}_{\mathrm{P}} f^{\mathrm{P}}) + \mathscr{R}_{\mathrm{P}}(u|_{\Gamma_{\mathrm{A}}}).$$

Since $\widetilde{u}$ solves (3.3.52), we must also have $[\![\nu^{-1}\partial_2\widetilde{u}]\!] = 0$, leading to

$$(\partial_2 u)|_{\Gamma_{\mathrm{A}}} = \nu_{\mathrm{P}}^{-1}\partial_2(\mathscr{L}_{\mathrm{P}} f^{\mathrm{P}}) + \mathscr{R}_{\mathrm{P}}(u|_{\Gamma_{\mathrm{A}}}). \tag{3.3.54}$$

Since $u \in H^1_\sharp(\Omega)$ and

$$-k^2\varepsilon u - \boldsymbol{\nabla}\cdot(\boldsymbol{A}\boldsymbol{\nabla}u) = \varepsilon f \text{ in } \Omega,$$

integration by parts shows that

$$-k^2(\varepsilon u, v)_\Omega - \langle \partial_2 u, v\rangle_{\Gamma_{\mathrm{A}}} + (\boldsymbol{A}\boldsymbol{\nabla}u, \boldsymbol{\nabla}v)_\Omega = (\varepsilon f, v)_\Omega \quad \forall v \in H^1_\sharp(\Omega).$$

Recalling (3.3.54), this leads to

$$b_{\mathrm{P}}(u,v) = (\varepsilon f, v)_\Omega + \nu^{-1}(\partial_2(\mathscr{L}_{\mathrm{P}} f^{\mathrm{P}}), v)_{\Gamma_{\mathrm{A}}},$$

for all $v \in H^1_\sharp(\Omega)$. □

The next tool we need is "mirror" operator.

Lemma 3.3.18 (Mirror operator through Γ_{A})**.** *Consider the change of coordinates*

$$\phi : \Omega_{\mathrm{P}} \ni \boldsymbol{x} \to \phi(\boldsymbol{x}) = \left(\boldsymbol{x}_1, \ell_2 + \frac{\ell_2}{\ell_{\mathrm{P}}}(\ell_2 - \boldsymbol{x}_2)\right) \in \Omega.$$

We then define the mirror operator $\mathscr{M} : H^1(\Omega) \to H^1(\Omega_{\mathrm{P}})$ *by* $\mathscr{M}v = v \circ \phi$ *and we have*

$$|\!|\!|\mathscr{M}v|\!|\!|_{k,\Omega_{\mathrm{P}}} \le \sqrt{\frac{\ell_{\mathrm{P}}}{\ell_2} + \frac{\ell_2}{\ell_P}}\frac{1}{\sqrt{\varepsilon_{\min}}}\,|\!|\!|v|\!|\!|_{k,\Omega}\,, \tag{3.3.55}$$

for all $v \in H^1(\Omega)$*. In addition, if* $\widetilde{u} \in H^1_\sharp(\widetilde{\Omega})$*, then* $\widetilde{u}|_{\Omega_{\mathrm{P}}} - \mathscr{M}(\widetilde{u}|_\Omega) \in H^1_{0,\sharp}(\Omega_{\mathrm{P}})$*.*

Proof. Let $v \in L^2(\Omega)$, then

$$\|\mathscr{M}v\|^2_{\Omega_{\mathrm{P}}} = \int_{\Omega_{\mathrm{P}}} |\mathscr{M}v|^2 = \int_{\Omega_{\mathrm{P}}} |v\circ\phi|^2,$$

and using the change of variables $\boldsymbol{y} = \phi(\boldsymbol{x})$ we get

$$\|\mathscr{M}v\|^2_{0,\Omega_{\mathrm{P}}} = \frac{\ell_{\mathrm{P}}}{\ell_2}\int_\Omega |v|^2 = \frac{\ell_{\mathrm{P}}}{\ell_2}\|v\|^2_\Omega.$$

In addition, if $v \in H^1(\Omega)$ we have

$$\|\boldsymbol{\nabla}(\mathscr{M}v)\|^2_{\Omega_{\mathrm{P}}} = \int_{\Omega_{\mathrm{P}}} |\partial_1(\mathscr{M}v)|^2 + \int_{\Omega_{\mathrm{P}}} |\partial_2(\mathscr{M}v)|^2 = \int_{\Omega_{\mathrm{P}}} |\mathscr{M}(\partial_1 v)|^2 + \frac{\ell_2^2}{\ell_{\mathrm{P}}^2}\int_{\Omega_{\mathrm{P}}} |\mathscr{M}(\partial_2 v)|^2$$

and using the same change of variables $\boldsymbol{y} = \phi(\boldsymbol{x})$ we obtain

$$\begin{aligned}\|\boldsymbol{\nabla}(\mathscr{M}v)\|^2_{0,\Omega_{\mathrm{P}}} &= \frac{\ell_{\mathrm{P}}}{\ell_2}\|\partial_1 v\|^2_{0,\Omega} + \frac{\ell_2}{\ell_{\mathrm{P}}}\|\partial_2 v\|^2_{0,\Omega} \\ &\leq \left(\frac{\ell_{\mathrm{P}}}{\ell_2} + \frac{\ell_2}{\ell_P}\right)\|\boldsymbol{\nabla} v\|^2_{\Omega}.\end{aligned}$$

Then, (3.3.55) simply follows from the fact that $A_{\min} = 1$. □

We are now ready to establish an explicit stability bound of $\widetilde{u}$ in Ω.

Theorem 3.3.19. *Suppose that*

$$\widetilde{b}(\widetilde{u},\widetilde{v}) = (\varepsilon\widetilde{f},\widetilde{v})_{\widetilde{\Omega}}, \quad \forall \widetilde{v} \in H^1_\sharp(\widetilde{\Omega}),$$

then, the PML weak formulation (3.1.4) is well-posed and

$$k\,|\!|\!|\widetilde{u}|\!|\!|_{k,\widetilde{\Omega}} \leq 4|\nu_{\mathrm{P}}|^5\left(\frac{\ell_2}{\ell_{\mathrm{P}}} + \frac{\ell_{\mathrm{P}}}{\ell_2}\right)^{1/2}\left(\sqrt{3} + k\ell_{\mathrm{P}}\right)^{3/2}(1+\mathscr{C}_{\mathrm{is,P}})\|\widetilde{f}\|_{\varepsilon,\widetilde{\Omega}}. \tag{3.3.56}$$

In particular, we have

$$\widetilde{\mathscr{C}_{\mathrm{st}}} \leq 20|\nu_{\mathrm{P}}|^5\frac{1}{\sqrt{\varepsilon_{\min}}}\left(\frac{\ell_2}{\ell_{\mathrm{P}}} + \frac{\ell_{\mathrm{P}}}{\ell_2}\right)^{1/2}\left(\sqrt{3} + k\ell_{\mathrm{P}}\right)^{3/2}(1+\mathscr{C}_{\mathrm{is,P}}).$$

Remark 3.3.20. *In the derivation of (3.3.56) we use the (in general) pessimistic estimate $k_\star \geq 0$, so that the estimate can be slightly improved in the absence of quasi-resonances. However, The resulting constant as a very involved expression, and we have chosen to present (3.3.56) in this form for the sake of simplicity.*

Proof. We start by employing the inf-sup condition for b_{P} together with (3.3.53) to show that

$$|\!|\!|\widetilde{u}|\!|\!|_{k,\Omega} \leq \mathscr{C}_{\mathrm{is,P}}\sup_{\substack{v\in H^1_\sharp(\Omega)\\ |\!|\!|v|\!|\!|_{k,\Omega}=1}} b_{\mathrm{P}}(u,v) = \mathscr{C}_{\mathrm{is,P}}\sup_{\substack{v\in H^1_\sharp(\Omega)\\ |\!|\!|v|\!|\!|_{k,\Omega}=1}}\left\{(\varepsilon\widetilde{f},v)_\Omega + \nu_{\mathrm{P}}^{-1}(\partial_2\mathscr{L}_{\mathrm{P}}f^{\mathrm{P}},v)_{\Gamma_{\mathrm{A}}}\right\}.$$

On the one hand, using (3.3.51b) we have

$$\begin{aligned}|\nu_{\mathrm{P}}^{-1}(\partial_2(\mathscr{L}_{\mathrm{P}}f^{\mathrm{P}}),v)_{\Gamma_{\mathrm{A}}}| &\leq |\nu_{\mathrm{P}}|^{-1}\|\partial_2(\mathscr{L}_{\mathrm{P}}f^{\mathrm{P}})\|_{\Gamma_{\mathrm{A}}}\|v\|_{\Gamma_{\mathrm{A}}} \\ &\leq |\nu_{\mathrm{P}}|^{-1}\left(5|\nu_{\mathrm{P}}|^2\ell_{\mathrm{P}}^{1/2}\|f^{\mathrm{P}}\|_{\varepsilon,\Omega_{\mathrm{P}}}\right)\left(k^{-1/2}\,|\!|\!|v|\!|\!|_{k,\Omega}\right) \\ &\leq 5|\nu_{\mathrm{P}}|(k\ell_{\mathrm{P}})^{-1/2}\ell_{\mathrm{P}}\|f^{\mathrm{P}}\|_{\varepsilon,\Omega_{\mathrm{P}}}\,|\!|\!|v|\!|\!|_{k,\Omega_{\mathrm{P}}},\end{aligned}$$

and on the other hand

$$(\varepsilon\widetilde{f},v)_\Omega \leq \|\widetilde{f}\|_{\varepsilon,\Omega}\|v\|_{\varepsilon,\Omega} \leq (k\ell_{\mathrm{P}})^{-1}\ell_{\mathrm{P}}\|f\|_{\varepsilon,\Omega}\,|\!|\!|v|\!|\!|_{k,\Omega},$$

so that

$$k\,\|\!|u_0|\!\|_{k,\Omega} \le 5|\nu_{\mathrm{P}}|\mathscr{C}_{\mathrm{is,P}}\left(1+(k\ell_{\mathrm{P}})^{1/2}\right)\|\widetilde{f}\|_{\varepsilon,\widetilde{\Omega}}. \tag{3.3.57}$$

We have $u^{\mathrm{P}} - \mathscr{M}u \in H^1_{0,\sharp}(\Omega_{\mathrm{P}})$, and

$$\widetilde{b}^{\mathrm{P}}(u^{\mathrm{P}} - \mathscr{M}u, v^{\mathrm{P}}) = (\varepsilon f^{\mathrm{P}}, v^{\mathrm{P}})_{\Omega_{\mathrm{P}}} - \widetilde{b}^{\mathrm{P}}(\mathscr{M}u, v^{\mathrm{P}}).$$

Since

$$\left|(\varepsilon f^{\mathrm{P}}, v^{\mathrm{P}})_{\Omega_{\mathrm{P}}} - \widetilde{b}^{\mathrm{P}}(\mathscr{M}u, v^{\mathrm{P}})\right| \le \left(\frac{1}{k}\|f^{\mathrm{P}}\|_{\varepsilon,\Omega_{\mathrm{P}}} + \|\!|\mathscr{M}u|\!\|_{k,\Omega_{\mathrm{P}}}\right)|\!|\!|v^{\mathrm{P}}|\!|\!|_{\varepsilon,\Omega_{\mathrm{P}}}$$

recalling (3.3.50), we have

$$|\!|\!|u^{\mathrm{P}} - \mathscr{M}u|\!|\!|_{k,\Omega_{\mathrm{P}}} \le 2|\nu|^4(1+(k\ell_{\mathrm{P}})^2)\left(\frac{1}{k}\|f^{\mathrm{P}}\|_{\varepsilon,\Omega_{\mathrm{P}}} + \|\!|\mathscr{M}u|\!\|_{k,\Omega_{\mathrm{P}}}\right),$$

and therefore

$$|\!|\!|u^{\mathrm{P}}|\!|\!|_{k,\Omega_{\mathrm{P}}} \le 2|\nu|^4(2+(k\ell_{\mathrm{P}})^2)\left(\frac{1}{k}\|f^{\mathrm{P}}\|_{\varepsilon,\Omega_{\mathrm{P}}} + \|\!|\mathscr{M}u|\!\|_{k,\Omega_{\mathrm{P}}}\right)$$

and

$$k^2\,|\!|\!|u^{\mathrm{P}}|\!|\!|^2_{k,\Omega_{\mathrm{P}}} \le 8|\nu|^8(2+(k\ell_{\mathrm{P}})^2)^2\left(\|f^{\mathrm{P}}\|^2_{\varepsilon,\Omega_{\mathrm{P}}} + k^2\frac{1}{\varepsilon_{\min}}\left(\frac{\ell_2}{\ell_{\mathrm{P}}}+\frac{\ell_{\mathrm{P}}}{\ell_2}\right)\|\!|u|\!\|^2_{k,\Omega}\right). \tag{3.3.58}$$

Summing the two equations (3.3.57) and (3.3.58), we have

$$\begin{aligned}
k^2\,\|\!|u|\!\|^2_{k,\widetilde{\Omega}} &\le 8|\nu|^8(3+(k\ell_{\mathrm{P}})^2)^2\left(\|f^{\mathrm{P}}\|^2_{\varepsilon,\Omega_{\mathrm{P}}} + k^2\frac{1}{\varepsilon_{\min}}\left(\frac{\ell_2}{\ell_{\mathrm{P}}}+\frac{\ell_{\mathrm{P}}}{\ell_2}\right)\|\!|u|\!\|^2_{k,\Omega}\right)\\
&\le 8|\nu|^8(3+(k\ell_{\mathrm{P}})^2)^2\left(\|f^{\mathrm{P}}\|^2_{\varepsilon,\Omega_{\mathrm{P}}} + \frac{1}{\varepsilon_{\min}}\left(\frac{\ell_2}{\ell_{\mathrm{P}}}+\frac{\ell_{\mathrm{P}}}{\ell_2}\right)25|\nu_{\mathrm{P}}|^2\mathscr{C}^2_{\mathrm{is,P}}\left(1+(k\ell_{\mathrm{P}})^{1/2}\right)^2\|\widetilde{f}\|^2_{\varepsilon,\widetilde{\Omega}}\right)\\
&\le 8|\nu|^8(3+(k\ell_{\mathrm{P}})^2)^2(1+\mathscr{C}_{\mathrm{is,P}})^2\frac{1}{\varepsilon_{\min}}\left(\frac{\ell_2}{\ell_{\mathrm{P}}}+\frac{\ell_{\mathrm{P}}}{\ell_2}\right)25|\nu_{\mathrm{P}}|^2\left(1+(k\ell_{\mathrm{P}})^{1/2}\right)^2\|\widetilde{f}\|^2_{\varepsilon,\widetilde{\Omega}}\\
&\le 200|\nu|^{10}(3+(k\ell_{\mathrm{P}})^2)\left(1+(k\ell_{\mathrm{P}})^{1/2}\right)^2(1+\mathscr{C}_{\mathrm{is,P}})^2\frac{1}{\varepsilon_{\min}}\left(\frac{\ell_2}{\ell_{\mathrm{P}}}+\frac{\ell_{\mathrm{P}}}{\ell_2}\right)\|\widetilde{f}\|^2_{\varepsilon,\widetilde{\Omega}}\\
&\le 400|\nu|^{10}(3+(k\ell_{\mathrm{P}})^2)\left(1+(k\ell_{\mathrm{P}})\right)(1+\mathscr{C}_{\mathrm{is,P}})^2\frac{1}{\varepsilon_{\min}}\left(\frac{\ell_2}{\ell_{\mathrm{P}}}+\frac{\ell_{\mathrm{P}}}{\ell_2}\right)\|\widetilde{f}\|^2_{\varepsilon,\widetilde{\Omega}}\\
&\le 400|\nu|^{10}(\sqrt{3}+k\ell_{\mathrm{P}})^3(1+\mathscr{C}_{\mathrm{is,P}})^2\frac{1}{\varepsilon_{\min}}\left(\frac{\ell_2}{\ell_{\mathrm{P}}}+\frac{\ell_{\mathrm{P}}}{\ell_2}\right)\|\widetilde{f}\|^2_{\varepsilon,\widetilde{\Omega}}
\end{aligned}$$

□

Theorem 3.3.21. *Assume that $\mathscr{C}_{\mathrm{is}}\mathscr{E}_{\mathrm{P}} < 1$. Then, we have*

$$\inf_{\substack{\widetilde{u}\in H^1_\sharp(\widetilde{\Omega})\\ \|\!|\widetilde{u}|\!\|_{k,\widetilde{\Omega}}=1}}\sup_{\widetilde{v}\in H^1_\sharp(\widetilde{\Omega})\atop \|\!|\widetilde{v}|\!\|_{k,\widetilde{\Omega}}=1}\widetilde{b}(\widetilde{u},\widetilde{v}) \ge \frac{1}{|\nu_{\mathrm{P}}|}\frac{1}{1+2\gamma_{\mathrm{r}}\widetilde{\mathscr{C}_{\mathrm{st}}}}. \tag{3.3.59}$$

Proof. Let $\widetilde{u} \in H^1_\sharp(\widetilde{\Omega})$ with $\|\!|\widetilde{u}|\!\|_{k,\widetilde{\Omega}} = 1$. We have

$$\begin{aligned}
\operatorname{Re}\widetilde{b}(\widetilde{u},\widetilde{u}) &= -k^2\|\widetilde{u}\|^2_{\varepsilon,\Omega} + \|\boldsymbol{\nabla}\widetilde{u}\|^2_{\boldsymbol{A},\Omega} - k^2\operatorname{Re}\nu_{\mathrm{P}}\|\widetilde{u}\|^2_{\varepsilon,\Omega_{\mathrm{P}}} + \operatorname{Re}\nu_{\mathrm{P}}\left\|\frac{\partial\widetilde{u}}{\partial\boldsymbol{x}_1}\right\|^2_{A_1,\Omega_{\mathrm{P}}} + \operatorname{Re}\nu_{\mathrm{P}}^{-1}\left\|\frac{\partial\widetilde{u}}{\partial\boldsymbol{x}_2}\right\|^2_{A_2,\Omega_{\mathrm{P}}} \\
&= -k^2\|\widetilde{u}\|^2_{\varepsilon,\Omega} + \|\boldsymbol{\nabla}\widetilde{u}\|^2_{\boldsymbol{A},\Omega} + \left\|\frac{\partial\widetilde{u}}{\partial\boldsymbol{x}_1}\right\|^2_{A_1,\Omega_{\mathrm{P}}} - k^2\gamma_{\mathrm{r}}\|\widetilde{u}\|^2_{\varepsilon,\Omega_{\mathrm{P}}} + \gamma_{\mathrm{r}}\left\|\frac{\partial\widetilde{u}}{\partial\boldsymbol{x}_1}\right\|^2_{A_1,\Omega_{\mathrm{P}}} + \frac{\gamma_{\mathrm{r}}}{|\nu_{\mathrm{P}}|}\left\|\frac{\partial\widetilde{u}}{\partial\boldsymbol{x}_2}\right\|^2_{A_2,\Omega_{\mathrm{P}}} \\
&\geq -\gamma_{\mathrm{r}}k^2\|\widetilde{u}\|^2_{\varepsilon,\widetilde{\Omega}} + \frac{1}{|\nu_{\mathrm{P}}|}\|\boldsymbol{\nabla}\widetilde{u}\|^2_{\boldsymbol{A},\widetilde{\Omega}}.
\end{aligned}$$

Let now $\xi \in H^1_\sharp(\widetilde{\Omega})$ solve

$$\widetilde{b}(w,\xi) = (\varepsilon w, \widetilde{u})_{\widetilde{\Omega}},$$

so that

$$\widetilde{b}(\widetilde{u},\xi) = \|\widetilde{u}\|^2_{\varepsilon,\widetilde{\Omega}}$$

and

$$k\,|\!|\!|\xi|\!|\!|_{k,\widetilde{\Omega}} \leq \widetilde{\mathscr{C}_{\mathrm{st}}}\|\widetilde{u}\|_{\varepsilon,\widetilde{\Omega}} \qquad k^2\,|\!|\!|\xi|\!|\!|_{k,\widetilde{\Omega}} \leq \widetilde{\mathscr{C}_{\mathrm{st}}}\,|\!|\!|\widetilde{u}|\!|\!|_{k,\widetilde{\Omega}}\,.$$

As a result

$$\operatorname{Re}\widetilde{b}(\widetilde{u},\widetilde{u}+2\gamma_{\mathrm{r}}k^2\xi) \geq k^2\gamma_{\mathrm{r}}\|\widetilde{u}\|^2_{\varepsilon,\widetilde{\Omega}} + \frac{1}{\nu_{\mathrm{P}}}\|\boldsymbol{\nabla}\widetilde{u}\|^2_{\boldsymbol{A},\widetilde{\Omega}} \geq \frac{1}{\nu_{\mathrm{P}}}\,|\!|\!|\widetilde{u}|\!|\!|^2_{k,\widetilde{\Omega}}\,.$$

Furthermore, we have

$$|\!|\!|\widetilde{u}+2\gamma_{\mathrm{r}}\xi|\!|\!|_{k,\widetilde{\Omega}} \leq |\!|\!|\widetilde{u}|\!|\!|_{k,\widetilde{\Omega}} + 2\gamma_{\mathrm{r}}k^2\,|\!|\!|\xi|\!|\!|_{k,\widetilde{\Omega}} \leq \left(1+2\gamma_{\mathrm{r}}\widetilde{\mathscr{C}_{\mathrm{st}}}\right)|\!|\!|\widetilde{u}|\!|\!|_{k,\widetilde{\Omega}}\,,$$

so that letting $\widetilde{v} := \widetilde{u} + 2\gamma_{\mathrm{r}}k^2\xi$, we have

$$\operatorname{Re}\widetilde{b}(\widetilde{u},\widetilde{v}) \geq \frac{1}{|\nu_{\mathrm{P}}|}\frac{1}{1+2\gamma_{\mathrm{r}}\widetilde{\mathscr{C}_{\mathrm{st}}}}\,|\!|\!|\widetilde{u}|\!|\!|_{k,\widetilde{\Omega}}\,|\!|\!|\widetilde{v}|\!|\!|_{k,\widetilde{\Omega}}\,.$$

□

Chapter 4

Periodic homogenization of finely textured layers

Contents

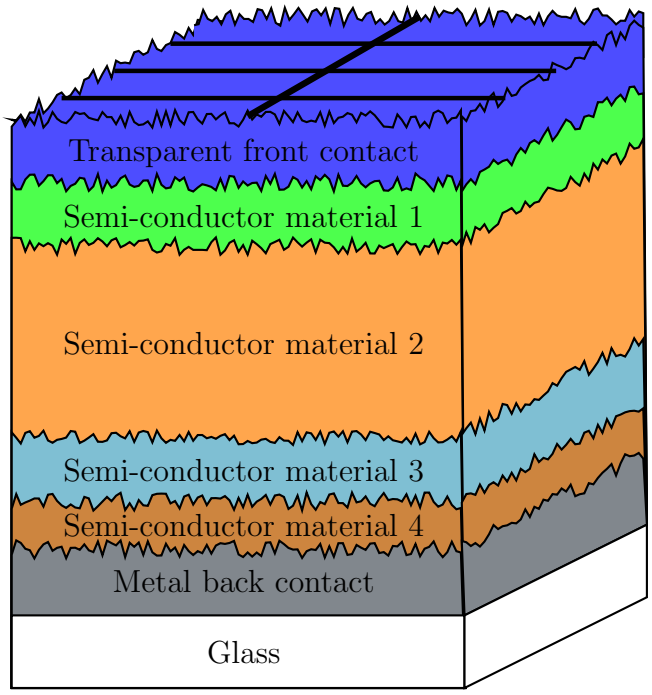


Figure 4.1: Structure of a silicon solar cell.

In this chapter, we consider the case of finely textured layered media. This study is physically motivated by solar cells similar to the one represented on Figure 4.1. Mathematically, we will work with a periodicity assumption and consider coefficients as described on Figure 4.2. Specifically, we want to analyze layers with oscillating interfaces of the form

$$\Gamma_\delta := \left\{ \boldsymbol{x} \in \Omega \mid \boldsymbol{x}_2 = \phi\left(\frac{\boldsymbol{x}_1}{\delta}\right) \right\},$$

where ϕ is a smooth function and $\delta > 0$ is small parameter representing the characteristic length of the layers texturation.

We will analyze this problem through the lens of periodic homogenization theory [48]. More precisely, similar to [35], we will observe that the stability bounds we established in chapters 2 and 3 are not only explicit in frequency, but also uniform in δ. In fact, more generally, our stability bounds are oblivious to variations of the coefficients along the $\boldsymbol{x}_1$ direction. As a result, we may be tempted to "pass to the limit" as $\delta \to 0$. This limit process is usually called "homogenization", as we will show that the solution u_δ converges to the solution u_0 of a Helmholtz problem with non-oscillating coefficients that are often called "effective" or "homogenized" coefficients. We will see that indeed, these homogenized coefficients correspond to some averaging of the oscillating coefficients. Following the standard theory in [48] and adapted to Helmholtz equation in [35], we will provide error estimates between u_0 and u_δ. Crucially, the fact that our stability bounds are uniform in δ will enable us to derive frequency-explicit error estimates controlling the difference between the solutions u_0 of the homogenized problem and u_δ of the oscillating problem.

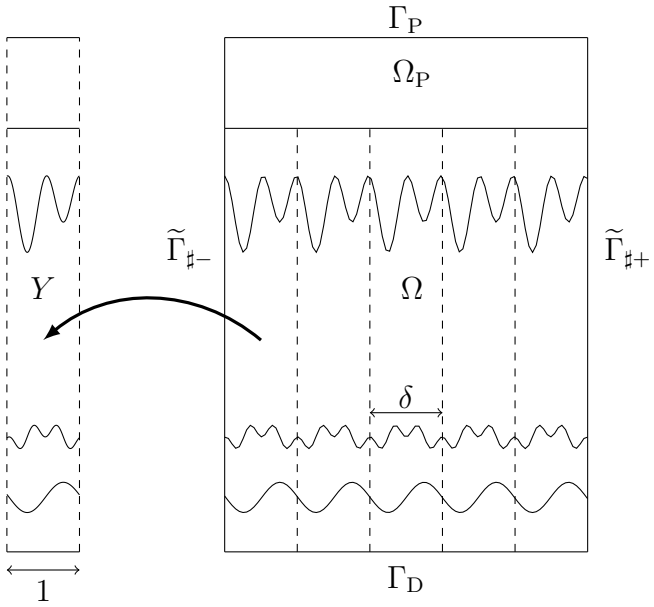


Figure 4.2: Example of the considered domain (modeling a cross section of a solar cell).

Beyond their intrinsic interest, the results we derive in this section are instrumental in the analysis of multiscale numerical methods. In fact, we will also heavily rely on the results on this chapter to analyze the properties of multiscale hybrid-mixed discretization of problems with finely textured layers in section 5.4.

4.1 Model problem

We start by introducing our model problem, and verifying that the stability results of chapters 2 and 3 do apply.

4.1.1 Oscillating coefficients

In the remainder of this chapter, $\delta > 0$ will denote a small real number. Since our problem correspond to one periodic slab of a solar cell structure, we will always implicitly assume that $\delta := \ell_1/M$ for some large integer M.

The concept of "fast" and "slow" variables will be very useful for our analysis. Specifically, we will often work with functions of the form

$$\phi^\delta(\boldsymbol{x}_1, \boldsymbol{x}_2) = \widehat{\phi}\left(\boldsymbol{x}_2, \left\{\frac{\boldsymbol{x}_1}{\delta}\right\}\right)$$

where $\phi(\boldsymbol{x}_2, y)$ is a 1-periodic function in y and $\{t\} := t - \lfloor t \rfloor$ denotes the fractional part of $t \in \mathbb{R}$. In particular, the functions describing the physical coefficients corresponding to Figure 4.2 are of this form. In this context, $\boldsymbol{x}_1$ is known as the slow variable, whereas we say that y is the fast variable. Thus, the following notations will be useful.

In the remainder of this chapter we set $Y := (0,1)$ for the periodic cell in which the fast variable y lives. If $\widehat{v} : (0, \ell_2 + \ell_{\mathrm{P}}) \times Y \to \mathbb{C}$, we employ the notations

$$\langle \widehat{v} \rangle_Y(\boldsymbol{x}_2) := \int_Y \widehat{v}(\boldsymbol{x}_2, y) dy \qquad \forall \boldsymbol{x}_2 \in (0, \ell_2 + \ell_{\mathrm{P}}),$$

and

$$\widehat{v}^\delta(\boldsymbol{x}) = \widehat{v}\left(\boldsymbol{x}_2, \left\{\frac{\boldsymbol{x}_1}{\delta}\right\}\right) \qquad \forall \boldsymbol{x} \in \widetilde{\Omega}.$$

The following identities

$$\frac{\partial \widehat{v}^\delta}{\partial \boldsymbol{x}_1} = \frac{1}{\delta}\left(\frac{\partial \widehat{v}}{\partial y}\right)^\delta \qquad \frac{\partial \widehat{v}^\delta}{\partial \boldsymbol{x}_2} = \left(\frac{\partial \widehat{v}}{\partial \boldsymbol{x}_2}\right)^\delta \tag{4.1.1}$$

are simple, and will often be useful.

In this chapter, we will work with coefficients that are highly oscillating in the $\boldsymbol{x}_1$ direction. Specifically, we consider functions $\widehat{\varepsilon}, \widehat{A}_1, \widehat{A}_2 : (0, \ell_2 + \ell_{\mathrm{P}}) \times Y \to \mathbb{R}$ such that: (i) $\varepsilon, \widehat{A}_1$ and $\widehat{A}_2$ are smooth, (ii) $\widehat{\varepsilon}$ is increasing with $\boldsymbol{x}_2$ and $\widehat{A}_1$ and $\widehat{A}_2$ are decreasing with $\boldsymbol{x}_2$ and (iii) $\varepsilon \equiv 1$, and $\widehat{A}_1 \equiv \widehat{A}_2 \equiv 1$ for $\boldsymbol{x}_2 \in (\ell_2, \ell_2 + \ell_{\mathrm{P}})$.

An important consequence of those assumptions is that the corresponding oscillating coefficients $\widehat{\varepsilon}^\delta$, and $\widehat{\boldsymbol{A}}^\delta$ satisfy the requirements of Assumption 2.1.1 with one layer ($N = 1$). We also notice that our assumptions on the coefficients imply that $\widehat{\varepsilon} \leq \varepsilon_{\max} := 1$ and $A_1, A_2 \geq A_{\min} =: 1$. We will denote by $\varepsilon_{\min}$ and $A_{\max}$ the other bounds of the coefficients.

Remark 4.1.1 (Periodicity assumption)**.** *Notice that due to our periodicity assumption, the whole solar cell structure is periodic with period δ. As explained in the introduction, it means that in principle, we could reformulate the whole problem on a periodic slab of size δ, i.e. with $\ell_1 = \delta$. However, the periodicity assumption we make in this chapter is only of a technical nature, and what we are striving for are layers with a rough texture that are part of a periodic pattern, as described for instance on Figure 4.1.*

Remark 4.1.2 (Smoothness assumption)**.** *For technical reasons, we need to assume that the coefficients are smooth. It means that they do not exactly represent layers with sharp boundaries as described on Figure 4.2 as they are continuous along the $\boldsymbol{x}_2$ direction. In fact, our coefficients $\widehat{\varepsilon}^\delta$ and $\widehat{\boldsymbol{A}}^\delta$ really correspond to a "blurry" version of what is represented on Figure 4.2.*

4.1.2 Homogenized coefficients

As $\delta \to 0$, one expects the solution to only see an homogenized version of $\widehat{\varepsilon}^\delta$ and $\widehat{\boldsymbol{A}}^\delta$. A naive guess would be that the corresponding effective coefficient is $\langle \varepsilon \rangle_Y$ and $\langle \widehat{\boldsymbol{A}} \rangle_Y$, but the reality is more complicated. In general, the expression of the effective parameters is not explicit, and involves a boundary value problem on the periodic cell Y as shown in [48,

sections 6.1 and 6.2]. Here, because the periodic cell is one dimensional, explicit expressions are actually available, and they are as follows

$$\varepsilon^{\mathrm{H}} := \langle \widehat{\varepsilon} \rangle_Y \qquad A_1^{\mathrm{H}} := 1/\langle \widehat{A}_1^{-1} \rangle_Y \qquad A_2^{\mathrm{H}} := \langle \widehat{A}_2 \rangle_Y. \tag{4.1.2}$$

The situation is in fact similar to what is considered in [48, section 5.4]. Notice that even in this simpler case, the homogenization formula is not obvious: it involves both the arithmetic and the harmonic means.

Since in this chapter we will be considering a family of Helmholtz problems (parameterized by δ), it will be convenient to work with a single energy norm. As a result, we introduce

$$|\!|\!| v |\!|\!|_{k,\widetilde{\Omega}}^2 := k^2 \|v\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}}^2 + \|\boldsymbol{\nabla} v\|_{\boldsymbol{A}^{\mathrm{H}},\widetilde{\Omega}}^2, \tag{4.1.3}$$

for all $v \in H^1_\sharp(\widetilde{\Omega})$.

In this chapter, we assume that the PML parameters are chosen such that $\mathscr{E}_{\mathrm{P}}$ is sufficiently small. In addition, in the remaining, ℓ will denote the diameter of $\widetilde{\Omega}$ and we will assume for simplicity that $k\ell \geq 1$. As tracking the dependency of multiplicative constants of this section will be tedious, we introduce the following notation: for $A, B \geq 0$, we will write $A \lesssim B$ if there exists a constant $C > 0$ which is independent of A, B, δ and k, but which possibly depends on $\widetilde{\Omega}$, $\varepsilon_{\min}$, $A_{\max}$, ν,

$$\ell \left\| \frac{\partial \widehat{\varepsilon}}{\partial \boldsymbol{x}_2} \right\|_{L^\infty(\widetilde{\Omega})}, \quad \ell \left\| \frac{\partial \widehat{A}_1}{\partial \boldsymbol{x}_2} \right\|_{L^\infty(\widetilde{\Omega})} \quad \text{and} \quad \ell \left\| \frac{\partial \widehat{A}_2}{\partial \boldsymbol{x}_2} \right\|_{L^\infty(\widetilde{\Omega})},$$

such that $A \leq CB$. We also write $A \gtrsim B$ when $B \lesssim A$.

We close this subsection with some simple properties of the homogenized coefficients.

Lemma 4.1.3 (Simple properties of the homogenized coefficients)**.** *We have $\varepsilon^{\mathrm{H}}, A_1^{\mathrm{H}}, A_2^{\mathrm{H}} \in C^{1,1}(\overline{\widetilde{\Omega}})$. In addition ε^{H}, A_1^{H} and A_2^{H} only depend on the $\boldsymbol{x}_2$ variable, and we have*

$$0 \leq \frac{\partial \varepsilon^{\mathrm{H}}}{\partial \boldsymbol{x}_2} \lesssim \frac{1}{\ell} \quad -\frac{1}{\ell} \lesssim \frac{\partial A_1^{\mathrm{H}}}{\partial \boldsymbol{x}_2} \leq 0 \quad -\frac{1}{\ell} \lesssim \frac{\partial A_2^{\mathrm{H}}}{\partial \boldsymbol{x}_2} \leq 0, \tag{4.1.4}$$

and

$$\varepsilon_{\min} \leq \varepsilon^{\mathrm{H}} \leq 1 \quad 1 \leq A_1^{\mathrm{H}} \leq A_{\max} \quad 1 \leq A_2^{\mathrm{H}} \leq A_{\max}.$$

Proof. Recalling the expressions (4.1.2), it is clear that the homogenized coefficients are independent of y which implies their independence from $\boldsymbol{x}_1$. Also, the $C^{1,1}$-smoothness of ε^{H}, A_1^{H} and A_2^{H} is a direct consequence of the smoothness assumption on $\widehat{\varepsilon}$, $\widehat{A}_1$ and $\widehat{A}_2$. In addition, we use the monotonicity condition on ε^{H}, A_1^{H} and A_2^{H} to obtain the bounds 0 at (4.1.4), and given that

$$\frac{\partial \varepsilon^{\mathrm{H}}}{\partial \boldsymbol{x}_2}(\boldsymbol{x}) = \left\langle \frac{\partial \widehat{\varepsilon}^\delta}{\partial \boldsymbol{x}_2} \right\rangle_Y (\boldsymbol{x}_1) = \left\langle \left(\frac{\partial \widehat{\varepsilon}}{\partial \boldsymbol{x}_2} \right)^\delta \right\rangle_Y (\boldsymbol{x}_1) \lesssim \left| \frac{\partial \widehat{\varepsilon}}{\partial \boldsymbol{x}_2}(\boldsymbol{x}_1) \right| \lesssim \frac{1}{\ell},$$

the other bounds at (4.1.4) follow since the hidden constants in $\lesssim$ may depends on

$$\ell \left\| \frac{\partial \widehat{\varepsilon}}{\partial \boldsymbol{x}_2} \right\|_{L^\infty(\widetilde{\Omega})}, \quad \ell \left\| \frac{\partial \widehat{A}_1}{\partial \boldsymbol{x}_2} \right\|_{L^\infty(\widetilde{\Omega})} \quad \text{and} \quad \ell \left\| \frac{\partial \widehat{A}_2}{\partial \boldsymbol{x}_2} \right\|_{L^\infty(\widetilde{\Omega})}.$$

On the other hand, since $\varepsilon_{\min} \leq \widehat{\varepsilon} \leq 1$ and $1 \leq A_1, A_2 \leq A_{\max}$, then by averaging over Y we have

$$\varepsilon_{\min} \leq \varepsilon^{\mathrm{H}} \leq 1 \quad 1 \leq A_1^{\mathrm{H}} \leq A_{\max} \quad 1 \leq A_2^{\mathrm{H}} \leq A_{\max}.$$

□

4.1.3 Oscillating problem

We are now ready to state our model problem. For $\delta > 0$ and $f \in L^2(\widetilde{\Omega})$, it consists in finding $u_\delta \in H^1_\sharp(\widetilde{\Omega})$ such that

$$b_\delta(u_\delta, v) = (\varepsilon^{\mathrm{H}} f, v)_{\widetilde{\Omega}} \qquad \forall v \in H^1_\sharp(\widetilde{\Omega}), \tag{4.1.5}$$

where the sesquilinear form is given by $b_\delta(\phi, v) := b_\delta^{\widetilde{\Omega}}(\phi, v)$ where

$$b_\delta^D(\phi, v) = -k^2(\nu \varepsilon^\delta \phi, v)_D + \left(\nu A_1^\delta \frac{\partial \phi}{\partial \boldsymbol{x}_1}, \frac{\partial v}{\partial \boldsymbol{x}_1} \right)_D + \left(\nu^{-1} A_2^\delta \frac{\partial \phi}{\partial \boldsymbol{x}_2}, \frac{\partial v}{\partial \boldsymbol{x}_2} \right)_D$$

for all $\phi, v \in H^1(D)$ and open sets $D \subset \widetilde{\Omega}$. We will refer to (4.1.5) as the "oscillating problem". The notation $b_\delta^D(\cdot, \cdot)$ will be useful in the proofs to work on subsets of $\widetilde{\Omega}$.

Notice that we employed the oscillating coefficients to define the sesquilinear form $b_\delta(\cdot, \cdot)$, but used the homogenized coefficients ε^{H} in the right-hand side. This choice is mostly done for technical reasons, in order to simplify the analysis. However, it is not an oversimplification in practice. Indeed, (i) in applications, the right-hand side f usually corresponds to an incoming field, and is supported near the top of Ω where $\widehat{\varepsilon}^\delta = \varepsilon^{\mathrm{H}} = 1$. Besides (ii), we will see in Chapter 5 that this setting is sufficient to derive optimal error estimates for multiscale numerical methods.

As we next show, the stability result derived in Theorem 2.3.3 applies straightforwardly here. In fact, a single slight adjustment is needed in order to account for the fact that we consider an energy norm with the homogenized coefficients instead of the oscillating coefficients.

Lemma 4.1.4 (Uniform stability of the oscillating problem)**.** *For all $\delta > 0$, and $f \in L^2(\widetilde{\Omega})$, problem (4.1.5) admits a unique solution $u_\delta \in H^1_\sharp(\widetilde{\Omega})$. In addition, the estimate*

$$k \, |\!|\!| u_\delta |\!|\!|_{k, \widetilde{\Omega}} \lesssim \mathscr{C}_{\mathrm{st}} \|f\|_{\varepsilon^{\mathrm{H}}, \widetilde{\Omega}},$$

holds uniformly in δ with

$$\mathscr{C}_{\mathrm{st}} := (k\ell)^{9/2}. \tag{4.1.6}$$

Proof. Assume that $\mathscr{C}_{\mathrm{is}}\mathscr{E}_{\mathrm{P}} < 1$, Theorem 3.3.19 from Chapter 3 implies that the problem (4.1.5) admits a unique solution $u_\delta \in H^1_\sharp(\widetilde{\Omega})$ and that

$$k\,|\!|\!|u_\delta|\!|\!|_{k,\widetilde{\Omega}} \lesssim (k\ell)^{9/2}\|f\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}},$$

the hidden constant depends on $\widetilde{\Omega}$, ν, $A_{\max}$ and $\varepsilon_{\min}$. □

4.1.4 Homogenized problem

In the homogenized problem, we replace the oscillating coefficients $\widehat{\varepsilon}^\delta$ and $\widehat{\boldsymbol{A}}^\delta$ by their effective counterparts ε^{H} and $\boldsymbol{A}^{\mathrm{H}}$. Namely, for $f \in L^2(\widetilde{\Omega})$, the "homogenized problem" consists in finding $u_0 \in H^1_\sharp(\widetilde{\Omega})$ such that

$$b_{\mathrm{H}}(u_0, v) = (\varepsilon^{\mathrm{H}} f, v)_{\widetilde{\Omega}} \qquad \forall v \in H^1_\sharp(\widetilde{\Omega}), \tag{4.1.7}$$

where the sesquilinear form is given by $b_{\mathrm{H}}(\phi, v) := b^{\widetilde{\Omega}}_{\mathrm{H}}(\phi, v)$ where

$$b^D_{\mathrm{H}}(\phi, v) := -k^2(\nu\varepsilon^{\mathrm{H}}\phi, v)_D + \left(\nu A^{\mathrm{H}}_1 \frac{\partial\phi}{\partial \boldsymbol{x}_1}, \frac{\partial v}{\partial \boldsymbol{x}_1}\right)_D + \left(\nu^{-1} A^{\mathrm{H}}_2 \frac{\partial\phi}{\partial \boldsymbol{x}_2}, \frac{\partial v}{\partial \boldsymbol{x}_2}\right)_D,$$

for all $\phi, v \in H^1(D)$ and open sets $D \subset \widetilde{\Omega}$.

Lemma 4.1.5 (Stability of the homogenized problem)**.** *For all $f \in L^2(\widetilde{\Omega})$, there exists a unique solution $u_0 \in H^1_\sharp(\widetilde{\Omega})$ to (4.1.7), and the estimate*

$$k\,|\!|\!|u_0|\!|\!|_{k,\widetilde{\Omega}} \lesssim \mathscr{C}_{\mathrm{st}}\|f\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}},$$

holds true with $\mathscr{C}_{\mathrm{st}}$ given by (4.1.6).

4.2 Convergence analysis

We now proceed to show that the solution u_δ to the oscillating problem converges to the solution u_0 of the homogenized problem in an appropriate sense as $\delta \to 0$. We will prove this results using the "multiple-scale" method described in [48, Chapter 7]. This approach has recently been employed to analyze a Helmholtz problem [35], but in a different setting. Specifically, the problem considered in [35] describes scattering by a bounded obstacle, so that no quasi-periodic boundary conditions are involved. Besides, fine layers are considered in [35], which turn out to leads to a different proof than the finely textured layers we consider here.

4.2.1 Correctors

Following [48, Section 7.1], we need to introduce specific functions $(0, \ell_2 + \ell_{\mathrm{P}}) \times Y \to \mathbb{R}$ to perform the convergence analysis. Such functions are often called "correctors" in this context. In general, these functions are defined as the solution to PDEs on the reference cell Y, and there analytical expression is not explicitly available. Here, because $\widehat{\varepsilon}^\delta$ and $\widehat{\boldsymbol{A}}^\delta$ only oscillate in the $\boldsymbol{x}_1$ direction, these PDEs are actually one-dimensional problems for which we can write down closed form formula.

We start by introducing the "first-order corrector".

Lemma 4.2.1 (First-order corrector)**.** *Define* $\widehat{\chi} : (0, \ell_2 + \ell_{\mathrm{P}}) \times Y \to \mathbb{R}$ *as*

$$\widehat{\chi}(\boldsymbol{x}_2, y) := y - A_1^{\mathrm{H}}(\boldsymbol{x}_2) \int_0^y \widehat{A}_1^{-1}(\boldsymbol{x}_2, z) dz \quad \forall (\boldsymbol{x}_2, y) \in (0, \ell_2 + \ell_{\mathrm{P}}) \times Y. \tag{4.2.1}$$

Then, we have

$$\widehat{A}_1 \frac{\partial \widehat{\chi}}{\partial y} = \widehat{A}_1 - A_1^{\mathrm{H}}. \tag{4.2.2}$$

In addition, we have $\widehat{\chi}^\delta \in C^{1,1}(\overline{\widetilde{\Omega}})$ *with*

$$\|\widehat{\chi}^\delta\|_{L^\infty(\widetilde{\Omega})} \lesssim 1, \quad \left\| \frac{\partial \widehat{\chi}^\delta}{\partial \boldsymbol{x}_1} \right\|_{L^\infty(\widetilde{\Omega})} \lesssim \delta^{-1}, \quad \left\| \frac{\partial \widehat{\chi}^\delta}{\partial \boldsymbol{x}_2} \right\|_{L^\infty(\widetilde{\Omega})} \lesssim \ell^{-1}. \tag{4.2.3}$$

Besides $\widehat{\chi}^\delta = 0$ *on* Ω_{P}.

Proof. The fact that (4.2.2) holds true is a direct consequence of the definition we chose for $\widehat{\chi}$ in (4.2.1). It is clear that $\widehat{\chi}$ is smooth, due to the smoothness assumption on $\widehat{A}_1$. To ensure that $\widehat{\chi}^\delta \in C^{1,1}(\overline{\widetilde{\Omega}})$, we must additionally check that

$$\widehat{\chi}(\boldsymbol{x}_2, 0) = \widehat{\chi}(\boldsymbol{x}_2, 1) \qquad \frac{\partial \widehat{\chi}}{\partial \boldsymbol{x}_2}(\boldsymbol{x}_2, 0) = \frac{\partial \widehat{\chi}}{\partial \boldsymbol{x}_2}(\boldsymbol{x}_2, 1),$$

for all $\boldsymbol{x}_2 \in (0, \ell_2 + \ell_{\mathrm{P}})$. The function

$$\phi : \boldsymbol{x}_2 \to A_1^{\mathrm{H}}(\boldsymbol{x}_2) \int_0^1 \widehat{A}^{-1}(\boldsymbol{x}_2, y) dy,$$

will be useful. Observe that $\phi = A_1^{\mathrm{H}} \langle \widehat{A}^{-1} \rangle_Y = 1$, due to the definition of A_1^{H}, so that

$$\phi \equiv 1 \qquad \frac{\partial \phi}{\partial \boldsymbol{x}_2} \equiv 0.$$

We then observe that

$$\widehat{\chi}(\boldsymbol{x}_2, 1) = 1 - \phi(\boldsymbol{x}_2) = 0 = \widehat{\chi}(\boldsymbol{x}_2, 0),$$

for all $\boldsymbol{x}_2 \in (0, \ell_2 + \ell_{\rm P})$. On the other hand,

$$\frac{\partial \widehat{\chi}}{\partial \boldsymbol{x}_2}(\boldsymbol{x}_2, 0) = \frac{\partial y}{\partial \boldsymbol{x}_2} = 0,$$

and

$$\frac{\partial \widehat{\chi}}{\partial \boldsymbol{x}_2}(\boldsymbol{x}_2, 1) = \frac{\partial \phi}{\partial \boldsymbol{x}_2}(\boldsymbol{x}_2) = 0,$$

so that

$$\frac{\partial \widehat{\chi}}{\partial \boldsymbol{x}_2}(\boldsymbol{x}_2, 0) = \frac{\partial \widehat{\chi}}{\partial \boldsymbol{x}_2}(\boldsymbol{x}_1, 1).$$

To show that $\widehat{\chi}^\delta = 0$ on $\Omega_{\rm P}$, it suffices to observe that

$$\widehat{\chi}(\boldsymbol{x}_2, y) = y - A_1^{\rm H}(\boldsymbol{x}_2) \int_0^y \widehat{A}_1^{-1}(\boldsymbol{x}_2, z) dz = y - \int_0^y dz = 0,$$

since $\widehat{A}_1^{-1}(\boldsymbol{x}_2, y) = \langle \widehat{A}^{-1} \rangle_Y(\boldsymbol{x}_2)$ for all $\boldsymbol{x}_2 \in (\ell_2, \ell_2 + \ell_{\rm P})$, as $\widehat{A}_1$ does not depend on y if $\boldsymbol{x}_2 > \ell_2$. □

We next introduce another key function $\widehat{\tau}$. The function $\widehat{\tau}$ is specifically required in our analysis because the coefficients only oscillate in one-direction instead of two.

Lemma 4.2.2 (Corrector for the non-oscillating direction)**.** *Define the function* $\widehat{\tau} : (0, \ell_2 + \ell_{\rm P}) \times Y \to \mathbb{R}$ *by*

$$\widehat{\tau}(\boldsymbol{x}_2, y) = \frac{1}{A_2^{\rm H}(\boldsymbol{x}_2)} \int_0^y \widehat{A}_2(\boldsymbol{x}_2, z) dz - y \quad \forall (\boldsymbol{x}_2, y) \in (0, \ell_2 + \ell_{\rm P}) \times Y. \tag{4.2.4}$$

Then, we have

$$A_2^{\rm H} \frac{\partial \widehat{\tau}}{\partial y} = \widehat{A}_2 - A_2^{\rm H}, \tag{4.2.5}$$

In addition, $\widehat{\tau}^\delta \in C^{1,1}(\overline{\widetilde{\Omega}})$ *and*

$$\|\widehat{\tau}^\delta\|_{L^\infty(\widetilde{\Omega})} \lesssim 1 \qquad \|\boldsymbol{\nabla} \widehat{\tau}^\delta\|_{L^\infty(\Omega)} \lesssim \delta^{-1}. \tag{4.2.6}$$

Proof. On can readily check that (4.2.5) holds true based on (4.2.4). Besides, it is clear that $\widehat{\tau}$ is smooth, and we only need to verify periodicity conditions to ensure that $\widehat{\tau}^\delta$ is smooth as well. Due to the definition of $A_2^{\rm H}$, the function $\phi(\boldsymbol{x}_2) := \langle \widehat{A}_2 \rangle_Y(\boldsymbol{x}_2) / A_2^{\rm H}(\boldsymbol{x}_2)$ satisfies $\phi \equiv 1$ on $(0, \ell_2 + \ell_{\rm P})$. We then observe that

$$\widehat{\tau}(\boldsymbol{x}_2, 1) = 1 - \phi(\boldsymbol{x}_2) = 0 = \widehat{\tau}(\boldsymbol{x}_2, 0).$$

On the other hand, we have

$$\frac{\partial \widehat{\tau}}{\partial \boldsymbol{x}_2}(\boldsymbol{x}_2, 1) = \phi'(\boldsymbol{x}_2) = 0,$$

and

$$\frac{\partial \widehat{\tau}}{\partial \boldsymbol{x}_2}(\boldsymbol{x}_2, 0) = -\frac{\partial y}{\partial \boldsymbol{x}_2} = 0.$$

□

Finally $\widehat{\eta}$ is the so-called "second-order" corrector. The properties of $\widehat{\eta}$ are obtained exactly as the ones of $\widehat{\tau}$, so that for the sake of shortness, we do not repeat the proof.

Lemma 4.2.3 (Second-order corrector)**.** *Let*

$$\widehat{\eta}(\boldsymbol{x}_2, y) := \frac{1}{\varepsilon^{\mathrm{H}}(\boldsymbol{x}_2)} \int_0^y \widehat{\varepsilon}(\boldsymbol{x}_2, z) dz - y, \tag{4.2.7}$$

for all $(\boldsymbol{x}_1, y) \in (0, \ell_2 + \ell_{\mathrm{P}}) \times Y$*. Then, we have*

$$\varepsilon^{\mathrm{H}} \frac{\partial \widehat{\eta}}{\partial y} = \widehat{\varepsilon} - \varepsilon^{\mathrm{H}}, \tag{4.2.8}$$

and

$$\|\widehat{\eta}^{\delta}\|_{L^\infty(\widetilde{\Omega})} \lesssim 1. \tag{4.2.9}$$

4.2.2 Technical results

We pursue our convergence analysis with some preliminary results that employs the functions $\widehat{\eta}$, $\widehat{\chi}$ and $\widehat{\tau}$ introduced and analyzed in subsection 4.2.1. In this subsection, $D \subset \widetilde{\Omega}$ is a fixed Lipschitz domain with diameter h_D. The situations we have in mind are when $D = \Omega$ or $D = K$ for an element $K \in \mathcal{T}_H$.

We recall that the following multiplicative trace inequality

$$\|v\|_{\partial D}^2 \leq C_{\mathrm{tr}}(D) \left(\frac{1}{h_D} \|v\|_D^2 + \|v\|_D \|\boldsymbol{\nabla} v\|_D \right), \tag{4.2.10}$$

always holds true [72, Theorem 1.5.1.10], and in the remaining of this subsection, we allow the hidden constant in the $\lesssim$ notation to depend on $C_{\mathrm{tr}}(D)$.

Then, we observe that from the definition of the coefficients, we have

$$\|\boldsymbol{\nabla} v\|_{\varepsilon^{\mathrm{H}}, D} \lesssim |\!|\!| v |\!|\!|_{k,D} \quad \forall v \in H^1(D). \tag{4.2.11}$$

Finally, an easy consequence of (4.2.10) and (4.2.11) is that

$$k \|v\|_{\varepsilon^{\mathrm{H}}, \partial D}^2 \lesssim \left(1 + \frac{1}{k h_D} \right) |\!|\!| v |\!|\!|_{k,D}^2 \quad \forall v \in H^1(D). \tag{4.2.12}$$

Our first technical result concerns the L^2 inner product arising in the sesquilinear forms $b_\delta(\cdot,\cdot)$ and $b_{\mathrm{H}}(\cdot,\cdot)$.

Lemma 4.2.4. *For all* $u_0, v \in H^1(D)$*, we have*

$$\left| k^2 \left((\widehat{\varepsilon}^{\delta} - \varepsilon^{\mathrm{H}}) u_0, v \right)_D \right| \lesssim \left(k\delta + \frac{\delta}{h_D} \right) |\!|\!| u_0 |\!|\!|_{k,D} \, |\!|\!| v |\!|\!|_{k,D} \, . \tag{4.2.13}$$

Proof. Let $u_0, v \in H^1(D)$. Recalling (4.2.8) from Lemma 4.2.3, we have

$$\widehat{\varepsilon}^\delta - \varepsilon^{\mathrm{H}} = (\widehat{\varepsilon} - \varepsilon^{\mathrm{H}})^\delta = \left(\varepsilon^{\mathrm{H}} \frac{\partial \widehat{\eta}}{\partial y}\right)^\delta = \delta \varepsilon^{\mathrm{H}} \frac{\partial \widehat{\eta}^\delta}{\partial \boldsymbol{x}_1},$$

and therefore

$$\left((\widehat{\varepsilon}^\delta - \varepsilon^{\mathrm{H}}) u_0, v\right)_D = \delta \left(\varepsilon^{\mathrm{H}} \frac{\partial \widehat{\eta}^\delta}{\partial \boldsymbol{x}_1} u_0, v\right)_D.$$

Recalling that ε^{H} does not depend on $\boldsymbol{x}_1$, and integrating by parts, we have

$$\begin{aligned}
\left(\varepsilon^{\mathrm{H}} \frac{\partial \widehat{\eta}^\delta}{\partial \boldsymbol{x}_1} u_0, v\right)_D &= \int_D \frac{\partial \widehat{\eta}^\delta}{\partial \boldsymbol{x}_1} (\varepsilon^{\mathrm{H}} u_0 \overline{v}) \\
&= \int_{\partial D} \varepsilon^{\mathrm{H}} \widehat{\eta}^\delta u_0 \overline{v} \boldsymbol{n}_1 - \int_D \widehat{\eta}^\delta \frac{\partial}{\partial \boldsymbol{x}_1} (\varepsilon^{\mathrm{H}} u_0 \overline{v}) \\
&= \int_{\partial D} \varepsilon^{\mathrm{H}} \widehat{\eta}^\delta u_0 \overline{v} \boldsymbol{n}_1 - \int_D \varepsilon^{\mathrm{H}} \widehat{\eta}^\delta \left(\frac{\partial u_0}{\partial \boldsymbol{x}_1} \overline{v} + u_0 \frac{\partial \overline{v}}{\partial \boldsymbol{x}_1}\right),
\end{aligned}$$

leading to

$$\left|k^2 \left((\widehat{\varepsilon}^\delta - \varepsilon^{\mathrm{H}}) u_0, v\right)_D\right| \lesssim k\delta \left(k \|u_0\|_{\varepsilon^{\mathrm{H}},\partial D} \|v\|_{\varepsilon^{\mathrm{H}},\partial D} + \left\|\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right\|_{\varepsilon^{\mathrm{H}},D} k \|v\|_{\varepsilon^{\mathrm{H}},D} + k \|u_0\|_{\varepsilon^{\mathrm{H}},D} \left\|\frac{\partial v}{\partial \boldsymbol{x}_1}\right\|_{\varepsilon^{\mathrm{H}},D}\right),$$

where we employed (4.2.9) to estimate $\|\eta^\delta\|_{L^\infty(D)}$. Then, (4.2.13) follows since (4.2.11) and (4.2.12) imply that

$$k \|u_0\|_{\varepsilon^{\mathrm{H}},\partial D} \|v\|_{\varepsilon^{\mathrm{H}},\partial D} + \left\|\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right\|_{\varepsilon^{\mathrm{H}},D} k \|v\|_{\varepsilon^{\mathrm{H}},D} + k \|u_0\|_{\varepsilon^{\mathrm{H}},D} \left\|\frac{\partial v}{\partial \boldsymbol{x}_1}\right\|_{\varepsilon^{\mathrm{H}},D} \lesssim \left(1 + \frac{1}{kh_D}\right) |\!|\!| u_0 |\!|\!|_{k,D} \, |\!|\!| v |\!|\!|_{k,D}.$$

□

Our next result deals with the inner product involving the $\boldsymbol{x}_1$ derivatives in the sesquilinear forms.

Lemma 4.2.5. *For all $u_0 \in H^2(D)$ and $v \in H^1(D)$, we have*

$$\begin{aligned}
&\left|b_\delta^D \left(\delta \widehat{\chi}^\delta \frac{\partial u_0}{\partial \boldsymbol{x}_1}, v\right) - \left(\nu (A_1^\delta - A_1^{\mathrm{H}}) \frac{\partial u_0}{\partial \boldsymbol{x}_1}, \frac{\partial v}{\partial \boldsymbol{x}_1}\right)_D\right| \\
&\qquad \lesssim \left(\frac{\delta}{\ell} |\!|\!| u_0 |\!|\!|_{k,D} + \delta \left|\!\left|\!\left| \frac{\partial u_0}{\partial \boldsymbol{x}_1} \right|\!\right|\!\right|_{k,D}\right) |\!|\!| v |\!|\!|_{k,D}. \quad (4.2.14)
\end{aligned}$$

Proof. Recall the function $\widehat{\chi}$ from Lemma 4.2.1. Using (4.1.1), we write that

$$\frac{\partial}{\partial \boldsymbol{x}_1} \left(\widehat{\chi}^\delta \frac{\partial u_0}{\partial \boldsymbol{x}_1}\right) = \widehat{\chi}^\delta \frac{\partial^2 u_0}{\partial \boldsymbol{x}_1^2} + \frac{\partial \widehat{\chi}^\delta}{\partial \boldsymbol{x}_1} \frac{\partial u_0}{\partial \boldsymbol{x}_1} = \widehat{\chi}^\delta \frac{\partial^2 u_0}{\partial \boldsymbol{x}_1^2} + \frac{1}{\delta} \left(\frac{\partial \widehat{\chi}}{\partial y}\right)^\delta \frac{\partial u_0}{\partial \boldsymbol{x}_1},$$

which, after multiplying by $\delta\nu\widehat{A}_\delta^1$, leads to

$$\delta\nu\widehat{A}_1^\delta\frac{\partial}{\partial \boldsymbol{x}_1}\left(\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right)=\delta\nu\widehat{A}_1^\delta\widehat{\chi}^\delta\frac{\partial^2 u_0}{\partial \boldsymbol{x}_1^2}+\nu\left(\widehat{A}_1\frac{\partial\widehat{\chi}}{\partial y}\right)^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1}=\delta\nu\widehat{A}_1^\delta\widehat{\chi}^\delta\frac{\partial^2 u_0}{\partial \boldsymbol{x}_1^2}+\nu(\widehat{A}_1^\delta-A_1^{\mathrm{H}})\frac{\partial u_0}{\partial \boldsymbol{x}_1},$$

where we have used (4.2.2) in the last equality. As a result, for all $v\in H^1(D)$, we have

$$\left(\nu(\widehat{A}_1^\delta-A_1^{\mathrm{H}})\frac{\partial u_0}{\partial \boldsymbol{x}_1},\frac{\partial v}{\partial \boldsymbol{x}_1}\right)_D=\left(\nu\widehat{A}_1^\delta\frac{\partial}{\partial \boldsymbol{x}_1}\left(\delta\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right),\frac{\partial v}{\partial \boldsymbol{x}_1}\right)_D-\delta\left(\nu\widehat{A}_1^\delta\chi^\delta\frac{\partial^2 u_0}{\partial \boldsymbol{x}_1^1},\frac{\partial v}{\partial \boldsymbol{x}_1}\right)_D,$$

and since

$$\begin{aligned}\left(\nu\widehat{A}_1^\delta\frac{\partial}{\partial \boldsymbol{x}_1}\left(\delta\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right),\frac{\partial v}{\partial \boldsymbol{x}_1}\right)_D&=b_\delta^D\left(\delta\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1},v\right)\\&\quad+k^2\left(\widehat{\varepsilon}^\delta\delta\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1},v\right)_D-\left(\widehat{A}_2^\delta\frac{\partial}{\partial \boldsymbol{x}_2}\left(\delta\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right),\frac{\partial v}{\partial \boldsymbol{x}_2}\right)_D,\end{aligned}$$

we arrive at

$$\begin{aligned}&\left(\nu(\widehat{A}_1^\delta-A_1^{\mathrm{H}})\frac{\partial u_0}{\partial \boldsymbol{x}_1},\frac{\partial v}{\partial \boldsymbol{x}_1}\right)_D-b_\delta^D\left(\delta\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1},v\right)=\\&\qquad k^2\delta\left(\widehat{\varepsilon}^\delta\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1},v\right)_D-\delta\left(\widehat{A}_2^\delta\frac{\partial}{\partial \boldsymbol{x}_2}\left(\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right),\frac{\partial v}{\partial \boldsymbol{x}_2}\right)_D-\delta\left(\nu\widehat{A}_1^\delta\chi^\delta\frac{\partial^2 u_0}{\partial \boldsymbol{x}_1^1},\frac{\partial v}{\partial \boldsymbol{x}_1}\right)_D,\end{aligned}$$

and it remains to bound the three terms in the right-hand side.

First, using (4.2.3), and by definition of the energy norm, we immediatly have

$$k^2\delta\left|\left(\widehat{\varepsilon}^\delta\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1},v\right)_D\right|\lesssim k^2\delta\left\|\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right\|_{\widehat{\varepsilon}^\delta,D}\|v\|_{\widehat{\varepsilon}^\delta,D}\lesssim\delta\left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right|\!\right|\!\right|_{k,D}|\!|\!|v|\!|\!|_{k,D}\,.$$

For the second term, we first write that

$$\delta\left|\left(\widehat{A}_2^\delta\frac{\partial}{\partial \boldsymbol{x}_2}\left(\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right),\frac{\partial v}{\partial \boldsymbol{x}_2}\right)_D\right|\lesssim\delta\left\|\frac{\partial}{\partial \boldsymbol{x}_2}\left(\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right)\right\|_{\widehat{A}_2^\delta}\left\|\frac{\partial v}{\partial \boldsymbol{x}_2}\right\|_{\widehat{A}_2^\delta}\lesssim\left\|\frac{\partial}{\partial \boldsymbol{x}_2}\left(\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right)\right\|_{\widehat{A}_2^\delta}|\!|\!|v|\!|\!|_{k,D}\,,$$

and then observe that

$$\begin{aligned}\left\|\frac{\partial}{\partial \boldsymbol{x}_2}\left(\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right)\right\|_{\widehat{A}_2^\delta}&\lesssim\left\|\widehat{\chi}^\delta\frac{\partial}{\partial \boldsymbol{x}_2}\left(\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right)\right\|_{\widehat{A}_2^\delta}+\left\|\frac{\partial\widehat{\chi}^\delta}{\partial \boldsymbol{x}_2}\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right\|_{\widehat{A}_2^\delta}\\&\lesssim\left\|\frac{\partial}{\partial \boldsymbol{x}_2}\left(\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right)\right\|_{\widehat{A}_2^\delta}+\ell^{-1}\left\|\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right\|_{\widehat{A}_2^\delta}\\&\lesssim\left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right|\!\right|\!\right|_{k,D}+\ell^{-1}|\!|\!|u_0|\!|\!|_{k,D}\,,\end{aligned}$$

leading to

$$\delta \left| \left(\widehat{A}_2^\delta \frac{\partial}{\partial \boldsymbol{x}_2} \left(\widehat{\chi}^\delta \frac{\partial u_0}{\partial \boldsymbol{x}_1} \right), \frac{\partial v}{\partial \boldsymbol{x}_2} \right)_D \right| \lesssim \left(\frac{\delta}{\ell} \left\| u_0 \right\|_{k,D} + \delta \left\| \frac{\partial u_0}{\partial \boldsymbol{x}_1} \right\|_{k,D} \right) \left\| v \right\|_{k,D}.$$

Finally, the last term is easily dealt with using (4.2.3)

$$\delta \left| \left(\nu \widehat{A}_1^\delta \chi^\delta \frac{\partial^2 u_0}{\partial \boldsymbol{x}_1^1}, \frac{\partial v}{\partial \boldsymbol{x}_1} \right)_D \right| \lesssim \delta \left\| \frac{\partial^2 u_0}{\partial \boldsymbol{x}_1^2} \right\|_{\widehat{A}_1^\delta, D} \left\| \frac{\partial v}{\partial \boldsymbol{x}_1} \right\|_{\widehat{A}_1^\delta, D} \lesssim \delta \left\| \frac{\partial u_0}{\partial \boldsymbol{x}_1} \right\|_{k,D} \left\| v \right\|_{k,D}.$$

□

We finally analyze the terms linked with the $\boldsymbol{x}_2$ derivatives. This last term requires more subtle arguments.

Lemma 4.2.6. *For all $u_0 \in H^2(D)$ and $v \in H^1(D)$, we have*

$$\left| \left((\widehat{A}_2^\delta - A_2^{\mathrm{H}}) \frac{\partial u_0}{\partial \boldsymbol{x}_2}, \frac{\partial v}{\partial \boldsymbol{x}_2} \right)_D \right| \lesssim \left\{ \left(\sqrt{\frac{\delta}{h_D}} + \frac{\delta}{\ell} \right) \left\| u_0 \right\|_{k,D} + (\delta + \sqrt{\delta h_D}) \left\| \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right\|_{k,D} \right\} \left\| v \right\|_{k,D}. \quad (4.2.15)$$

Proof. Let $u_0 \in H^2(D)$ and $v \in H^1(D)$. We will use the function $\widehat{\tau}$ from Lemma 4.2.2. Invoking (4.1.1) and (4.2.5), we start by writing that

$$I := \left((\widehat{A}_2^\delta - A_2^{\mathrm{H}}) \frac{\partial u_0}{\partial \boldsymbol{x}_2}, \frac{\partial v}{\partial \boldsymbol{x}_2} \right)_D = \left(A_2^{\mathrm{H}} \left(\frac{\partial \widehat{\tau}}{\partial y} \right)^\delta \frac{\partial u_0}{\partial \boldsymbol{x}_2}, \frac{\partial v}{\partial \boldsymbol{x}_2} \right)_D = \delta \left(A_2^{\mathrm{H}} \frac{\partial \widehat{\tau}^\delta}{\partial \boldsymbol{x}_1} \frac{\partial u_0}{\partial \boldsymbol{x}_2}, \frac{\partial v}{\partial \boldsymbol{x}_2} \right)_D.$$

Since we also have

$$\frac{\partial \widehat{\tau}^\delta}{\partial \boldsymbol{x}_1} \frac{\partial u_0}{\partial \boldsymbol{x}_2} = \frac{\partial}{\partial \boldsymbol{x}_1} \left(\widehat{\tau}^\delta \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right) - \widehat{\tau}^\delta \frac{\partial^2 u_0}{\partial \boldsymbol{x}_1 \partial \boldsymbol{x}_2},$$

we end up with the identity $I = I_1 + I_2$, where

$$I_1 := \delta \left(A_2^{\mathrm{H}} \frac{\partial}{\partial \boldsymbol{x}_1} \left(\widehat{\tau}^\delta \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right), \frac{\partial v}{\partial \boldsymbol{x}_2} \right)_D \qquad I_2 := \delta \left(A_2^{\mathrm{H}} \widehat{\tau}^\delta \frac{\partial^2 u_0}{\partial \boldsymbol{x}_1 \partial \boldsymbol{x}_2}, \frac{\partial v}{\partial \boldsymbol{x}_2} \right)_D.$$

Actually, the second term is the simplest to deal with, so we shall address it first. In fact, using (4.2.6) to bound $\|\widehat{\tau}^\delta\|_{L^\infty(D)}$, we have

$$|I_2| \lesssim \delta \left\| \frac{\partial^2 u_0}{\partial \boldsymbol{x}_1 \partial \boldsymbol{x}_2} \right\|_{A_2^{\mathrm{H}}, D} \left\| \frac{\partial v}{\partial \boldsymbol{x}_2} \right\|_{A_2^{\mathrm{H}}, D} \lesssim \delta \left\| \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right\|_{k,D} \left\| v \right\|_{k,D}.$$

For the other term, since $A_2^{\rm H}$ does not depend on $\boldsymbol{x}_1$, we have

$$I_1 = \delta \left(\frac{\partial}{\partial \boldsymbol{x}_1} \left(A_2^{\rm H} \widehat{\tau}^\delta \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right), \frac{\partial v}{\partial \boldsymbol{x}_2} \right)_D = I_{1,1} + I_{1,2},$$

with

$$I_{1,1} := \delta \left(\frac{\partial}{\partial \boldsymbol{x}_2} \left(A_2^{\rm H} \widehat{\tau}^\delta \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right), \frac{\partial v}{\partial \boldsymbol{x}_1} \right)_D, \qquad I_{1,2} := \delta \left(\boldsymbol{\nabla} \left(A_2^{\rm H} \widehat{\tau}^\delta \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right) \times \boldsymbol{n}, v \right)_{\partial D},$$

where we have employed the integration by part formula

$$\left(\frac{\partial \phi}{\partial \boldsymbol{x}_1}, \frac{\partial w}{\partial \boldsymbol{x}_2} \right)_D = \left(\frac{\partial \phi}{\partial \boldsymbol{x}_2}, \frac{\partial w}{\partial \boldsymbol{x}_1} \right)_D + (\boldsymbol{\nabla} \phi \times \boldsymbol{n}, w)_{\partial D},$$

valid for all $\phi \in H^2(D)$ and $w \in H^1(D)$.

For the $I_{1,1}$ term, we expand

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{x}_2} \left(A_2^{\rm H} \widehat{\tau}^\delta \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right) &= \frac{\partial}{\partial \boldsymbol{x}_2} \left(A_2^{\rm H} \widehat{\tau}^\delta \right) \frac{\partial u_0}{\partial \boldsymbol{x}_2} + A_2^{\rm H} \widehat{\tau}^\delta \frac{\partial^2 u_0}{\partial \boldsymbol{x}_2^2} \\ &= \left(\widehat{\tau}^\delta \frac{\partial A_2^{\rm H}}{\partial \boldsymbol{x}_2} + \frac{\partial \widehat{\tau}^\delta}{\partial \boldsymbol{x}_2} A_2^{\rm H} \right) \frac{\partial u_0}{\partial \boldsymbol{x}_2} + A_2^{\rm H} \widehat{\tau}^\delta \frac{\partial^2 u_0}{\partial \boldsymbol{x}_2^2}, \end{aligned}$$

leading to

$$I_{1,1} = \delta \left\{ \left(A_2^{\rm H} \widehat{\tau}^\delta \left\{ \left(\frac{1}{A_2^{\rm H}} \frac{\partial A_2^{\rm H}}{\partial \boldsymbol{x}_2} \right) + \frac{\partial \widehat{\tau}^\delta}{\partial \boldsymbol{x}_2} \right\} \frac{\partial u_0}{\partial \boldsymbol{x}_2}, \frac{\partial v}{\partial \boldsymbol{x}_1} \right)_D + \left(A_2^{\rm H} \widehat{\tau}^\delta \frac{\partial^2 u_0}{\partial \boldsymbol{x}_2^2}, \frac{\partial v}{\partial \boldsymbol{x}_1} \right)_D \right\},$$

and

$$|I_{1,1}| \lesssim \delta \left(\ell^{-1} \left\| u_0 \right\|_{k,D} + \left\| \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right\|_{k,D} . \right) \left\| v \right\|_{k,D}$$

We now analyze $I_{1,2}$, which is the most technical. Let us formally introduce the operator $\mathscr{L} : H^1(\partial D) \to L^2(\partial D)$ by

$$\mathscr{L}(\psi) := \boldsymbol{\nabla} \left(A_2^{\rm H} \widehat{\tau}^\delta \psi \right) \times \boldsymbol{n}.$$

Notice that we also have $\mathscr{L} : L^2(\partial D) \to H^{-1}(\partial D)$. We have

$$\|\mathscr{L}(\psi)\|_{L^2(\partial D)} \lesssim A_{\max} \left(\delta^{-1} \|\psi\|_{L^2(\partial D)} + |\psi|_{H^1(D)} \right),$$

and

$$\|\mathscr{L}(\psi)\|_{H^{-1}(\partial D)} \lesssim A_{\max} \|\psi\|_{L^2(D)},$$

so that, see [129], by interpolation

$$\|\mathscr{L}(\psi)\|_{H^{-1/2}(\partial D)} \lesssim A_{\max} \left(\delta^{-1/2} \|\psi\|_{L^2(\partial D)} + |\psi|_{H^{1/2}(\partial D)} \right).$$

We finally observe that

$$(\mathscr{L}(\psi), 1)_{\partial D} = 0.$$

It follows that

$$\begin{aligned}
|I_{1,2}| &= \delta \left| \left(\mathscr{L}\left(\frac{\partial u_0}{\partial \boldsymbol{x}_2} \right), v \right)_{\partial D} \right| \\
&= \delta \left| \left(\mathscr{L}\left(\frac{\partial u_0}{\partial \boldsymbol{x}_2} \right), v - v_{\partial D} \right)_{\partial D} \right| \\
&\lesssim \delta A_{\max} \left(\delta^{-1/2} \left\| \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right\|_{L^2(\partial D)} + \left| \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right|_{H^{1/2}(\partial D)} \right) \|v - v_{\partial D}\|_{H^{1/2}(\partial D)} \\
&\lesssim A_{\max} \left(\delta^{1/2} \left\| \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right\|_{L^2(\partial D)} + \delta \left| \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right|_{H^{1/2}(\partial D)} \right) |v|_{H^{1/2}(\partial D)} \\
&\lesssim (A_{\max})^{1/2} \left(\delta^{1/2} \left\| \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right\|_{L^2(\partial D)} + \delta \left| \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right|_{H^1(D)} \right) |\!|\!| v |\!|\!|_{k,D} .
\end{aligned}$$

We then have

$$\left\| \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right\|_{L^2(\partial D)} \lesssim h_D^{-1/2} \left\| \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right\|_{L^2(D)} + h_D^{1/2} \left| \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right|_{H^1(D)},$$

leading to

$$|I_{1,2}| \lesssim \left(\sqrt{\frac{\delta}{h_D}} |\!|\!| u_0 |\!|\!|_{k,D} + (\sqrt{\delta h_D} + \delta) \left| \!\left| \!\left| \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right| \!\right| \!\right|_{k,D} \right) |\!|\!| v |\!|\!|_{k,D},$$

and (4.2.15) follows. □

4.2.3 Error estimates

In this subsection, we state our main convergence result. We start by showing that the homogenized solution is regular.

Lemma 4.2.7 (Regularity of the homogenized solution)**.** *Let $f \in L^2(\widetilde{\Omega})$. For the associated solution $u_0 \in H^1_\sharp(\widetilde{\Omega})$ to (4.1.7), we have $u_0 \in H^2(\Omega) \cup H^2(\Omega_{\mathrm{P}})$ with*

$$\begin{aligned}
k |\!|\!| u_0 |\!|\!|_{k,\Omega} + \left| \!\left| \!\left| \frac{\partial u_0}{\partial \boldsymbol{x}_1} \right| \!\right| \!\right|_{k,\Omega} + \left| \!\left| \!\left| \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right| \!\right| \!\right|_{k,\Omega} + \\
k |\!|\!| u_0 |\!|\!|_{k,\Omega_{\mathrm{P}}} + \left| \!\left| \!\left| \frac{\partial u_0}{\partial \boldsymbol{x}_1} \right| \!\right| \!\right|_{k,\Omega_{\mathrm{P}}} + \left| \!\left| \!\left| \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right| \!\right| \!\right|_{k,\Omega_{\mathrm{P}}} \lesssim \mathscr{C}_{\mathrm{st}} \|f\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}}. \quad (4.2.16)
\end{aligned}$$

Proof. We first use 4.1.5 to obtain

$$k |\!|\!| u_0 |\!|\!|_{k,\Omega} + k |\!|\!| u_0 |\!|\!|_{k,\Omega_{\mathrm{P}}} \lesssim \mathscr{C}_{\mathrm{st}} \|f\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}}. \quad (4.2.17)$$

Now, for the H^2 regularity, we observe that

$$-\boldsymbol{\nabla}\cdot\left(\boldsymbol{D}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{\nabla}u\right)=\varepsilon^{\mathrm{H}}f+k^2\nu\varepsilon^{\mathrm{H}}u \text{ in } \widetilde{\Omega}.$$

Then, the coefficients (piecewise) smoothness and the standard elliptic regularity results (see, e.g. [20, Section 9.6]) imply that $u_0\in H^2(\Omega)\cup H^2(\Omega_{\mathrm{P}})$ and

$$\left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right|\!\right|\!\right|_{k,\Omega}+\left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_2}\right|\!\right|\!\right|_{k,\Omega}+\left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right|\!\right|\!\right|_{k,\Omega_{\mathrm{P}}}+\left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_2}\right|\!\right|\!\right|_{k,\Omega_{\mathrm{P}}}\lesssim \|f\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}}+k^2\|u\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}}\lesssim \mathscr{C}_{\mathrm{st}}\|f\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}}.$$

□

We are now ready to provide the main result of this chapter.

Theorem 4.2.8 (Error estimate)**.** *For all $\delta>0$ and $f\in L^2(\widetilde{\Omega})$, we have*

$$k\left|\!\left|\!\left|u_\delta-u_0-\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right|\!\right|\!\right|_{k,\widetilde{\Omega}}\lesssim \mathscr{C}_{\mathrm{st}}^2\left(\sqrt{k\ell}\sqrt{k\delta}+k\delta\right)\|f\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}}.$$

Proof. Recalling that $\widehat{\chi}(\boldsymbol{x}_2,y)=0$ whenever $\boldsymbol{x}_2>\ell_2$, we have

$$b_\delta\left(\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1},v\right)=b_\delta^\Omega\left(\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1},v\right).$$

Therefore

$$\begin{aligned}b_\delta\left(u_\delta-u_0-\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1},v\right)&=b_0(u_0,v)-b_\delta(u_0,v)-b_\delta\left(\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1},v\right)\\&=-k^2\left((\varepsilon^{\mathrm{H}}-\varepsilon^\delta)u_0,v\right)_\Omega\\&\quad+\left((A_1^{\mathrm{H}}-A_1^\delta)\frac{\partial u_0}{\partial \boldsymbol{x}_1},\frac{\partial v}{\partial \boldsymbol{x}_1}\right)_\Omega-b_\delta^\Omega\left(\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1},v\right)\\&\quad+\left((A_2^{\mathrm{H}}-A_2^\delta)\frac{\partial u_0}{\partial \boldsymbol{x}_2},\frac{\partial v}{\partial \boldsymbol{x}_2}\right)_\Omega.\end{aligned}$$

It follows that for all $v\in H^1_\sharp(\widetilde{\Omega})$ with $\|v\|_{k,\widetilde{\Omega}}=1$, we have

$$\begin{aligned}\left|b_\delta\left(u_\delta-u_0-\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1},v\right)\right|&\lesssim\left(k\delta+\frac{\delta}{\ell}\right)\|u_0\|_{k,\Omega}\\&\quad+\delta\left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right|\!\right|\!\right|_{k,\Omega}\\&\quad+\sqrt{\frac{\delta}{\ell}}\|u_0\|_{k,\Omega}+(\delta+\sqrt{\delta\ell})\left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_2}\right|\!\right|\!\right|_{k,\Omega}\\&\lesssim\left(k\delta+\sqrt{\frac{\delta}{\ell}}+\frac{\delta}{\ell}\right)\|u_0\|_{k,\Omega}+(\delta+\sqrt{\delta\ell})\left(\left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right|\!\right|\!\right|_{k,\Omega}+\left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_2}\right|\!\right|\!\right|_{k,\Omega}\right)\\&\lesssim\left(k\delta+\sqrt{\frac{\delta}{\ell}}\right)\|u_0\|_{k,\Omega}+\sqrt{\delta\ell}\left(\left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right|\!\right|\!\right|_{k,\Omega}+\left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_2}\right|\!\right|\!\right|_{k,\Omega}\right).\end{aligned}$$

As a result

$$
\begin{aligned}
k\left|b_\delta\left(u_\delta-u_0-\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1},v\right)\right| &\lesssim \left(k\delta+\sqrt{\frac{\delta}{\ell}}\right)k\,|\!|\!|u_0|\!|\!|_{k,\Omega}+k\sqrt{\delta\ell}\left(\left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right|\!\right|\!\right|_{k,\Omega}+\left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_2}\right|\!\right|\!\right|_{k,\Omega}\right)\\
&\lesssim \left(k\delta+\left(\sqrt{k\ell}\right)^{-1}\sqrt{k\delta}\right)k\,|\!|\!|u_0|\!|\!|_{k,\Omega}\\
&\quad+\sqrt{k\delta}\sqrt{k\ell}\left(\left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right|\!\right|\!\right|_{k,\Omega}+\left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_2}\right|\!\right|\!\right|_{k,\Omega}\right)\\
&\lesssim \left(k\delta+\sqrt{k\delta}\sqrt{k\ell}\right)\left(k\,|\!|\!|u_0|\!|\!|_{k,\Omega}+\left(\left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right|\!\right|\!\right|_{k,\Omega}+\left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_2}\right|\!\right|\!\right|_{k,\Omega}\right)\right)
\end{aligned}
$$

Then, (4.2.16) implies that

$$
k\left|b_\delta\left(u_\delta-u_0-\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1},v\right)\right|\lesssim \mathscr{C}_{\rm st}\left(k\delta+\sqrt{k\delta}\sqrt{k\ell}\right)\|f\|_{\varepsilon^{\rm H},\widetilde{\Omega}},
$$

and since

$$
u_\delta-u_0-\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1}\in H^1_\sharp(\widetilde{\Omega}),
$$

the result follows from the inf-sup condition (3.3.59) of the sesquilinear form b_δ over the space $H^1_\sharp(\widetilde{\Omega})$. □

Chapter 5

Discretization with a multiscale hybrid-mixed method

Contents

Numerical simulations of wave propagation in realistic two- or three-dimensional problems are often characterized by multiscale structures. However the numerical approximation of these multiscale problems by standard tools is extremely expensive, as resolving the fine scales result in a very large number of degrees of freedom coupled together. Henceforth, using multiscale methods represents an appropriate and effective option for dealing with such problems. The common feature of these multiscale methods is the use of a coarse mesh coupled with special basis functions adapted to local medium properties in each coarse mesh "macro element". These basis functions are constructed as solutions to completely independent element-wise local problems. Therefore, for most of these multiscale techniques, it is interesting to introduce a second-level algorithm in which the multiscale basis functions are computed by solving the local problems. Consequently, this feature makes the algorithms derived from these multiscale methods particularly interesting for use in parallel computing environments.

This chapter focuses on a multiscale finite element method called the multiscale hybrid-mixed (MHM) method. First introduced and presented in [76] for the heterogeneous Darcy equation, the MHM method derives from the primal hybridization of the original equation (see [121]). In this hybrid formulation, the regularity of the unknown is relaxed using an element-wise Sobolev space (piecewise H^1 regularity). Then, the solution continuity is weakly imposed by the action of Lagrange multipliers space. Subsequently, the MHM method relies on this hybrid formulation, and it characterizes the solution by decomposition into a global formulation posed on the skeleton of a (coarse) mesh of the domain and solutions to independent local problems. As we will see, unlike other multiscale methods, the local problems responsible for the multiscale basis functions are embedded in a natural way, and their numerical approximation corresponds to the second-level MHM method.

After introducing the method in [76], F. Valentin and his collaborators studied the convergence of the method for the Darcy model in [3]. First, they showed that it produces accurate numerical primary and dual variables with respect to the mesh size H. Then, considering the same model, they demonstrated its convergence and robustness with respect to the (small) characteristic of the physical coefficients for highly oscillatory cases [116]. With the same multiscale formulation steps, the MHM method was further extended to various operators, linear elasticity [75], advective-reactive [77], Maxwell equations [95, 69], and recently Oseen equation [2].

The method was extended and adapted to the highly heterogeneous acoustic Helmholtz equation by T. Chaumont-frelet and F. Valentin in [36]. There, they considered the case of a bounded domain with a boundary composed of two parts: a Dirichlet condition in one part and a first-order radiation condition in the other. For their model, they proposed a numerical analysis showing the well-posedness of the MHM formulation and the quasi-optimality of its numerical solution. In addition, they show that the MHM can produce exact solutions for precise propagation directions under certain conditions.

This chapter is dedicated to the analysis of the MHM method applied to the two-dimensional PML Helmholtz problem given in (5.1.1). We start by proposing a hybrid formulation of our model problem in Sections 5.1 followed by the derivation of the MHM formulation and the study of its well-posedness in and 5.2. Next, we present a conver-

gence with first-order polynomial discretizations showing the performance of the MHM method in the presence of the quasi-resonant frequencies in Section 5.3. We then turn to the convergence analysis of the MHM with respect to the (small) characteristic length of oscillations, assuming highly oscillatory coefficients in Section 5.4.

5.1 An hybrid reformulation

The starting point to establish the MHM method is the so-called "primal hybrid formulation" introduced in [121] for the Poisson problem. The goal of this section is to establish such a formulation for a Helmholtz problem in periodic structures. The primal hybrid formulation of a Helmholtz problem has already been introduced in [36]. Here, the key novelties are the treatment of perfectly matched layers and quasi-periodic boundary conditions.

5.1.1 The model problem

For the reader's convenience, we recall the statement of the PML Helmholtz problem here. $\widetilde{\Omega} := (0,\ell_1)\times(0,\ell_2+\ell_{\rm P})$ is a rectangular domain composed of two parts. The physical domain is $\Omega := (0,\ell_1)\times(0,\ell_2)$, and the materials contained in Ω are characterized by the coefficients coefficient ε and $\boldsymbol{A}$. The absorbing layer $\Omega_{\rm P} := (0,\ell_1)\times(\ell_2,\ell_{\rm P})$ is an artificial device used to bound the computational domain. The properties of the absorbing layer are described by its damping coefficient $\nu_{\rm P} = \gamma_{\rm r} + i\gamma_{\rm i}$, where $\gamma_{\rm r},\gamma_{\rm i}\geq 1$. We also let $\ell := \sqrt{\ell_1^2+(\ell_2+\ell_{\rm P})^2}$ denote the diameter of $\widetilde{\Omega}$.

As defined in (3.1.2), the domain boundary is made of three parts $\partial\widetilde{\Omega} = \Gamma_{\rm P}\cup\Gamma_{\rm D}\cup\widetilde{\Gamma}_\sharp$. We consider the PML problem consists in finding $u\in H^1_\sharp(\widetilde{\Omega})$ such that

$$\left\{\begin{array}{rcll} -k^2\nu\varepsilon u - \boldsymbol{\nabla}\cdot(\boldsymbol{D}\boldsymbol{A}\boldsymbol{\nabla}u) &=& \varepsilon f, & \text{in } \widetilde{\Omega}\\ u &=& 0 & \text{on } \Gamma_{\rm P},\\ u &=& 0 & \text{on } \Gamma_{\rm D},\\ u_+ - e^{i\alpha\ell_1}u_- &=& 0 & \text{on } \widetilde{\Gamma}_\sharp, \end{array}\right. \tag{5.1.1}$$

where

$$\boldsymbol{A} = \begin{pmatrix} A_1 & 0\\ 0 & A_2\end{pmatrix} \quad\text{and}\quad \boldsymbol{D} = \begin{pmatrix} \nu & 0\\ 0 & \nu^{-1}\end{pmatrix}.$$

Let us recall that assuming $f\in L^2(\widetilde{\Omega})$, the above PML problem is equivalent to: Find $u\in H^1_\sharp(\widetilde{\Omega})$ such that

$$b(u,v) = (\varepsilon f, v)_{\widetilde{\Omega}} \quad \forall v\in H^1_\sharp(\widetilde{\Omega}), \tag{5.1.2}$$

where

$$b(u,v) := -k^2(\nu\varepsilon u, v)_{\widetilde{\Omega}} + \left(\nu A_1\frac{\partial u}{\partial \boldsymbol{x}_1}, \frac{\partial v}{\partial \boldsymbol{x}_1}\right)_{\widetilde{\Omega}} + \left(\nu^{-1}A_2\frac{\partial u}{\partial \boldsymbol{x}_2}, \frac{\partial v}{\partial \boldsymbol{x}_2}\right)_{\widetilde{\Omega}},$$

and

$$H^1_\sharp(\widetilde{\Omega}) := \left\{ v\in H^1(\widetilde{\Omega},\mathbb{C}) \mid v_{|\Gamma_{\rm D}} = v_{|\Gamma_{\rm P}} = 0 \text{ and } v_{|\widetilde{\Gamma}_{\sharp+}} = e^{i\alpha\ell_1}v_{|\widetilde{\Gamma}_{\sharp-}}\right\}.$$

Assumption 5.1.1 (Stability of the PML problem)**.** *For all $f \in L^2(\widetilde{\Omega})$, there exists a unique $u \in H^1_\sharp(\widetilde{\Omega})$ solution to (5.1.2). In addition we have*

$$|\!|\!| u |\!|\!|_{k,\widetilde{\Omega}} \leq \frac{\mathscr{C}_{\rm st}}{k} \|f\|_{\varepsilon,\widetilde{\Omega}}, \tag{5.1.3}$$

for a constant $\mathscr{C}_{\rm st} > 1$ independent of f and u.

Remark 5.1.2 (Tilda notation)**.** *In chapter 3, we employed the notations $\widetilde{b}$, $\widetilde{u}$ and $\widetilde{f}$ instead of b, u and f for the solution and data of (5.1.2). We made this choice in chapter 3 in order to distinguish between the DtN and PML problems. Here, we shall only work with the PML problem, and as a result, we drop the $\widetilde{\cdot}$ superscript in order to lighten the notations.*

Remark 5.1.3 (Validity of Assumption 5.1.1)**.** *We have established in chapter 3 that assumption 5.1.1 is valid in large number of relevant situations, where we can in fact provide an explicit estimate for $\mathscr{C}_{\rm st}$.*

5.1.2 Functional spaces for hybridization

We will need additional function spaces linked with the divergence operator to rigorously write down the primal hybrid formulation. We shall describe these spaces here.

5.1.2.1 Piecewise smooth functions

In the primal hybrid formulation, we want to relax the continuity of functions at the interfaces of the mesh. As a result, our core energy space will incorporate the mesh in its definition. Thus, in the remainder of this section, we consider a mesh $\mathcal{T}_H$ of $\widetilde{\Omega}$. We assume that the elements $K \in \mathcal{T}_H$ are open triangles and that either entirely lie in Ω or in $\Omega_{\rm P}$. For the sake of simplicity, we also require that the mesh $\mathcal{T}_H$ is conforming, meaning if the intersection $\partial K_+ \cap \partial K_-$ of two elements $K_\pm \in \mathcal{T}_H$ is not empty, it is either a single vertex or a full face of both K_+ and K_-. The space

$$H^1(\mathcal{T}_H) := \Big\{ v \in L^2(\widetilde{\Omega}) \mid v|_K \in H^1(K) \quad \forall K \in \mathcal{T}_H \Big\}$$

collects functions that have piecewise H^1 regularity onto the mesh $\mathcal{T}_H$. It will be the space where we seek the solution to the Helmholtz problem.

When dealing with functions in $H^1(\mathcal{T}_H)$, we will employ the notation

$$(\cdot,\cdot)_{\mathcal{T}_H} := \sum_{K\in\mathcal{T}_H} (\cdot,\cdot)_K, \qquad \|\cdot\|^2_{\mathcal{T}_H} := (\cdot,\cdot)_{\mathcal{T}_H},$$

for broken inner-products and norms (we also employ the same notations for weighted norms). In particular, the energy norm

$$|\!|\!| v |\!|\!|^2_{k,\mathcal{T}_H} := k^2 \|v\|^2_{\varepsilon,\mathcal{T}_H} + \|\boldsymbol{\nabla} v\|^2_{\boldsymbol{A},\mathcal{T}_H} \qquad \forall v \in H^1(\mathcal{T}_H),$$

will be useful.

5.1.2.2 Vector-valued Sobolev space

As the continuity of functions in $H^1(\mathcal{T}_H)$ is relaxed, it should be weakly enforced by some other means to obtain an equivalent formulation of the Helmholtz problem. To do so, the key concept we need is the Sobolev space

$$\boldsymbol{H}(\mathrm{div},\widetilde{\Omega}) := \left\{ \boldsymbol{q} \in \boldsymbol{L}^2(\widetilde{\Omega}) \mid \boldsymbol{\nabla}\cdot\boldsymbol{q} \in L^2(\widetilde{\Omega}) \right\},$$

of vector-valued function with square-integrable divergence [65].

The following expression

$$(\boldsymbol{q}\cdot\boldsymbol{n}, v)_{\partial\widetilde{\Omega}} = (\boldsymbol{\nabla}\cdot\boldsymbol{q}, v)_{\widetilde{\Omega}} + (\boldsymbol{q}, \boldsymbol{\nabla} v)_{\widetilde{\Omega}},$$

is valid for all smooth functions $\boldsymbol{q}$ and v. Since the right-hand side is continuous for $\boldsymbol{q} \in \boldsymbol{H}(\mathrm{div},\widetilde{\Omega})$ and $v \in H^1(\widetilde{\Omega})$ in the sense that

$$|(\boldsymbol{\nabla}\cdot\boldsymbol{q}, v)_{\widetilde{\Omega}} + (\boldsymbol{q}, \boldsymbol{\nabla} v)_{\widetilde{\Omega}}| \leq \|\boldsymbol{\nabla}\cdot\boldsymbol{q}\|_{\widetilde{\Omega}}\|v\|_{\widetilde{\Omega}} + \|\boldsymbol{q}\|_{\widetilde{\Omega}}\|\boldsymbol{\nabla} v\|_{\widetilde{\Omega}},$$

it appears that the normal trace $\boldsymbol{q}\cdot\boldsymbol{n}$ belongs to the dual of the trace space of $H^1(\widetilde{\Omega})$. This is in fact true, and we have $\boldsymbol{q}\cdot\boldsymbol{n} \in H^{-1/2}(\partial\widetilde{\Omega}) := (H^{1/2}(\partial\widetilde{\Omega}))'$ for all $\boldsymbol{q} \in \boldsymbol{H}(\mathrm{div},\widetilde{\Omega})$. In addition, the normal trace mapping is surjective from $\boldsymbol{H}(\mathrm{div},\widetilde{\Omega})$ onto $H^{-1/2}(\partial\widetilde{\Omega})$, see [65].

5.1.2.3 Characterization of $H^1_0(\widetilde{\Omega})$

As a result of the above discussion, we can actually write that

$$H^1_0(\widetilde{\Omega}) := \left\{ v \in H^1(\widetilde{\Omega}) \mid (\boldsymbol{\nabla}\cdot\boldsymbol{q}, v)_{\widetilde{\Omega}} + (\boldsymbol{q}, \boldsymbol{\nabla} v)_{\widetilde{\Omega}} = 0 \quad \forall \boldsymbol{q} \in \boldsymbol{H}(\mathrm{div},\widetilde{\Omega}) \right\}.$$

In other words, we can characterize $H^1_0(\widetilde{\Omega})$, the space of $H^1(\widetilde{\Omega})$ function with vanishing trace using $\boldsymbol{H}(\mathrm{div},\widetilde{\Omega})$ functions. In the primal hybrid formulation, we want to work with piecewise H^1 function, instead of globally H^1 function. A key idea introduced in [121] is that we can also use $\boldsymbol{H}(\mathrm{div},\widetilde{\Omega})$ functions to characterize piecewise H^1 functions that are actually globally H^1.

A key result established in [121] is then that

$$H^1_0(\widetilde{\Omega}) = \left\{ v \in H^1(\mathcal{T}_H) \mid (\boldsymbol{\nabla}\cdot\boldsymbol{q}, v)_{\mathcal{T}_H} + (\boldsymbol{q}, \boldsymbol{\nabla} v)_{\mathcal{T}_H} = 0 \quad \forall \boldsymbol{q} \in \boldsymbol{H}(\mathrm{div},\widetilde{\Omega}) \right\}. \tag{5.1.4}$$

The characterization of $H^1_0(\widetilde{\Omega})$ in (5.1.4) is the starting point of the primal hybrid formulation, and therefore, is central in the design of the MHM method.

Here, our energy space is not $H^1_0(\widetilde{\Omega})$. Instead, $H^1_\sharp(\widetilde{\Omega})$ incorporates quasi-periodic boundary conditions. The remaining of this section is thus dedicated to a characterization similar to (5.1.4) of $H^1_\sharp(\widetilde{\Omega})$.

5.1.2.4 Characterization of quasi-periodic functions

As we demonstrate below, the correct space we need to weakly impose continuity is a Sobolev space of vector valued functions that incorporates quasi-periodic boundary itself. Specifically, we introduce

$$\boldsymbol{H}_\sharp(\mathrm{div},\widetilde{\Omega}) := \left\{ \boldsymbol{q} \in \boldsymbol{H}(\mathrm{div},\widetilde{\Omega}) \;\middle|\; \boldsymbol{q}_+ \cdot \boldsymbol{n}_+ + e^{i\alpha\ell_1}\boldsymbol{q}_- \cdot \boldsymbol{n}_- \;=\; 0 \text{ on } \Gamma_\sharp \right\}.$$

The rigorous mathematical definition of the traces spaces involved in the definition of $\boldsymbol{H}_\sharp(\mathrm{div},\widetilde{\Omega})$ is subtle, and we refer the reader to [60] for more details.

Theorem 5.1.4 (Characterization of $H^1_\sharp(\widetilde{\Omega})$)**.** *We have*

$$H^1_\sharp(\widetilde{\Omega}) = \left\{ v \in H^1(\mathcal{T}_H) \mid (\boldsymbol{\nabla}\cdot\boldsymbol{q}, v)_{\mathcal{T}_H} + (\boldsymbol{q}, \boldsymbol{\nabla} v)_{\mathcal{T}_H} = 0 \quad \forall \boldsymbol{q} \in \boldsymbol{H}_\sharp(\mathrm{div},\widetilde{\Omega}) \right\}.$$

Proof. This proof employs delicate specific spaces for the traces of functions on part of the boundary. For the sake of simplicity, we do not report these definition here, but they can be found in [60].

Let us first assume that $v \in H^1_\sharp(\widetilde{\Omega})$. Then, for all $\boldsymbol{q} \in \boldsymbol{H}_\sharp(\mathrm{div},\widetilde{\Omega})$, we have

$$(\boldsymbol{\nabla}\cdot\boldsymbol{q}, v)_{\mathcal{T}_H} + (\boldsymbol{q}, \boldsymbol{\nabla} v)_{\mathcal{T}_H} = \langle \boldsymbol{q}\cdot\boldsymbol{n}, v\rangle_{\partial\widetilde{\Omega}} = \langle \boldsymbol{q}\cdot\boldsymbol{n}, v\rangle_{\Gamma_{\sharp_+}} + \langle \boldsymbol{q}\cdot\boldsymbol{n}, v\rangle_{\Gamma_{\sharp_-}} \tag{5.1.5}$$

where the second identity follows from the fact that $v|_{\Gamma_\mathrm{D}} = 0$ and $v|_{\Gamma_\mathrm{D}} = 0$. We also note that duality the pairings in (5.1.5) are well-defined over $H^{-1/2}(\Gamma_\sharp{}^{\pm})$ and $H^{1/2}_{00}(\Gamma_\sharp{}^{\pm})$. Then it remains to observe that using the quasi-periodicity properties of $\boldsymbol{q}$ and v, we have

$$\begin{aligned}\langle \boldsymbol{q}\cdot\boldsymbol{n}, v\rangle_{\Gamma_{\sharp_+}} &= \langle (\boldsymbol{q}_1)_+, v_+\rangle_{\Gamma_{\sharp_+}} \\ &= \langle e^{i\alpha\ell_1}(\boldsymbol{q}_1)_-, e^{i\alpha\ell_1} v_-\rangle_{\Gamma_{\sharp_+}} = \langle (\boldsymbol{q}_1)_-, v_-\rangle_{\Gamma_{\sharp_-}} = -\langle \boldsymbol{q}\cdot\boldsymbol{n}, v_-\rangle_{\Gamma_{\sharp_-}}\end{aligned}$$

because the normal vector changes sign. This shows that we indeed have

$$(\boldsymbol{\nabla}\cdot\boldsymbol{q}, v) + (\boldsymbol{q}, \boldsymbol{\nabla} v) = 0.$$

On the other hand, let us now consider $v \in H^1(\mathcal{T}_H)$ such that

$$0 = (v, \boldsymbol{\nabla}\cdot\boldsymbol{q})_{\mathcal{T}_H} + (\boldsymbol{\nabla} v, \boldsymbol{q})_{\mathcal{T}_H}.$$

We start by observing that if $\psi : \widetilde{\Omega} \to \mathbb{C}$ is smooth compactly supported, then the $(\psi, 0), (0, \psi) \in \boldsymbol{H}_\sharp(\widetilde{\Omega})$. Thus, if we let $\boldsymbol{q} = (\psi, 0)$, $\boldsymbol{\nabla}\cdot\boldsymbol{q} = \partial\psi/\partial\boldsymbol{x}_1$, and we see that

$$\left(v, \frac{\partial\psi}{\partial\boldsymbol{x}_1}\right)_{\widetilde{\Omega}} = -(\boldsymbol{\nabla} v, \boldsymbol{q})_{\mathcal{T}_H} = -\left(\frac{\partial v}{\partial\boldsymbol{x}_1}, \psi\right)_{\mathcal{T}_H} = -\,(G, \psi)_{\widetilde{\Omega}}\,,$$

where $G \in L^2(\widetilde{\Omega})$ is defined element wise as $\partial v/\partial\boldsymbol{x}_1$. It follows that the distributional derivative of v in fact coincides with its broken derivative, and thus $\partial v/\partial\boldsymbol{x}_1 \in L^2(\widetilde{\Omega})$ in the sense of distribution. The same argument shows that $\partial v/\partial\boldsymbol{x}_2 \in L^2(\widetilde{\Omega})$, so that $v \in H^1(\widetilde{\Omega})$.

We now know that $v \in H^1(\widetilde{\Omega})$, and we need to check the boundary conditions. To do so, we first observe that any smooth vector functions $\boldsymbol{q}$ vanishing in a neighborhood of ${\Gamma_\sharp}^+$ and ${\Gamma_\sharp}^-$ belongs to $\boldsymbol{H}_\sharp(\mathrm{div}, \widetilde{\Omega})$. It follows that for such $\boldsymbol{q}$

$$0 = (\boldsymbol{\nabla}\cdot\boldsymbol{q}, v)_{\mathcal{T}_H} + (\boldsymbol{q}, \boldsymbol{\nabla} v)_{\mathcal{T}_H} = (\boldsymbol{\nabla}\cdot\boldsymbol{q}, v)_{\widetilde{\Omega}} + (\boldsymbol{q}, \boldsymbol{\nabla} v)_{\widetilde{\Omega}} = \langle \boldsymbol{q}\cdot\boldsymbol{n}, v\rangle_{\partial\widetilde{\Omega}} = \langle \boldsymbol{q}\cdot\boldsymbol{n}, v\rangle_{\Gamma_{\mathrm{P}}} + \langle \boldsymbol{q}\cdot\boldsymbol{n}, v\rangle_{\Gamma_{\mathrm{D}}},$$

where the duality pairing are understood between $\left(H^{1/2}(\Gamma_{\mathrm{D}})\right)'$ and $H^{1/2}(\Gamma_{\mathrm{D}})$ (and similarly for Γ_{P}). Since $\boldsymbol{q}$ is an arbitrary smooth function vanishing in a neighborhood of ${\Gamma_\sharp}^+$ and ${\Gamma_\sharp}^-$, it follows by density that $v|_{\Gamma_{\mathrm{D}}} = 0$ and $v|_{\Gamma_{\mathrm{P}}} = 0$. We now only need to check the quasi-periodic conditions. Since v vanishes on $v|_{\Gamma_{\mathrm{D}}} = 0$ and $v|_{\Gamma_{\mathrm{P}}} = 0$, we can now write that

$$0 = (\boldsymbol{\nabla}\cdot\boldsymbol{q}, v)_{\widetilde{\Omega}} + (\boldsymbol{q}, \boldsymbol{\nabla} v)_{\widetilde{\Omega}} = \langle \boldsymbol{q}\cdot\boldsymbol{n}, v_+\rangle_{{\Gamma_\sharp}^+} + \langle \boldsymbol{q}\cdot\boldsymbol{n}, v_-\rangle_{{\Gamma_\sharp}^-},$$

with duality pairing between $H^{-1/2}({\Gamma_\sharp}^\pm)$ and $H^{1/2}_{00}({\Gamma_\sharp}^\pm)$. Using the quasi-periodic boundary condition on $\boldsymbol{q}$, we thus have

$$0 = \langle \boldsymbol{q}\cdot\boldsymbol{n}, v_+\rangle_{{\Gamma_\sharp}^+} - \langle e^{i\alpha\ell_1}\boldsymbol{q}\cdot\boldsymbol{n}, v_-\rangle_{{\Gamma_\sharp}^+} = \langle \boldsymbol{q}\cdot\boldsymbol{n}, v_+ - e^{-i\alpha\ell_1} v_-\rangle_{{\Gamma_\sharp}^+},$$

so that $v_+ = e^{i\alpha\ell_1} v_-$ in the sense of $H^{1/2}_{00}$. This completes the proof. □

5.1.3 The primal hybrid formulation

We are going to look for the solution in $V := H^1(\mathcal{T}_H)$. Currently, the sesquilinear form in the weak formulation (5.1.2) of the Helmholtz problem is only defined on $H^1_\sharp(\widetilde{\Omega})$. To be able to work in $H^1(\mathcal{T}_H)$, we extend the definition of the sesquilinear form to $H^1(\mathcal{T}_H)$ by introducing

$$b_{\mathcal{T}_H}(\phi, v) = -k^2\,(\nu\varepsilon\phi, v)_{\mathcal{T}_H} + (\boldsymbol{D}\boldsymbol{A}\boldsymbol{\nabla}\phi, \boldsymbol{\nabla} v)_{\mathcal{T}_H} \qquad \forall \phi, v \in V.$$

Next, we want to use the space $\boldsymbol{H}_\sharp(\mathrm{div}, \widetilde{\Omega})$ to weakly impose the continuity and boundary conditions in $H^1(\mathcal{T}_H)$. An important remark is that in fact, only the normal traces of the field $\boldsymbol{q} \in \boldsymbol{H}_\sharp(\mathrm{div}, \widetilde{\Omega})$ on the elements boundary are used, not the actual values inside the element. As a result, we introduce the quotient space

$$\Lambda := \left\{ \mu \in H^{-1/2}(\partial K) \mid \exists \boldsymbol{q} \in \boldsymbol{H}_\sharp(\mathrm{div}, \widetilde{\Omega});\ \mu|_{\partial K} = \boldsymbol{q}\cdot\boldsymbol{n}_K \quad \forall K \in \mathcal{T}_H \right\},$$

and the duality pairing

$$\langle \mu, v\rangle_{\partial\mathcal{T}_H} = \sum_{K\in\mathcal{T}_H} \langle \mu, v\rangle_{\partial K} \qquad \forall (\mu, v) \in \Lambda \times H^1(\mathcal{T}_H).$$

It is shown in [3, Lemma 8.3] that the application

$$\|\mu\|_{\Lambda,k} := \sup_{\substack{v\in V \\ \|v\|_{k,\mathcal{T}_H}=1}} \langle \mu, v\rangle_{\partial\mathcal{T}_H},$$

is a Hilbertian norm on Λ. In fact, it shown that it is a norm on the space Λ used without quasi-periodic boundary conditions, but the proof easily carries over to the situation considered here.

Lemma 5.1.5 (Characterization of $H^1_\sharp(\widetilde{\Omega})$)**.** *We have*

$$H^1_\sharp(\widetilde{\Omega}) = \{v \in V \mid \langle \mu, v\rangle_{\partial\mathcal{T}_H} = 0 \quad \forall \mu \in \Lambda\}\,.$$

Proof. Let $\boldsymbol{q} \in \boldsymbol{H}_\sharp(\mathrm{div}, \widetilde{\Omega})$, and define $\mu \in \Lambda$ by setting $\mu|_{\partial K} := \boldsymbol{q}\cdot\boldsymbol{n}_K$ for all $K \in \mathcal{T}_H$. Then, we have

$$0 = \langle \mu, v\rangle_{\partial\mathcal{T}_H} = \langle \boldsymbol{q}\cdot\boldsymbol{n}, v\rangle_{\partial\mathcal{T}_H} = (\boldsymbol{\nabla}\cdot\boldsymbol{q}, v)_{\mathcal{T}_H} + (\boldsymbol{q}, \boldsymbol{\nabla} v)_{\mathcal{T}_H}.$$

Hence, the result follows from Theorem 5.1.4. □

The primal hybrid formulation consists in finding a couple $(\lambda, u) \in \Lambda \times V$ such that

$$\left\{\begin{array}{rcll} b_{\mathcal{T}_H}(u,v) + \langle \lambda, v\rangle_{\partial\mathcal{T}_H} & = & (\varepsilon f, v)_{\widetilde{\Omega}} & \forall v \in V, \\ \langle \mu, v\rangle_{\partial\mathcal{T}_H} & = & 0 & \forall \mu \in \Lambda. \end{array}\right. \tag{5.1.6}$$

As we show next, (5.1.6) is indeed an equivalent reformulation of (5.1.2).

Theorem 5.1.6. *Problem* (5.1.6) *has a unique solution* $(\lambda, u) \in \Lambda \times V$*. Moreover* $u \in H^1_\sharp(\widetilde{\Omega})$ *is the solution of problem* (5.1.1) *and*

$$\lambda = -\boldsymbol{D}\boldsymbol{A}\boldsymbol{\nabla} u\cdot\boldsymbol{n}_K \text{ on } \partial K \ \ \forall K \in \mathcal{T}_H.$$

Proof. Let $(\lambda, u) \in \Lambda \times V$, be a solution to (5.1.6). The second equation of (5.1.6) is

$$\langle \mu, u\rangle_{\partial\mathcal{T}_H} = 0 \quad \forall \mu \in \Lambda,$$

and Lemma 5.1.5 then implies that $u \in H^1_\sharp(\widetilde{\Omega})$. On the other hand, picking a test function $v \in H^1_\sharp(\widetilde{\Omega}) \subset V$ in the first equation of (5.1.6), we see that

$$b_{\mathcal{T}_H}(u,v) = b(u,v) = (\varepsilon f, v)_{\widetilde{\Omega}},$$

since $\langle \lambda, v\rangle_{\partial\mathcal{T}_H} = 0$, due to Lemma 5.1.5 again. Since this is the usual weak formulation of the PML Helmholtz problem, we conclude that u is the weak solution to (5.1.1).

Then, considering a function $v \in H^1(K) \subset V$ in the first equation of (5.1.6), we have

$$b_{\mathcal{T}_H}(u,v) + \langle \lambda, v\rangle_{\partial K} = (\varepsilon f, v)_K,$$

and integrating by part the weak formulation in a, we obtain that

$$(\varepsilon f, v)_K + \langle \boldsymbol{D}\boldsymbol{A}\boldsymbol{\nabla} u\cdot\boldsymbol{n}_K, v\rangle + \langle \lambda, v\rangle_{\partial K} = (\varepsilon f, v)_K,$$

so that $\lambda = -\boldsymbol{D}\boldsymbol{A}\boldsymbol{\nabla} u\cdot\boldsymbol{n}_K$ on ∂K.

This shows that the solution to (5.1.6) is unique, if it exists. To show the existence, we let $u \in H^1_\sharp(\widetilde{\Omega})$ be the solution to (5.1.1). Firstly, Lemma 5.1.5 gives that

$$\langle \mu, u \rangle_{\partial \mathcal{T}_H} = 0 \quad \forall \mu \in \Lambda.$$

Secondly, taking $v \in V$ and applying a elementwise integration by parts to (5.1.1), we obtain

$$b_{\mathcal{T}_H}(u, v) + \langle \lambda, v \rangle_{\partial \mathcal{T}_H} = (\varepsilon f, v)_{\widetilde{\Omega}} \qquad \forall v \in V,$$

where $\lambda = -\boldsymbol{D}\boldsymbol{A}\boldsymbol{\nabla} u \cdot \boldsymbol{n}_K$ on ∂K, for all $K \in \mathcal{T}_H$. □

Remark 5.1.7 (Link with optimization)**.** *When the sesquilinar form $b_{\mathcal{T}_H}(\cdot,\cdot)$ is coercive, the original formulation of the problem (5.1.2) can be viewed as the Euler-Lagrange equations to the minimization problem*

$$\min_{u \in H^1_\star(\widetilde{\Omega})} \operatorname{Re}\left\{ \frac{1}{2} b_{\mathcal{T}_H}(u,u) - (\varepsilon f, u)_{\widetilde{\Omega}} \right\}. \tag{5.1.7}$$

We can think of the looking for $u \in V$ instead of $H^1_\star(\widetilde{\Omega})$ as a unconstrained minimization problem. The theory of Lagrange multipliers then allows us to equivalently rewrite (5.1.7) as

$$\max_{\lambda \in \Lambda} \min_{u \in V} \operatorname{Re}\left\{ \frac{1}{2} b_{\mathcal{T}_H}(u,u) - (\varepsilon f, u)_{\widetilde{\Omega}} + \langle \lambda, u \rangle_{\mathcal{T}_H} \right\}. \tag{5.1.8}$$

In this context, (5.1.8) is known as a saddle point problem, and λ as a Lagrange multiplier.

In the context of Helmholtz problems, $b(\cdot,\cdot)$ is not coercive, so that the primal hybrid formulation (5.1.6) does not correspond to the Euler-Lagrange equations of (5.1.8). Nevertheless, (5.1.6) still characterizes a critical point of the functional appearing in (5.1.8). For the reason, we will sometimes refer to (5.1.6) as a saddle point problem and to λ as a Lagrange multiplier.

5.2 The multiscale hybrid-mixed method

In the MHM formulation, we substitute u for λ in order to obtain a problem set on the skeleton $\partial\mathcal{T}_H$ of the mesh involving only λ as unknown. We will see that it amounts to characterize the solution of (5.1.6) as a collection of local solutions that are tied together through a global problem.

To do so, we start with the first equation of the hybrid formulation (5.1.6), and write that

$$b_{\mathcal{T}_H}(u, v) = (\varepsilon f, v)_{\widetilde{\Omega}} - \langle \lambda, v \rangle_{\partial \mathcal{T}_H}.$$

We remark that at least formally, by linearity, the solution can be written as

$$u = \widehat{T} f + T\lambda, \tag{5.2.1}$$

where $T : \Lambda \to X$ and $\hat{T} : L^2(\widetilde{\Omega}) \to X$ are two linear bounded operators defined by

$$b_{\mathcal{T}_H}(\widehat{T}f, v) = (\varepsilon f, v)_{\widetilde{\Omega}} \quad \forall v \in V, \tag{5.2.2}$$

and

$$b_{\mathcal{T}_H}(T\lambda, v) = -\langle \lambda, v \rangle_{\partial\mathcal{T}_H} \quad \forall v \in V. \tag{5.2.3}$$

Thus, using the definition of the applications $b_{\mathcal{T}_H}$ and $\langle \cdot, \cdot \rangle_{\partial\mathcal{T}_H}$ in the broken space V, we find that the operators T and $\widehat{T}$ are defined locally in each element $K \in \mathcal{T}_H$ as the solutions to the following local Neumann Helmholtz problems

$$\begin{cases} -k^2\nu\varepsilon T_K\lambda - \boldsymbol{\nabla} \cdot (\boldsymbol{DA}\boldsymbol{\nabla} T_K\lambda) &= 0 & \text{in } K, \\ \boldsymbol{DA}\boldsymbol{\nabla}\,(T_K\lambda) \cdot \boldsymbol{n} &= -\lambda & \text{on } \partial K. \end{cases} \tag{5.2.4a}$$

and

$$\begin{cases} -k^2\nu\varepsilon \hat{T}_K f - \boldsymbol{\nabla} \cdot \left(\boldsymbol{DA}\boldsymbol{\nabla} \hat{T}_K f\right) &= \varepsilon f & \text{in } K, \\ \boldsymbol{DA}\boldsymbol{\nabla}\left(\hat{T}_K f\right) \cdot \boldsymbol{n} &= 0 & \text{on } \partial K. \end{cases} \tag{5.2.4b}$$

Assuming that the above problems lead to a sound definition of the operators T and $\widehat{T}$, we can substitute the decomposition $u = \widehat{T}f + T\lambda$ in the second equation of the hybrid formulation (5.1.6), leading to the global MHM problem: Find $\lambda \in \Lambda$ such that

$$\langle \mu, T\lambda \rangle_{\partial\mathcal{T}_H} = -\langle \mu, \widehat{T}f \rangle_{\partial\mathcal{T}_H} \quad \forall \mu \in \Lambda, \tag{5.2.5}$$

where λ is the only unknown.

The MHM solution splitting (5.2.1) and the definition (5.2.4) suggest that the solution can be expressed in each element as a sum of two operators which are solutions to a element-wise Helmholtz problem. The global problem (5.2.5) then ties together these local contributions. This is illustrated in Figure 5.1.

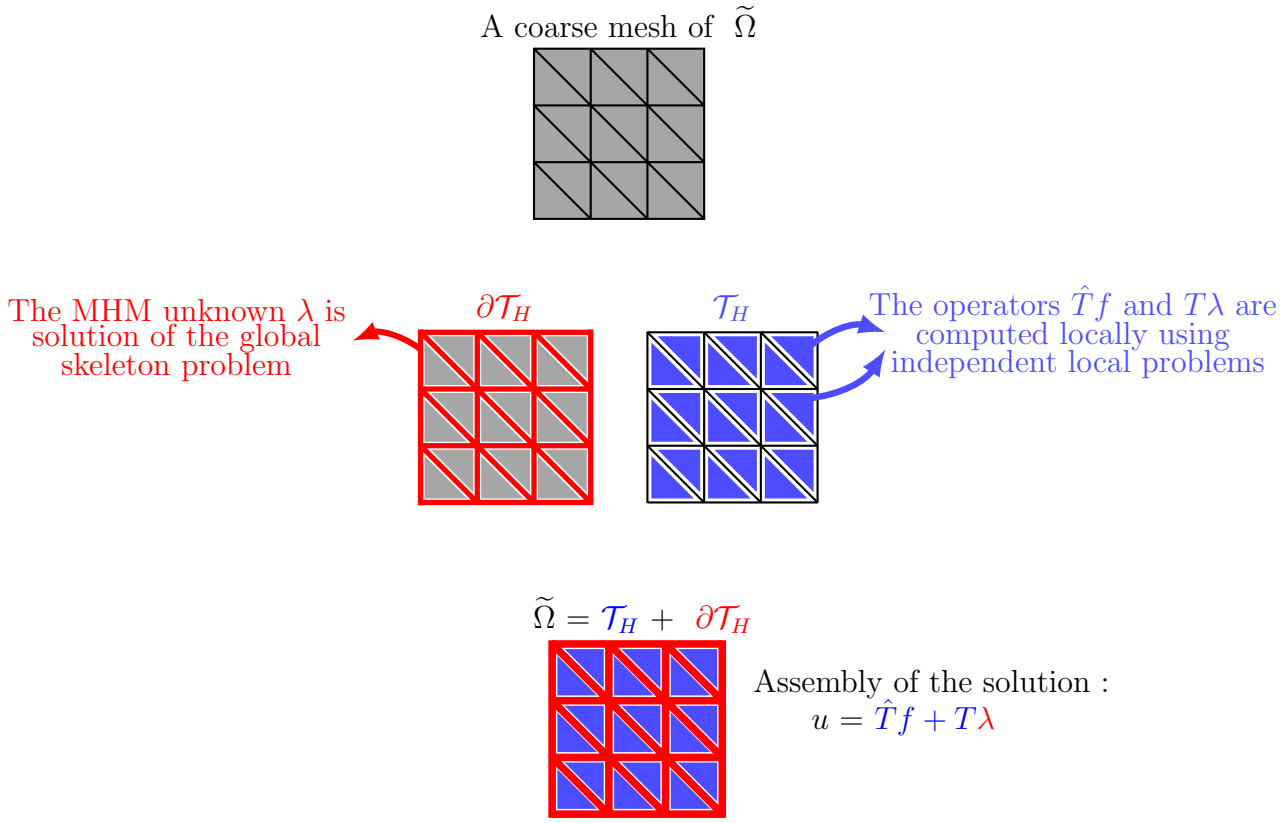


Figure 5.1: Support of the global and local MHM formulations.

In the remainder of this section, we will address the well-posedness of the global MHM problem (5.2.5) and of its discretization. Specifically, in subsection 5.2.1, we will show that the local problem (5.2.4a) and (5.2.4b) are well-posed, which in turns guarantee that the operators T and $\widehat{T}$ are well-defined. We will then show in subsection 5.2.2 an inf-sup condition for the sesquilinear form $\langle \cdot, T\cdot \rangle_{\partial\mathcal{T}_H}$ appearing in (5.2.5). Subsection 5.2.3 then deals with stability and convergence of discretizations to (5.2.5). Finally, we present some implementation details in subsection 5.2.4.

Remark 5.2.1. *The local problems* (5.2.4) *are very similar to the ones derived in [36] where the authors study the Helmholtz equation, but without quasi-periodic boundary conditions. The main differences between [36] and the present work is the presence of PML (which changes the coefficients) and quasi-periodic boundary conditions (which change the space* Λ*).*

5.2.1 Elementwise problems

We start by analyzing the well-posedness of the local MHM problems (5.2.4) defining the operators T and $\widehat{T}$. As indicated in Section 2.1.10, the well-posedness of these Helmholtz problems is equivalent to establishing an ins-sup condition the sesquilinear form.

The local problems (5.2.4) defining the operators are Helmholtz problems with Neumann boundary conditions. Such problems are not always well-posed for all frequencies. Specifically, the problem is not well-posed at resonant frequencies, which correspond to (the square of) an eigenvalue of the Laplace operator with Neumann boundary condition on ∂K. As observed in [36], however, resonant frequencies can be ruled out if the mesh $\mathcal{T}_H$ is sufficiently fine. This leads us to the following assumption.

Assumption 5.2.2. *All the elements $K \in \mathcal{T}_H$ are convex, and there exists $\tau > 0$ such that*

$$\frac{kH_K}{\vartheta_K} \leq \left(|\nu_K|^{-2} - \tau\right)^{1/2} \pi, \tag{5.2.6}$$

where $\nu_K := \nu|_K \in \mathbb{C}$ and

$$\vartheta_K := \min_K \min(A_1, A_2) / \max_K \varepsilon.$$

In our local problems well-posedness analysis, we proceed in the same way as in [36] for a Helmholtz equation with a first-order absorbing condition. Here, considering a PML layer, the derivation of the global-local MHM formulations seems to be similar. However, the PML damping coefficient ν affects the definition of the sesquilinear form "$b_{\mathcal{T}_H}$", and condition (5.2.6) shows that we need slightly smaller elements in the PML region to get the local problems coercivity.

Below, we use the following notation: for two positive real numbers $A, B \geq 0$, we will write $A \lesssim B$ if there exists a constant $C > 0$ which is independent of A, B, H and k, but which possibly depends on $\widetilde{\Omega}$, $\varepsilon_{\min}$, $\varepsilon_{\max}$, $A_{\min}$, $A_{\max}$, γ_{r}, γ_{i}, and τ such that $A \leq CB$. We also write $A \gtrsim B$ when $B \lesssim A$.

We start by observing that the continuity of $b_{\mathcal{T}_H}$ on $V \times V$ can be straightforwardly established by a triangle inequality.

Lemma 5.2.3 (Continuity)**.** *For all $\phi, v \in V$, it holds that*

$$|b_{\mathcal{T}_H}(\phi, v)| \lesssim |\!|\!|\phi|\!|\!|_{k,\mathcal{T}_H} \, |\!|\!|v|\!|\!|_{k,\mathcal{T}_H} \,. \tag{5.2.7}$$

For convenience, we introduce the local sesquilinear from

$$b_K(\phi, v) = -k^2(\nu\varepsilon\phi, v)_K + (\boldsymbol{D}\boldsymbol{A}\boldsymbol{\nabla}\phi, \boldsymbol{\nabla}v)_K \qquad \forall \phi, v \in H^1(K)$$

Straightforward computations reveal that identities

$$\operatorname{Re} b_K(v, v) = \gamma_{\mathrm{r}} \left(-k^2 \|v\|_{\varepsilon,K} + \left\| \frac{\partial u}{\partial \boldsymbol{x}_1} \right\|_{A_1,K} + |\nu|^{-2} \left\| \frac{\partial u}{\partial \boldsymbol{x}_2} \right\|_{A_2,K} \right),$$

and

$$\operatorname{Im} b_K(v, v) = \gamma_{\mathrm{i}} \left(-k^2 \|v\|_{\varepsilon,K} + \left\| \frac{\partial v}{\partial \boldsymbol{x}_1} \right\|_{A_1,K} - |\nu|^{-2} \left\| \frac{\partial u}{\partial \boldsymbol{x}_2} \right\|_{A_2,K} \right).$$

We then show that these local sesquilinear forms satisy inf-sup conditions.

Lemma 5.2.4 (Local inf-sup constants)**.** *Let $K \in \mathcal{T}_H$. For $u \in H^1(K)$, there exists an element $u^\star \in H^1(K)$ such that*

$$\operatorname{Re} b_K(u, u^\star) \gtrsim |\!|\!| u |\!|\!|_{k,K}^2 \qquad \text{and} \qquad |\!|\!| u |\!|\!|_{k,K} \gtrsim |\!|\!| u^\star |\!|\!|_{k,K}. \tag{5.2.8}$$

Proof. Fix $K \in \mathcal{T}_H$ and $u \in H^1(K)$. Let

$$u_0 = \frac{1}{|K|}\int_K u \qquad \text{and} \qquad u^\perp = u - u_0.$$

Using the function $v^\star = u - 2u_0$ and that $\boldsymbol{\nabla} u_0 = 0$, we have

$$\begin{aligned}
\operatorname{Re} b_K(u, v^\star) &= \operatorname{Re} b_K(u^\perp + u_0, u^\perp - u_0) \\
&= \operatorname{Re} b_K(u^\perp, u^\perp) - \operatorname{Re} b_K(u_0, u_0) \\
&= \operatorname{Re} b_K(u^\perp, u^\perp) + k^2 \int_K \varepsilon \gamma_{\mathrm{r}} |u_0|^2,
\end{aligned}$$

so that

$$\operatorname{Re} b_K(u, v^\star) \geq \operatorname{Re} b_K(u^\perp, u^\perp). \tag{5.2.9}$$

Using that $u^\perp$ has zero mean value and recalling that K is convex with diameter H_K, we apply the Poincare-Wirtinger inequality [117] to get

$$\|u^\perp\|_K \leq \frac{H_K}{\pi} \|\boldsymbol{\nabla} u^\perp\|_K,$$

hence

$$\|u^\perp\|_{\varepsilon,K}^2 \leq \frac{H^2}{\pi^2} \frac{\varepsilon_{K,\max}}{A_{K,\min}} \|\boldsymbol{\nabla} u^\perp\|_{\boldsymbol{A},K}^2,$$

and we deduce

$$\begin{aligned}
\operatorname{Re} b_K(u^\perp, u^\perp) &\geq -\gamma_{\mathrm{r}} k^2 \|u^\perp\|_{\varepsilon,K}^2 + \gamma_{\mathrm{r}} \left(\|\partial_1 u^\perp\|_{A_1,K}^2 + |\nu|^{-2} \|\partial_2 u^\perp\|_{A_2,K}^2 \right) \\
&\geq \gamma_{\mathrm{r}} \left(|\nu|^{-2} - \frac{\varepsilon_{K,\max}}{A_{K,\min}} \frac{k^2 H_K^2}{\pi^2} \right) \|\boldsymbol{\nabla} u^\perp\|_{\boldsymbol{A},K}^2 \\
&= \gamma_{\mathrm{r}} \left(|\nu|^{-2} - \frac{k^2 H_K^2}{\pi^2 \vartheta_K^2} \right) \|\boldsymbol{\nabla} u\|_{\boldsymbol{A},K}^2.
\end{aligned}$$

Assumption (5.2.6) gives

$$\frac{H^2 k^2}{\pi^2 \vartheta_K^2} \leq |\nu|^{-2} - \tau,$$

and therefore

$$b_K(u^\perp, u^\perp) \geq \tau \gamma_{\mathrm{r}} \|\boldsymbol{\nabla} u\|_{\boldsymbol{A},K}^2, \tag{5.2.10}$$

Thus, using (5.2.9) and (5.2.10) we obtain

$$\operatorname{Re} b_K(u, v^\star) \geq \tau \gamma_{\mathrm{r}} \|\boldsymbol{\nabla} u\|_{\boldsymbol{A},K}^2. \tag{5.2.11}$$

On the other hand, observe that

$$\begin{aligned}\operatorname{Re} b_K(u,-u) &= \gamma_{\mathrm{r}} k^2 \|u\|_{\varepsilon,K}^2 - \gamma_{\mathrm{r}} \|\partial_1 u\|_{A_1,K}^2 - \frac{\gamma_{\mathrm{r}}}{|\nu|^2}\|\partial_2 u\|_{\mu_2,K}^2 \\ &\geq \gamma_{\mathrm{r}} k^2 \|u\|_{\varepsilon,K}^2 - \gamma_{\mathrm{r}} \|\boldsymbol{\nabla} u\|_{\boldsymbol{A},K}^2,\end{aligned}$$

as a result

$$\begin{aligned}\operatorname{Re} b_K\left(u, 2\tau^{-1}v^\star - u\right) &\geq \gamma_{\mathrm{r}} k^2 \|u\|_{\varepsilon,K}^2 + \gamma_{\mathrm{r}} |\boldsymbol{\nabla} u|_{\boldsymbol{A},K}^2 \\ &\geq \gamma_{\mathrm{r}} \left(k^2 \|u\|_{\varepsilon,K}^2 + |\boldsymbol{\nabla} u|_{\boldsymbol{A},K}^2\right) \\ &\geq \gamma_{\mathrm{r}} \|u\|_{k,K}^2,\end{aligned}$$

and (5.2.8) follows by taking $u^\star := 2\tau^{-1}v^\star - u$. □

Lemma 5.2.5 (Inf-sup condition)**.** *For $u \in V$, it holds that*

$$\sup_{v\in V} \frac{\operatorname{Re} b_{\mathcal{T}_H}(u,v)}{\|v\|_{k,\mathcal{T}_H}} \gtrsim \|u\|_{k,\mathcal{T}_H}. \tag{5.2.12}$$

Proof. Let $u \in V$. Then, for each element $K \in \mathcal{T}_H$, we have $u|_K \in H^1(K)$. We can then define $u^\star \in V$ by defining $(u^\star)|_K = (u|_K)^\star$ as in Lemma 5.2.4. Then, we have

$$\operatorname{Re} b_K(u,u^\star) \gtrsim \|u\|_{k,K}^2 \qquad \text{and} \qquad \|u\|_{k,K}^2 \gtrsim \|u^\star\|_{k,K}^2,$$

and it follows by summation over $K \in \mathcal{T}_H$ that

$$\operatorname{Re} b_{\mathcal{T}_H}(u,u^\star) \gtrsim \|u\|_{k,\mathcal{T}_H}^2 \qquad \text{and} \qquad \|u\|_{k,\mathcal{T}_H}^2 \gtrsim \|u^\star\|_{k,\mathcal{T}_H}^2,$$

so that (5.2.12) follows from

$$\sup_{v\in V} \frac{\operatorname{Re} b_{\mathcal{T}_H}(u,v)}{\|v\|_{k,\mathcal{T}_H}} \geq \frac{\operatorname{Re} b_{\mathcal{T}_H}(u,u^\star)}{\|u^\star\|_{k,\mathcal{T}_H}} \gtrsim \frac{\|u\|_{k,\mathcal{T}_H}^2}{\|u^\star\|_{k,\mathcal{T}_H}} \gtrsim \frac{\|u\|_{k,\mathcal{T}_H}\|u^\star\|_{k,\mathcal{T}_H}}{\|u^\star\|_{k,\mathcal{T}_H}} = \|u\|_{k,\mathcal{T}_H}.$$

□

As a result, the local problems defining T and $\widehat{T}$ are well-posed, and we can obtain explicit stability estimates in Theorem 5.2.6 below.

Theorem 5.2.6 (Well-posedness of the local problems)**.** *For all $\mu \in \Lambda$ and $f \in L^2(\widetilde{\Omega})$, there exist unique elements $T\lambda, \widehat{T}f \in V$ such that*

$$b_{\mathcal{T}_H}(T\lambda, v) = \langle \lambda, v\rangle_{\partial\mathcal{T}_H} \qquad b_{\mathcal{T}_H}(\hat{T}f, v) = (\varepsilon f, v)_{\widetilde{\Omega}} \qquad \forall v \in V.$$

In addition, we have

$$\|T\lambda\|_{k,\mathcal{T}_H} \lesssim \|\lambda\|_{\Lambda,k} \quad \text{and} \quad \left\|\widehat{T}f\right\|_{k,\mathcal{T}_H} \lesssim k^{-1}\|f\|_{\varepsilon,\widetilde{\Omega}}. \tag{5.2.13}$$

Proof. For each $\mu \in \Lambda$ and $f \in L^2(\widetilde{\Omega})$, the existence and uniqueness of $T\lambda$ and $\widehat{T}f$ is ensured by the continuity of the sesquilinear form $b_{\mathcal{T}_H}$ given in (5.2.7) together with the inf-sup condition (5.2.12). Furthermore, if $\lambda \in \Lambda$, we have

$$\|\!|T\lambda|\!\|_{k,\mathcal{T}_H} \lesssim \sup_{v\in V} \frac{\operatorname{Re} b_{\mathcal{T}_H}(T\lambda, v)}{\|\!|v|\!\|_{k,\mathcal{T}_H}} = -\sup_{v\in V} \frac{\operatorname{Re}\langle \lambda, v\rangle_{\partial\mathcal{T}_H}}{\|\!|v|\!\|_{k,\mathcal{T}_H}} = \|\lambda\|_{\Lambda,k}.$$

Similarly, for $f \in L^2(\widetilde{\Omega})$, we can write

$$\operatorname{Re} b_{\mathcal{T}_H}(\widehat{T}f, v) = \operatorname{Re}(\varepsilon f, v)_{\widetilde{\Omega}} \lesssim \|f\|_{\varepsilon,\widetilde{\Omega}} \|v\|_{\varepsilon,\widetilde{\Omega}} \lesssim k^{-1}\|f\|_{\varepsilon,\widetilde{\Omega}} \|\!|v|\!\|_{k,\mathcal{T}_H},$$

so that

$$\left\|\!\left|\widehat{T}f\right|\!\right\|_{k,\mathcal{T}_H} \lesssim \sup_{v\in V} \frac{\operatorname{Re} b_{\mathcal{T}_H}(\widehat{T}f, v)}{\|\!|v|\!\|_{k,\mathcal{T}_H}} \lesssim k^{-1}\|f\|_{\varepsilon,\widetilde{\Omega}}.$$

□

We close this subsection by the following norm equivalence result.

Lemma 5.2.7 (Norm equivalence)**.** *The norms $\|\cdot\|_{\Lambda,k}$ and $\|\!|T\cdot|\!\|_{k,\mathcal{T}_H}$ are equivalent on Λ. Specifically, we have*

$$\|\mu\|_{\Lambda,k} \lesssim \|\!|T\mu|\!\|_{k,\mathcal{T}_H} \lesssim \|\mu\|_{\Lambda,k} \qquad \forall \mu \in \Lambda. \tag{5.2.14}$$

Proof. Let $\mu \in \Lambda$. On the one hand, we have

$$\begin{aligned}
\|\!|T\mu|\!\|_{k,\mathcal{T}_H} &\lesssim \sup_{v\in V} \frac{\operatorname{Re} b_{\mathcal{T}_H}(T\mu, v)}{\|\!|v|\!\|_{k,\mathcal{T}_H}} && \text{using inf-sup condition (5.2.12)}\\
&= \sup_{v\in V} \left\{ -\frac{\operatorname{Re}\langle \mu, v\rangle_{\partial\mathcal{T}_H}}{\|\!|v|\!\|_{k,\mathcal{T}_H}} \right\} && \text{using the definition of } T \text{ in (5.2.3)}\\
&\lesssim \|\mu\|_{\Lambda,k} && \text{using the norm } \|\cdot\|_{\Lambda,k} \text{ definition.}
\end{aligned}$$

On the other hand

$$\begin{aligned}
\|\mu\|_{\Lambda,k} &\lesssim \sup_{v\in V} \frac{\operatorname{Re}\langle \mu, v\rangle_{\partial\mathcal{T}_H}}{\|\!|v|\!\|_{k,\mathcal{T}_H}} && \text{using the norm } \|\cdot\|_{\Lambda,k} \text{ definition}\\
&= \sup_{v\in V} \left\{ -\frac{\operatorname{Re} b_{\mathcal{T}_H}(T\mu, v)}{\|\!|v|\!\|_{k,\mathcal{T}_H}} \right\} && \text{using the definition of } T \text{ in (5.2.3)}\\
&\lesssim \|\!|T\mu|\!\|_{k,\mathcal{T}_H} && \text{using the continuity of } b_{\mathcal{T}_H}(\cdot,\cdot).
\end{aligned}$$

□

5.2.2 The MHM formulation

We now show that the global MHM problem is well-posed. We will largely follow the analysis in [36] with slight modifications to accommodate for the absorbing layer and the quasi-periodic boundary conditions.

Theorem 5.2.8. *There exists a unique solution $\lambda \in \Lambda$ solution to (5.2.5). In addition, we have*

$$k\|\lambda\|_{\Lambda,k} \lesssim \mathscr{C}_{\mathrm{st}} \|f\|_{\varepsilon,k}. \tag{5.2.15}$$

Proof. Proposition 5.1.6 affirms that there exists a unique couple $(u,\lambda) \in V \times \Lambda$ solution of the hybrid problem (5.1.6), where u is the usual solution to the Helmholtz equation and λ is defined as $\boldsymbol{DA}$ multiplied by the normal derivative of u on the boundary of each elements $K \in \mathcal{T}_H$.

Furthermore, recalling Theorem 5.2.6, the operators T and $\widehat{T}$ are well-defined and invertible. As a result, the first equation of (5.1.6) shows that

$$u = \widehat{T} f + T\lambda. \tag{5.2.16}$$

Injecting 5.2.16 into the second equation of (5.1.6) shows that $\lambda = -\boldsymbol{DA}\boldsymbol{\nabla} u \cdot \boldsymbol{n}_K$ is the solution to the continuous MHM formulation (5.2.5), and existence follows.

Furthermore, uniqueness follows as the couple (u,λ) is unique. Recalling (5.2.16), we have

$$\begin{aligned}
\|\lambda\|_{\Lambda,k} &\lesssim |\!|\!| T\lambda |\!|\!|_{k,\mathcal{T}_H} && \text{using the norm equivalence (5.2.14)} \\
&\lesssim |\!|\!| u |\!|\!|_{k,\mathcal{T}_H} + \left|\!\left|\!\left| \widehat{T} f \right|\!\right|\!\right|_{k,\mathcal{T}_H} && \text{using the solution splitting (5.2.16)} \\
&\lesssim \left(\frac{\mathscr{C}_{\mathrm{st}}}{k} + k^{-1} \right) \|f\|_{\varepsilon,\widetilde{\Omega}} && \text{using (5.2.13) and (5.1.3)},
\end{aligned}$$

and the result follows since $\mathscr{C}_{\mathrm{st}} \geq 1$. □

The next goal of this subsection is to derive an inf-sup condition of the sesquilinear form $\langle \cdot, T\cdot \rangle_{\partial\mathcal{T}_H}$ and find the frequency dependence of its inf-sup constant. The first step to do so is to show a symmetry property of $\langle \cdot, T\cdot \rangle_{\partial\mathcal{T}_H}$.

Lemma 5.2.9 (Symmetry of $\langle T\cdot, \cdot \rangle_{\partial\mathcal{T}_H}$)**.** *For all $\mu, \lambda \in \Lambda$, we have*

$$\langle \mu, T\lambda \rangle_{\partial\mathcal{T}_H} = \left\langle \overline{\lambda}, T\overline{\mu} \right\rangle_{\partial\mathcal{T}_H}.$$

Proof. Let $\mu, \lambda \in \Lambda$. We have

$$\overline{\langle \mu, T\lambda \rangle_{\partial\mathcal{T}_H}} = \langle \overline{\mu}, \overline{T}\lambda \rangle_{\partial\mathcal{T}_H} = -b_{\mathcal{T}_H}(T\overline{\mu}, \overline{T}\lambda) = -b_{\mathcal{T}_H}(T\lambda, \overline{T}\overline{\mu}) = \langle \lambda, \overline{T}\overline{\mu} \rangle_{\partial\mathcal{T}_H},$$

and the result follows by taking the complex conjugate. □

We then introduce a key function used in duality argument.

Lemma 5.2.10. *For $\lambda \in \Lambda$, define $\eta_\lambda \in \Lambda$ as the unique solution to*

$$\langle \mu, T\overline{\eta_\lambda}\rangle_{\partial\mathcal{T}_H} = -\langle \mu, \widehat{T}\big(\overline{T}\lambda\big)\rangle_{\partial\mathcal{T}_H} \qquad \mu \in \Lambda. \tag{5.2.17}$$

Then we have

$$\langle \eta_\lambda, T\lambda\rangle_{\partial\mathcal{T}_H} = \|T\lambda\|^2_{\varepsilon,\widetilde{\Omega}}, \tag{5.2.18}$$

and

$$|\!|\!| T\eta_\lambda |\!|\!|_{k,\mathcal{T}_H} \lesssim \frac{\mathscr{C}_{\mathrm{st}}}{k}\|T\lambda\|_{\varepsilon,\widetilde{\Omega}}. \tag{5.2.19}$$

Proof. Let $\lambda \in \Lambda$. According to Theorem 5.2.8 the definition (5.2.17) is well-posed and η_λ exists and it is well-defined. Then, recalling Lemma 5.2.9, it holds that

$$\langle \eta_\lambda, T\lambda\rangle_{\partial\mathcal{T}_H} = \langle \overline{\lambda}, T\overline{\eta_\lambda}\rangle_{\partial\mathcal{T}_H} = -\langle \overline{\lambda}, \hat{T}\big(\overline{T}\lambda\big)\rangle_{\partial\mathcal{T}_H}.$$

Next, using the definitions of T and $\hat{T}$, we show that

$$\begin{aligned} -\overline{\langle \overline{\lambda}, \hat{T}(\overline{T}\lambda)\rangle}_{\partial\mathcal{T}_H} = -\langle \lambda, \overline{\hat{T}}(\overline{T}\lambda)\rangle_{\partial\mathcal{T}_H} &= b_{\mathcal{T}_H}(T\lambda, \overline{\hat{T}}(\overline{T}\lambda)) \\ &= b_{\mathcal{T}_H}(\hat{T}(\overline{T}\lambda), \overline{T}\lambda) = (\varepsilon\overline{T}\lambda, \overline{T}\lambda)_{\widetilde{\Omega}} = \|\overline{T}\lambda\|^2_{\varepsilon,\widetilde{\Omega}}, \end{aligned}$$

and taking the complex conjugate we get

$$-\langle \overline{\lambda}, \hat{T}(T\lambda)\rangle_{\partial\mathcal{T}_H} = \overline{\|\overline{T}\lambda\|^2_{\varepsilon,\widetilde{\Omega}}} = \|\overline{T}\lambda\|^2_{\varepsilon,\widetilde{\Omega}} = \|T\lambda\|^2_{\varepsilon,\widetilde{\Omega}}$$

and (5.2.18) follows. Finally, the estimate (5.2.19) follows from the definition of η_λ and Theorem 5.2.8. □

We are now ready to derive an inf-sup condition of the MHM global problem (5.2.5).

Theorem 5.2.11. *We have*

$$\inf_{\substack{\lambda\in\Lambda \\ \|\lambda\|_{\Lambda,k}=1}} \sup_{\substack{\mu\in\Lambda \\ \|\mu\|_{\Lambda,k}=1}} \operatorname{Re}\langle \mu, T\lambda\rangle_{\partial\mathcal{T}_H} \gtrsim \frac{1}{\mathscr{C}_{\mathrm{st}}}. \tag{5.2.20}$$

Proof. Let $\lambda \in \Lambda$, we have

$$\begin{aligned} -\operatorname{Re}\langle \lambda, T\lambda\rangle_{\partial\mathcal{T}_H} &= \operatorname{Re} b_{\mathcal{T}_H}(T\lambda, T\lambda) \\ &= \gamma_{\mathrm{r}}\left(\|\partial_1(T\lambda)\|^2_{A_1,\widetilde{\Omega}} + |\nu|^{-2}\|\partial_2(T\lambda)\|^2_{A_2,\widetilde{\Omega}} - k^2\|T\lambda\|^2_{\varepsilon,\widetilde{\Omega}}\right) \\ &\geq \gamma_{\mathrm{r}}\left(|\nu|^{-2}\|\boldsymbol{\nabla}(T\lambda)\|^2_{\boldsymbol{A},\widetilde{\Omega}} - k^2\|T\lambda\|^2_{\varepsilon,\widetilde{\Omega}}\right). \end{aligned}$$

Next, we define $\mu = \gamma_{\mathrm{r}}^{-1}\lambda - 2k^2\eta_\lambda$, where η_λ is defined as in Lemma 5.2.10 we obtain

$$\begin{aligned} -\operatorname{Re}\langle \mu, T\lambda\rangle_{\partial\mathcal{T}_H} &= -\gamma_{\mathrm{r}}^{-1}\operatorname{Re}\langle \lambda, T\lambda\rangle_{\partial\mathcal{T}_H} + 2k^2\operatorname{Re}\langle \eta_\lambda, T\lambda\rangle_{\partial\mathcal{T}_H} \\ &\gtrsim |\nu|^{-2}\|\boldsymbol{\nabla}(T\lambda)\|^2_{\boldsymbol{A},\widetilde{\Omega}} - k^2\|T\lambda\|^2_{\varepsilon,\widetilde{\Omega}} + 2k^2\|T\lambda\|^2_{\varepsilon,\widetilde{\Omega}} \\ &\gtrsim |\nu|^{-2}\,|\!|\!|T\lambda|\!|\!|^2_{k,\mathcal{T}_H}. \end{aligned}$$

Hence, it remains to show that $\|\!|T\mu|\!\|_{k,\mathcal{T}_H} \lesssim \mathscr{C}_{\mathrm{st}} \|\!|T\lambda|\!\|_{k,\mathcal{T}_H}$

$$\begin{aligned}
\|\!|T\mu|\!\|_{k,\mathcal{T}_H} &\lesssim \|\!|T\lambda|\!\|_{k,\mathcal{T}_H} + 2k^2 \|\!|T\eta_\lambda|\!\|_{k,\mathcal{T}_H} \\
&\lesssim \|\!|T\lambda|\!\|_{k,\mathcal{T}_H} + 2k^2 \frac{\mathscr{C}_{\mathrm{st}}}{k} \|T\lambda\|_{\varepsilon,\widetilde{\Omega}} \\
&\lesssim (1 + 2\mathscr{C}_{\mathrm{st}}) \|\!|T\lambda|\!\|_{k,\mathcal{T}_H} \\
&\lesssim \mathscr{C}_{\mathrm{st}} \|\!|T\lambda|\!\|_{k,\mathcal{T}_H} .
\end{aligned}$$

□

5.2.3 Discretization

This subsection deals with the discretization of the MHM formulation given in (5.2.5). Specifically, the MHM formulation involves the infinite-dimensional space Λ that we need to discretize to obtain a square-linear system. Although several options are possible, in this work, we focus on the simplest discretization space, and we set

$$\Lambda_H := \left\{ \mu_H \in \Lambda \cap L^2(\partial\mathcal{T}_H) \mid \mu_H|_K \in \mathcal{P}_0(\mathcal{F}_K) \; \forall K \in \mathcal{T}_H \right\}, \tag{5.2.21}$$

where $\mathcal{P}_0(\mathcal{F}_K)$ stands for the subset of $L^2(\partial K)$ of functions that take a constant value on each face $F \in \mathcal{F}_K$ of K.

The discrete problem then consists in finding $\lambda_H \in \Lambda_H$ such that

$$\langle \mu_H, T\lambda_H \rangle_{\partial\mathcal{T}_H} = -\langle \mu_H, \widehat{T} f \rangle_{\partial\mathcal{T}_H} \quad \forall \mu_H \in \Lambda_H. \tag{5.2.22}$$

Once λ_H is computed as the solution to (5.2.22), an approximation to u is obtained by setting

$$u_H := \widehat{T} f + T\lambda_H. \tag{5.2.23}$$

5.2.3.1 Abstract convergence analysis

We first provide an abstract stability and convergence analysis for (5.2.22). We will follow convergence studies based on Shatz argument of the finite element method in [101] and the MHM method in [36]. In such analysis, a central concept is the approximation factor: a real number that helps characterizing the approximation properties of $\Lambda_H \subset \Lambda$. Speficially, we set

$$\mathscr{C}_{\mathrm{app}} := k \sup_{\substack{f \in L^2(\widetilde{\Omega}) \\ \|f\|_{\varepsilon,\widetilde{\Omega}} = 1}} \inf_{\mu_H \in \Lambda_H} \|\lambda_f - \mu_H\|_{\Lambda,k}, \tag{5.2.24}$$

where λ_f is the unique element of Λ such that

$$\langle \mu, T\lambda_f \rangle_{\partial\mathcal{T}_H} = -\langle \mu, \widehat{T} f \rangle_{\partial\mathcal{T}_H},$$

for all $\mu \in \Lambda$. As a consequence of (5.2.24), for all $f \in L^2(\widetilde{\Omega})$ there exists an element $\mu_H \in \Lambda_H$ such that

$$k\|\lambda_f - \mu_H\|_{\Lambda,k} \leq \mathscr{C}_{\mathrm{app}} \|f\|_{\varepsilon,\widetilde{\Omega}}. \tag{5.2.25}$$

The following Theorem presents a well-posedness result of the discrete global MHM problem (5.2.22).

Theorem 5.2.12 (Discrete inf-sup condition)**.** *There exists a constant $\mathscr{C}^\star$ solely depending on the extremal values of the coefficients, such that, if $\mathscr{C}_{\text{app}} \leq \mathscr{C}^\star$, then*

$$\inf_{\substack{\lambda_H \in \Lambda_H \\ \|\lambda_H\|_{\Lambda,k}=1}} \sup_{\substack{\mu_H \in \Lambda_H \\ \|\mu_H\|_{\Lambda,k}=1}} \operatorname{Re} \langle \mu_H, T\lambda_H \rangle_{\partial \mathcal{T}_H} \gtrsim \frac{1}{\mathscr{C}_{\text{st}}}, \tag{5.2.26}$$

with a hidden constant wish is independent of $\mathscr{C}_{\text{app}}$.

Proof. Let $\lambda_H \in \Lambda_H$. Recalling the proof of Theorem 5.2.11, we have

$$-\operatorname{Re} \langle \mu, T\lambda_H \rangle_{\partial \mathcal{T}_H} \gtrsim |\nu|^{-2} \left\| T\lambda_H \right\|^2_{k,\mathcal{T}_H},$$

where $\mu = \gamma_{\text{r}}^{-1}\lambda_H - 2k^2\eta_{\lambda_H}$. So ,we define $\mu_H \in \Lambda_H$ as

$$\mu_H = \gamma_{\text{r}}^{-1}\lambda_H - 2k^2\eta_H,$$

where η_H is the best approximation of η_{λ_H}. It follows that

$$\mu - \mu_H = 2k^2(\eta_{\lambda_H} - \eta_H),$$

and, recalling the definitions of η_{λ_H} from Lemma 5.2.10 and the definition of $\mathscr{C}_{\text{app}}$, we see that

$$\begin{aligned} \left\| T(\mu - \mu_H) \right\|_{k,\mathcal{T}_H} &\lesssim 2k^2 \left\| T(\eta_{\lambda_H} - \eta_H) \right\|_{k,\mathcal{T}_H} \\ &\lesssim 2k\mathscr{C}_{\text{app}} \|T\lambda_H\|_{\varepsilon,\widetilde{\Omega}} \\ &\lesssim 2\mathscr{C}_{\text{app}} \left\| T\lambda_H \right\|_{k,\mathcal{T}_H}. \end{aligned}$$

Thus, there exist two constants B_1, B_2 such that

$$\begin{aligned} -\operatorname{Re} \langle \mu_H, T\lambda_H \rangle_{\partial \mathcal{T}_H} &= -\operatorname{Re} \langle \mu, T\lambda_H \rangle_{\partial \mathcal{T}_H} + \operatorname{Re} \langle \mu - \mu_H, T\lambda_H \rangle_{\partial \mathcal{T}_H} \\ &\geq B_1|\nu|^{-2} \left\| T\lambda_H \right\|^2_{k,\mathcal{T}_H} - B_2 \left\| T(\mu - \mu_H) \right\|_{k,\mathcal{T}_H} \left\| T\lambda_H \right\|_{k,\mathcal{T}_H} \\ &\geq B_1|\nu|^{-2} \left\| T\lambda_H \right\|^2_{k,\mathcal{T}_H} - 2B_2\mathscr{C}_{\text{app}}) \left\| T\lambda_H \right\|^2_{k,\mathcal{T}_H} \\ &\geq B_2 \left(\frac{B_1}{B_2}|\nu|^{-2} - 2\mathscr{C}_{\text{app}} \right) \left\| T\lambda_H \right\|^2_{k,\mathcal{T}_H} \end{aligned}$$

As a resut, defining

$$\mathscr{C}^\star := \frac{1}{4}\frac{B_1}{B_2}|\nu|^{-2},$$

we have

$$\frac{B_1}{B_2}|\nu|^{-2} - 2\mathscr{C}_{\text{app}} \geq \frac{B_1}{B_2}|\nu|^{-2} - 2\mathscr{C}^\star \geq \frac{1}{2}\frac{B_1}{B_2}|\nu|^{-2} - 2\mathscr{C}_{\text{app}} \gtrsim 1$$

so that

$$-\operatorname{Re}\langle \mu_H, T\lambda_H\rangle_{\mathcal{T}_H} \gtrsim |\!|\!| T\lambda_H |\!|\!|^2_{k,\mathcal{T}_H}.$$

Thus, it remains to show that

$$|\!|\!| T\mu_H |\!|\!|_{k,\mathcal{T}_H} \lesssim \mathscr{C}_{\rm st}\, |\!|\!| T\lambda_H |\!|\!|_{k,\mathcal{T}_H}.$$

We have

$$|\!|\!| T\mu_H |\!|\!|_{k,\mathcal{T}_H} \le |\!|\!| T\mu |\!|\!|_{k,\mathcal{T}_H} + |\!|\!| T(\mu-\mu_H) |\!|\!|_{k,\mathcal{T}_H} \lesssim (\mathscr{C}_{\rm st} + 2\mathscr{C}_{\rm app})\, |\!|\!| T\lambda_H |\!|\!|_{k,\mathcal{T}_H} \lesssim \mathscr{C}_{\rm st}\, |\!|\!| T\lambda_H |\!|\!|_{k,\mathcal{T}_H},$$

using again the smallness assumption on $\mathscr{C}_{\rm app}$. □

Lemma 5.2.13 (Aubin-Nitsche trick)**.** *Let $\lambda \in \Lambda$ solve (5.2.5) and $\lambda_H \in \Lambda_H$ satisfy (5.2.22). It holds that*

$$k\|T(\lambda-\lambda_H)\|_{\varepsilon,\widetilde{\Omega}} \lesssim \mathscr{C}_{\rm app}\, |\!|\!| T(\lambda-\lambda_H) |\!|\!|_{k,\mathcal{T}_H}. \tag{5.2.27}$$

Proof. As in Lemma 5.2.10, we define $\eta \in \Lambda$

$$\langle \mu, T\bar{\eta}\rangle_{\partial\mathcal{T}_H} = -\left\langle \mu, \hat{T}\left(\overline{T}(\lambda-\lambda_H)\right)\right\rangle_{\partial\mathcal{T}_H},$$

so that

$$\langle \eta, T(\lambda-\lambda_H)\rangle_{\partial\mathcal{T}_H} = \|T(\lambda-\lambda_H)\|^2_{\varepsilon,\widetilde{\Omega}}.$$

Then, for all $\eta_H \in \Lambda_H$, it holds that

$$\begin{aligned}
\|T(\lambda-\lambda_H)\|^2_{\varepsilon,\widetilde{\Omega}} &= \langle \eta, T(\lambda-\lambda_H)\rangle_{\partial\mathcal{T}_H} \\
&\lesssim \langle \eta-\overline{\eta_H}, T(\lambda-\lambda_H)\rangle_{\partial\mathcal{T}_H} && \text{by Galerkin's orthogonality} \\
&\lesssim \left\langle \bar{\eta}-\eta_H, \overline{T}(\lambda-\lambda_H)\right\rangle_{\partial\mathcal{T}_H} && \text{by Lemma 5.2.9} \\
&\lesssim -b_{\mathcal{T}_H}(T(\overline{\eta}-\eta_H), \overline{T}(\lambda-\lambda_H)) && \text{by the definition of } T \\
&\lesssim |\!|\!| T(\bar{\eta}-\eta_H) |\!|\!|_{k,\mathcal{T}_H}\, |\!|\!| \overline{T}(\lambda-\lambda_H) |\!|\!|_{k,\mathcal{T}_H} && \text{by the continuity of } b_{\mathcal{T}_H}(\cdot,\cdot).
\end{aligned}$$

In addition, by definition of $\mathscr{C}_{\rm app}$ in (5.2.24), there exists an $\eta_H \in \Lambda_H$ such that

$$k\, |\!|\!| T(\bar{\eta}-\eta_H) |\!|\!|_{k,\mathcal{T}_H} \le \mathscr{C}_{\rm app}\|T(\lambda-\lambda_H)\|_{\varepsilon,\widetilde{\Omega}}$$

and (5.2.27) follows. □

Theorem 5.2.14 (Quasi-optimality)**.** *There exists a constant $\mathscr{C}^\star$ solely depending on the extremal values of the coefficients, such that, if $\mathscr{C}_{\rm app} \le \mathscr{C}^\star$, then there exists a unique $\lambda_H \in \Lambda_H$ solution to (5.2.22), and we have*

$$\|\lambda-\lambda_H\|_{\Lambda,k} \lesssim \min_{\mu_H\in\Lambda_H} \|\lambda-\mu_H\|_{\Lambda,k}, \tag{5.2.28}$$

where the hidden constant does not depend on $\mathscr{C}_{\rm app}$.

Proof. By the definitions of $b_{\mathcal{T}_H}(\cdot,\cdot)$ and $\langle\cdot,\cdot\rangle_{\partial\mathcal{T}_H}$, we have that

$$\begin{aligned}
&-\operatorname{Re}\langle(\lambda-\lambda_H),T(\lambda-\lambda_H)\rangle_{\partial\mathcal{T}_H}\\
&=\operatorname{Re} b_{\mathcal{T}_H}\left(T(\lambda-\lambda_H),T(\lambda-\lambda_H)\right)\\
&=\gamma_{\mathrm{r}}\left(\|\partial_1(T(\lambda-\lambda_H))\|^2_{\mu_1,\widetilde{\Omega}}+|\nu|^{-2}\|\partial_2(T(\lambda-\lambda_H))\|^2_{\mu_2,\widetilde{\Omega}}-k^2\|T(\lambda-\lambda_H)\|^2_{\varepsilon,\widetilde{\Omega}}\right)\\
&\geq\gamma_{\mathrm{r}}\left(-k^2\|T(\lambda-\lambda_H)\|^2_{\varepsilon,\widetilde{\Omega}}+|\nu|^{-2}|T(\lambda-\lambda_H)|^2_{\boldsymbol{A},\widetilde{\Omega}}\right)\\
&\geq\gamma_{\mathrm{r}}\left(|\nu|^{-2}\,|\!|\!|T(\lambda-\lambda_H)|\!|\!|^2_{k,\mathcal{T}_H}-2k^2\|T(\lambda-\lambda_H)\|^2_{\varepsilon,\widetilde{\Omega}}\right).
\end{aligned}$$

Recalling Lemma 5.2.13, there exist a constant B_3 such that

$$2k^2\|T(\lambda-\lambda_H)\|^2_{\varepsilon,\widetilde{\Omega}}\lesssim 2B_3\mathscr{C}^2_{\mathrm{app}}\,|\!|\!|T(\lambda-\lambda_H)|\!|\!|^2_{k,\mathcal{T}_H},$$

it follows that

$$-\operatorname{Re} b_{\mathcal{T}_H}\left((\lambda-\lambda_H),T(\lambda-\lambda_H)\right)\gtrsim\left(|\nu|^{-2}-2B_3\mathscr{C}^2_{\mathrm{app}}\right)|\!|\!|T(\lambda-\lambda_H)|\!|\!|^2_{k,\mathcal{T}_H}.$$

As a result, selecting $\mathscr{C}^\star$ such that

$$2B_3(\mathscr{C}^\star)^2=\frac{1}{2}|\nu|^{-2},$$

we obtain that

$$-\operatorname{Re} b_{\mathcal{T}_H}\left((\lambda-\lambda_H),T(\lambda-\lambda_H)\right)\gtrsim|\!|\!|T(\lambda-\lambda_H)|\!|\!|^2_{k,\mathcal{T}_H}.$$

We can now end the proof of error estimate (5.2.28) using Galerkin's orthogonality. Indeed, using Galerkin's orthogonality, it holds that

$$\begin{aligned}
|\!|\!|T(\lambda-\lambda_H)|\!|\!|^2_{k,\mathcal{T}_H}&\lesssim|\langle(\lambda-\lambda_H),T(\lambda-\lambda_H)\rangle_{\partial\mathcal{T}_H}|\\
&\lesssim|\langle(\lambda-\mu_H),T(\lambda-\lambda_H)\rangle_{\partial\mathcal{T}_H}|\\
&\lesssim\|\lambda-\mu_H\|_{\Lambda,k}\,|\!|\!|T(\lambda-\lambda_H)|\!|\!|_{k,\mathcal{T}_H},
\end{aligned}$$

for all $\mu_H\in\Lambda_H$, and (5.2.28) follows from norm equivalence (5.2.14). □

5.2.3.2 Convergence rates

We now use the abstract stability analysis to provide convergence rates for the method. Specifically, we will provide qualitative estimate on the dependence of $\mathscr{C}_{\mathrm{app}}$ on the frequency k and the mesh size H. To do so, following [34, 121], we define an "interpolation" operator $\pi_H:\Lambda\cap L^2(\partial\mathcal{T}_H)\to\Lambda_H$ by requiring that

$$(\pi_H\mu,q)_{\partial K}=(\mu,q)_{\partial K}\quad\forall q\in\mathcal{P}_0(\mathcal{F}_K),$$

for all $K\in\mathcal{T}_H$. Error estimate for the interpolation operator π_H, specifically, the proof of the estimates (5.2.29) and (5.2.30) below can respectively be found in [121, Lemma 9] and [34]

Lemma 5.2.15 (Interpolation error estimate)**.** *Let $\mu \in \Lambda$ and assume that there exists $u \in H^2(\mathcal{T}_H)$ such that $\mu|_K = \boldsymbol{DA\nabla} u \cdot \boldsymbol{n}|_K$ for all $K \in \mathcal{T}_H$. Then, the error estimate*

$$\|\mu - \pi_H \mu\|_{\Lambda,k} \lesssim H |u|_{H^2(\mathcal{T}_H)}, \tag{5.2.29}$$

holds true.

Besides, if $\mu \in \Lambda \cap H^1(\mathcal{F}_H)$, then we have

$$\|\mu - \pi_H \mu\|_{\Lambda,k} \lesssim H^{3/2} |\mu|_{H^1(\mathcal{F}_H)}. \tag{5.2.30}$$

Equipped with Lemma 5.2.15, we can provide a first coarse convergence result.

Theorem 5.2.16 (Basic convergence result)**.** *Assume that for all $f \in L^2(\widetilde{\Omega})$, the solution $u \in H^1_\sharp(\widetilde{\Omega})$ to (5.1.2) belongs to $H^2(\mathcal{T}_H)$ with*

$$\left\| \frac{\partial u}{\partial \boldsymbol{x}_1} \right\|_{k,\mathcal{T}_H} + \left\| \frac{\partial u}{\partial \boldsymbol{x}_2} \right\|_{k,\mathcal{T}_H} \lesssim \mathscr{C}_{\mathrm{st}} \|f\|_{\varepsilon,\widetilde{\Omega}}. \tag{5.2.31}$$

Then, we have

$$\mathscr{C}_{\mathrm{app}} \lesssim \mathscr{C}_{\mathrm{st}} kH.$$

In particular, if $\mathscr{C}_{\mathrm{st}} kH$ is small enough, there is a unique solution $\lambda_H \in \Lambda_H$ to the discrete MHM formulation (5.2.22), and we have

$$k \, |\!|\!| u - u_H |\!|\!|_{k,\mathcal{T}_H} \lesssim \mathscr{C}_{\mathrm{st}} kH \|f\|_{\varepsilon,\widetilde{\Omega}}.$$

Proof. Let $u \in H^1_\sharp(\widetilde{\Omega})$ be the solution to (5.1.2) and λ its associated one-level MHM solution. Then, the definition of $\mathscr{C}_{\mathrm{app}}$ given in (5.2.24) implies that

$$\begin{aligned}
\mathscr{C}_{\mathrm{app}} &= k \sup_{\substack{f \in L^2(\widetilde{\Omega}) \\ \|f\|_{\varepsilon,\widetilde{\Omega}} = 1}} \inf_{\mu_H \in \Lambda_H} \|\lambda - \mu_H\|_{\Lambda,k} \\
&\leq k \|\lambda - \pi_H \lambda\|_{\Lambda,k} \\
&\lesssim kH |u|_{H^2(\mathcal{T}_H)} \\
&\lesssim kH \left(\left\| \frac{\partial u}{\partial \boldsymbol{x}_1} \right\|_{k,\mathcal{T}_H} + \left\| \frac{\partial u}{\partial \boldsymbol{x}_2} \right\|_{k,\mathcal{T}_H} \right) \\
&\lesssim \mathscr{C}_{\mathrm{st}} kH,
\end{aligned}$$

where we have used (5.2.29) and (5.2.31).

On the other hand, since $\mathscr{C}_{\mathrm{app}} \lesssim \mathscr{C}_{\mathrm{st}} kH$, we may assume that $\mathscr{C}_{\mathrm{st}} kH$ is small enough that $\mathscr{C}_{\mathrm{app}} \leq |\nu|^{-2}/2$. Then, Theorems 5.2.12 and 5.2.14 imply that there exists a unique $\lambda_H \in \Lambda_H$ solution to (5.2.22), and we have

$$\begin{aligned}
k \, |\!|\!| u - u_H |\!|\!|_{k,\mathcal{T}_H} = k \, |\!|\!| T\lambda - T\lambda_H |\!|\!|_{k,\mathcal{T}_H} \lesssim k \|\lambda - \lambda_H\|_{\Lambda,k} &\lesssim k \inf_{\mu_H \in \Lambda_H} \|\lambda - \mu_H\|_{\Lambda,k} \\
&\lesssim k \|\lambda - \pi_H \lambda\|_{\Lambda,k} \lesssim \mathscr{C}_{\mathrm{st}} kH \|f\|_{\varepsilon,\widetilde{\Omega}},
\end{aligned}$$

where we use Theorem 5.2.14 and the definition of $\mathscr{C}_{\mathrm{app}}$ in (5.2.24). □

While Theorem 5.2.16 ensures the convergence of the MHM method in a rather general setting, it is not fully satisfactory. Specifically, in the case of homogeneous media, the constant $\mathscr{C}_{\text{st}}$ may be large at quasi-resonances, whereas we expect the method to be robust in this case. Besides, in the case of finely textured layered, the hidden constant would blow up as $1/\delta$ when the characteristic length δ of the texturation goes to zero. It is the goal of sections 5.3 and 5.4 to derive sharper error estimates in these two cases.

5.2.4 Implementation details

We conclude this section with some comments on the computer implementation of the MHM method.

5.2.4.1 One- and two-level MHM method

The MHM method is based on the discretization of the coupled global-local problems (5.2.4)-(5.2.5). In practice, this discretization can actually be performed in two steps.

Indeed, although problem (5.2.22) correspond to a finite-dimensional linear system, it still employs the operators T and $\widehat{T}$ in its definition. It is important to realize that these operators correspond to the solve of PDE problems, and therefore, they are not explicitly available in general. This leads to the two-level MHM method, where the operators T and $\widehat{T}$ are replaced with corresponding Galerkin approximations. Interestingly, because the PDE problems defining T and $\widehat{T}$ are independent and element-wise, a fine mesh can usually be employed, leading to a multiscale procedure.

Generally, the MHM algorithm is rather flexible and can accept a large variety of numerical discretization schemes to approximate the local problems. For instance, mixed finite element [51] and discontinuous Galerkin [95] discretizations have been employed in the past. However, the simplest setting is to consider a second-level Galerkin discretization, whereby we introduce a discretization space $V_h \subset V$, leading to the following definitions of $\hat{T}_h, T_h \in V_h$,

$$b_{\mathcal{T}_H}(\widehat{T}_h f, v_h) = (\varepsilon f, v_h)_{\widetilde{\Omega}}, \qquad b_{\mathcal{T}_H}(T_h \lambda, v_h) = -\langle \lambda, v_h \rangle_{\partial \mathcal{T}_H}, \qquad \forall v_h \in V_h.$$

for all $f \in L^2(\widetilde{\Omega})$ and $\lambda \in \Lambda$. Then, the two-level MHM solution reads as follows: Find $\lambda_{H,h} \in \Lambda_H$ such that

$$\langle \mu_H, T_h \lambda_{H,h} \rangle_{\partial \mathcal{T}_H} = \left\langle \mu_H, \widehat{T}_h f \right\rangle_{\partial \mathcal{T}_H},$$

and set

$$u_{H,h} = \widehat{T}_h f + T_h \lambda_{H,h}.$$

In this work, we will only analyze the one-level MHM discretization. In practice, it means that the second-level mesh defining the space V_h has to be sufficiently refined for our result to apply.

5.2.4.2 Practical MHM algorithm

Here, we present a brief overview of the MHM algorithm. To this end, we start by considering a mesh $\mathcal{T}_H$ and a finite-dimensional discretization space $\Lambda_H \in \Lambda$. It is important to realize that the MHM discretization allows to completely decouple the global problem (5.2.22) "first level MHM" from the local problems (5.2.4) "second level MHM ". If $(\mu_j)_{j=1}^n$ is a basis of Λ_H, the MHM algorithm can be described as follows:

- The "second level MHM " method is a pre-processing step before solving the global MHM problem and it is responsible for computing the MHM multiscale basis functions $(\psi_j)_{j=1}^n$ and ψ_f images of $(\mu_j)_{j=1}^n$ and f by the operators T_h and $\widehat{T}_h$, respectively. To clarify, $\psi_j = T_h\mu_j$ and $\psi_f = \widehat{T}_h f$ are solution to

$$b_{\mathcal{T}_H}(\psi_f, v_h) = (\varepsilon f, v_h)_{\widetilde{\Omega}}, \qquad \text{and} \qquad b_{\mathcal{T}_H}(\psi_j, v_h) = -\left\langle \mu_j, v_h \right\rangle_{\partial\mathcal{T}_H}, \qquad \forall v_h \in V_h.$$

  These computations correspond to a collection of local problems that can be solved in parallel.

- Once the multiscale basis functions $(\psi_j)_{j=1}^n$ and ψ_f are available from the local problems, we can build the first level MHM method designed to solving the global skeleton problem (5.2.22). We first note that the MHM approximation λ_H of λ is given by $\lambda_H = \sum_{j=1}^n c_j\mu_j$, where $c_j \in \mathbb{C}$. Then, using the linearity of operator T_h we have

$$T_h\lambda_H = \sum_{j=1}^{n} c_j T_h\mu_j = \sum_{j=1}^{n} c_j\psi_j.$$

  Henceforth, the global formulation (5.2.22) allows to compute the degrees of freedom c_j by solving the following $n \times n$ linear system

$$\sum_{j=1}^{n} \left\langle \mu_p, \psi_j \right\rangle_{\partial\mathcal{T}_H} \bar{c}_j = -\left\langle \mu_p, \psi_f \right\rangle_{\partial\mathcal{T}_H}, \qquad \forall p \in \{1, \ldots, n\}.$$

- Then, the MHM approximate solution follows from (5.2.1) as

$$u_{H,h} = \sum_{j=1}^{n} c_j\psi_j + \psi_f.$$

5.2.4.3 Quasi-periodic meshes

Quasi-periodic boundary conditions are used to simulate periodic geometries and allow the study to be restricted to a single periodic cell. Then, in order to simulate the connection between the studied cell and its neighboring cells, an identical discretization of the periodic boundaries $\Gamma_{\sharp-}$ and $\Gamma_{\sharp+}$ is necessary to obtain a conforming finite element mesh. To do this, each edge of the periodic boundary $\Gamma_{\sharp-}$ must match its opposite face element in the other

periodic boundary $\Gamma_{\sharp+}$. Therefore, each mesh element has two neighbors in the periodic direction, and the mesh of a periodic cell will therefore be similar to the mesh of a vertical cylinder. See Figure 5.2 for an illustration.

After designing the periodic mesh, the desired MHM discretization space Λ_H can be easily constructed. Thereby, quasi-periodic Lagrange multipliers for the MHM formulation can be considered as follows: for every two elements K^- and K^+ connected on the periodic edge $\Gamma_\sharp$ (which have a shared face belonging to $\Gamma_\sharp$), if the first element K^- (on the side $\Gamma_{\sharp-}$) receives the basis function $\mu_j = \mu_H$ as a data to solve the local problem, then the opposite element K^+ (on the side $\Gamma_{\sharp-}$) will receive the basis function $\mu_j = \mu_H e^{i\alpha\ell_1}$.

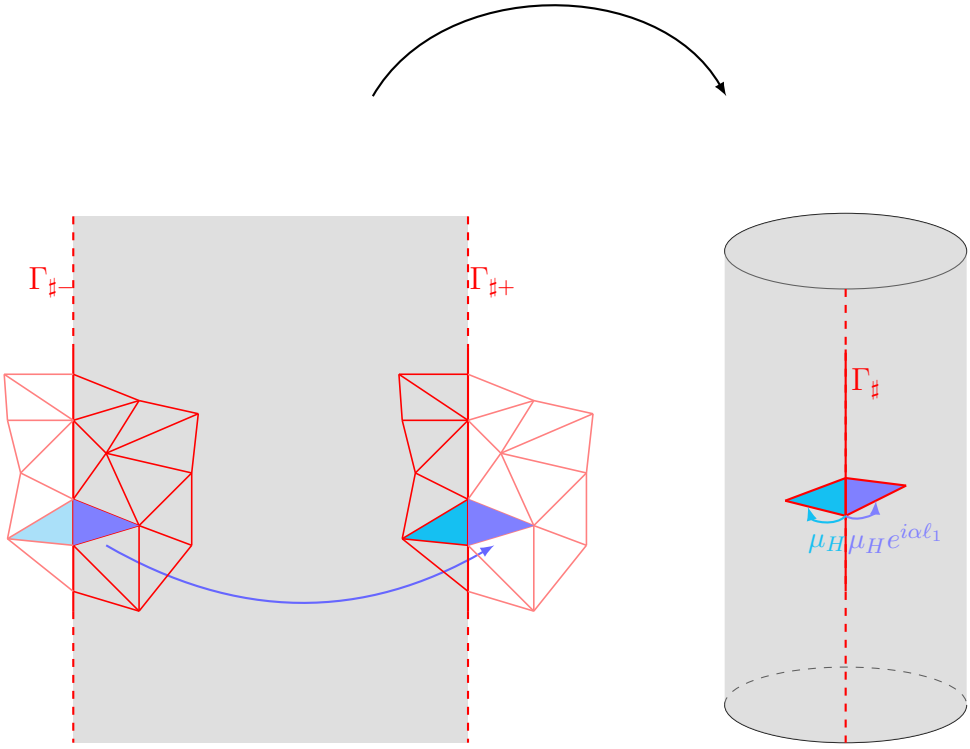


Figure 5.2: Periodic mesh

5.3 Convergence in homogeneous media

As shown in the first chapter, the presence of quasi-resonant modes affects the numerical methods performances. This effect has been illustrated in the numerical examples for the finite element method in Section 2.2.4. Besides, it is due to the relative lack of stability in the presence of these quasi-resonant frequencies. In this section, we present a one-level convergence analysis of the MHM method showing its robustness in facing these quasi-resonances. We will use a one-level Cartesian mesh to achieve our goal.

We restrict our analysis to the homogeneous case. However, we believe that the robust-

ness of MHM remains valid for a larger class of layered media. This restriction is due in the first place to the optimal stability constant that can be established for the homogeneous case and which is not the case for the multilayer media.

5.3.1 Exact representation of planewaves

The field of application, which is that of photovoltaic cells, and the rectangular medium model considered in this work allow us to use Cartesian meshes. Furthermore, on a Cartesian mesh, we have two sets of faces $\mathcal{F}_H$; the set of vertical faces $\mathcal{F}_H^{\mathrm{v}}$ and the set of horizontal faces $\mathcal{F}_H^{\mathrm{h}}$. Specifically, one easily sees that

$$\boldsymbol{n}_F = \pm(0,1) \quad \forall F \in \mathcal{F}_H^{\mathrm{h}},$$

and

$$\boldsymbol{n}_F = \pm(1,0) \quad \forall F \in \mathcal{F}_H^{\mathrm{v}}.$$

On the other hand, for a plane wave function u traveling in the mesh directions, we have

$$u = e^{ik\boldsymbol{x}_1} \quad \text{for a plane wave traveling in } x_1 \text{ direction},$$

and

$$u = e^{ik\boldsymbol{x}_2} \quad \text{for a plane wave traveling in } x_2 \text{ direction}.$$

As a result , we actually see that for all $F \in \mathcal{F}_H$

$$\lambda|_F = \boldsymbol{D}\boldsymbol{\nabla} u \cdot \boldsymbol{n}_F|_F \in \mathcal{P}_0(F).$$

Consequently, using a Cartesian mesh and the first-order polynomial approximation space given by (5.2.21), the one-level MHM scheme is expected to produce exact solution for plane waves traveling in the mesh directions.

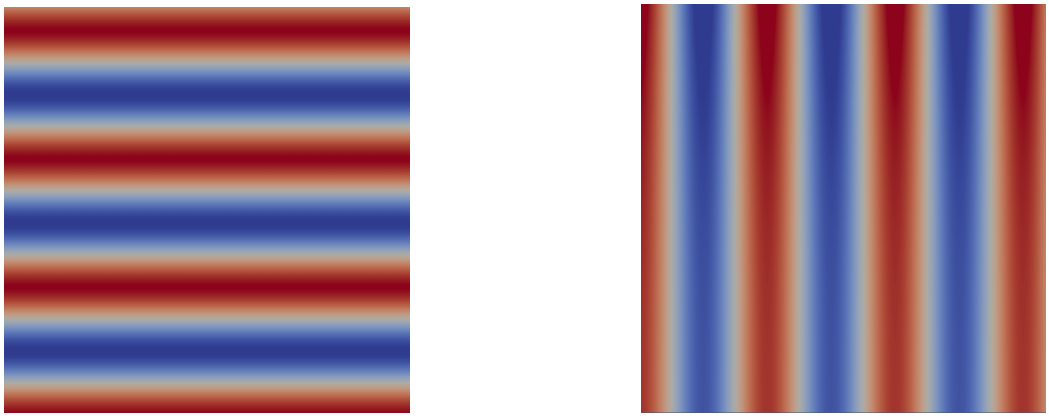


Figure 5.3: Plane waves traveling in the mesh directions; x_1-direction (right) and x_2-direction (left).

Recall from Section 2.1.7, that quasi-resonant modes are Fourier modes $\widehat{u}_n$ corresponding to a value of n for which $k_n = 0$. Hence, $\widehat{u}_n(\boldsymbol{x}_2)$ is generally constant and the associated solution $u = \widehat{u}_n e^{i(\alpha+\alpha_n)\boldsymbol{x}_1}$ is a plane wave traveling in $\boldsymbol{x}_1$ direction. Coupled with the result concerning the MHM scheme exactness for plane waves, elaborated just before, the MHM method is expected to perform well in the presence of quasi-resonant frequencies.

5.3.2 The solution splitting

As mentioned above, the stability constant and its frequency dependence are key ingredients for the convergence analysis of a numerical method. Moreover, the relative lack of stability for our model problem is due to the quasi-resonant modes and it is represented by the loss of a frequency half power in the stability constant. For this reason, our analysis will rely on the Fourier expansion (2.1.8) to avoid this difficulty. Using the linearity of the Helmholtz operator, we will split our quasi-periodic solution into two Fourier parts. On the one hand, one part consists of Fourier modes far from quasi-resonances and may therefore have the best frequency dependence for its stability constant. For this part, the usual convergence analysis gives the desired optimal error estimate. On the other hand, a second part consists of Fourier modes close to quasi-resonances. Therefore, its stability constant loses in its frequency-order dependence, and the usual convergence analysis using the stability constant must be ignored. Henceforth, some particularities of the considered MHM scheme will be used to show its expected performance in the case of quasi-resonances.

5.3.2.1 Splitting

Actually, using the Fourier expansion (2.1.8), we split the solution as $u = \widetilde{u} + \breve{u}$, where

$$\breve{u} := \sum_{2|k_n|^2<k^2} \widehat{u}_n e^{i(\alpha+\alpha_n)x_1} \quad \text{and} \quad \widetilde{u} := \sum_{2|k_n|^2\geq k^2} \widehat{u}_n e^{i(\alpha+\alpha_n)x_1}. \tag{5.3.1}$$

The discussion in Section 2.1.6 imply that $\breve{u}$ and $\widetilde{u}$ respectively solve of the Helmholtz problem (5.1.1) with the right hand sides

$$\breve{f} = \sum_{2|k_n|^2<k^2} \widehat{f}_n e^{i(\alpha+\alpha_n)x_1} \quad \text{and} \quad \widetilde{f} = \sum_{2|k_n|^2\geq k^2} \widehat{f}_n e^{i(\alpha+\alpha_n)x_1}.$$

Before going further, applying the same steps of the MHM formulation to the two functions $\widetilde{u}$ and $\breve{u}$, we will note $\widetilde{\lambda}$ and $\breve{\lambda}$ the solutions associated with their global MHM problems

$$\begin{cases} \text{Find } \widetilde{\lambda} \in \Lambda \text{ such that:} \\ \langle \mu, T\widetilde{\lambda} \rangle_{\partial\mathcal{T}_H} = -\langle \mu, \widehat{T}\widetilde{f} \rangle_{\partial\mathcal{T}_H} \quad \forall \mu \in \Lambda. \end{cases} \tag{5.3.2}$$

and

$$\begin{cases} \text{Find } \breve{\lambda} \in \Lambda \text{ such that:} \\ \langle \mu, T\breve{\lambda} \rangle_{\partial\mathcal{T}_H} = -\langle \mu, \widehat{T}\breve{f} \rangle_{\partial\mathcal{T}_H} \quad \forall \mu \in \Lambda. \end{cases} \tag{5.3.3}$$

Proposition 5.3.1. *Let $f \in L^2(\widetilde{\Omega})$ and $u \in H^1_\sharp(\widetilde{\Omega})$ be the associated solution to (5.1.1). If $\breve{f}$ and $\widetilde{f}$ are splitting of f and $\breve{u}$ and $\widetilde{u}$ are the splitting of u, the estimates*

$$k\|\widetilde{u}\|_{\varepsilon,\widetilde{\Omega}} \lesssim (1+k\ell_2)\frac{1}{k}\|\widetilde{f}\|_{\varepsilon,\widetilde{\Omega}} \tag{5.3.4}$$

and

$$k\|\breve{u}\|_{\varepsilon,\widetilde{\Omega}} \lesssim (1+(k\ell_2)^2)\frac{1}{k}\|\breve{f}\|_{\varepsilon,\widetilde{\Omega}} \tag{5.3.5}$$

hold true.

Proof. Let $u \in H^1_\sharp(\widetilde{\Omega})$ be the solution to (5.1.1) associated to the right hand side $f \in L^2(\widetilde{\Omega})$. If $\breve{f}$ and $\widetilde{f}$ are splitting of f and $\breve{u}$ and $\widetilde{u}$ are the splitting of u. Then, it is clear that $\breve{u}$ and $\widetilde{u}$ respectively solve the Helmholtz problem (5.1.1) with the right hand sides $\breve{f}$ and $\widetilde{f}$. In addition, we have

$$\|\breve{u}\|^2_{0,\widetilde{\Omega}} = \ell_1 \sum_{2|k_n|^2\geq k^2} \|\widehat{u}_n\|^2 \quad \text{and} \quad \|\widetilde{u}\|^2_{0,\widetilde{\Omega}} = \ell_1 \sum_{2|k_n|^2\leq k^2} \|\widehat{u}_n\|^2,$$

where for each $n \in \mathbb{Z}$, $\widehat{u}_n$ is the only element of $H^1_0(0,\ell+\ell_{\mathrm{P}})$ such that

$$-k_n^2(\nu\widehat{u}_n,\widehat{v}) + (\nu^{-1}\widehat{u}_n',\widehat{v}') = (\widehat{f_n},\widehat{v}) \quad \forall \widehat{v} \in H^1_0(0,\ell+\ell_{\mathrm{P}}).$$

Focus first on the estimate (5.3.4). On the one hand, if k_n is a real wave number ($k_n \in \mathbb{R}_+$), the one-dimensional stability result (3.3.17) shows that

$$\|\widehat{u}_n\| \lesssim (1+(k_n\ell_2)^{-1})\frac{1}{k_n}\ell_2\|\widehat{f}_n\|.$$

On the other hand, if k_n is a imaginary wave number ($k_n \in i\mathbb{R}_+$), the one-dimensional stability estimate (3.3.6) shows that

$$\|\widehat{u}_n\| \lesssim \min\left(4,(|k_n|\ell_2)^{-2}\right)\ell_2^2\|\widehat{f}_n\| \lesssim |k_n|^{-2}\|\widehat{f}_n\| \lesssim (1+(|k_n|\ell_2)^{-1})\frac{1}{|k_n|\ell_2}\ell_2^2\|\widehat{f}_n\|.$$

And since $2|k_n|^2 \geq k^2$ we get

$$\begin{aligned}
k^2\|\widetilde{u}\|^2_{0,\widetilde{\Omega}} &= k^2\ell_1 \sum_{2|k_n|^2\geq k^2} \|\widehat{u}_n\|^2 \\
&\lesssim k^2\ell_1 \sum_{2|k_n|^2\geq k^2} (1+(|k_n|\ell_2)^{-1})^2 \frac{1}{(|k_n|\ell_2)^2}\ell_2^4\|\widehat{f}_n\|^2 \\
&\lesssim (1+(k\ell_2)^{-1})^2(k\ell_2)^2\frac{1}{k^2}\ell_1 \sum_{2|k_n|^2\geq k^2} \|\widehat{f}_n\|^2 \\
&\lesssim (1+k\ell_2)^2\frac{1}{k^2}\|\widetilde{f}\|^2_{\varepsilon,\widetilde{\Omega}},
\end{aligned}$$

and (5.3.4) follows. Finally, to establish the estimate (5.3.5), we use the fact that $2|k_n|^2 \leq k^2$ and combine the one-dimensional stability results (3.3.11) and (3.3.22) to obtain

$$\|\widehat{u}_n\| \lesssim \ell_2^2 \|\widehat{f}_n\|,$$

therefore

$$\begin{aligned} k^2\|\breve{u}\|_{0,\widetilde{\Omega}}^2 &= k^2\ell_1 \sum_{2|k_n|^2\geq k^2} \|\widehat{u}_n\|^2 \\ &\lesssim k^2\ell_1 \sum_{2|k_n|^2\geq k^2} \ell_2^4 \|\widehat{f}_n\|^2 \\ &\lesssim (k\ell_2)^4 \frac{1}{k^2}\ell_1 \sum_{2|k_n|^2\geq k^2} \|\widehat{f}_n\|^2 \\ &\lesssim (k\ell_2)^4 \frac{1}{k^2}\|\breve{f}\|_{\varepsilon,\widetilde{\Omega}}^2. \end{aligned}$$

□

In the following, we study the convergence of the two functions $\breve{u}$ and $\widetilde{u}$ separately.

5.3.2.2 High-frequency component

Since the high-frequency component $\widetilde{u}$ has a favorable estimate in (5.3.4), we can follow the standard convergence proof.

Lemma 5.3.2. *Consider $f \in L^2(\widetilde{\Omega})$, and let $\widetilde{\lambda} \in \Lambda$ be the solution to (5.3.3). Then, we have*

$$k\|\widetilde{\lambda} - \pi_H\widetilde{\lambda}\|_{\Lambda,k} \lesssim (1+k\ell)kH\|\widetilde{f}\|_{\varepsilon,\widetilde{\Omega}}. \tag{5.3.6}$$

Proof. Using the estimate (5.2.29) from Lemma 5.2.15, we have

$$\|\widetilde{\lambda} - \pi_H\widetilde{\lambda}\|_{\Lambda,k} \lesssim H|\widetilde{u}|_{2,\widetilde{\Omega}}.$$

On the other hand, recalling (5.1.1), we have

$$\nu^2\frac{\partial^2\tilde{u}}{\partial x_1^2} + \frac{\partial^2\tilde{u}}{\partial x_2^2} = \nu\tilde{f} + k^2\nu^2\tilde{u},$$

and standard elliptic regularity (see, e.g. [20, Section 9.6]) shows that

$$|\widetilde{u}|_{2,\widetilde{\Omega}} = \lesssim \|\widetilde{f}\|_{\varepsilon,\widetilde{\Omega}} + k^2\|\widetilde{u}\|_{\varepsilon,\widetilde{\Omega}} \lesssim (1+k\ell_2)\|\widetilde{f}\|_{\varepsilon,\widetilde{\Omega}},$$

thanks to (5.3.4). As a result

$$\|\widetilde{\lambda} - \pi_H\widetilde{\lambda}\|_{\Lambda,k} \lesssim H(1+k\ell_2)\|\widetilde{f}\|_{\varepsilon,\widetilde{\Omega}}.$$

□

5.3.2.3 Low-frequency component

We now focus on the convergence analysis of the low-frequency solution part $\breve{u}$, whose Fourier modes are close to quasi-resonant modes ($k_n \approx 0$). As shown in (5.3.5), the naive stability estimate for those modes has the worst frequency dependence, making it a more subtle case to study. To avoid the unfavorable stability constant, our analysis combines the improved stability estimates of Theorem 3.3.12, the alternative interpolation estimate (5.2.30) and the fact that we are using a Cartesian mesh. This actually reflects the point discussed in subsection 5.3.1, that the MHM method exactly reproduced plane waves traveling in the mesh direction.

Lemma 5.3.3. *Consider $f \in L^2(\widetilde{\Omega})$ and let $\breve{\lambda} \in \Lambda$ be the solution to (5.3.2), then*

$$k\|\breve{\lambda} - \pi_H \breve{\lambda}\|_{\Lambda,k} \lesssim (1 + k\ell)kH\|\breve{f}\|_{\varepsilon,\widetilde{\Omega}}. \tag{5.3.7}$$

Proof. We have

$$\sum_{F\in\mathcal{F}_H} |\breve{\lambda}|^2_{H^1(F)} = \sum_{F\in\mathcal{F}_H} \left\| \frac{\partial^2 \breve{u}}{\partial x_1 \partial x_2} \right\|^2_F = \sum_{j=0}^{N_1} \left\| \frac{\partial^2 \breve{u}}{\partial x_1 \partial x_2} \right\|_{\Gamma_j^{\mathrm{v}}} + \sum_{j=0}^{N_2} \left\| \frac{\partial^2 \breve{u}}{\partial x_1 \partial x_2} \right\|_{\Gamma_j^{\mathrm{h}}},$$

where N_1 and N_2 are respectively the number of the vertical and the horizontal lines of the Cartesian mesh.

On the one hand, fixing $j \in \{0, \ldots, N_2\}$, we have

$$\begin{aligned}\left\| \frac{\partial^2 u}{\partial x_1 \partial x_2} \right\|^2_{\Gamma_j^{\mathrm{h}}} &= \sum_{\substack{n\in\mathbb{Z} \\ 2|k_n|^2<k^2}} \left\| (\alpha+\alpha_n)\widehat{u}_n'(x_2) e^{i(\alpha+\alpha_n)x_1} \right\|^2_{\Gamma_j^{\mathrm{h}}} \\ &\leq \ell_1 \sum_{\substack{n\in\mathbb{Z} \\ 2|k_n|^2<k^2}} |\alpha+\alpha_n|^2 \|\widehat{u}_n'\|^2_{L^\infty(0,\ell_2+\ell_{\mathrm{P}})} \lesssim k^2\ell_1 \sum_{\substack{n\in\mathbb{Z} \\ 2|k_n|^2<k^2}} \|\widehat{u}_n'\|^2_{L^\infty(0,\ell_2+\ell_{\mathrm{P}})},\end{aligned}$$

since $|\alpha+\alpha_n|^2 \lesssim k^2$ for the considered set of indices n. We now recall from (3.3.37) in Chapter 3 that

$$\|\widehat{u}_n'\|_{L^\infty(0,\ell_2+\ell_{\mathrm{P}})} \lesssim \sqrt{\ell_2}\|\widehat{f}_n\|_{L^2(0,\ell_2+\ell_{\mathrm{P}})},$$

leading to

$$\left\| \frac{\partial^2 u}{\partial x_1 \partial x_2} \right\|^2_{\Gamma_j^{\mathrm{h}}} \leq k^2\ell_1\ell_2 \sum_{\substack{n\in\mathbb{Z} \\ 2|k_n|^2<k^2}} \|\widehat{f}_n\|^2_{L^2(0,\ell_2+\ell_{\mathrm{P}})} \leq k^2\ell_2\|\breve{f}\|^2_{\varepsilon,\widetilde{\Omega}}$$

since $N_2 \sim \ell/H$, we arrive at

$$\sum_{j=0}^{N_2} \left\| \frac{\partial^2 \breve{u}}{\partial x_1 \partial x_2} \right\|^2_{\Gamma_j^{\mathrm{h}}} \lesssim k^2\ell^2 H^{-1}\|\breve{f}\|^2_{\varepsilon,\widetilde{\Omega}}$$

and

$$k^2H^3\sum_{\ell=0}^{N_2}\left\|\frac{\partial^2\breve{u}}{\partial x_1\partial x_2}\right\|^2_{\Gamma^{\rm h}_\ell}\lesssim k^4\ell^2H^2\|\breve{f}\|^2_{\varepsilon,\widetilde{\Omega}}=(k\ell)^2(kH)^2\|\breve{f}\|^2_{\varepsilon,\widetilde{\Omega}}.$$

On the other hand, if we fix $j\in\{0,\ldots,N_1\}$, we have

$$\begin{aligned}\left\|\frac{\partial^2\breve{u}}{\partial x_1\partial x_2}\right\|^2_{\Gamma^{\rm v}_j}&=\sum_{\substack{n\in\mathbb{Z}\\ 2|k_n|^2<k^2}}\left\|(\alpha+\alpha_n)\widehat{u}_n'(x_2)e^{i(\alpha+\alpha_n)x_1}\right\|^2_{\Gamma^{\rm v}_j}\\&=\ell_1\sum_{\substack{n\in\mathbb{Z}\\ 2|k_n|^2<k^2}}|\alpha+\alpha_n|^2\|\widehat{u}_n'\|^2_{L^2(0,\ell_2+\ell_{\rm P})}\\&\lesssim(k\ell)^2\sum_{\substack{n\in\mathbb{Z}\\ 2|k_n|^2<k^2}}\|\widehat{f}_n\|^2_{L^2(0,\ell_2+\ell_{\rm P})}\\&=k^2\ell\|\breve{f}\|^2_{\varepsilon,\widetilde{\Omega}},\end{aligned}$$

where we have used the inequality

$$\|\widehat{u}_n'\|_{L^2(0,\ell_2+\ell_{\rm P})}\lesssim\ell_2\|\widehat{f}_n\|_{L^2(0,\ell_2+\ell_{\rm P})},$$

derived from (3.3.37) in Chapter 3. Now, since $N_1\sim\ell_1/H$, we arrive at

$$k^2H^3\sum_{j=0}^{N_1}\left\|\frac{\partial^2\breve{u}}{\partial x_1\partial x_2}\right\|^2_{\Gamma^{\rm v}_j}\lesssim(k\ell)^2(kH)^2\|\breve{f}\|^2_{\varepsilon,\widetilde{\Omega}}.$$

We have thus established that

$$k^2H^3\sum_{F\in\mathcal{F}_H}|\breve{\lambda}|^2_{H^1(F)}\lesssim(k\ell)^2(kH)^2\|\breve{f}\|^2_{\varepsilon,\widetilde{\Omega}}.$$

and (5.3.7) follows from the interpolation error estimates in (5.2.30). □

5.3.3 Stability and convergence

We can now present the key result of this section by combining the estimates we obtained for each part of the splitting.

Theorem 5.3.4 (Control of the approximation factor)**.** *We have*

$$\mathscr{C}_{\rm app}\lesssim(1+k\ell)kH.$$

Proof. Let $u \in H^1_\sharp(\widetilde{\Omega})$ be the solution to (5.1.2) and λ its associated one-level MHM solution. Then, the definition of $\mathscr{C}_{\text{app}}$ given in (5.2.24) implies that

$$\begin{aligned}
\mathscr{C}_{\text{app}} &= k \sup_{\substack{f\in L^2(\widetilde{\Omega}) \\ \|f\|_{\varepsilon,\widetilde{\Omega}}=1}} \inf_{\mu_H\in\Lambda_H} \|\lambda-\mu_H\|_{\Lambda,k} \\
&\leq k\|\lambda - (\pi_H\breve{\lambda} + \pi_H\widetilde{\lambda})\|_{\Lambda,k} \\
&\leq k\left(\|\breve{\lambda}-\pi_H\breve{\lambda}\|_{\Lambda,k} + \|\widetilde{\lambda}-\pi_H\widetilde{\lambda}\|_{\Lambda,k}\right) \\
&\lesssim (1+k\ell)kH,
\end{aligned}$$

where we have used (5.3.6) and (5.3.7) in the last inequality. □

Corollary 5.3.5 (Error estimate)**.** *Assume that $(1+k\ell)kH$ is small enough, then there exists a unique $\lambda_H \in \Lambda_H$ solution to (5.2.22), and we have*

$$|\!|\!| u-u_H |\!|\!|_{k,\mathcal{T}_H} \lesssim (1+k\ell)H\|f\|_{\varepsilon,\widetilde{\Omega}}.$$

Proof. Since $\mathscr{C}_{\text{app}} \lesssim (1+k\ell)kH$, we may assume that $(1+k\ell)kH$ is small enough that $\mathscr{C}_{\text{app}} \leq |\nu|^{-2}/2$.

Then, Theorems 5.2.12 and 5.2.14 imply that there exists a unique $\lambda_H \in \Lambda_H$ solution to (5.2.22), and we have

$$\begin{aligned}
|\!|\!| u-u_H |\!|\!|_{k,\mathcal{T}_H} &= |\!|\!| T\lambda - T\lambda_H |\!|\!|_{k,\mathcal{T}_H} && \text{by (5.2.1) and (5.2.23)} \\
&\lesssim \|\lambda-\lambda_H\|_{\Lambda,k} && \text{by norm equivalence (5.2.14)} \\
&\lesssim \inf_{\mu_H\in\Lambda_H} \|\lambda-\mu_H\|_{\Lambda,k} && \text{by Theorm 5.2.14} \\
&\lesssim \|\lambda-\pi_H\lambda\|_{\Lambda,k} && \text{by the minimum definition} \\
&\lesssim (1+k\ell)H\|f\|_{\varepsilon,\widetilde{\Omega}} && \text{by Lemmas 5.3.2 and 5.3.3 .}
\end{aligned}$$

□

The results of the previous two Lemmas 5.3.2 and 5.3.3 show that the MHM method supports the same frequency-orders of convergence $(1+k\ell)kH$, for both low- ($\breve{u}$) and high- $\widetilde{u}$ frequency component. Furthermore, the quasi-optimality of the MHM method is uniformly established if $(1+k\ell)kH$ is small enough ($\mathscr{C}_{\text{app}} \lesssim (1+k\ell)kH$). Theoretically, the current results suggest that the MHM method will not suffer from the pollution effect close to the quasi-resonances. In contrast, the error curves of the finite element method (see the numerical illustrations in Section 2.2.4) show the important effect of quasi-resonances on the finite element scheme. Moreover, the standard finite element analysis of our model Helmholtz problem shows that the finite element solution is quasi-optimal provided that $\mathscr{C}_{\text{app}} \lesssim (\mathscr{C}_{\text{st}})kH \approx (1+(k\ell)^2)kH$. This difference of one frequency-power in the quasi-optimal condition allows the MHM method to remain performant in the presence of anomalous frequencies.

5.3.4 Numerical experiments

In this subsection, we evaluate our theoretical convergence results and show the performance of the MHM method on periodic domains. In particular, its robustness to quasi-resonant frequencies will be nicely illustrated. To this end, we will compare the first-order MHM and finite element schemes through a sequence of numerical tests.

For the following numerical examples, we use Cartesian meshes $\mathcal{T}_H$ made of square elements and the first-order polynomial space Λ_H defined in (5.2.21) as a MHM approximation space.

The analytical function that we seek to approximate is solution of the following 2D Helmholtz problem posed in a homogeneous unite square $\Omega := (0,1)^2$:

$$\begin{cases} -k^2\nu\widetilde{u} - \dfrac{\partial}{\partial x_1}\left(\nu\dfrac{\partial\widetilde{u}}{\partial x_2}\right) - \dfrac{\partial}{\partial x_2}\left(\nu^{-1}\dfrac{\partial\widetilde{u}}{\partial x_2}\right) &= f \quad \text{in } \widetilde{\Omega} \\ \widetilde{u} &= 0 \quad \text{on } \Gamma_{\mathrm{P}} \\ \widetilde{u} &= 0 \quad \text{on } \Gamma_{\mathrm{D}} \\ \widetilde{u}_+ - e^{i\alpha\ell_1}\widetilde{u}_- &= 0 \quad \text{on } \widetilde{\Gamma}_\sharp, \end{cases} \tag{5.3.8}$$

where the source $f \in L^2(\widetilde{\Omega})$ (illustrated in Figure 2.7) is chosen so that

$$u(\boldsymbol{x}) = \chi(\boldsymbol{x})e^{ik\boldsymbol{d}_{in}\cdot\boldsymbol{x}} + e^{ik\boldsymbol{d}_{out}\cdot\boldsymbol{x}},$$

where $\boldsymbol{d}^{\text{in}}\cdot\boldsymbol{d}^{\text{in}} = \boldsymbol{d}^{\text{out}}\cdot\boldsymbol{d}^{\text{out}} = 1$, $d_1^{\text{in}} = d_1^{\text{in}} = \alpha + m\pi/\ell_1$ for $m \in \mathbb{N}$, $d_2^{\text{in}} < 0$ and $d_2^{\text{out}} = -d_2^{\text{in}}$, and and the cutoff function χ defined in (2.2.30).

As for the previous numerical examples, we will choose a source term f so that the solution u is composed of one Fourier mode $\widehat{u}_j$. In this case, there exists an integer $j \in \mathbb{N}$ such that the one-dimensional wave number is:

$$k_j^2 = k^2 - (k\sin(\theta) + 2\pi j)^2.$$

Thus, for a fixed integer j and frequency k, the quasi-resonance cases ($k_j = 0$) will depend on the angle of incidence θ.

In the following experiments, we plot the relative H^1 errors with respect to the angle of incidence θ. In particular, we are interested in cases close to quasi-resonances. Therefore, we expect that approaching a quasi-resonance, the MHM error remains controlled and much smaller than the FEM error.

Our goal here is to illustrate the effect of the quasi-resonant mode (characterized by its angle of incidence θ) on the finite element and MHM methods. For this purpose, we fix a frequency k a mode $j \in N$, and a mesh of size H, and plot the relative H^1 FEM and MHM errors versus the angle θ on Figures 5.4 and 5.5.

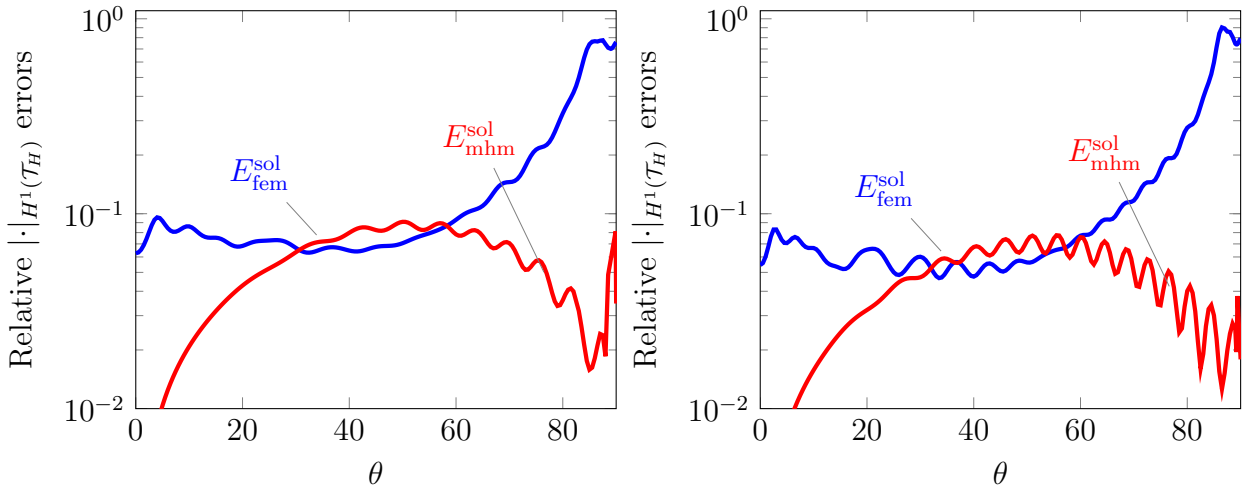


Figure 5.4: MHM and FEM errors for $j = 0$ with $k = 10\pi$ (left) and $k = 15\pi$ (right).

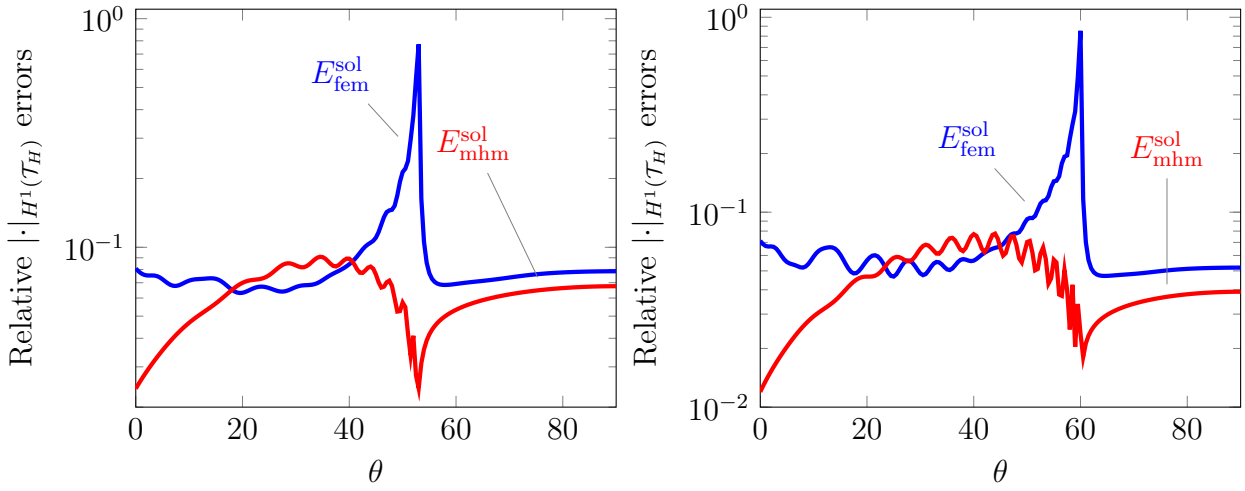


Figure 5.5: MHM and FEM errors for $j = 1$ with $k = 10\pi$ (left) and $k = 15\pi$ (right).

Figures 5.4 and 5.5 show that quasi-resonant modes have a considerable effect on finite element errors. This effect is manifested by a large increase in the error value close to quasi-resonant modes. Besides, we notice that the MHM method maintains its optimality, and its error value does not produce almost any increase near the quasi-resonant frequencies. As can be seen, there is at least a ten percent difference between the two H^1 relative MHM and FEM errors. Especially, this situation is particularly for $\theta = 90°$ and $k = 10\pi$ and $k = 15\pi$ when $j = 0$ (Figure 5.4) and for $\theta = 53.13°$, and $\theta = 60.07°$ for $k = 10\pi$ and $k = 15\pi$, respectively when $j = 1$ (Figure 5.5).

5.4 Convergence in finely textured layered media

As a multiscale method, the MHM method is designed to capture small-scale heterogeneities using coarse meshes. In this section, we will focus on the case of finely textured layers where $\delta \ll H$, and obtain error estimates that are robust in δ. Such a situation was out of the scope of the work in [36], so that those results are entirely new.

5.4.1 Settings

For the remainder of this section, we assume that the coefficients ε, A_1 and A_2 satisfy the assumptions of Chapter 4. Specifically, we assume that for a given $\delta > 0$, the coefficients are given by

$$\varepsilon(\boldsymbol{x}) = \widehat{\varepsilon}^{\delta}(\boldsymbol{x}) := \widehat{\varepsilon}\left(\boldsymbol{x}_2, \left\{\frac{\boldsymbol{x}_1}{\delta}\right\}\right), \quad A_j(\boldsymbol{x}) = \widehat{A}_j^{\delta}(\boldsymbol{x}) := \widehat{A}_j\left(\boldsymbol{x}_2, \left\{\frac{\boldsymbol{x}_1}{\delta}\right\}\right),$$

where $\widehat{\varepsilon}, \widehat{A}_j : (0, \ell_2 + \ell_{\mathrm{P}}) \times (0,1) \to \mathbb{R}$, for $j = 1$ and 2, are smooth functions periodic in the second variable.

Our model problem is to approximate the solution u_δ associated with the oscillating coefficients $\widehat{\varepsilon}^{\delta}$ and $\widehat{\boldsymbol{A}}^{\delta}$. For $\delta > 0$ and $f \in L^2(\widetilde{\Omega})$, the global MHM problem then consists in finding $\lambda_\delta \in \Lambda$ such that

$$\langle \mu, T_\delta \lambda_\delta \rangle_{\partial \mathcal{T}_H} = -\langle \mu, \widehat{T}_\delta f \rangle_{\partial \mathcal{T}_H} \qquad \forall \mu \in \Lambda, \tag{5.4.1}$$

where the local operators are defined by

$$b_\delta(\widehat{T}_\delta f, v) = (\varepsilon^{H} f, v)_{\widetilde{\Omega}} \quad b_\delta(T_\delta \lambda_\delta, v) = -\langle \lambda_\delta, v \rangle_{\partial \mathcal{T}_H} \qquad \forall v \in H^1(\mathcal{T}_H).$$

Our analysis will also rely on the homogenized problem

$$\langle \mu, T_0 \lambda_0 \rangle_{\partial \mathcal{T}_H} = -\langle \mu, \widehat{T}_0 f \rangle_{\partial \mathcal{T}_H} \qquad \forall \mu \in \Lambda, \tag{5.4.2}$$

with the local operators

$$b_{\mathrm{H}}(\widehat{T}_0 f, v) = (\varepsilon^{\mathrm{H}} f, v)_{\widetilde{\Omega}} \quad b_{\mathrm{H}}(T_0 \lambda_0, v) = -\langle \lambda_0, v \rangle_{\partial \mathcal{T}_H} \qquad \forall v \in H^1(\mathcal{T}_H).$$

Notice that the meshing condition for the well-posedness of the local problems we obtained in section 5.2.1 only involves the maximum and minimum values of the coefficients. In particular it does not depend on δ.

5.4.2 Technical results

This subsection presents a few key preliminary results we employ in the convergence proof. We start by showing that $\widehat{T}_0 f$ is always piecewise H^2.

Lemma 5.4.1 (Piecewise regularity for $\widehat{T}_0$)**.** *For all $f \in L^2(\widetilde{\Omega})$, we have $\widehat{T}f \in H^2(\mathcal{T}_H)$ with*

$$k \left|\!\left|\!\left| \widehat{T}_0 f \right|\!\right|\!\right|_{k,\mathcal{T}_H} + \left|\!\left|\!\left| \frac{\partial}{\partial \boldsymbol{x}_1}(\widehat{T}_0 f) \right|\!\right|\!\right|_{k,\mathcal{T}_H} + \left|\!\left|\!\left| \frac{\partial}{\partial \boldsymbol{x}_2}(\widehat{T}_0 f) \right|\!\right|\!\right|_{k,\mathcal{T}_H} \lesssim \|f\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}}. \tag{5.4.3}$$

Proof. The first part of the estimate, namely

$$k \left|\!\left|\!\left| \widehat{T}_0 f \right|\!\right|\!\right|_{k,\mathcal{T}_H} \lesssim \|f\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}} \tag{5.4.4}$$

has already been established at (5.2.13) in Theorem 5.2.6. For the piecewise H^2, we proceed element by element. Thus, let us consider $K \in \mathcal{T}_H$. We then observe that

$$\left\{ \begin{array}{rcll} -\boldsymbol{\nabla} \cdot \left(\boldsymbol{D} \boldsymbol{A}^{\mathrm{H}} \boldsymbol{\nabla} (\widehat{T}_0 f) \right) & = & \varepsilon^{\mathrm{H}} f + k^2 \varepsilon^{\mathrm{H}} \widehat{T} f & \text{in } K \\ \boldsymbol{\nabla}(\widehat{T}_0 f) \cdot \boldsymbol{n} & = & 0 & \text{on } \partial K. \end{array} \right.$$

Then, since K is convex, standard elliptic regularity results [72, Theorems 3.1.3.1 and 3.2.1.2] implies that $\widehat{T}_0 f \in H^2(K)$, with

$$\left|\!\left|\!\left| \frac{\partial}{\partial \boldsymbol{x}_1}(\widehat{T}_0 f) \right|\!\right|\!\right|_{k,K} + \left|\!\left|\!\left| \frac{\partial}{\partial \boldsymbol{x}_2}(\widehat{T}_0 f) \right|\!\right|\!\right|_{k,K} \lesssim \|f + k^2 u\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}}$$

Then, (5.4.3) follows from (5.4.4) since

$$\|f + k^2 u\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}} \leq \|f\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}} + k^2 \|u\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}} \lesssim \|f\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}} + k \left|\!\left|\!\left| u \right|\!\right|\!\right|_{k,K} \lesssim \|f\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}}.$$

□

We then provide an homogenization error estimate similar to the one we derived in Chapter 4, but for the local problems instead of the global Helmholtz problem. We start with the local problems corresponding with the operator $\widehat{T}_\delta$

Lemma 5.4.2 (Homogenization for the operator $\widehat{T}_\delta$)**.** *For all $\delta > 0$ and $f \in L^2(\widetilde{\Omega})$, we have*

$$k \left|\!\left|\!\left| \widehat{T}_\delta f - \widehat{T}_0 f - \delta \widehat{\chi}^\delta \frac{\partial}{\partial \boldsymbol{x}_1}(\widehat{T}_0 f) \right|\!\right|\!\right|_{k,\mathcal{T}_H} \lesssim \left(kH + k\delta + \sqrt{\frac{\delta}{H}} + \frac{\delta}{H} \right) \|f\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}} \tag{5.4.5}$$

Proof. Fix an element $K \in \mathcal{T}_H$. Let us set $u_\delta := (\widehat{T}_\delta f)|_K \in H^1(K)$ and $u_0 := (\widehat{T}_0 f)|_K \in H^1(K)$. For all $v \in H^1(K)$, we have Let $\left|\!\left|\!\left| v \right|\!\right|\!\right|_{k,D} = 1$. Since ν is constant in K, we have

$$\begin{aligned} b_\delta^K(u_\delta - u_0, v) = b_{\mathrm{H}}^K(u_0, v) - b_\delta^K(u_0, v) = & -k^2 \nu_K \left((\varepsilon^{\mathrm{H}} - \widehat{\varepsilon}^\delta) u_0, v \right)_K \\ & + \left(\nu (A_1^{\mathrm{H}} - \widehat{A}_1^\delta) \frac{\partial u_0}{\partial \boldsymbol{x}_1}, \frac{\partial v}{\partial \boldsymbol{x}_1} \right)_K \\ & + \nu_K^{-1} \left((A_2^{\mathrm{H}} - \widehat{A}_2^\delta) \frac{\partial u_0}{\partial \boldsymbol{x}_2}, \frac{\partial v}{\partial \boldsymbol{x}_2} \right)_K, \end{aligned}$$

and

$$b_\delta^K\left(u_\delta - u_0 - \delta\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1}, v\right) = -k^2\nu_K\left((\varepsilon^{\mathrm{H}} - \widehat{\varepsilon}^\delta)u_0, v\right)_K + \left(\nu(A_1^{\mathrm{H}} - \widehat{A}_1^\delta)\frac{\partial u_0}{\partial \boldsymbol{x}_1}, \frac{\partial v}{\partial \boldsymbol{x}_1}\right)_K - b_\delta^K\left(\delta\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1}, v\right) + \nu_K^{-1}\left((A_2^{\mathrm{H}} - \widehat{A}_2^\delta)\frac{\partial u_0}{\partial \boldsymbol{x}_2}, \frac{\partial v}{\partial \boldsymbol{x}_2}\right)_K,$$

and therefore

$$\operatorname{Re} b_\delta^K\left(u_\delta - u_0 - \delta\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1}, v\right) \lesssim k^2\left|\left((\varepsilon^{\mathrm{H}} - \widehat{\varepsilon}^\delta)u_0, v\right)_K\right| + \left|\left(\nu(A_1^{\mathrm{H}} - \widehat{A}_1^\delta)\frac{\partial u_0}{\partial \boldsymbol{x}_1}, \frac{\partial v}{\partial \boldsymbol{x}_1}\right)_K - b_\delta^K\left(\delta\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1}, v\right)\right| + \left|\left((A_2^{\mathrm{H}} - \widehat{A}_2^\delta)\frac{\partial u_0}{\partial \boldsymbol{x}_2}, \frac{\partial v}{\partial \boldsymbol{x}_2}\right)_K\right|.$$

Recalling Lemma 5.4.1, we actually have $u_0 \in H^2(K)$. As a result, we can apply (4.2.13), (4.2.14) and (4.2.15) from subsection 4.2.2 to show that

$$k\operatorname{Re} b_\delta^K\left(u_\delta - u_0 - \delta\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1}, v\right) \lesssim \left\{\left(k\delta + \sqrt{\frac{\delta}{H_K}} + \frac{\delta}{H_K}\right)k\,|\!|\!|u_0|\!|\!|_{k,D} + k(\delta + \sqrt{\delta H_K})\left(\left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right|\!\right|\!\right|_{k,K} + \left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_2}\right|\!\right|\!\right|_{k,K}\right)\right\}|\!|\!|v|\!|\!|_{k,K},$$

which, using that $\delta + \sqrt{\delta H_K} \lesssim \delta + H_K$, we simplify as

$$k\operatorname{Re} b_\delta^K\left(u_\delta - u_0 - \delta\widehat{\chi}^\delta\frac{\partial u_0}{\partial \boldsymbol{x}_1}, v\right) \lesssim \left(k\delta + kH_K + \sqrt{\frac{\delta}{H_K}} + \frac{\delta}{H_K}\right)\left\{k\,|\!|\!|u_0|\!|\!|_{k,D} + \left(\left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right|\!\right|\!\right|_{k,K} + \left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_2}\right|\!\right|\!\right|_{k,K}\right)\right\}|\!|\!|v|\!|\!|_{k,K}.$$

After summing up this identity over all $K \in \mathcal{T}_H$, we have

$$k\operatorname{Re} b_\delta\left(\widehat{T}_\delta f - \widehat{T}_0 f - \delta\widehat{\chi}^\delta\frac{\partial}{\partial \boldsymbol{x}_1}(\widehat{T}_0 f), v\right) \lesssim \left(k\delta + kH_K + \sqrt{\frac{\delta}{H_K}} + \frac{\delta}{H_K}\right)\left\{k\left|\!\left|\!\left|\widehat{T}_0 f\right|\!\right|\!\right|_{k,\mathcal{T}_H} + \left(\left|\!\left|\!\left|\frac{\partial}{\partial \boldsymbol{x}_1}(\widehat{T}_0 f)\right|\!\right|\!\right|_{k,\mathcal{T}_H} + \left|\!\left|\!\left|\frac{\partial}{\partial \boldsymbol{x}_2}(\widehat{T}_0 f)\right|\!\right|\!\right|_{k,\mathcal{T}_H}\right)\right\}|\!|\!|v|\!|\!|_{k,\mathcal{T}_H},$$

and we arrive at

$$k\operatorname{Re} b_\delta\left(\widehat{T}_\delta f - \widehat{T}_0 f - \delta\widehat{\chi}^\delta\frac{\partial}{\partial \boldsymbol{x}_1}(\widehat{T}_0 f), v\right) \lesssim \left(k\delta + kH_K + \sqrt{\frac{\delta}{H_K}} + \frac{\delta}{H_K}\right)\|f\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}}\,|\!|\!|v|\!|\!|_{k,\mathcal{T}_H},$$

using (5.4.3), and (5.4.5) follows from inf-sup condition (5.2.12) for the sesquilinear form $b_\delta(\cdot,\cdot)$ over $V \times V$. □

We then provide a similar homogenization error estimate for the operator T_δ.

Lemma 5.4.3 (Homogenization for the operator T_δ)**.** *Let $f \in L^2(\widetilde{\Omega})$, and consider the solution $\lambda_0 \in V$ to the homogenized MHM formulation (5.4.2). Then*

$$\left|\!\left|\!\left| T_\delta \lambda_0 - T_0 \lambda_0 - \delta \widehat{\chi}^\delta \frac{\partial}{\partial \boldsymbol{x}_1}(T_0\lambda) \right|\!\right|\!\right|_{k,\mathcal{T}_H} \lesssim \mathscr{C}_{\mathrm{st}} \left(kH + k\delta + \sqrt{\frac{\delta}{H}} + \frac{\delta}{H} \right) \|f\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}}$$

Proof. We start by looking at a single element $K \in \mathcal{T}_H$. We let $v_0 = (T_0\lambda_0)_K \in H^1(K)$, and $v_\delta = (T_\delta \lambda_0)$. We know that $u_0 := T_0\lambda_0 + \widehat{T}_0 f \in H^2(\widetilde{\Omega})$ since it is the original solution ot the problem. As a result, since we previously showed that $\widehat{T} f \in H^2(\mathcal{T}_H)$, we have $v_0 = u_0|_K - (\widehat{T}f)|_K \in H^2(K)$. In addition, using (4.2.16) and (5.4.3), we have

$$\begin{aligned} k \left|\!\left|\!\left| v_0 \right|\!\right|\!\right|_{k,K} + \left|\!\left|\!\left| \frac{\partial v_0}{\partial \boldsymbol{x}_1} \right|\!\right|\!\right|_{k,K} + \left|\!\left|\!\left| \frac{\partial v_0}{\partial \boldsymbol{x}_2} \right|\!\right|\!\right|_{k,H} &\lesssim k \left|\!\left|\!\left| u_0 \right|\!\right|\!\right|_{k,K} + \left|\!\left|\!\left| \frac{\partial u_0}{\partial \boldsymbol{x}_1} \right|\!\right|\!\right|_{k,K} + \left|\!\left|\!\left| \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right|\!\right|\!\right|_{k,H} \\ &\quad + k \left|\!\left|\!\left| \widehat{T} f \right|\!\right|\!\right|_{k,K} + \left|\!\left|\!\left| \frac{\partial}{\partial \boldsymbol{x}_1}(\widehat{T} f) \right|\!\right|\!\right|_{k,K} + \left|\!\left|\!\left| \frac{\partial}{\partial \boldsymbol{x}_2}(\widehat{T} f) \right|\!\right|\!\right|_{k,H} \\ &\lesssim k \left|\!\left|\!\left| u_0 \right|\!\right|\!\right|_{k,K} + \left|\!\left|\!\left| \frac{\partial u_0}{\partial \boldsymbol{x}_1} \right|\!\right|\!\right|_{k,K} + \left|\!\left|\!\left| \frac{\partial u_0}{\partial \boldsymbol{x}_2} \right|\!\right|\!\right|_{k,H} + \|f\|_{\varepsilon^{\mathrm{H}},K}. \end{aligned}$$

Then, the proof follows exactly the lines of the one of Lemma 5.4.2, with u_0 and u_δ replaced by v_0 and v_δ. □

5.4.3 Error estimate

We are now ready to establish our key result concerning the MHM in highly heterogeneous media.

Theorem 5.4.4. *Let $f \in L^2(\widetilde{\Omega})$, and let $\lambda_\delta \in \Lambda$ be the solution to the oscillating MHM formulation (5.4.1). Then, we have*

$$k\|\lambda_\delta - \pi_H \lambda_0\|_{\Lambda,k} \lesssim \left\{ \mathscr{C}_{\mathrm{st}} \left(kH + k\delta + \sqrt{\frac{\delta}{H}} + \frac{\delta}{H} \right) + \mathscr{C}_{\mathrm{st}}^2 \left(\sqrt{k\ell}\sqrt{k\delta} + k\delta \right) \right\} \|f\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}}. \tag{5.4.6}$$

In particular,

$$\mathscr{C}_{\mathrm{app}} \lesssim \mathscr{C}_{\mathrm{st}} \left(kH + k\delta + \sqrt{\frac{\delta}{H}} + \frac{\delta}{H} \right) + \mathscr{C}_{\mathrm{st}}^2 \left(\sqrt{k\ell}\sqrt{k\delta} + k\delta \right) \tag{5.4.7}$$

Proof. We start with the triangular inequality

$$\|\lambda_\delta - \pi_H\lambda_0\|_{\Lambda,k} \leq \|\lambda_\delta - \lambda_0\|_{\Lambda,k} + \|\lambda_0 - \pi_H\lambda_0\|_{\Lambda,k}.$$

To deal with the second term, we cans imply employed the usual MHM interpolation estimate (5.2.29) together with the $H^2(\mathcal{T}_H)$ estimate (4.2.16) for $u_0 = T_0\lambda_0 + \widehat{T}_0 f$ as follows:

$$k\|\lambda_0 - \pi_H\lambda_0\|_{\Lambda,k} \lesssim kH\left(\left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_1}\right|\!\right|\!\right|_{k,\mathcal{T}_H} + \left|\!\left|\!\left|\frac{\partial u_0}{\partial \boldsymbol{x}_2}\right|\!\right|\!\right|_{k,\mathcal{T}_H}\right) \lesssim \mathscr{C}_{\mathrm{st}} kH\|f\|_{\varepsilon^{\mathrm{H}},\widetilde{\Omega}}.$$

For the first term, we write that

$$\begin{aligned}
\|\lambda_\delta - \lambda_0\|_{\Lambda,k} &\lesssim |\!|\!|T_\delta\lambda_\delta - T_\delta\lambda_0|\!|\!|_{k,\mathcal{T}_H} \\
&\lesssim \left|\!\left|\!\left|T_\delta\lambda_\delta - T_0\lambda_0 - \delta\widehat{\chi}^\delta\frac{\partial}{\partial \boldsymbol{x}_1}(T\lambda_0)\right|\!\right|\!\right|_{k,\mathcal{T}_H} + \left|\!\left|\!\left|T_\delta\lambda_0 - T_0\lambda_0 - \delta\widehat{\chi}^\delta\frac{\partial}{\partial \boldsymbol{x}_1}(T\lambda_0)\right|\!\right|\!\right|_{k,\mathcal{T}_H} \\
&\lesssim |\!|\!|u_\delta - u_0 - \delta\widehat{\chi}^\delta u_0|\!|\!|_{k,\mathcal{T}_H} + \left|\!\left|\!\left|\widehat{T}_\delta f - \widehat{T}_0 f - \delta\widehat{\chi}^\delta\frac{\partial}{\partial \boldsymbol{x}_1}(\widehat{T}f)\right|\!\right|\!\right|_{k,\mathcal{T}_H} \\
&\quad + \left|\!\left|\!\left|T_\delta\lambda_0 - T_0\lambda_0 - \delta\widehat{\chi}^\delta\frac{\partial}{\partial \boldsymbol{x}_1}(T\lambda_0)\right|\!\right|\!\right|_{k,\mathcal{T}_H},
\end{aligned}$$

so that (5.4.6) follows from Theorem 4.2.8 and Lemmas 5.4.2 and 5.4.3. Then, (5.4.7) is a direct consequence of the definition of $\mathscr{C}_{\mathrm{app}}$ in (5.2.24). □

5.4.4 Numerical examples

The purpose of this section is to assess the theoretical convergence results of the previous section. To do this, we consider a problem with a highly-oscillatory medium coefficient depicted in Figure 5.6. The domain is a unit square with prescribed quasi-periodic boundary conditions and PML layer.

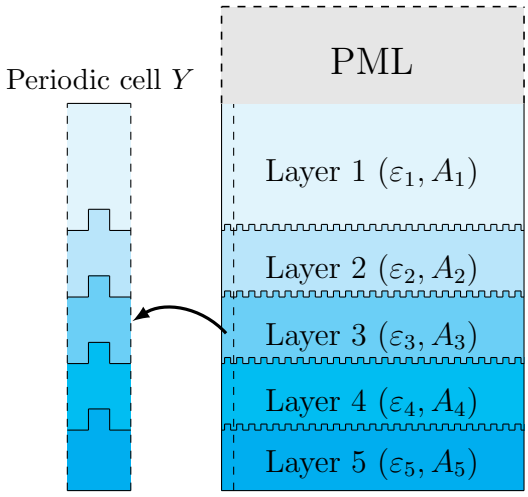


Figure 5.6: A highly-oscillatory medium coefficient

The oscillation are assumed to be periodic with a period δ, as depicted in Figure 5.6. Furthermore, we use the same source f as in the previous numerical tests (illustrated in Figure 2.7).

In this example, no exact solution is available. As a result, we construct a reference solution using a mesh size $H = \frac{1}{2560}$. The error is calculated on five horizontal lines of size $\ell_1 = 1$, located in $\boldsymbol{x}_2 = \frac{2}{8}$, $\boldsymbol{x}_2 = \frac{3}{8}$, $\boldsymbol{x}_2 = \frac{4}{8}$, $\boldsymbol{x}_2 = \frac{5}{8}$, $\boldsymbol{x}_2 = \frac{6}{8}$.

In the following experiments, we investigate the convergence with respect to the mesh size H for a small fixed value of δ. In particular, we plot the relative H^1 errors with respect to H, and we are interested on the effect of the resonance error term $\frac{\delta}{H}$. To do this end, we fix a frequency $k = 12.6\pi$, an incident angle $\theta = 23°$, and the PML parameters $\ell_{\mathrm{P}} = 0.5$ and $\gamma_{\mathrm{r}} = \gamma_{\mathrm{i}} = 4$. The length of the texturation teeth is also fixed at $\frac{1}{16}$. We have also chosen to fix the physical coefficient $\boldsymbol{A} = 1$ and choose the following permittivity values: $\varepsilon_1 = 1$, $\varepsilon_2 = 0.6$, $\varepsilon_3 = 0.3$, $\varepsilon_4 = 0.2$, $\varepsilon_5 = 0.1$.

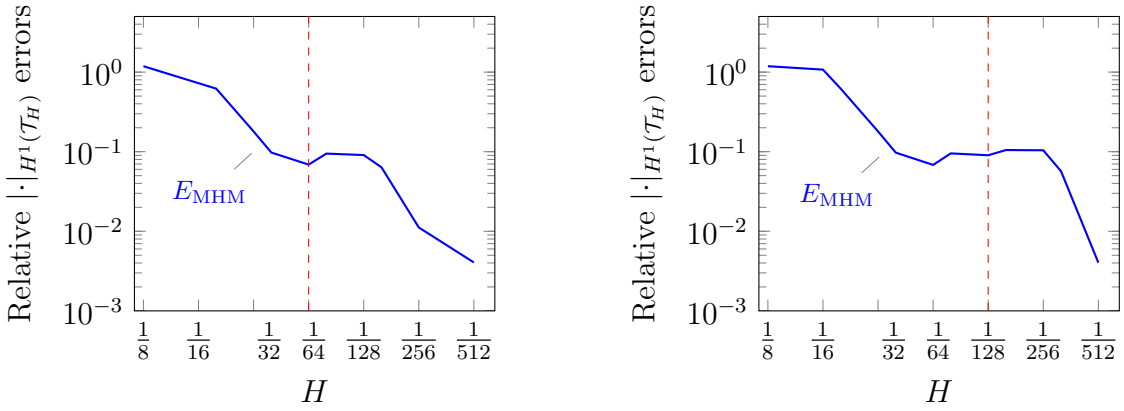


Figure 5.7: MHM errors for with $\delta = \frac{1}{64}$ (left) and $\delta = \frac{1}{128}$ $k = 15\pi$ (right).

Figure 5.7 shows that the error curve is composed of three regimes. In the two regimes where $H < \frac{\delta}{2}$ or $H > 2\delta$, we observe that the error decreases with an order of H. However, in the other regime, the behavior of the error curves when $H \approx \delta$ is different. In particular, we observe a stabilization of the error curve corresponding to the resonance term $\frac{\delta}{H}$, which validates the theoretical convergence results found in the previous section.

Chapter 6

Conclusion

Summary

This thesis was concerned with a mathematical and numerical study of the two-dimensional Helmholtz equation with quasi-periodic boundary conditions. The properties of the model problem have been chosen in the context of solar cell applications often characterized by periodic and textured multilayer devices.

Having in mind its particular interest in stability and convergence analysis of finite element discretizations, we have started with frequency-explicit stability analysis of the continuous Helmholtz solution. Our stability estimates, presented in Chapter 2, are frequency-explicit, and they were obtained for two cases of physical coefficients. In the first case, we considered a homogeneous model, and we were interested in the effect of the quasi-periodic boundary conditions. Our main achievements, in this case, are the derivation of stability estimates that are explicit and optimal with respect to the frequency and that clearly show the effect of the quasi-periodic boundary conditions on the problem's well-posedness. We illustrated then these optimal stability results by highlighting how they impact the stability of finite element discretizations. For the second case, our results are valid for more general physical coefficients satisfying a monotonicity hypothesis. The estimates obtained in this case are less sharp in terms of frequency but are uniform to the variation of the coefficients in the periodic direction, so they are valid for layers with highly oscillating interfaces.

In the second contribution, we replace the non-local absorption conditions with a PML layer, keeping in mind efficient numerical schemes. We then focused on two central results. Firstly, for the case of the right-hand sides contained in the original domain, we proposed an explicit convergence analysis showing that the PML problem's solution converges to the original problem's solution when the PML parameters are correctly chosen. The proposed error estimates are explicit and clearly show the effect of the quasi-resonant modes.

Numerical experiments illustrating our theoretical convergence results are then presented. Secondly, we focused on the PML problem's well-posedness and considered the case of the right-hand sides contained in the absorbing layer, which is paramount in the stability and convergence study of numerical methods. We have presented our well-posedness analysis for two different cases. On the one hand, we provided optimal stability results for the case of a homogeneous medium. On the other hand, we provided a well-posedness study for the PML problem with general physical coefficients. Our results, in this case, rely on the well-posedness results of the original problem obtained in Chapter 2. As a result, they are sub-optimal compared to those stability results obtained for the original DtN problem.

In the fourth chapter, we considered physical coefficients with small periodic oscillations. Then, we analyzed our Helmholtz problem using the homogenization theory. Indeed, we used that the stability bounds achieved in chapters 2 and 3 are uniform to the variation of the physical coefficients in the periodic direction. We therefore derived frequency-explicit error estimates controlling for the difference between the oscillating and homogenized solutions. The homogenization results obtained in this chapter appeared in our study as an essential pre-step to analyze the properties of the MHM method. However, these results are interesting for homogenization theory and can be applied to the analysis of a wide range of multiscale numerical methods.

Regarding the analysis of the MHM method, we have shown the effect of the PML and the quasi-periodic boundary conditions on the MHM formulation. In particular, their effect on the primal and dual MHM variables. We have also studied the well-posedness of local and global MHM formulations. Then, we presented two convergence analyses to show the performance of the MHM method both in the presence of the quasi-resonant frequencies and when considering highly oscillatory coefficients. In the first part, we have built on the optimal stability results obtained in Chapters 2 and 3 and have been able to gain one frequency-power in the MHM quasi-optimal condition compared to the FEM quasi-optimal condition, which allows the MHM method to remain efficient in the presence of anomalous frequencies. A sequence of numerical tests was presented at the end of this part, illustrating the robustness of the MHM method compared to the FEM method in the presence of quasi-resonant frequencies. The second part dealt with the case of finely textured layers. There, we presented an MHM multiscale convergence analysis based on the homogenization results obtained in chapter 4. The acquired results show that the MHM method is robust in this case and can capture small-scale heterogeneities using coarse meshes. Furthermore, these theoretical results have been evaluated via examples with highly oscillating medium coefficients.

Perspectives

Concerning the stability results presented for the homogeneous case in Chapter 2 for Helmholtz DtN problems and in Chapter 3 for PML problems, the derived estimates are optimal, and they lose a half power of frequency compared to the standard case without quasi-periodic boundary condition. A possible extension of the approach used in this case, namely Fourier mode analysis, can be applied to layered media with flat interfaces. We expect that with additional theoretical developments for this case, optimal estimates can be obtained. However, the stability analysis of the textured layer case is more delicate to manage. By adjusting the chosen "Morawetz multipliers," we can follow [30] and remove the monotonicity hypothesis. In addition, to avoid the deterioration of the stability constant obtained for the general PML problem, a possible application of the technique is possible with further computations to control the norms of the solution in the PML layer. On the other hand, we can also expect an extension of the error analysis between the original solution and the solution of the PML problem for the case of non-constant PML damping functions.

In the homogenization analysis presented in this thesis, we considered smooth coefficients, which means that they do not exactly represent textured layers. This hypothesis has been used in the analysis of homogenization correctors. As a result, a desired extension path of the homogenization analysis is the consideration of physical coefficients representing the textured multilayered cases.

As the MHM method is a relatively new multiscale method, several extensions are imaginable. Although, concerning our problem model, we have seen that the parameters of the PML can impact the MHM formulation. Therefore, a fully-explicit well-posedness analysis is desired for the local and global MHM formulations. Other extension possibilities are either to perform a detailed second-level analysis or to provide a multiscale convergence analysis, including the impact of face partitions. Furthermore, an evident numerical extension of the method for 3D Helmholtz equations is expected. Finally, from an implementation point of view, the MHM code used in this work only accepts Cartesian meshes. Thus, generalizing this restriction can allow the numerical analysis of the method to progress.

Bibliography

[1] G. Alessandrini. Strong unique continuation for general elliptic equations in 2D. *Journal of Mathematical Analysis and Applications*, 386(2):669–676, 2012.

[2] R. Araya, C. Cárcamo, A. H. Poza, and F. Valentin. An adaptive multiscale hybrid-mixed method for the Oseen equations. *Advances in Computational Mathematics*, 47(1):1–36, 2021.

[3] Rodolfo Araya, Christopher Harder, Diego Paredes, and Frédéric Valentin. Multiscale Hybrid-Mixed Method. *SIAM Journal on Numerical Analysis*, 51(6):3505–3531, January 2013. Publisher: Society for Industrial and Applied Mathematics.

[4] F. Assous, P. Ciarlet, and S. Labrunie. *Mathematical foundations of computational electromagnetism*. Springer, 2018.

[5] R. J. Astley. Infinite elements for wave problems: a review of current formulations and an assessment of accuracy. *International Journal for Numerical Methods in Engineering*, 49(7):951–976, 2000.

[6] J. P. Aubin. Behavior of the error of the approximate solutions of boundary value problems for linear elliptic operators by galerkin's and finite difference methods. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 21(4):599–637, 1967.

[7] A. K. Aziz. *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations*. Academic Press, May 2014. Google-Books-ID: EMLiBQAAQBAJ.

[8] A. K. Aziz, R. B. Kellogg, and A. B. Stephens. A two point boundary value problem with a rapidly oscillating solution. *Numerische Mathematik*, 53(1-2):107–121, 1988.

[9] G. Bao. Diffractive optics in periodic structures: the TM polarization. 1994.

[10] G. Bao, Z. Chen, and H. Wu. Adaptive finite-element method for diffraction gratings. *Journal of the Optical Society of America A*, 22(6):1106, 2005.

[11] G. Bao, D. C. Dobson, and J. A. Cox. Mathematical studies in rigorous grating theory. *Journal of the Optical Society of America A*, 12(5):1029, 1995.

[12] G. Bao, P. Li, and H. Wu. An adaptive edge element method with perfectly matched absorbing layers for wave scattering by biperiodic structures. *Mathematics of Computation*, 79(269):1–34, 2010.

[13] Gang Bao and David Dobson. On the scattering by a biperiodic structure. *Proceedings of the American Mathematical Society*, 128(9):2715–2723, 2000.

[14] G. R. Barrenechea, F. Jaillet, D. Paredes, and F. Valentin. The multiscale hybrid mixed method in general polygonal meshes. *Numerische Mathematik*, 145(1):197–237, May 2020.

[15] H. Barucq, T. Chaumont-Frelet, and C. Gout. Stability analysis of heterogeneous Helmholtz problems and finite element solution based on propagation media approximation. *Mathematics of Computation*, 86(307):2129–2157, 2017.

[16] D. Baskin, E. A. Spence, and J. Wunsch. Sharp high-frequency estimates for the helmholtz equation and applications to boundary integral equations. *SIAM Journal on Mathematical Analysis*, 48(1):229–267, 2016.

[17] J. P. Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of computational physics*, 114(2):185–200, 1994.

[18] J. P. Berenger. Perfectly matched layer for the FDTD solution of wave-structure interaction problems. *IEEE Transactions on Antennas and Propagation*, 44(1):110–117, 1996.

[19] A. Bermúdez, L. Hervella-Nieto, A. Prieto, and R. Rodrı. An optimal perfectly matched layer with unbounded absorbing function for time-harmonic acoustic scattering problems. *Journal of computational Physics*, 223(2):469–488, 2007.

[20] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*, volume 2. Springer, 2011.

[21] M. L Brongersma, Y. Cui, and S. Fan. Light management for photovoltaics using high-index nanostructures. *Nature materials*, 13(5):451–460, 2014.

[22] D. L. Brown, D. Gallistl, and D. Peterseim. Multiscale Petrov-Galerkin Method for High-Frequency Heterogeneous Helmholtz Equations. In *Meshfree methods for partial differential equations VIII*, pages 85–115. Springer, 2017.

[23] N. Burq. Décroissance de l'énergie locale de l'équation des ondes pour le problème extérieur et absence de résonance au voisinage du réel. *Acta mathematica*, 180(1):1–29, 1998.

[24] N. Burq. Smoothing effect for schrodinger boundary value problems. *Duke Mathematical Journal*, 123(2):403, 2004.

[25] P. Campbell and M. A. Green. Light trapping properties of pyramidally textured surfaces. *Journal of Applied Physics*, 62(1):243–249, 1987.

[26] S.N. Chandler-Wilde, I.G. Graham, S. Langdon, and M. Lindner. Condition number estimates for combined potential boundary integral operators in acoustic scattering. *Journal of Integral Equations and Applications*, 21(2), 2009.

[27] S.N. Chandler-Wilde and P. Monk. Existence, Uniqueness, and Variational Methods for Scattering by Unbounded Rough Surfaces. *SIAM Journal on Mathematical Analysis*, 37(2):598–618, 2005.

[28] S.N. Chandler-Wilde and P. Monk. Wave-Number-Explicit Bounds in Time-Harmonic Scattering. *SIAM Journal on Mathematical Analysis*, 39(5):1428–1455, 2008.

[29] S.N. Chandler-Wilde, E.A. Spence, A. Gibbs, and V.P. Smyshlyaev. High-frequency bounds for the Helmholtz equation under parabolic trapping and applications in numerical analysis. *SIAM J. Math. Anal.*, 52(1):845–893, 2020.

[30] T. Chaumont-Frelet. *Finite element approximation of Helmholtz problems with application to seismic wave propagation*. PhD thesis, INSA Rouen and Inria project-team Magique3D, 2015.

[31] T. Chaumont-Frelet, D. Gallistl, S. Nicaise, and J. Tomezyk. Wavenumber-explicit convergence analysis for finite element discretizations of time-harmonic wave propagation problems with perfectly matched layers. *Communications in Mathematical Sciences*, 20(1):1–52, 2022.

[32] T. Chaumont-Frelet and S. Nicaise. High-frequency behaviour of corner singularities in Helmholtz problems. *ESAIM: Mathematical Modelling and Numerical Analysis*, 52(5):1803–1845, 2018.

[33] T. Chaumont-Frelet and S. Nicaise. Wavenumber explicit convergence analysis for finite element discretizations of general wave propagation problems. *IMA Journal of Numerical Analysis*, 40(2):1503–1543, 2020.

[34] T. Chaumont-Frelet, D. Paredes, and F. Valentin. Flux approximation on unfitted meshes and application to multiscale hybrid-mixed methods. 2022. hal-id: 03834748.

[35] T. Chaumont-Frelet and E.A. Spence. Scattering by finely-layered obstacles: frequency-explicit bounds and homogenization. 2021.

[36] T. Chaumont-Frelet and F. Valentin. A Multiscale Hybrid-Mixed Method for the Helmholtz Equation in Heterogeneous Domains. *SIAM Journal on Numerical Analysis*, 58(2):1029–1067, 2020.

[37] H.-T. Chen, A. J. Taylor, and N. Yu. A review of metasurfaces: physics and applications. *Reports on Progress in Physics*, 79(7):076401, 2016.

[38] X. Chen and A. Friedman. Maxwell's equations in a periodic structure. *Transactions of the American Mathematical Society*, 323(2):465–507, 1991.

[39] Z. Chen and X. Liu. An Adaptive Perfectly Matched Layer Technique for Time-harmonic Scattering Problems. *SIAM Journal on Numerical Analysis*, 43(2):645–671, 2005.

[40] Z. Chen and H. Wu. An Adaptive Finite Element Method with Perfectly Matched Absorbing Layers for the Wave Scattering by Periodic Structures. *SIAM Journal on Numerical Analysis*, 41(3):799–826, 2003.

[41] Zhiming Chen and Xueshuang Xiang. A source transfer domain decomposition method for helmholtz equations in unbounded domain. *SIAM Journal on Numerical Analysis*, 51(4):2331–2356, 2013.

[42] F. Collino and P. Monk. The perfectly matched layer in curvilinear coordinates. *SIAM Journal on Scientific Computing*, 19(6):2061–2090, 1998.

[43] F. Collino and C. Tsogka. Application of the perfectly matched absorbing layer model to the linear elastodynamic problem in anisotropic heterogeneous media. *Geophysics*, 66(1):294–307, 2001.

[44] V. Depauw, C. Trompoukis, I. Massiot, W. Chen, A. Dmitriev, P. R. i Cabarrocas, I. Gordon, and J. Poortmans. Sunlight-thin nanophotonic monocrystalline silicon solar cells. *Nano Futures*, 1(2):021001, 2017.

[45] J. Diaz. *Approches analytiques et numériques de problèmes de transmission en propagation d'ondes en régime transitoire. Application au couplage fluide-structure et aux méthodes de couches parfaitement adaptées*. PhD thesis, ENSTA ParisTech, 2005.

[46] D. Dobson and A. Friedman. The time-harmonic maxwell equations in a doubly periodic structure. *Journal of Mathematical Analysis and Applications*, 166(2):507–528, 1992.

[47] D. C. Dobson. Optimal design of periodic antireflective structures for the helmholtz equation. *European Journal of Applied Mathematics*, 4(4):321–339, 1993.

[48] P. Donato and D. Cioranescu. *An introduction to homogenization*, volume 17. Oxford university press, 1999.

[49] J. Douglas Jr, J. E. Santos, D. Sheen, and L. S. Bennethum. Frequency domain treatment of one-dimensional scalar waves. *Mathematical models and methods in applied sciences*, 3(02):171–194, 1993.

[50] R. B. Dunbar, H. C. Hesse, D. S. Lembke, and L. Schmidt-Mende. Light-trapping plasmonic nanovoid arrays. *Phys. Rev. B*, 85:035301, 2012.

[51] O. Duran, P. R. B. Devloo, S. M. Gomes, and F. Valentin. A multiscale hybrid method for darcy's problems using mixed finite element local solvers. *Computer methods in applied mechanics and engineering*, 354:213–244, 2019.

[52] S. Dyatlov and M. Zworski. *Mathematical Theory of Scattering Resonances*, volume 200 of *Graduate Studies in Mathematics*. American Mathematical Society, 2019.

[53] Weinan E, B. Engquist, X. Li, W. Ren, and E. Vanden-Eijnden. The heterogeneous multiscale method: A review. *Commun. Comput. Phys*, 2007.

[54] Y. Efendiev, J. Galvis, and T. Y. Hou. Generalized multiscale finite element methods (GMsFEM). *Journal of computational physics*, 251:116–135, 2013.

[55] Y. Efendiev and T. Y. Hou. *Multiscale finite element methods: theory and applications*, volume 4. Springer Science & Business Media, 2009.

[56] B. Engquist and A. Majda. Absorbing boundary conditions for numerical simulation of waves. *Proceedings of the National Academy of Sciences*, 74(5):1765–1766, 1977.

[57] B. Engquist and A. Majda. Radiation boundary conditions for acoustic and elastic wave calculations. *Communications on Pure and Applied Mathematics*, 32(3):313–357, 1979.

[58] X. Feng and H. Wu. Discontinuous Galerkin methods for the Helmholtzequation with large wave number. *SIAM Journal on Numerical Analysis*, 47(4):2872–2896, 2009.

[59] X. Feng and H. Wu. *hp*-discontinuous Galerkin methods for the Helmholtz equation with large wave number. *Math. Comp.*, 80(276):1997–2024, 2011.

[60] P. Fernandes and G. Gilardi. Magnetostatic and electrostatic problems in inhomogeneous anisotropic media with irregular boundary and mixed boundary conditions. *Math. Meth. Appl. Sci.*, 47(4):2872–2896, 1997.

[61] J. Galkowski, D. Lafontaine, and E. A. Spence. Perfectly-matched-layer truncation is exponentially accurate at high frequency. 2022.

[62] J. Galkowski, E. H. Müller, and E. A. Spence. Wavenumber-explicit analysis for the Helmholtz h-BEM: error estimates and iteration counts for the Dirichlet problem. *Numerische Mathematik*, 142(2):329–357, June 2019.

[63] D. Gallistl and D. Peterseim. Stable multiscale petrov-galerkin finite element method for high frequency acoustic scattering. *Computer Methods in Applied Mechanics and Engineering*, 295:1–17, 2015.

[64] Q. Gan, F. J. Bartoli, and Z. H. Kafafi. Plasmonic-enhanced organic photovoltaics: Breaking the 10% efficiency barrier. *Advanced materials*, 25(17):2385–2396, 2013.

[65] V. Girault and P.A. Raviart. *Finite element methods for Navier-Stokes equations: theory and algorithms*. Springer-Verlag, 1986.

[66] D. Givoli. *Exact and High-Order Non-Reflecting Computational Boundaries*. Springer Berlin Heidelberg, 2003.

[67] D. Givoli. High-order local non-reflecting boundary conditions: a review. *Wave motion*, 39(4):319–326, 2004.

[68] D. Givoli, I. Patlashenko, and J. B. Keller. High-order boundary conditions and finite elements for infinite domains. *Computer Methods in Applied Mechanics and Engineering*, 143(1):13–39, 1997.

[69] Alexis Gobé. *Discontinuous Galerkin methods for the simulation of multiscale nanophotonic problems with application to light trapping in solar cells*. PhD thesis, Université Côte d'Azur, 2020.

[70] I. G. Graham, M. Löhndorf, J. M. Melenk, and E. A. Spence. When is the error in the h-bem for solving the helmholtz equation bounded independently of k? *BIT Numerical Mathematics*, 55(1):171–214, 2015.

[71] D. J. Griffiths. *Introduction to electrodynamics*. American Association of Physics Teachers, 2014.

[72] P. Grisvard. *Elliptic problems in nonsmooth domains*. Pitman, 1985.

[73] M. J. Grote and J. B. Keller. Exact nonreflecting boundary conditions for the time dependent wave equation. *SIAM Journal on Applied Mathematics*, 55(2):280–297, 1995.

[74] T. Ha and I. Kim. Analysis of one-dimensional Helmholtz equation with PML boundary. *Journal of computational and applied mathematics*, 206(1):586–598, 2007.

[75] C. Harder, A. L. Madureira, and F. Valentin. A hybrid-mixed method for elasticity. *ESAIM: Mathematical Modelling and Numerical Analysis*, 50(2):311–336, 2016.

[76] C. Harder, D. Paredes, and F. Valentin. A family of multiscale hybrid-mixed finite element methods for the Darcy equation with rough coefficients. *Journal of Computational Physics*, 245:107–130, 2013.

[77] C. Harder, D. Paredes, and F. Valentin. On a multiscale hybrid-mixed method for advective-reactive dominated problems with heterogeneous coefficients. *Multiscale Modeling & Simulation*, 13(2):491–518, 2015.

[78] Christopher Harder and Frédéric Valentin. Foundations of the MHM method. In *Building Bridges: Connections and Challenges in Modern Approaches to Numerical Partial Differential Equations*, pages 401–433. Springer, 2016.

[79] P. Henning and D. Peterseim. Oversampling for the multiscale finite element method. *Multiscale Modeling and Simulation*, 11(4):1149–1175, 2013.

[80] U. Hetmaniuk. Stability estimates for a class of Helmholtz problems. *Communications in Mathematical Sciences*, 5(3):665–678, 2007.

[81] R. Hiptmair and P. Meury. Stabilized FEM-BEM coupling for Helmholtz transmission problems. *SIAM journal on numerical analysis*, 44(5):2107–2130, 2006.

[82] T. Hohage, F. Schmidt, and L. Zschiedrich. Solving time-harmonic scattering problems based on the pole condition II: convergence of the PML method. *SIAM journal on mathematical analysis*, 35(3):547–560, 2003.

[83] T. Hou, X.-H. Wu, and Z. Cai. Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients. *Mathematics of computation*, 68(227):913–943, 1999.

[84] T. Y. Hou and X.-H. Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *Journal of computational physics*, 134(1):169–189, 1997.

[85] F. Q. Hu. On absorbing boundary conditions for linearized Euler equations by a perfectly matched layer. *Journal of computational physics*, 129(1):201–219, 1996.

[86] F. Q. Hu. A stable, perfectly matched layer for linearized Euler equations in unsplit physical variables. *Journal of Computational Physics*, 173(2):455–480, 2001.

[87] F. Ihlenburg. Finite element analysis of acoustic scattering. 1998.

[88] F. Ihlenburg and I. Babuška. Finite element solution of the Helmholtz equation with high wave number Part I: The h-version of the FEM. *Computers & Mathematics with Applications*, 30(9):9–37, 1995.

[89] F. Ihlenburg and I. Babuška. Finite Element Solution of the Helmholtz Equation with High Wave Number Part II: The $h - p$ Version of the FEM. *SIAM Journal on Numerical Analysis*, 34(1):315–358, 1997.

[90] M. Ikawa. Decay of solutions of the wave equation in the exterior of several convex bodies. *Annales de l'institut Fourier*, 38(2):113–146, 1988.

[91] T. Jiang, T. Jiao, and Y. Li. A low mutual coupling mimo antenna using periodic multi-layered electromagnetic band gap structures. *The Applied Computational Electromagnetics Society Journal (ACES)*, pages 305–311, 2018.

[92] K. Kempa, B. Kimball, J. Rybczynski, Z. P. Huang, P. F. Wu, D. Steeves, M. Sennett, M. Giersig, D. V. G. L. N. Rao, D. L. Carnahan, D. Z. Wang, J. Y. Lao, W. Z. Li, and Z. F. Ren. Photonic crystals based on periodic arrays of aligned carbon nanotubes. *Nano Letters*, 3(1):13–18, 2003.

[93] J. Krč, F. Smole, and M. Topič. Potential of light trapping in microcrystalline silicon solar cells with textured substrates. *Progress in photovoltaics: Research and applications*, 11(7):429–436, 2003.

[94] D. Lafontaine, E.A. Spence, and J. Wunsch. Wavenumber-explicit convergence analysis of the *hp*-fem for the full-space heterogeneous Helmholtz equation with smooth coefficients. *Comput. Math. Appl.*, 113:59–69, 2022.

[95] S. Lanteri, D. Paredes, C. Scheid, and F. Valentin. The multiscale hybrid-mixed method for the Maxwell equations in heterogeneous media. *Multiscale Modeling & Simulation*, 16(4):1648–1683, 2018.

[96] M. Lassas, J. Liukkonen, and E. Somersalo. Complex Riemannian metric and absorbing boundary conditions. *Journal de mathématiques pures et appliquées*, 80(7):739–768, 2001.

[97] S.-F. Leung, M. Yu, Q. Lin, K. Kwon, K.-L. Ching, L. Gu, K. Yu, and Z. Fan. Efficient photon capturing with ordered three-dimensional nanowell arrays. *Nano letters*, 12(7):3682–3689, 2012.

[98] Y. Li and H. Wu. Fem and cip-fem for helmholtz equation with high wave number and perfectly matched layer truncation. *SIAM Journal on Numerical Analysis*, 57(1):96–126, 2019.

[99] C. Makridakis, F. Ihlenburg, and I. Babuška. Analysis and finite element methods for a fluid-solid interaction problem in one dimension. *Mathematical Models and Methods in Applied Sciences*, 6(08):1119–1141, 1996.

[100] N. Meinzer, W. L. Barnes, and I. R. Hooper. Plasmonic meta-atoms and metasurfaces. *Nature Photonics*, 8(12):889–898, 2014.

[101] J. M. Melenk and S. Sauter. Wavenumber explicit convergenceanalysis for galerkin discretizations of the helmholtz equation. *SIAM Journal on Numerical Analysis*, 49(3):1210–1243, 2011.

[102] J.M. Melenk. *On generalized finite element methods.* PhD thesis, University of Maryland, 1995.

[103] J.M. Melenk and S. Sauter. Convergence analysis for finite element discretizations of the Helmholtz equation with Dirichlet-to-Neumann boundary conditions. *Mathematics of Computation*, 79(272):1871–1914, 2010.

[104] C. Min, J. Li, G. Veronis, JY. Lee, S. Fan, and P. Peumans. Enhancement of optical absorption in thin-film organic photovoltaic solar cells through the excitation of plasmonic modes in metallic gratings. *Appl. Phys. Lett*, 96:133302–1, 2010.

[105] R. Mittra, C.H. Chan, and T. Cwik. Techniques for analyzing frequency selective surfaces-a review. *Proceedings of the IEEE*, 76(12):1593–1615, 1988.

[106] A. Moiola and E.A Spence. Acoustic transmission problems: wavenumber-explicit bounds and resonance-free regions. *Math. Meth. Appl. Sci.*, 29(2):317–354, 2019.

[107] S. Mokkapati, FJ. Beck, A. Polman, and KR. Catchpole. Designing periodic arrays of metal nanoparticles for light-trapping applications in solar cells. *Applied Physics Letters*, 95(5):053115, 2009.

[108] C.S. Morawetz. The decay of solutions of the exterior initial-boundary value problem for the wave equation. *Communications on Pure and Applied Mathematics*, 14(3):561–568, 1961.

[109] C.S. Morawetz. Decay for solutions of the exterior problem for the wave equation. *Communications on Pure and Applied Mathematics*, 28(2):229–264, 1975.

[110] C.S. Morawetz and D. Ludwig. An inequality for the reduced wave operator and the justification of geometrical optics. *Communications on pure and applied mathematics*, 21(2):187–203, 1968.

[111] C.S. Morawetz, J.V. Ralston, and W.A. Strauss. Decay of solutions of the wave equation outside nontrapping obstacles. *Communications on Pure and Applied Mathematics*, 30(4):447–508, 1977.

[112] A. Målqvist and D. Peterseim. Localization of elliptic multiscale problems. *Mathematics of Computation*, 83(290):2583–2603, 2014.

[113] J.-C. Nedelec and F. Starling. Integral equation methods in a quasi-periodic diffraction problem for the time-harmonic Maxwell's equations. *SIAM Journal on Mathematical Analysis*, 22(6):1679–1701, 1991.

[114] J. Nitsche. Ein kriterium für die quasi-optimalität des ritzschen verfahrens. *Numerische Mathematik*, 11(4):346–348, 1968.

[115] M. Ohlberger and B. Verfurth. A new heterogeneous multiscale method for the helmholtz equation with high contrast. *Multiscale Modeling and Simulation*, 16(1):385–411, 2018.

[116] D. Paredes, F. Valentin, and H. Versieux. On the robustness of multiscale hybrid-mixed methods. *Mathematics of Computation*, 86(304):525–548, 2017.

[117] L. E. Payne and H. F. Weinberger. An optimal Poincaré inequality for convex domains. *Archive for Rational Mechanics and Analysis*, 5(1):286–292, 1960.

[118] B. Perthame and L. Vega. Morrey–Campanato estimates for Helmholtz equations. *Journal of Functional Analysis*, 164(2):340–355, 1999.

[119] D. Peterseim. Eliminating the pollution effect in Helmholtz problems by local subscale correction. *Math. Comp.*, 86, 2017.

[120] D. Peterseim and B. Verfürth. Computational high frequency scattering from high contrast heterogeneous media. *Mathematics of Computation*, 89(326):2649–2674, 2020.

[121] P-A Raviart and J.-M. Thomas. Primal hybrid finite element methods for 2nd order elliptic equations. *Mathematics of computation*, 31(138):391–413, 1977.

[122] F. Rellich. Darstellung der eigenwerte von $\nabla u + \lambda u = 0$ durch ein randintegral. *Mathematische Zeitschrift*, 46:635–636, 1940.

[123] C. Rivas, R. Rodríguez, and M.E. Solano. A perfectly matched layer for finite-element calculations of diffraction by metallic surface-relief gratings. *Wave Motion*, 78:68–82, 2018.

[124] S. Sauter and C. Torres. Stability estimate for the Helmholtz equation with rapidly jumping coefficients. *Zeitschrift für angewandte Mathematik und Physik*, 69(6):139, 2018.

[125] A.H. Schatz. An observation concerning ritz-galerkin methods with indefinite bilinear forms. *Mathematics of Computation*, 28(128):959–962, 1974.

[126] M. A. Sefunc, A. K. Okyay, and H. V. Demir. Plasmonic backcontact grating for p3ht: Pcbm organic solar cells enabling strong optical absorption increased in all polarizations. *Optics express*, 19(15):14200–14209, 2011.

[127] E.A. Spence. Wavenumber-explicit bounds in time-harmonic acoustic scattering. *SIAM Journal on Mathematical Analysis*, 46(4):2987–3024, 2014.

[128] Y. Sugawara, T. A. Kelf, J. J. Baumberg, M. E. Abdelsalam, and P. N. Bartlett. Strong coupling between localized plasmons and organic excitons in metal nanovoids. *Phys. Rev. Lett.*, 97:266808, 2006.

[129] Luc Tartar. *An introduction to Sobolev spaces and interpolation spaces*, volume 3. Springer Science and Business Media, 2007. Lecture Notes of the Unione Matematica Italiana.

[130] E. Weinan and E. Björn. The heterogeneous multi-scale method for homogenization problems. pages 89–110, 2005.

[131] K. Yamamoto, D. Adachi, H. Uzu, M. Ichikawa, T. Terashita, T. Meguro, N. Nakanishi, M. Yoshimi, and J. L. Hernández. High-efficiency heterojunction crystalline si solar cell and optical splitting structure fabricated by applying thin-film si technology. *Japanese Journal of Applied Physics*, 54(8S1):08KD15, 2015.

[132] Z. Yu, A. Raman, and S. Fan. Fundamental limit of nanophotonic light trapping in solar cells. *Proceedings of the National Academy of Sciences*, 107(41):17491–17496, 2010.

[133] L. Zhao and A. C. Cangellaris. A general approach for the development of unsplit-field time-domain implementations of perfectly matched layers for FDTD grid truncation. *IEEE Microwave and Guided Wave Letters*, 6(5):209–211, 1996.

[134] W. Zhou and H. Wu. An adaptive finite element method for the diffraction grating problem with pml and few-mode dtn truncations. *Journal of Scientific Computing*, 76(3):1813–1838, 2018.

[135] J. Zhu, C.-M. Hsu, Z. Yu, S. Fan, and Y. Cui. Nanodome solar cells with efficient light management and self-cleaning. *Nano letters*, 10(6):1979–1984, 2010.

[136] L. W. Zschiedrich. *Transparent boundary conditions for Maxwell's equations: Numerical concepts beyond the PML method*. PhD Thesis, 2009.

[137] L. W. Zschiedrich, S. Burger, B. Kettner, and F. Schmidt. Advanced finite element method for nano-resonators. 6115:164–174, 2006.

[138] L. W. Zschiedrich, S. Burger, R. Klose, A. Schaedle, and F. Schmidt. Jcmmode: an adaptive finite element solver for the computation of leaky modes. 5728:192–202, 2005.

[139] M. Zworski. *Semiclassical analysis*, volume 138. American Mathematical Society, 2022.