Académie universitaire Wallonie–Europe
Université de Liège — Faculté des Sciences Appliquées
Collège de doctorat en Électricité, électronique et informatique

Absorbing Layers for Wave-Like Time-Dependent Problems

Design, Discretization and Optimization

Doctoral Dissertation presented by

Axel MODAVE

in fullfilment of the requirements for the degree of
Docteur en Sciences de l'Ingénieur

October 2013

Thesis committee:

Prof. Xavier ANTOINE, Université de Lorraine
Prof. Jean-Marie BECKERS, Université de Liège
Prof. Éric DELHEZ, Université de Liège *(co-advisor)*
Dr Julien DIAZ, Inria Bordeaux Sud-Ouest
Prof. Jean-Jacques EMBRECHTS, Université de Liège *(president)*
Prof. Christophe GEUZAINE, Université de Liège *(advisor)*
Dr Lionel PICHON, Laboratoire de Génie Électrique de Paris – Supélec
Prof. Jean-François REMACLE, Université catholique de Louvain
Prof. Ruth V. SABARIEGO, Katholieke Universiteit Leuven

Version generated on 6th October 2013.

Abstract

The numerical simulation of wave-like phenomena occurring in large or infinite domains is a great challenge for a wide range of technological and scientifical problems. A classical way consists in considering only a limited computational domain with an artificial boundary that requires a specific treatment. In this thesis, *absorbing layers* are developed and studied for time-dependent problems in order to deal with such artificial boundary.

A large part of this thesis is dedicated to the *perfectly matched layers* (PMLs), which exhibit appealing properties. They are first studied in a fundamental case with non-dispersive linear scalar waves. A procedure for building PMLs is proposed for convex domains with regular boundary. It permits a great flexibility when choosing the shape of the computational domain. After, the issue of choosing PML parameters is addressed with the aim of optimizing the PML effectiveness in discrete contexts. The role of each parameter, including the so-called *absorption function*, is highlighted by means of analytical and numerical results. A systematic comparison of different kinds of absorption functions is performed for several classical numerical schemes (based on finite differences, finite volumes or finite elements). Then, while the PMLs do not *a priori* account for incoming signals generated outside the computational domain, different problem formulations that account for such forcing are detailed and compared. The interest of the whole approach is finally illustrated with two- and three-dimensional numerical examples in electromagnetism and acoustics, using a discontinuous finite element scheme.

In regional oceanic models, modeling open-sea boundaries brings new difficulties. Indeed, additional linear/nonlinear dynamics are involved and the external forcing is generally poorly known. In this context, different absorbing layers and the widely used Flather boundary condition are compared by means of classical benchmarks. The choice of the absorption function and the way of prescribing the external forcing are discussed in specific marine cases.

Résumé

La simulation numérique de phénomènes de propagation d'ondes dans des domaines très grands, voire infinis, reste un défi pour nombre de problèmes technologiques et scientifiques. Un moyen classique consiste à utiliser un domaine de calcul limité dont la frontière nécessite un traitement spécifique. Cette thèse est dédiée au développement et à l'étude de *couches absorbantes* pour la modélisation de ce type de frontière artificielle pour des problèmes en temps.

Les couches absorbantes de type *"perfectly matched layer"* (PML) présentent des propriétés intéressantes. Dans un premier temps, on considère des problèmes de base supportant des ondes scalaires linéaires non-dispersives. On propose une procédure de construction de PML pour des domaines convexes à frontière régulière. Les formulations obtenues permettent une grande souplesse dans le choix de la forme du domaine de calcul. La question du choix des paramètres de la PML est alors abordée avec pour objectif d'optimiser son efficacité dans des contextes discrets. Le rôle de chaque paramètre, dont celui de la *fonction d'absorption*, est étudié au moyen de résultats analytiques et numériques. Une comparaison systématique de différentes fonctions d'absorption est réalisée pour plusieurs schémas numériques classiques (basés sur des différences finies, des volumes finis ou des éléments finis). Ensuite, alors que les PMLs ne permettent *a priori* pas de forcer des signaux générés à l'extérieur du domaine de calcul, différentes formulations du problème permettant un tel forçage sont présentées et comparées. Enfin, l'intérêt de l'approche globale est illustré par des simulations numériques à deux et trois dimensions, en électromagnétisme et en acoustique. Le schéma numérique utilisé est basé sur des éléments finis discontinus.

Dans les modèles océaniques régionaux, les frontières en mer ouverte apportent de nouvelles difficultés. En effet, des dynamiques supplémentaires linéaires et/ou non-linéaires doivent être considérées, et le forçage extérieur est généralement mal connu. Dans ce contexte, différentes couches absorbantes et la condition aux limites de Flather sont comparées au moyen de cas-tests classiques. Le choix de la fonction d'absorption et la façon de tenir compte d'un éventuel forçage extérieur sont également discutés.

Remerciements

La réalisation de cette thèse fût une véritable aventure, au cours de laquelle j'ai pu découvrir le monde de la recherche, contribuer à des enseignements et côtoyer d'excellents collègues. Par ces quelques lignes, je voudrais remercier celles et ceux qui ont contribué, de près ou de loin, à cette aventure.

Je tiens avant tout à remercier Christophe Geuzaine, pour m'avoir accueilli dans son équipe de recherche il y a 3 ans, en me permettant de continuer un projet entamé ailleurs. Il m'a donné les moyens de mener à bien cette thèse, en me laissant une grande liberté, tout en me prodiguant d'excellents conseils. Son enthousiasme, ses encouragements et son support ont énormément compté. Au delà de la recherche, travailler dans son service aura été un vrai plaisir, notamment grâce la bonne ambiance qui y règne et aux enseignements dans lesquels il m'a impliqué. Je le remercie pour tout.

Je remercie Eric Delhez pour avoir guidé mes premiers pas dans le monde de la recherche, d'abord à travers le travail de fin d'études qu'il m'a proposé à la fin de ma formation d'ingénieur, et ensuite lors de la préparation de mon premier article et de ma première conférence. Au final, réaliser un travail de thèse dans deux services aux thématiques très différentes aura été une vraie difficulté, mais aussi un atout que j'ai tâché d'exploiter de mon mieux. Cela a été possible notamment grâce aux conseils dont il m'a fait bénéficié tout au long de ces années, même après mon départ de son service. Je l'en remercie.

Xavier Antoine, Jean-Marie Beckers, Julien Diaz, Jean-Jacques Embrechts, Lionel Pichon, Jean-François Remacle et Ruth Sabariego ont accepté de faire partie de mon jury de thèse. Je les en remercie, ainsi que pour les commentaires qu'ils m'ont pu me faire parvenir pour améliorer le manuscrit. Je tiens particulièrement à remercier Ruth pour les milles et une astuces au sujet de Gmsh et GetDP, ainsi que pour toutes nos nombreuses discussions, scientifiques ou non. Un merci particulier aussi à Xavier pour les invitations aux réunions de l'ANR "Microwaves" et aux journées de Metz 2012.

Merci à mes collègues liégeois, passés et présents. Merci à Christophe M. pour les nombreux moments passés ensemble et pour ses conseils au début de mes recherches. Merci à Bertrand pour ses conseils pour la rédaction, pour les bons moments (même par mails interposés) et pour son amitié. Merci à Marlène, Natalie et Patricia pour

les nombreux temps de midi passés à réinventer la faculté ... mais pas seulement ! Merci à mon voisin de bureau d'ACE, Alex, pour la bonne ambiance du I.157. Merci également à mes autres collègues du service de mathématiques générales, Sébastien, Frank et Francine, ainsi qu'à mes autres collègues (et assimilés) d'ACE, Amaury, Ayoub, Dave, Frédéric, Innocent, Isabel, Jean de Dieu, Jeannot, Nicolas, Patrick, Pierre, Riccardo, Sabine, Sandrine, Simon, Valera, Véronique, Vincent F., Vincent N. et Vuong.

Je remercie très chaleureusement Abelin et Mohamed pour le bon accueil qu'ils m'ont réservé à chacun de mes passages au LGEP, ainsi que pour notre collaboration. Merci à Jon et aux collègues de Louvain-la-Neuve qui ont contribué au code DG que j'ai utilisé au cours de cette thèse. Merci également à Eric Deleersnijder pour m'avoir impliqué dans le projet SLIM au début de ma thèse.

Puisqu'il y a ces petites choses qui comptent pour beaucoup ... Merci à Lilly, Joey, Anncha et Oz pour les moments de décompression. Merci aux (ex-)étudiants qui m'ont permis de découvrir la Californie. Re-merci à Christophe G. pour la glace sur le Pier de Santa Monica. Re-merci à Christophe M. pour avoir été mon juke-box pendant les 2 ans où j'ai été son presque-voisin de bureau. Re-merci à Bertrand pour le chocolat et le clafouti. Re-merci à Ruth pour me donner constamment envie de faire du sport. Re-merci aux collègues pour les bons moments en conférence.

Last but not least ...

J'ai toujours pu compter sur mes parents et mon frère en toutes circonstances. Pendant mes études et la réalisation de cette thèse, ils ont été un socle sur lequel j'ai pu me (re)poser. Merci de tout cœur pour tout ! Je tiens également à remercier ma belle-sœur et ma belle-famille pour les bons moments en famille et pour leur soutien.

Enfin, je remercie mon compagnon, Cédric, pour son soutien, ses conseils, ses relectures du manuscrit ... mais surtout pour son affection et sa patience. C'est une page très particulière qui se tourne pour nous deux, parfois difficile à cause de la distance géographique. Aujourd'hui, j'en suis au point où il en était lorsque nous nous sommes rencontré: en fin de thèse. Je suis très fier de lui, très fier de présenter ma thèse devant lui, et impatient de savoir ce que nous réserve la suite.

A. Modave
Octobre 2013

Contents

Introduction

Numerical simulations are intensively used in both industry and academia for studying and solving both technological and scientifical problems. The complexity of problems of interest is ever increasing, as well as the accuracy requirements on the numerical solutions. In order to reduce the computational costs, researchers are permanently looking for methods that are both cheap and accurate.

In this work, we are interested in the numerical resolution of wave-like problems set on spatial domains that are infinite or very large in comparison with the regions of interest. Such kinds of problems are encountered in various branches of science and engineering. For example, one purpose of aero-acoustics is the study of the noise generated around aircraft engines. Even if we are only interested in what happens in a limited region around the engines, the physical domain that should be ideally considered is the whole space. In electromagnetic compatibility, electronic devices should ideally be tested in open space to remove exterior influences, which is approached experimentally by performing the measurements in anechoic chambers. Finally, many oceanographic and atmospheric models deal with a limited geographic zone instead of the whole globe. All these examples have in common the propagation of waves (*i.e.* sound waves, electromagnetic waves and oceanic/atmospheric waves, respectively) in infinite or large regions.

A major challenge for the numerical simulation of such problems is to compute the solution on a small domain — the *truncated domain* — instead of the large or infinite physical one, without altering the original solution. The boundary of the truncated domain is purely artificial, and requires a specific treatment. At this boundary, the model is supposed to describe accurately the outward propagation of signals and perturbations of all kinds generated inside the truncated domain, even if they are *a priori* not known. For this purpose, a lot of boundary conditions, artificial layers and alternative techniques have been developed and used in various physical and numerical contexts (see *e.g.* [9, 27, 41, 71, 73, 79, 164] and references therein).

Artificial layers, generally called *absorbing layers*, have been used for a long time to deal with truncated domains. Inside these layers, the model equations are modified in order to damp (or 'absorb') outgoing waves. Initially, such absorbing layers were built based on physical reasoning, and were mainly dedicated to improving the efficicency of artificial boundary conditions. The introduction of

the *perfectly matched layer* (PML) concept by Bérenger [21] in 1994 changed the paradigm. With his PML, Bérenger showed that it is possible to design layers that perfectly simulate the truncation of the physical domain in electromagnetism. Afterwards, the concept has been extended and used in a large range of wave-like problems.

While the efficiency of PMLs has already been demonstrated in many cases, further developements are required and new absorbing layers must be developed to deal with increasingly complex problems. This notably emcompasses problems involving nonlinear dynamics, complicated geometries or large-scale computation.

Scope and goals of this work

This work contributes to the improvement of truncation strategies for simulating wave-like phenomena that occur in large or infinite areas. In particular, it is focused on the design, the discretization and the optimization of absorbing layers for time-dependent problems.

For this purpose, three topics are treated in the context of scalar-waves time-dependent problems:

- *The design of PMLs for truncated domains of general shape.*

  Since the choice of the truncated domains is *a priori* arbitrary, it is advantageous to take a truncated domain as small as possible in order to reduce the computational cost. This can lead to domains with non-conventional shapes (*i.e.* that are not cuboid, cylindrical or spherical). However, only a few time-dependent PML formulations have been investigated in the literature for these kinds of domains [52, 107]. In this thesis, we are looking for a simple procedure to build such formulations, which can be easily implemented in existing finite difference, finite volume or finite element codes.

- *The optimal choice of PML parameters.*

  In discrete contexts, the efficiency of a PML essentially depends on the layer thickness, the discretization and a spatially varying function σ, called *absorption function*. Increasing the layer thickness and refining the discretization generally improve the PML effectiveness, but at the price of an increasing computational cost. Therefore, a common approach is to first choose them, and then tune the absorption function σ. However, this tuning is generally made by using costly case-dependent optimization procedures or empirical rules. In this thesis, we study in detail the influence of the problem parameters on the PML efficiency at the discrete level, and compare different absorption functions. We aim at reaching a conclusion as general as possible on optimum parameters in order to make any tuning unnecessary for a large range of applications.

- *The external forcing of incoming signals through PMLs.*

  For several problems, such as scattering problems, wave sources are placed outside the truncated domain. The incoming signals generated by these sources must then be prescribed through the layer. However, PML formulations do not naturally account for such external forcing. We are therefore looking for strategies to prescribed incoming signals without altering the efficiency of the PML.

Even if all developments are made in the specific context of a basic scalar wave system, the scope of our conclusions is larger. The applicability of our approach is first tested on acoustic and two-dimensional electromagnetic cases. We then deal with regional oceanographic models where additional difficulties are considered: multiple dynamics and the poor knowledge of external forcing.

Outline

This work is divided into five chapters. The chapter following this introduction presents the mathematical and numerical frameworks of this thesis. Fundamental systems of equations for wave-like problems are defined: the hyperbolic system and the scalar wave system. Their properties related to wave propagation are highlighted. Then, different numerical schemes are derived with classical methods: finite difference method, finite volume method and finite element method with both continuous and discontinuous elements.

In chapter 2, we propose a review of the main kinds of boundary conditions and absorbing layers that deal with truncated domains. Bérenger's PML is presented and its fundamental properties at the continuous level are highlighted. Some intepretations of PMLs are explained and illustrated in the time-harmonic domain. After, we propose a procedure for building novel PMLs for general convex truncated domains with regular boundary. The PML equations are explicitly written for the time-dependent scalar wave system in Cartesian coordinates.

Chapter 3 deals with the optimization of PMLs at the discrete level, *i.e.* when numerical methods are used. First, we study in detail the influence of the problem parameters on the efficiency of discretized PMLs in one dimension. We derive and interpret new analytical results for the finite difference scheme. Numerical results are then provided to extend the analysis to other numerical schemes. Then, different kinds of absorption functions are optimized and compared by means of numerical results in both one and two dimensions. In order to draw conclusions as general as possible, all numerical simulations are systematically performed with the different numerical schemes.

In chapter 4, we present different strategies to account for incoming signals in scattering problems when PMLs are used. Then, we propose a discontinuous Galerkin scheme for the general PML equations of the scalar wave system. Its nu-

merical convergence is shown. Two illustrations of electromagnetic and acoustic scattering applications are presented in two and three dimensions, respectively. They show the efficiency of the proposed PML formulations with generally shaped truncated domains, and confirm the conclusion of the previous chapter.

Chapter 5 is dedicated to the study of absorbing layers in the perspective of marine applications. Different absorbing layers (which are perfectly matched or not) are presented in the framework of the shallow water model. These layers are compared, and the choice of the absorption function (a common constraint for all these layers) is discussed using benchmarks that are strongly influenced by the dynamics of oceanography. The methods for prescribing an external solution through an absorbing layer are discussed in the case of the shallow water model.

Finally, general conclusions are drawn and research prospectives are pointed out.

Original contributions and communications

Hereafter, a list of main contributions that we believe original is drawn up:

- A complete procedure for building PMLs for convex truncated domains with regular boundary, applicable in time-harmonic and time-dependent problems (section 2.4). The implementation of a PML for a non-conventionally shaped truncated domain in a three-dimensional benchmark (section 4.4.2);
- Exact formulas of discrete reflection coefficients associated to a PML in one dimension, discretized with a finite difference method, for a constant absorption function (section 3.2.1);
- The systematic comparison of optimized absorption functions in both one and two dimensions, with four different numerical methods in the time domain (section 3.3). The extension of this comparison to cases dominated by ocean dynamics (sections 5.3 and 5.4);
- An original interpretation of the optimized hyperbolic absorption functions (section 3.3.3.3 and end of section 5.4);
- Two novel layers for shallow water equations (section 5.2). A comparison between these novel layers, classical layers and a classical boundary condition by means of different benchmarks (chapter 5);
- The use of an interface condition for forcing incoming signals through layers in the context of the shallow water model (section 5.5).

An important part of the time devoted to the realization of this thesis has been spent on implementing the proposed methods in different research codes. The majority of simulations, performed with discontinuous Galerkin schemes, have been performed using a C++/Python code developed in collaboration with the research group MEMA of the University of Louvain. Our specific contribution in this code consists in new physical modules with supplementary conservation laws, interface conditions and boundary conditions. Then, the continuous finite element simulations has been done using the open-source GetDP software [54]. Finally, the finite difference and finite volume schemes have been implemented in original Matlab codes, except for the nonlinear oceanographic simulation of section 5.4, which has been implemented in the Fortan code developped by C. Mercier in the course of his PhD thesis [120].

Some parts of this work have been published in peer-reviewed journals and conference proceedings:

A. Modave, E. Deleersnijder, E. Delhez, On the parameters of absorbing layers for shallow water models, *Ocean Dynamics* (2010), 60(1), 65-79. 15 pages.

A. Modave, E. Delhez, C. Geuzaine, On the parameters of the PML in discrete contexts, *Proceedings of Waves 2011 (Canada)*. 4 pages.

A. Modave, A. Kameni, J. Lambrechts, E. Delhez, L. Pichon, C. Geuzaine, An optimum PML for scattering problems in the time domain, *The European Physical Journal - Applied Physics*. 6 pages. (accepted for publication)

A. Modave, J. Lambrechts, E. Delhez, C. Geuzaine, A PML for Convex Truncated Domains in Time-Dependent Acoustics with a Discontinuous Galerkin Finite Element Discretization, *Proceedings of Waves 2013 (Tunisia)*. 2 pages.

This work has been presented at conferences: JONSMOD 2010 (The Netherlands), ICIAM 2011 (Canada), WAVES 2011 (Canada), Journées de Metz 2012 (France), CANUM 2012 (France), NUMELEC 2012 (France), EMF 2013 (Belgium) and WAVES 2013 (Tunisia).

CHAPTER 1

Some Fundamentals of Wave-Like Problems

1.1 Introduction

In this work, we investigate the numerical resolution of time-dependent problems defined with hyperbolic partial differential equations. These equations are informally interpreted as supporting wave-like phenomena. Important laws of physics are described by hyperbolic equations, such as the Euler equations, the Maxwell equations and the shallow water equations. Since the properties of these equations are similar, common numerical methods have been developed and used for the various application contexts involving propagation of waves.

The aim of this chapter is to introduce the mathematical framework and the numerical methods used in this work. In section 1.2, the equations of wave-like problems are defined: the hyperbolic system (general case) and the scalar wave system (fundamental case). For both cases, mathematical tools highlighting some fundamental properties of the equations are presented: the plane-wave solution and the characteristic analysis. In section 1.3, the numerical methods used in this thesis are presented: the *finite difference method* (FD method), the *finite volume method* (FV method) and both *continuous* and *discontinuous finite element methods* (FE methods). The numerical schemes are given in the context of the scalar wave system, which is enough to sketch the key ideas of each method.

The first part of this chapter is based on the reference books of Evans [61], Hirsch [87], Toro [163] and Whitham [166]. For the numerical methods we also refer to LeVeque [109] and Hesthaven and Warburton [86].

1.2 Definition and properties of wave-like systems

In this section, two systems of partial differential equations (PDEs) describing wave-like phenomena are presented and studied: the general hyperbolic system and the scalar wave system. Their main properties are highlighted using the plane-wave analysis and the characteristic analysis.

1.2.1 Hyperbolic system

In the most general case, let us consider the system of m first-order PDEs with the quasi-linear form

$$\frac{\partial \boldsymbol{v}}{\partial t} + \sum_{i=1}^{d} \mathbf{A}_i \frac{\partial \boldsymbol{v}}{\partial x_i} = \boldsymbol{s}, \quad \forall (\boldsymbol{x}, t) \in \Omega \times \mathbb{R}^+, \tag{1.1}$$

where the unknown $\boldsymbol{v}(\boldsymbol{x}, t)$ is a vector field with m real components. It is defined in a bounded spatial domain $\Omega \subset \mathbb{R}^d$ with the spatial dimension $d = 1$, 2 or 3. The vector $\boldsymbol{x} = (x_1, \ldots, x_d)^T$ denotes the spatial position and t is the time. The $m \times m$ matrices $\mathbf{A}_i$ and the m-component vector $\boldsymbol{s}$ are real. They may depend on $\boldsymbol{v}$, $\boldsymbol{x}$ and t, but not on the derivatives of $\boldsymbol{v}$. The hyperbolicity condition then reads [61, 87, 163, 166]:

> The system (1.1) is *hyperbolic* if, for every vector $\boldsymbol{n} = (n_1, \ldots, n_d)^T \in \mathbb{R}^d$, the eigenvalues of the $m \times m$ matrix $\mathbf{B} = \sum_{i=1}^{d} n_i \mathbf{A}_i$ are real and its eigenvectors form a basis of $\mathbb{R}^m$, $\forall (\boldsymbol{x}, t) \in \Omega \times \mathbb{R}^+$.

As a consequence, the matrix $\mathbf{B}$ is diagonalizable.

1.2.2 Scalar wave system

One fundamental PDE is the *scalar wave equation*,

$$\frac{\partial^2 p}{\partial t^2} - a \nabla \cdot (b \nabla p) = 0, \quad \forall (\boldsymbol{x}, t) \in \Omega \times \mathbb{R}^+, \tag{1.2}$$

where the unknown $p(\boldsymbol{x}, t)$ is a real scalar field and both a and b are constant nonzero real parameters. Introducing a real vector field $\boldsymbol{u}(\boldsymbol{x}, t)$, this second-order equation can be rewritten as the first-order *scalar wave system*

$$\begin{cases} \dfrac{\partial p}{\partial t} + a \nabla \cdot \boldsymbol{u} = 0, \quad \forall (\boldsymbol{x}, t) \in \Omega \times \mathbb{R}^+, \\ \dfrac{\partial \boldsymbol{u}}{\partial t} + b \nabla p = 0, \quad \forall (\boldsymbol{x}, t) \in \Omega \times \mathbb{R}^+, \end{cases} \tag{1.3}$$

which is a hyperbolic system. Both equation (1.2) and equivalent system (1.3) are used for modeling various physical waves, such as acoustic waves, linear gravity waves or transverse electric/magnetic waves.

1.2.3 Plane-wave analysis

The elementary solution in the form of a complex harmonic plane wave

$$e^{\imath(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}$$

is frequently used to study wave-like problems [74, 85, 103]. The real vector $\boldsymbol{k}$ is the *wave vector* and the real scalar ω is the *angular frequency*, assumed to be positive. The wave vector indicates the direction of propagation of the plane wave. This solution is an ideal representation of many physical phenomena. The Fourier theory allows arbitrary wave forms to be constructed from harmonic plane waves.

Plane-wave solution of the general hyperbolic system

Hyperbolic systems are informally interpreted as systems supporting wave-like solutions [61, 87]. It can be illustrated considering the harmonic plane-wave solution of the system (1.1),

$$\boldsymbol{v}(\boldsymbol{x},t) = \boldsymbol{V}e^{\imath(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}, \tag{1.4}$$

where $\boldsymbol{V}$ is the amplitude, assumed to be constant. The homogeneous part of the system (1.1) admits solutions of form (1.4) if the system

$$\left(\sum_{i=1}^{d} k_i\mathbf{A}_i - \omega\mathbf{I}\right)\boldsymbol{V} = 0$$

admits non-trivial solutions. This equality asserts that $\boldsymbol{V}$ is an eigenvector of the matrix $\sum_{i=1}^{d} k_i\mathbf{A}_i$ corresponding to the eigenvalue ω. Therefore, the hyperbolicity condition means that there are m distinct plane-wave solutions of (1.1) in each direction $\boldsymbol{n}$. The propagation velocities of these plane waves is $\lambda_i / \|\boldsymbol{n}\|$, where λ_i is the i^{th} eigenvalue of $\mathbf{B}$, corresponding to the i^{th} mode of propagation.

Plane-wave solution of the scalar wave system

For the scalar wave system (1.3), the elementary harmonic plane-wave solution reads

$$\begin{pmatrix} p(\boldsymbol{x},t) \\ \boldsymbol{u}(\boldsymbol{x},t) \end{pmatrix} = \begin{pmatrix} P \\ \boldsymbol{U} \end{pmatrix} e^{\imath(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}, \tag{1.5}$$

where P and $\boldsymbol{U}$ are the *amplitudes*. This plane wave is a solution of the system (1.3) if two important relations are satisfied: the *dispersion relation*

$$\omega = \sqrt{ab}\,\|\boldsymbol{k}\| \tag{1.6}$$

and the *amplitude relation*

$$P\frac{\boldsymbol{k}}{\|\boldsymbol{k}\|} = \sqrt{\frac{a}{b}}\,\boldsymbol{U}, \tag{1.7}$$

where $\|\boldsymbol{k}\| = \sqrt{\boldsymbol{k}\cdot\boldsymbol{k}}$ is the norm of the vector $\boldsymbol{k}$. Indeed, injecting the solution (1.5) in the system (1.3) gives the system

$$\begin{cases} -\imath\omega P + a\imath\boldsymbol{k}\cdot\boldsymbol{U} = 0, \\ -\imath\omega\boldsymbol{U} + b\imath\boldsymbol{k}P = 0, \end{cases}$$

which is true if and only if the relations (1.6) and (1.7) are satisfied.

Two important quantities are obtained from the relations (1.6) and (1.7): the *phase velocity* c, which is the propagation velocity of the plane waves in the direction $\boldsymbol{k}$,

$$c \overset{\text{def.}}{=} \frac{\omega}{\|\boldsymbol{k}\|} = \sqrt{ab},$$

and the *impedance* Z, which is the ratio of the amplitudes of fields,

$$Z \overset{\text{def.}}{=} \frac{|P|}{\|\boldsymbol{U}\|} = \sqrt{\frac{a}{b}}.$$

They characterise the wave propagation in the medium.

1.2.4 Characteristic analysis

The characteristic analysis provides an efficient way to highlight properties of quasi-linear first-order PDEs. In the most fundamental cases, the solution of initial boundary value problems can be obtained with this method (see *e.g.* [87, 163]).

Characteristic analysis of the general hyperbolic system

In order to introduce the concepts, let us consider the hyperbolic system (1.1) in a one-dimensional homogeneous case with constant coefficients. The system then reduces to

$$\frac{\partial \boldsymbol{v}}{\partial t} + \mathbf{A}\frac{\partial \boldsymbol{v}}{\partial x} = 0, \tag{1.8}$$

Since this system is hyperbolic, the constant matrix $\mathbf{A}$ is diagonalizable. Define the eigenvalue matrix $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_m)$ and the eigenvector matrix $\mathbf{T} = [\boldsymbol{t}_1 \cdots \boldsymbol{t}_m]$, which are connected by the relation $\mathbf{\Lambda} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$. The system (1.8) can then be rewritten

$$\frac{\partial \boldsymbol{w}}{\partial t} + \mathbf{\Lambda}\frac{\partial \boldsymbol{w}}{\partial x} = 0, \tag{1.9}$$

where $\boldsymbol{w} = \mathbf{T}^{-1}\boldsymbol{v} = (w_1, \ldots, w_m)^T$ is a new set of unknowns called *characteristic variables*. The equations of this system are the *characteristic equations*. Since the matrix $\mathbf{\Lambda}$ is diagonal, the system is decoupled and each equation can be interpreted separately.

Each equation of the system (1.9) is a *transport equation*. Indeed, the j^{th} equation

$$\frac{\partial w_j}{\partial t} + \lambda_j \frac{\partial w_j}{\partial x} = 0 \tag{1.10}$$

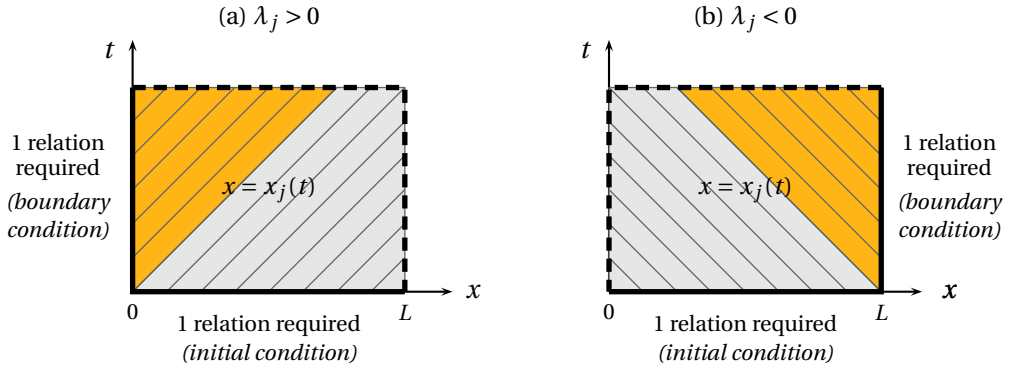


Figure 1.1: Characteristic lines in the $x-t$ plane for the one-dimensional problem with $\lambda_j > 0$ (a) and $\lambda_j < 0$ (b). For both cases, the gray zone indicates the area influenced by the initial value of the characteristic variable w_j, *i.e.* the value prescribed at $t = 0$ on the segment $[0, L]$, while the orange zone is influenced by the value prescribed at the boundary.

describes the transport of the quantity w_j at the velocity λ_j, for $j = 1, \ldots, m$. A useful interpretation is obtained considering the set of *characteristic lines* $x = x_j(t)$ in the $x-t$ plane defined by

$$\frac{\mathrm{d}x_j}{\mathrm{d}t} = \lambda_j. \tag{1.11}$$

The rate of change of w_j along the line $x = x_j(t)$ is

$$\left.\frac{\mathrm{d}w_j}{\mathrm{d}t}\right|_{x=x_j(t)} = \frac{\partial w_j}{\partial t} + \frac{\mathrm{d}x_j}{\mathrm{d}t}\frac{\partial w_j}{\partial x}.$$

Considering the equations (1.10) and (1.11), this rate is equal to zero. Therefore, along a characteristic line $x = x_j(t)$, the corresponding characteristic variable w_j stays constant. It is called *Riemann invariant*.

Considering now the complete system (1.9), the interpretation above means that, as time increases, the value of the m characteristic variables are preserved along m sets of characteristic lines. Therefore, information (the value of the variables w_i) is transported (or 'propagated') in the spatial domain at velocities given by the eigenvalues λ_j. This interpretation exactly corresponds to a wave propagation phenomenon.

In the context of initial-boundary value problems, the characteristic analysis provides the number of relations to prescribe as initial or boundary conditions. This can be shown for the system (1.8) with $x \in [0, L]$, by considering the zones of influence of the characteristic variables, illustrated in Figure 1.1. Since the value of a characteristic variable $w_j(x,t)$ is preserved along the corresponding characteristic lines, $w_j(x,t)$ is entirely defined by its initial value in the gray zone and by a

boundary value in the yellow zone. Therefore, for both cases shown in Figure 1.1, one relation is required on each bold segment, where characteristic lines start. On the contrary, no relation must be prescribed on the dashed segment, where w_j is already determined by the values propagated along characteristic lines. For the complete problem (1.8), the number of initial conditions then corresponds to the number of fields, while the number of boundary conditions depends on the number of sets of characteristic lines that start at the considered boundary.

The characteristic analysis can be performed considering a more general non-constant matrix $\mathbf{A}(\boldsymbol{v},\boldsymbol{x},t)$ and a source term $\boldsymbol{s}(\boldsymbol{v},\boldsymbol{x},t)$ in the one-dimensional system (1.8). In this case, the characteristic lines become curves satisfying

$$\frac{\mathrm{d}x_j}{\mathrm{d}t} = \lambda_j(\boldsymbol{v},\boldsymbol{x},t)$$

and the characteristic variables w_j can vary along thereof. Nonetheless, the global idea of the analysis remains the same, especially concerning the number of boundary conditions [87, 166].

Unfortunately, the extension of this analysis to the multidimensional case is not straightforward (see *e.g.* [87, 166]). Indeed, the general system (1.1) cannot be diagonalized, except if the matrices $\mathbf{A}_i$ are simultaneously diagonalizable, which happens in only very specific cases. Nevertheless, useful information is obtained assuming $\boldsymbol{v}$ varies in only one direction $\boldsymbol{n}$. Depending on the interpretation, this direction can be the propagation direction of a plane wave or the normal direction to a boundary. The general system (1.1) then reduces to

$$\frac{\partial \boldsymbol{v}}{\partial t} + \mathbf{B}\frac{\partial \boldsymbol{v}}{\partial x_n} = \boldsymbol{s},$$

where $x_n = \boldsymbol{x}\cdot\boldsymbol{n}$ and $\mathbf{B} = \sum_{i=1}^{d} n_i\mathbf{A}_i$. Then, a characteristic analysis of this system indicates the information that is propagated upstream and downstream the direction $\boldsymbol{n}$ and the number of boundary conditions to prescribe for initial-boundary value problems.

Characteristic analysis of the scalar wave system

Following the procedure presented for the general hyperbolic system, let us assume that the fields vary only along a direction $\boldsymbol{n}$. Therefore, the scalar wave system (1.3) reduces to the one-dimensional system

$$\begin{cases} \dfrac{\partial p}{\partial t} + a\dfrac{\partial u_n}{\partial x_n} = 0, \\ \dfrac{\partial u_n}{\partial t} + b\dfrac{\partial p}{\partial x_n} = 0, \end{cases} \tag{1.12}$$

where u_n and x_n are the projections of the vectors $\boldsymbol{u}$ and $\boldsymbol{x}$ in the direction $\boldsymbol{n}$. Some algebraic manipulations allow to rewrite this system with the decoupled charac-

teristic equations

$$\begin{cases} \dfrac{\partial w_1}{\partial t} + \sqrt{ab}\,\dfrac{\partial w_1}{\partial x_n} = 0, \\ \dfrac{\partial w_2}{\partial t} - \sqrt{ab}\,\dfrac{\partial w_2}{\partial x_n} = 0, \end{cases}$$

where the characteristic variables

$$w_1 = p + \sqrt{\frac{a}{b}}\,u_n, \tag{1.13}$$

$$w_2 = p - \sqrt{\frac{a}{b}}\,u_n, \tag{1.14}$$

are the Riemann invariants. These equations are transport equations. Interpreting this system, the information contained in the characteristic variables w_1 and w_2 is transported in the spatial domain respectively to the directions $\boldsymbol{n}$ (downstream) and $-\boldsymbol{n}$ (upstream) at the velocity $\sqrt{ab}$.

In the case of an initial-boundary value problem, one boundary condition must be prescribed at each boundary, whatever the spatial dimension. Indeed, considering $\boldsymbol{n}$ is the outward normal direction to the boundary, one characteristic variable (w_1) is propagated along characteristic lines from the boundary (one relation required), while the other (w_2) is propagated towards the boundary (no relation required).

1.3 Numerical resolution

The goal of numerical methods is to provide approximate solutions, called *numerical solutions*, for problems that cannot (or can only hardly) be solved analytically. These methods involve two steps: the choice of a *discrete space*, that determines the shape of the numerical solution, and the design of a *numerical scheme* that governs this solution (instead of continuous laws).

In this section, the numerical methods used in this thesis are presented in the context of the wave system (1.3), which is sufficient to sketch the key ideas of each method. Explanations are focused on the numerical resolution in space. The numerical resolution in time (or *time-stepping*) is not a critical issue in this work and is performed using classical schemes. Therefore, we refer to the reference books [87, 155] for further details.

1.3.1 Finite difference method

The *finite difference method* (FD method) is the oldest and simplest numerical method used for solving problem with PDEs. It was already known by Leonhard Euler in 1768. Today, it is widely used in both research and industry [*e.g.* 44, 56, 77, 87, 153, 155].

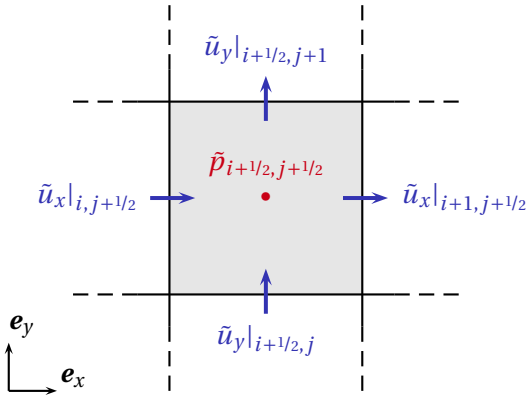


Figure 1.2: FD method. Discretization of the fields on the mesh.

The idea of the FD method is quite simple. The continuous fields are directly approximated by discrete ones. Each discrete field has then values at different points of a regular grid. They are governed by discrete equations that are built by replacing the partial derivatives of the PDEs with finite differences.

The FD schemes used in this thesis are built considering a spatial discretization of the fields on staggered grids. For the two-dimensional problem, the spatial domain is partitioned into regular non-overlapping cells of size $\Delta x \times \Delta y$. As illustrated in Figure 1.2, the scalar field p is approximated by the spatially discrete field $\tilde{p}$, that is defined at the center of each cell. The Cartesian components of the vector field $\boldsymbol{u} = \left(u_x, u_y\right)^T$ are approximated by $\tilde{u}_x$ and $\tilde{u}_y$, that have values at the center of the interfaces. This can be written

$$\tilde{p}_{i+1/2,j+1/2}(t) \approx p(\boldsymbol{x}_{i+1/2,j+1/2}, t), \tag{1.15}$$

$$\tilde{u}_x|_{i,j+1/2}(t) \approx u_x(\boldsymbol{x}_{i,j+1/2}, t), \tag{1.16}$$

$$\tilde{u}_y|_{i+1/2,j}(t) \approx u_y(\boldsymbol{x}_{i+1/2,j}, t), \tag{1.17}$$

where $\boldsymbol{x}_{i,j} = (i\Delta x, j\Delta y)$ is the position of the discrete point (i, j), and i and j are the indices corresponding to the spatial directions. Such staggered grids have been introduced by Yee [168] in the electromagnetic context and by Arakawa and Lamb [10] in the oceanographic context. They are now considered classical [44, 155].

The equations governing the semi-discrete fields (1.15)-(1.17) are simply obtained by remplacing the spatial partial derivatives of the PDEs (1.3) with central

finite differences of fields. The numerical scheme is then written

$$\begin{cases} \dfrac{\mathrm{d}\tilde{p}_{i+1/2,j+1/2}}{\mathrm{d}t} + a\left(\dfrac{\tilde{u}_x|_{i+1,j+1/2} - \tilde{u}_x|_{i,j+1/2}}{\Delta x} + \dfrac{\tilde{u}_y|_{i+1/2,j+1} - \tilde{u}_y|_{i+1/2,j}}{\Delta y}\right) = 0, \\ \dfrac{\mathrm{d}\tilde{u}_x|_{i,j+1/2}}{\mathrm{d}t} + b\dfrac{\tilde{p}_{i+1/2,j+1/2} - \tilde{p}_{i-1/2,j+1/2}}{\Delta x} = 0, \\ \dfrac{\mathrm{d}\tilde{u}_y|_{i+1/2,j}}{\mathrm{d}t} + b\dfrac{\tilde{p}_{i+1/2,j+1/2} - \tilde{p}_{i+1/2,j-1/2}}{\Delta y} = 0. \end{cases}$$

At the border of the computational domain, boundary conditions are easily prescribed. The value of the discrete field (for a Dirichlet condition) or the finite difference (for a Neumann condition) is simply replaced at the border.

This spatial scheme is dispersive and non-dissipative (see section A.1 and [87, 155]). The final scheme conserves the total energy if this spatial scheme is used together with a non-dissipative time-stepping scheme, such as the Crank-Nicolson scheme or the Leapfrog scheme with staggered grids in time.

1.3.2 Finite volume method

The *finite volume method* (FV method) is based on the discretization of the integral form of PDEs instead of the direct discretization of the differential form. This method presents some major advantages: unstructured meshes (and thus complex geometries) and media with discontinuous properties are easily considered. In addition, since integral forms are used for expressing conservation properties, FV numerical schemes that preserve these properties can be naturally built [87, 109].

The fields are discretized using a partition $\mathscr{M}$ of the spatial domain Ω into non-overlapping polyhedral cells. In this work, each unknown field is approximated by its average value over each cell — this strategy is sometimes called *cell-centered approach* —, *i.e.*

$$\begin{aligned} \tilde{p}(\boldsymbol{x},t) &= \tilde{p}_e(t), \quad \text{for } \boldsymbol{x} \in \Omega_e, \quad \forall \Omega_e \in \mathscr{M}, \\ \tilde{\boldsymbol{u}}(\boldsymbol{x},t) &= \tilde{\boldsymbol{u}}_e(t), \quad \text{for } \boldsymbol{x} \in \Omega_e, \quad \forall \Omega_e \in \mathscr{M}, \end{aligned}$$

with

$$\begin{aligned} \tilde{p}_e(t) &\approx \frac{1}{V_e}\int_{\Omega_e} p(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}, \\ \tilde{\boldsymbol{u}}_e(t) &\approx \frac{1}{V_e}\int_{\Omega_e} \boldsymbol{u}(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}, \end{aligned}$$

where Ω_e denotes the e^{th} cell and V_e is its area (in two dimensions) or its volume (in three dimensions). Therefore, the numerical solution is piecewise-constant in the domain.

A numerical FV scheme for the system (1.3) is obtained by discretizing its integral form instead of its differential form for each cell. The integral form corresponding to the cell Ω_e reads

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\displaystyle\int_{\Omega_e} p \,\mathrm{d}\Omega_e + \int_{\Gamma_e} (a\boldsymbol{u})\cdot\boldsymbol{n}_e \,\mathrm{d}\Gamma_e = 0, \\ \dfrac{\mathrm{d}}{\mathrm{d}t}\displaystyle\int_{\Omega_e} \boldsymbol{u} \,\mathrm{d}\Omega_e + \int_{\Gamma_e} (bp)\boldsymbol{n}_e \,\mathrm{d}\Gamma_e = 0, \end{cases} \tag{1.18}$$

where Γ_e is the boundary of the e^{th} cell and $\boldsymbol{n}_e$ is its outward unit normal on Γ_e. For both equations, the approximation of the volume integrals is trivial since the numerical solution is constant in each cell. However, it is more complicated for the surface integrals: the solution is discontinuous at interfaces between cells and boundary conditions must be taken into account at the border of the domain Ω. Therefore, specific *numerical fluxes* must be carefully defined in order to correctly evaluate these surface integrals. Finally, the numerical scheme reads

$$\begin{cases} \dfrac{\mathrm{d}\tilde{p}_e}{\mathrm{d}t} + \dfrac{1}{V_e}\displaystyle\sum_{i=1}^{N_e^{\text{faces}}} f_{e,i}^{\text{num}}\, S_{e,i} = 0, \\ \dfrac{\mathrm{d}\tilde{\boldsymbol{u}}_e}{\mathrm{d}t} + \dfrac{1}{V_e}\displaystyle\sum_{i=1}^{N_e^{\text{faces}}} \boldsymbol{f}_{e,i}^{\text{num}}\, S_{e,i} = 0, \end{cases} \tag{1.19}$$

where N_e^{faces} is the number of faces of the cell Ω_e, $f_{e,i}^{\text{num}}$ and $\boldsymbol{f}_{e,i}^{\text{num}}$ are the numerical fluxes across the i^{th} face and $S_{e,i}$ is the area of this face.

Numerical fluxes at the interface between two cells

At the interface between two cells, a first set of numerical fluxes, called *centered fluxes*, are obtained by simply considering the mean value of the fields, *i.e.*

$$f_{e,i}^{\text{num}} = a\{\tilde{\boldsymbol{u}}\}\cdot\boldsymbol{n}_e, \tag{1.20}$$

$$\boldsymbol{f}_{e,i}^{\text{num}} = b\{\tilde{p}\}\boldsymbol{n}_e, \tag{1.21}$$

where $\{\aleph\}$ denotes the mean value of the quantity $\aleph$ at the i^{th} face of Ω_e, *i.e.*

$$\{\aleph\} = \frac{\aleph^+ + \aleph^-}{2}.$$

The superscripts $^+$ and $^-$ correspond to the value of $\aleph$ in the adjacent cell and the current cell respectively, as illustrated in Figure 1.3 for the two-dimensional case. The resulting numerical scheme is non-dissipative [163].

As an alternative, the so-called *upwind fluxes* are obtained by using a one-dimensional Riemann solver. At the interface, consider the one-dimensional system (1.12), where $\boldsymbol{n}$ is here the outward unit normal $\boldsymbol{n}_e$. With the Riemann solver,

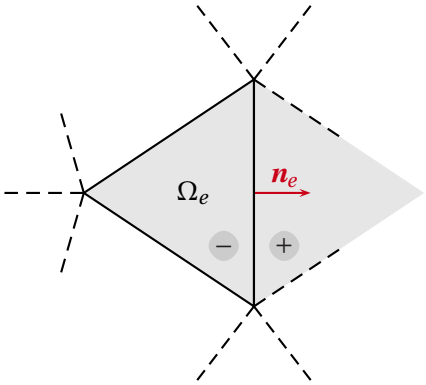


Figure 1.3: FV method. Notation used on both sides of an interface for discontinuous quantities. The sign '−' corresponds to the current cell (Ω_e), while '+' is for the adjacent cell.

an upwind scheme is used for each characteristic variable. The characteristic variable that is transported outward (w_1) takes its value in the current cell, while the other (w_2), which is transported inward, takes its value in the adjacent cell, *i.e.*

$$\begin{cases} w_1^\star = w_1^+, \\ w_2^\star = w_2^-. \end{cases}$$

Using the definition of the characteristic variables (1.13)–(1.14), this system is rewritten

$$\begin{cases} \tilde{p}^\star + \sqrt{\dfrac{a}{b}}\tilde{u}_n^\star = \tilde{p}^- + \sqrt{\dfrac{a}{b}}\tilde{u}_n^-, \\ \tilde{p}^\star - \sqrt{\dfrac{a}{b}}\tilde{u}_n^\star = \tilde{p}^+ - \sqrt{\dfrac{a}{b}}\tilde{u}_n^+, \end{cases}$$

where $\tilde{p}^\star$ and $\tilde{u}_n^\star$ are the values to use in the numerical fluxes. Solving this system gives

$$\tilde{p}^\star = \{\tilde{p}\} - \sqrt{\frac{a}{b}}[\![\tilde{u}_n]\!],$$
$$\tilde{u}_n^\star = \{\tilde{u}_n\} - \sqrt{\frac{b}{a}}[\![\tilde{p}]\!],$$

where $[\![\aleph]\!]$ denotes the jump value of the quantity $\aleph$ at the interface, *i.e.*

$$[\![\aleph]\!] = \frac{\aleph^+ - \aleph^-}{2}.$$

Finally, the numerical fluxes read

$$f_{e,i}^{\text{num}} = a\{\tilde{\boldsymbol{u}}\} \cdot \boldsymbol{n}_e - \sqrt{ab}[\![\tilde{p}]\!], \tag{1.22}$$
$$\boldsymbol{f}_{e,i}^{\text{num}} = b\{\tilde{p}\}\boldsymbol{n}_e - \sqrt{ab}\left([\![\tilde{\boldsymbol{u}}]\!] \cdot \boldsymbol{n}\right)\boldsymbol{n}_e. \tag{1.23}$$

In comparison with centered fluxes, these upwind fluxes are penalized by the jump of the fields in the second terms. These additional terms introduce dissipation in the numerical scheme [163].

When the coefficients a and b are piecewise constant, the integral form (1.18) remains valid, as well as the numerical scheme (1.19), but the numerical fluxes must be adapted to take into account the jump of these coefficients at the interface between cells. The numerical fluxes are then written

$$\begin{aligned} f_{e,i}^{\text{num}} &= a^- u_n^\star, \\ \boldsymbol{f}_{e,i}^{\text{num}} &= b^- p^\star \boldsymbol{n}_e, \end{aligned}$$

where the values $p^\star$ and $u_n^\star$ are given by an adapted Riemann solver [109],

$$\begin{aligned} \tilde{p}^\star &= \frac{\{Y\tilde{p}\}}{\{Y\}} - \frac{\alpha}{\{Y\}}[\![\tilde{u}_n]\!], \\ \tilde{u}_n^\star &= \frac{\{Z\tilde{u}_n\}}{\{Z\}} - \frac{\alpha}{\{Z\}}[\![\tilde{p}]\!], \end{aligned}$$

with $Z = \sqrt{a/b}$ and $Y = \sqrt{b/a}$. The decentering parameter α is equal to 1 for upwind fluxes and 0 for centered fluxes.

Numerical fluxes at the boundary of the domain

At the border of the domain, the boundary condition must be prescribed through the numerical fluxes. A strategy consists in extending the domain to include an additional *ghost cell* [109, 163]. Therefore, the numerical fluxes defined above are reused with the ghost cells as adjacent cells.

In this work, a Dirichlet condition on one field is prescribed by replacing the value of this field in the ghost cell. For the other field, the value of the current cell is reused in the ghost cell. For example, the Dirichlet condition $\boldsymbol{u}\cdot\boldsymbol{n} = g(t)$ is prescribed using

$$\begin{aligned} \tilde{p}^+ &= \tilde{p}^-, \\ \tilde{u}_n^+ &= g(t). \end{aligned}$$

The upwind fluxes (1.22)–(1.23) then become

$$\begin{aligned} f_{e,i}^{\text{num}} &= a\frac{g + (\tilde{\boldsymbol{u}}^- \cdot \boldsymbol{n}_e)}{2}, \\ \boldsymbol{f}_{e,i}^{\text{num}} &= b\tilde{p}^- \boldsymbol{n}_e - \sqrt{ab}\left(\frac{g - (\tilde{\boldsymbol{u}}^- \cdot \boldsymbol{n}_e)}{2}\right)\boldsymbol{n}_e. \end{aligned}$$

1.3.3 Finite element method

The *finite element method* (FE method) offers a general framework to build numerical schemes. This method is very appealing for solving PDEs on complex geometries, especially if curved boundaries must be considered. Indeed, unstructured meshes are naturally taken into account and curved cells can be used. Moreover, various techniques have been developed to improve the representation of the solution, such as high-order elements, mixed elements [67, 99, 128, 145], discontinuous elements [37, 86] and the extended finite element method (X-FEM) [124].

In this work, we consider a *continuous Galerkin method* (CG method) and a *discontinuous Galerkin method* (DG method). For both methods, the solution is approximated by polynomial functions over each cell (called *element*) of the mesh. The fundamental difference is the continuity of the numerical solution at the interfaces between elements, which is enforced with CG and not with DG. In the case of first-order elements, where the discrete unknowns of the discrete problems are the values of the fields at the nodes of the mesh, each field has one single value per node with CG, while it has different values per node (one for each neighboring element) with DG, as illustrated in Figure 1.4.

The DG method is generally considered as a clever combination of FV and CG methods: from the former, it naturally deals with discontinuous solutions; from the latter, high-order and curved elements can be used. However, since the discrete unknowns have multiple values at the interfaces between elements, the number of unknowns and the size of the algebraic system to solve are larger with the DG method than with the CG method. This disadvantage is offset considering the DG scheme has an ideal form to perform a parallel computation, which explains its success.

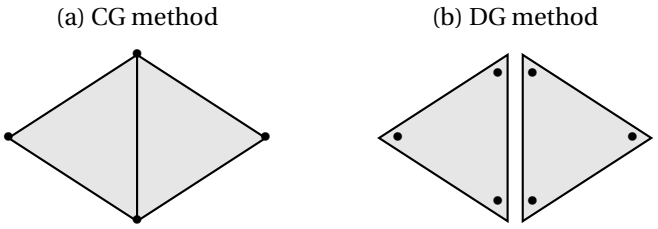


Figure 1.4: Illustration of the discrete unknowns in first-order elements for both CG and DG methods. At an interface between two elements, each field has one single value per node with the continuous method (a), and one value per node for each neighboring element with the discontinuous method (b).

Two finite element solvers have been used during the realisation of this thesis. The GetDP software [54] for the CG schemes, and a DG code developed in collaboration with the research group MEMA of the University of Louvain. All meshes are made using the Gmsh software [68, 69].

The CG and DG numerical schemes are detailed hereafter. For more clarity, inner products over the domain $\mathscr{D}$ and its boundary $\mathscr{B} = \partial\mathscr{D}$ are denoted

$$\left(f, g\right)_{\mathscr{D}} = \int_{\mathscr{D}} f\, g \, \mathrm{d}\mathscr{D},$$
$$\left\langle f, g\right\rangle_{\mathscr{B}} = \int_{\mathscr{B}} f\, g \, \mathrm{d}\mathscr{B},$$

where f and g are scalar fields. Similar notations are used for inner products of vectors.

1.3.3.1 Continuous Galerkin method

Let us consider a partition $\mathscr{M}$ of the domain Ω into non-overlapping cells. Over the e^{th} cell, denoted Ω_e, the *local numerical solution* is written as a truncated polynomial expansion

$$\tilde{p}_e(\boldsymbol{x}, t) = \sum_{i=1}^{N_e} \tilde{p}_{e,i}(t)\, \varphi_{e,i}(\boldsymbol{x}), \quad \text{for } \boldsymbol{x} \in \Omega_e,$$
$$\tilde{\boldsymbol{u}}_e(\boldsymbol{x}, t) = \sum_{i=1}^{N_e} \tilde{\boldsymbol{u}}_{e,i}(t)\, \varphi_{e,i}(\boldsymbol{x}), \quad \text{for } \boldsymbol{x} \in \Omega_e,$$

where N_e is the number of discrete unknowns associated to the cell, $\tilde{p}_{e,i}$ and $\tilde{\boldsymbol{u}}_{e,i}$ are the discrete unknowns, and the function $\varphi_{e,i}(\boldsymbol{x})$ is a *local basis function*. Hierarchical functions [53, 55, 67] are used as local basis functions. Since the solution is continuous throughout the domain, the *global numerical solution* can be written

$$\tilde{p}(\boldsymbol{x}, t) = \sum_{i=1}^{N} \tilde{p}_i(t)\, \psi_i(\boldsymbol{x}), \tag{1.24}$$
$$\tilde{\boldsymbol{u}}(\boldsymbol{x}, t) = \sum_{i=1}^{N} \tilde{\boldsymbol{u}}_i(t)\, \psi_i(\boldsymbol{x}), \tag{1.25}$$

where N is the total number of discrete unknowns associated to each field, i is now a global index, $\tilde{p}_i$ and $\tilde{\boldsymbol{u}}_i$ are the discrete unknowns with the global numerotation, and the function $\psi_i(\boldsymbol{x})$ is a *global basis function*. Each global basis function is obtained by combining the local basis functions associated to the same discrete unknown in different elements. Resulting functions are therefore continuous.

The numerical scheme is built from a weak form of the equations (1.3) over the domain Ω. This weak form is obtained by multiplying the equations by test

functions and integrating them over the domain Ω. Using integration by parts, one has

$$\begin{cases} \left(\dfrac{\partial p}{\partial t}, \hat{p}\right)_\Omega - \left(a\boldsymbol{u}, \nabla \hat{p}\right)_\Omega + \left\langle (a\boldsymbol{u})\cdot \boldsymbol{n}, \hat{p}\right\rangle_\Gamma = 0, \quad \forall \hat{p}, \\ \left(\dfrac{\partial \boldsymbol{u}}{\partial t}, \hat{\boldsymbol{u}}\right)_\Omega + \left(b\nabla p, \hat{\boldsymbol{u}}\right)_\Omega = 0, \quad \forall \hat{\boldsymbol{u}}, \end{cases} \tag{1.26}$$

where the vector $\boldsymbol{n}$ is the outward unit normal on the boundary Γ. The system (1.26) must stay valid for all *test functions* $\hat{p}$ and $\hat{\boldsymbol{u}}$ that are in the same spatial function space as the unknown fields p and $\boldsymbol{u}$, respectively. In the discrete problem, both the numerical solution (1.24)–(1.25) and the test functions must belong to the finite basis $\{\psi_i\}_{i=1}^N$, which is complete. Therefore, injecting the numerical solution in the weak form, and using the basis functions as test functions, one obtains successively, for $j = 1, \ldots, N$,

$$\begin{cases} \left(\dfrac{\partial \tilde{p}}{\partial t}, \psi_j\right)_\Omega - \left(a\tilde{\boldsymbol{u}}, \nabla \psi_j\right)_\Omega + \left\langle (a\tilde{\boldsymbol{u}})\cdot \boldsymbol{n}, \psi_j\right\rangle_\Gamma = 0, \\ \left(\dfrac{\partial \tilde{\boldsymbol{u}}}{\partial t}, \psi_j\right)_\Omega + \left(b\nabla \tilde{p}, \psi_j\right)_\Omega = 0, \end{cases}$$

and

$$\begin{cases} \displaystyle\sum_{i=1}^N \left[\left(\psi_i, \psi_j\right)_\Omega \frac{\mathrm{d}\tilde{p}_i}{\mathrm{d}t} - \left(a\psi_i, \nabla\psi_j\right)_\Omega \cdot \tilde{\boldsymbol{u}}_i + \left\langle (a\psi_i)\boldsymbol{n}, \psi_j\right\rangle_\Gamma \cdot \tilde{\boldsymbol{u}}_i \right] = 0, \\ \displaystyle\sum_{i=1}^N \left[\left(\psi_i, \psi_j\right)_\Omega \frac{\mathrm{d}\tilde{\boldsymbol{u}}_i}{\mathrm{d}t} + \left(b\nabla\psi_i, \psi_j\right)_\Omega \tilde{p}_i \right] = 0. \end{cases} \tag{1.27}$$

This system can be written in the matrix form

$$\mathbf{M}\frac{\mathrm{d}\tilde{\boldsymbol{V}}}{\mathrm{d}t} + \mathbf{K}\tilde{\boldsymbol{V}} = 0, \tag{1.28}$$

where $\mathbf{M}$ is the *global mass matrix*, and $\mathbf{K}$ is the *global stiffness matrix*. The vector $\tilde{\boldsymbol{V}}$ contain all the discrete unknowns of the problem. The elements of matrices are computed using numerical integration.

A Dirichlet boundary condition on a field is strongly incorporated in the formulation by fixing the value of the corresponding discrete field at the nodes of the boundary Γ. A Neumann boundary condition on $\boldsymbol{u}$ is weakly taken into account by modifying the boundary term in the system (1.27).

Stabilization with the PSPG method

Unfortunately, spurious numerical oscillations are generated by the formulation (1.27) [49]. Stabilization techniques have been developed to avoid them (see *e.g.* [49, 81, 111]).

In this work, we consider the *Pressure-Stabilization Petrov-Galerkin* method (PSPG method) proposed by Hughes, Franca and Balestra [94]. It consists in adding

a stabilization term in the left-hand side of the first equation of the weak form (1.26),

$$-\left(\kappa \boldsymbol{R}^{\boldsymbol{u}}, \nabla \hat{p}\right)_{\Omega},$$

where κ is the numerical stabilization parameter and $\boldsymbol{R}^{\boldsymbol{u}}$ is the residue associated to the second equation of the differential form (1.3), *i.e.*

$$\boldsymbol{R}^{\boldsymbol{u}} = \frac{\partial \boldsymbol{u}}{\partial t} + b\nabla p.$$

1.3.3.2 Discontinuous Galerkin method

With the DG method, the procedure for building the numerical solution and the numerical scheme is similar to the one employed above with the CG method. The fundamental difference is that the global basis functions are here discontinuous. An additional step is then introduced in the procedure in order to account for the discontinuities of the numerical solution.

Over the e^{th} cell of the mesh $\mathscr{M}$, the *local numerical solution* reads, with a truncated polynomial expansion,

$$\tilde{p}_e(\boldsymbol{x}, t) = \sum_{i=1}^{N_e} \tilde{p}_{e,i}(t)\, \phi_{e,i}(\boldsymbol{x}), \quad \text{for } \boldsymbol{x} \in \Omega_e,$$
$$\tilde{\boldsymbol{u}}_e(\boldsymbol{x}, t) = \sum_{i=1}^{N_e} \tilde{\boldsymbol{u}}_{e,i}(t)\, \phi_{e,i}(\boldsymbol{x}), \quad \text{for } \boldsymbol{x} \in \Omega_e,$$

where N_e is the number of discrete unknowns associated to each field on the e^{th} cell, $\tilde{p}_{e,i}$ and $\tilde{\boldsymbol{u}}_{e,i}$ are the discrete unknowns, and the function $\phi_{e,i}(\boldsymbol{x})$ is a *local basis function*. Here, the local basis functions are the multivariate Lagrange interpolation polynomials associated to points $\{\boldsymbol{x}_{e,i}\}_{i=1}^{N_e}$ (called *nodes*) of the cell (see *e.g.* [86]). They are such that

$$\phi_{e,i}(\boldsymbol{x}_{e,j}) = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j, \end{cases} \qquad \text{for } i, j = 1, \ldots, N_e.$$

Therefore, N_e represents the number of nodes associated to the e^{th} element, and $\tilde{p}_{e,i}$ and $\tilde{\boldsymbol{u}}_{e,i}$ are the values of numerical fields at the node $\boldsymbol{x}_{e,i}$. The *global numerical solution* writtes

$$\tilde{p}(\boldsymbol{x}, t) = \tilde{p}_e(\boldsymbol{x}, t), \quad \text{for } \boldsymbol{x} \in \Omega_e, \quad \forall \Omega_e \in \mathscr{M},$$
$$\tilde{\boldsymbol{u}}(\boldsymbol{x}, t) = \tilde{\boldsymbol{u}}_e(\boldsymbol{x}, t), \quad \text{for } \boldsymbol{x} \in \Omega_e, \quad \forall \Omega_e \in \mathscr{M}.$$

This solution is discontinuous at the interfaces between elements.

The numerical scheme is now built from a weak form of the equations (1.3) over each element Ω_e,

$$\begin{cases} \left(\dfrac{\partial p}{\partial t}, \hat{p}\right)_{\Omega_e} - \left(a\boldsymbol{u}, \nabla \hat{p}\right)_{\Omega_e} + \left\langle (a\boldsymbol{u}) \cdot \boldsymbol{n}_e, \hat{p} \right\rangle_{\Gamma_e} = 0, \quad \forall \hat{p}, \\ \left(\dfrac{\partial \boldsymbol{u}}{\partial t}, \hat{\boldsymbol{u}}\right)_{\Omega_e} - \left(bp, \nabla \cdot \hat{\boldsymbol{u}}\right)_{\Omega_e} + \left\langle (bp)\boldsymbol{n}_e, \hat{\boldsymbol{u}} \right\rangle_{\Gamma_e} = 0, \quad \forall \hat{\boldsymbol{u}}, \end{cases} \tag{1.29}$$

where the vector $\boldsymbol{n}_e$ is the outward unit normal on the boundary Γ_e. Again, the system (1.29) must stay valid for all *test functions* $\hat{p}$ and $\hat{\boldsymbol{u}}$ that are in the same spatial function space as the unknown fields p and $\boldsymbol{u}$, respectively. Injecting the numerical solution in the weak form (1.29), and using the local basis functions as test functions, one obtains successively, for $j = 1, \ldots, N_e$,

$$\begin{cases} \left(\dfrac{\partial \tilde{p}_e}{\partial t}, \phi_{e,j}\right)_{\Omega_e} - \left(a\tilde{\boldsymbol{u}}_e, \nabla \phi_{e,j}\right)_{\Omega_e} + \left\langle f_e^{\text{num}}, \phi_{e,j} \right\rangle_{\Gamma_e} = 0, \\ \left(\dfrac{\partial \tilde{\boldsymbol{u}}_e}{\partial t}, \phi_{e,j}\right)_{\Omega_e} - \left(b\tilde{p}_e, \nabla \phi_{e,j}\right)_{\Omega_e} + \left\langle \boldsymbol{f}_e^{\text{num}}, \phi_{e,j} \right\rangle_{\Gamma_e} = 0, \end{cases}$$

and

$$\begin{cases} \displaystyle\sum_{i=1}^{N_e} \left[\left(\phi_{e,i}, \phi_{e,j}\right)_{\Omega_e} \frac{\mathrm{d}\tilde{p}_{e,i}}{\mathrm{d}t} - \left(a\phi_{e,i}, \nabla \phi_{e,j}\right)_{\Omega_e} \cdot \tilde{\boldsymbol{u}}_{e,i} \right] + \left\langle f_e^{\text{num}}, \phi_{e,j} \right\rangle_{\Gamma_e} = 0, \\ \displaystyle\sum_{i=1}^{N_e} \left[\left(\phi_{e,i}, \phi_{e,j}\right)_{\Omega_e} \frac{\mathrm{d}\tilde{\boldsymbol{u}}_{e,i}}{\mathrm{d}t} - \left(b\phi_{e,i}, \nabla \phi_{e,j}\right)_{\Omega_e} \cdot \tilde{p}_{e,i} \right] + \left\langle \boldsymbol{f}_e^{\text{num}}, \phi_{e,j} \right\rangle_{\Gamma_e} = 0. \end{cases} \tag{1.30}$$

where f_e^{num} and $\boldsymbol{f}_e^{\text{num}}$ are numerical fluxes. The system (1.30) can be written in the matrical form

$$\begin{cases} \mathbf{M}_e^p \dfrac{\mathrm{d}\tilde{\boldsymbol{P}}_e}{\mathrm{d}t} + \mathbf{K}_e^p \tilde{\boldsymbol{U}}_e + \displaystyle\sum_{k}^{N_e^{\text{faces}}} \boldsymbol{F}_{e,k}^p = 0 \\ \mathbf{M}_e^{\boldsymbol{u}} \dfrac{\mathrm{d}\tilde{\boldsymbol{U}}_e}{\mathrm{d}t} + \mathbf{K}_e^{\boldsymbol{u}} \tilde{\boldsymbol{P}}_e + \displaystyle\sum_{k}^{N_e^{\text{faces}}} \boldsymbol{F}_{e,k}^{\boldsymbol{u}} = 0 \end{cases} \tag{1.31}$$

where the vectors $\tilde{\boldsymbol{P}}_e$ and $\tilde{\boldsymbol{U}}_e$ contain the discrete unknowns associated to the e^{th} element, $\mathbf{M}_e^p$ and $\mathbf{M}_e^{\boldsymbol{u}}$ are *local mass matrices*, $\mathbf{K}_e^p$ and $\mathbf{K}_e^{\boldsymbol{u}}$ are *local stiffness matrices*, $\boldsymbol{F}_{e,k}^p$ and $\boldsymbol{F}_{e,k}^{\boldsymbol{u}}$ are *flux vectors*, and N_e^{faces} is the number of faces of the e^{th} element.

As for the FV scheme, the way of evaluating the fluxes vectors is a key element of the scheme. The strategy employed in section 1.3.2 for the FV scheme is reused, at both the boundary of the domain and the interfaces between elements: numerical fluxes, such as the centered fluxes or the upwind fluxes, are inserted in the scheme (1.30). Since these numerical fluxes involve discrete unknowns on both sides of an interface between two elements, the fluxes vectors connect the systems corresponding to the different elements together. Information is then transmitted from one element to its neighboring elements through these terms.

Assembling all local systems of the form (1.31) with the numerical fluxes provides a global system having the form (1.28). The final system connects all unknowns together.

The DG scheme has an ideal form to perform a parallel computation. Indeed, the global mass matrix, obtained by assembling local mass matrices $\mathbf{M}_e^p$ and $\mathbf{M}_e^{\boldsymbol{u}}$, is block-diagonal and can then be fastly inverted. Moreover, the coupling between discrete unknowns of different neighboring elements is only made through interface terms.

The DG method is frequently seen as a generalization of the FV method. As with the FV scheme, the dissipation and dispersion errors of the DG scheme depend on the numerical fluxes [93]: centered fluxes provide a dispersive scheme, while the numerical scheme is dissipative with upwind fluxes. When the local basis functions are constant (the order of elements is then 0), the DG scheme exactly reduces to the FV scheme.

CHAPTER 2

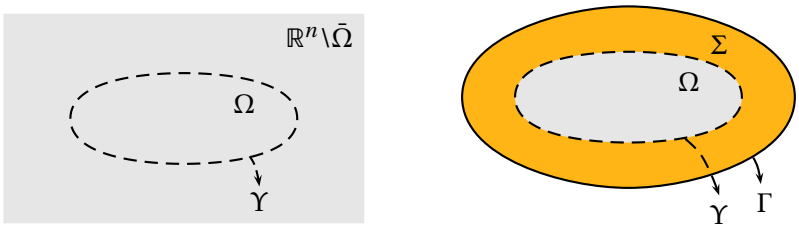
Design of Perfectly Matched Layers for Unbounded Domains

2.1 Introduction

Many wave-like problems are posed on unbounded spatial domains. However, the classical numerical methods presented in the previous chapter require a bounded computational domain. The numerical resolution is then based on a modified version of the physical problem, posed on a bounded domain (called *truncated domain*). As a result of the truncation, the model domain is bounded by an artificial boundary. Although unphysical, this boundary must be carefully modeled in order to recover the solution corresponding to the original problem.

For many decades, artificial boundary treatments have been proposed, studied and used in the literature. Their aim is to reproduce as accurately as possible the correct physical dynamics (*i.e.* those of the original problem) at the artificial boundary. A first strategy consists in using a specific *boundary condition*, *i.e.* a prescribed relation between the fields at the boundary. The so-called *layer techniques* can be considered as a second category of strategies. The truncated domain is surrounded by a layer (Figure 2.1), where the fields are subject to a particular treatment. A general review of these methods is sketched in this chapter.

Layer techniques have received great attention after the introduction of the *perfectly matched layer* (PML) concept by Bérenger in 1994 [21]. Layers designed with Bérenger's technique have a great advantage: their medium is both *dissipative* and *perfectly matched* with the original medium of the truncated domain. This means that every outgoing wave is perfectly transmitted from the truncated domain to the layer, and is damped in the layer, whatever the angle of incidence. This was not possible with previously proposed layers. Due to its interesting properties, Bérenger's PML, originally introduced for two-dimensional electromagnetic prob-



(a) Unbounded domain (b) Truncated domain with layer

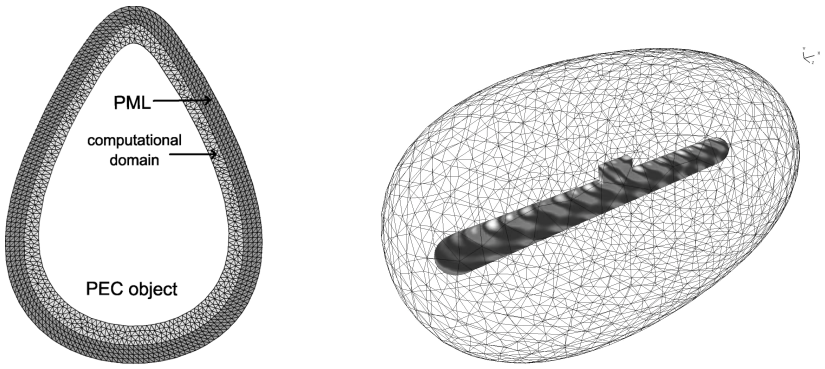
Figure 2.1: Geometry of the problem with the unbounded domain (a) and the truncated domain (b). In the second case, the truncated domain Ω is surrounded by a layer Σ.

lems [21], was quickly extended to three dimensions [22, 100] and other wave-like equations [45, 84, 89]. Unfortunately, some instabilities have been observed in PMLs built with Bérenger's technique (see [19] and references herein). The design of stabilized versions and the extension for cases of increasing complexity (*e.g.* general geometries, dispersive waves, nonlinear problems, ...) are not straightforward. Today, numerous versions have been proposed in various physical contexts. This research field remains very active [6, 13, 17, 57, 80, 92, 102, 110, 130, 138].

In order to minimize the computational cost of the numerical resolution, the truncated domain should be taken as small as possible. This can lead to domains with non-conventional shape, as illustrated in Figure 2.2. It is then advantageous to havea general PML formulation for arbitrarily-shaped truncated domains. However, most PML formulations are designed for cuboids, cylinders and spheres using corresponding coordinate systems. Some versions dealing with general convex domains have been proposed in time-harmonic contexts [64, 105, 106, 159, 161]. A first application in the time domain has been proposed by Donderici and Teixeira [52] with a specific numerical scheme. PML formulations have been investigated for non-convex domains [107, 125]. Nevertheless, no general formulation that could be used with classical numerical methods, has been proposed for time-dependent problems.

The aim of this chapter is to describe a method to build PMLs for convex truncated domains with a regular boundary (*i.e.* without corner). The PML formulation is derived for the scalar wave system (1.3). It is adapted to physical applications in chapters 4 and 5. The PML equations are written in Cartesian coordinates in order to facilitate their implementation in computational codes.

This chapter is organized as follows. In section 2.2, we present a general review of methods used to truncate unbounded domains. In section 2.3, we review some key ideas about the PMLs: the Bérenger's method to design a PML, fundamental properties of this layer, interpretations of the method and alternative ways to design PMLs. Finally, in section 2.4, we detail a procedure to derive PMLs for convex truncated domains in the time domain.



(a) Electromagnetic device (b) Submarine

Figure 2.2: Examples of wave-like problems where a general convex truncation boundary is advantageous: (a) electromagnetic scattering by a perfectly conducting device and (b) acoustic or electromagnetic scattering by a submarine. In both cases, the truncation boundary should be close to the scatterer in order to reduce the computational domain, and thus the computational cost. *Reproduced from [52] (© 2008 IEEE) and [162] (© 2001 John Wiley & Sons, Ltd), respectively.*

2.2 General review of boundary treatments for truncated domains

The main kinds of methods employed to truncate unbounded domains are reviewed in this section. They are written in the context of the scalar wave problem. The initial condition of the original problem (corresponding to the unbounded domain $\mathbb{R}^d$) is assumed to have compact support in the truncated domain $\Omega \subset \mathbb{R}^d$. Therefore, the presented boundary treatments deal only with outgoing waves. We assume that no waves come from the exterior domain, through the artificial boundary. The first part of this review (section 2.2.1), dedicated to boundary conditions, is mainly based on the recent paper of Givoli [73].

2.2.1 Boundary conditions

Artificial boundary conditions dealing with truncated domains are called *transparent*, *non-reflecting* or *absorbing boundary conditions* (ABCs). Generally, transparent and non-reflecting conditions refer to boundary conditions that provide a perfect treatment of the artificial boundary, while absorbing boundary conditions are only approximate.

Radiation boundary conditions

The simplest artificial conditions derive from the well known *radiation condition* proposed by Sommerfeld [154] in the harmonic context. Considering the Helmholtz equation in the unbounded domain $\mathbb{R}^d$, *i.e.*

$$\Delta p + k^2 p = 0, \quad \forall \boldsymbol{x} \in \mathbb{R}^d,$$

the radiation condition reads

$$\boxed{\lim_{\|\boldsymbol{x}\| \to \infty} \|\boldsymbol{x}\|^{(d-1)/2} \left(\frac{\partial p}{\partial \|\boldsymbol{x}\|} - \imath k p \right) = 0,} \tag{2.1}$$

where d is the spatial dimension and k is the wave number. It is based on a physical principle: *"The sources must be sources, not sinks, of energy. The energy which is radiated from the sources must scatter to infinity; no energy may be radiated from infinity into the prescribed singularities of the field."* (Arnold Sommerfeld) [154, p. 189]. The radiation condition must be prescribed at infinity, and is not directly applicable when the domain is truncated.

In the one-dimensional case, the radiation condition leads to a boundary condition that perfectly simulates the truncation of the unbounded domain. For the time-dependent problem with the wave equation

$$\frac{\partial^2 p}{\partial t^2} - c^2 \frac{\partial^2 p}{\partial x^2} = 0, \quad \forall (x,t) \in [0, L] \times \mathbb{R}^+,$$

the condition (2.1), evaluated at the boundary $x = L$, becomes:

$$\boxed{\frac{\partial p}{\partial t} + c \frac{\partial p}{\partial x} = 0,} \quad \forall t \in \mathbb{R}^+. \tag{2.2}$$

Given that no incoming waves from infinity are allowed, the solution at the boundary $x = L$ must contain waves moving only to the right. The most general form of this solution is then $p(x,t) = P(x - ct)$. Since this general solution satisfies the relation (2.2), this boundary condition perfectly simulates the truncation. Let us note that this condition can be obtained by an alternative way. Indeed, writing the second-order wave equation as

$$\left[\frac{\partial}{\partial t} - c \frac{\partial}{\partial x} \right] \left[\frac{\partial}{\partial t} + c \frac{\partial}{\partial x} \right] p = 0, \quad \forall t \in \mathbb{R}^+,$$

it appears that the relation (2.2) is obtained by removing the first differential operator in the left-hand side of the equation. Therefore, this boundary condition is sometimes considered as a modified version of the wave equation, with simplified dynamics: only the outgoing waves are accounted for. For a boundary condition at the other side of the domain ($x = 0$), the sign + is replaced with − in the relation (2.2).

The boundary condition (2.2) still holds for the problem described with the wave system, *i.e.*

$$\begin{cases} \dfrac{\partial p}{\partial t} + a\dfrac{\partial u}{\partial x} = 0, & \forall (x,t) \in [0,L] \times \mathbb{R}^+, \\ \dfrac{\partial u}{\partial t} + b\dfrac{\partial p}{\partial x} = 0, & \forall (x,t) \in [0,L] \times \mathbb{R}^+, \end{cases}$$

where $c = \sqrt{ab}$. Another convenient version of the boundary condition is obtained by injecting the second equation of this system into (2.2):

$$\begin{aligned} & \frac{\partial p}{\partial t} - c\left(\frac{1}{b}\frac{\partial u}{\partial t}\right) = 0, \\ \Leftrightarrow \quad & \frac{\partial}{\partial t}\left(p - \sqrt{\frac{a}{b}}u\right) = 0. \end{aligned}$$

The last equation implies that the quantity $p - \sqrt{a/b}\,u$ remains constant at the boundary. If this quantity is initially equal to zero, the condition becomes

$$\boxed{p - \sqrt{\frac{a}{b}}u = 0,} \quad \forall t \in \mathbb{R}^+. \tag{2.3}$$

This condition can be interpreted by the characteristic analysis (see section 1.2.4). The quantity in the left-hand side of relation (2.3) is the Riemann invariant of the wave system, transported to the left. The condition means therefore that no information (or a 'zero' information) is transported inward, along the incoming characteristic lines.

When FV or DG methods are used with upwind fluxes, the boundary condition (2.3) is easily implemented. The value of the incoming Riemann invariant in the Riemann solver is simply prescribed to zero. This is achieved by replacing the values of the variables in the ghost cells with zero (see the end of section 1.3.2).

Hierarchical and exact boundary conditions

Both radiation conditions (2.2) and (2.3) can be applied to multidimensional cases. However, they are only approximate for more than one spatial dimension, and errors are generated [71]. Different kinds of accurate boundary conditions have been derived in the literature.

First, sequences of ABCs have been proposed in the late 70's and 80's. These conditions are hierarchical: an order is associated with each condition of these sequences. The higher the order of the condition, the better its accuracy. The most well-known are the Engquist-Majda conditions [59] and the Bayliss-Turkel conditions [15]. The formers are for plane boundaries, while the latters are for circular and spherical truncated domains. In two dimensions, for a straight boundary with

the outward normal $\boldsymbol{n} = \boldsymbol{e}_x$, the first-order Engquist-Majda condition is equivalent to the radiation condition (2.2), and the second-order one reads

$$\left[\frac{1}{c}\frac{\partial^2}{\partial x \partial t} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \frac{1}{2}\frac{\partial^2}{\partial y^2}\right] p = 0.$$

For a spherical domain of radius R, the J^{th}-order Bayliss-Turkel condition is

$$\left[\prod_{j=1}^{J}\left(\frac{1}{c}\frac{\partial}{\partial t} + \frac{\partial}{\partial r} + \frac{2j-1}{R}\right)\right] p = 0,$$

where r is the radial coordinate. Unfortunately, because high-order conditions of these sequences require the computation of high-order partial derivatives, only low-order conditions are usable for numerical resolutions [73]. Further details and other similar boundary conditions can be found in papers [8, 70, 71, 78, 164].

Since 90's, new kinds of hierarchical boundary conditions, that can really be used with high-orders, have been developed. They involve the definition of auxiliary equations at the boundary, instead of using high-order derivatives. The higher the order, the greater the number of auxiliary equations. A sequence of this type has been first introduced by Collino [38]. In two dimensions, for a straight boundary with outward normal $\boldsymbol{n} = \boldsymbol{e}_x$, the J^{th}-order condition of Collino reads

$$\begin{cases} \dfrac{\partial p}{\partial t} + c\dfrac{\partial p}{\partial x} - \displaystyle\sum_{j=1}^{J} \beta_j \frac{\partial \phi_j}{\partial t} = 0, \\ \dfrac{1}{c^2}\dfrac{\partial^2 \phi_j}{\partial t^2} - \alpha_j \dfrac{\partial^2 \phi_j}{\partial y^2} - \dfrac{\partial^2 p}{\partial y^2} = 0, \quad \text{for } j = 1, \ldots, J, \end{cases}$$

where ϕ_j is an auxiliary variable and the coefficients α_j and β_j are defined in [38]. Reviews and further developments can be found in [72, 73].

Finally, exact boundary conditions providing a perfect treatment of fields at the truncated boundary have been proposed (see *e.g.* [71, 101]). Inside the truncated domain, the solution of the problem is then identical to the solution of the original problem with the unbounded domain. However, such boundary conditions are nonlocal: they involve a boundary integral operator, instead of only differential operators. All points of the boundary are then directly coupled. In a numerical resolution, it generally leads to dense algebraic systems that are expensive to solve.

2.2.2 Layer techniques

With layer techniques, the artificial boundary Υ is modeled by using a particular treatment of the fields inside an artificial layer Σ that surrounds the truncated domain Ω. In the layer, the original equations are then modified considering either physical or mathematical arguments.

Geometric layers

A first way to design layers consists in using geometric transformations: the layer Σ is mapped onto the exterior domain $\mathbb{R}^n \backslash \bar{\Omega}$ of the original problem [75, 95, 129]. The external problem is then fully solved inside the layer. In the domain Ω, the solutions obtained with both the original problem and the resulting one are theoretically identical.

However, the numerical solution of the resulting problem fails for oscillatory solutions [75]. Indeed, the apparent wavelength of plane waves in the layer decreases due to the spatial mapping. A finer spatial mesh is then required and the computational cost increases.

Absorbing layers

Efficient layers are obtained by adding dissipation. Then, travelling waves that leave the domain Ω are damped inside the layer. Since waves disappear and seem to be absorbed by the medium, these layers are called *absorbing* or *sponge layers*.

The most intuitive method to design an absorbing layer is to introduce physical dissipation in the medium. A damped version of the system (1.3) is easily obtained by adding a viscous term and/or a friction term in the first equation,

$$\begin{cases} \dfrac{\partial p}{\partial t} + a\nabla\cdot\boldsymbol{u} = \mu\dfrac{\partial^2 p}{\partial t^2} - \sigma p, \\ \dfrac{\partial \boldsymbol{u}}{\partial t} + b\nabla p = 0, \end{cases}$$

where $\mu(\boldsymbol{x})$ is the viscous coefficient and $\sigma(\boldsymbol{x})$ is the friction coefficient. In that case, the plane waves are damped in the layer, but the medium becomes dispersive and spurious transient waves appear [96].

A more efficient absorbing layer is obtained by introducing an unphysical friction term in the second equation of (1.3), in addition to the friction term of the first equation [88], *i.e.*

$$\begin{cases} \dfrac{\partial p}{\partial t} + a\nabla\cdot\boldsymbol{u} = -\sigma p, \\ \dfrac{\partial \boldsymbol{u}}{\partial t} + b\nabla p = -\tilde{\sigma}\boldsymbol{u}, \end{cases} \tag{2.4}$$

where $\tilde{\sigma}(\boldsymbol{x})$ is an unphysical coefficient. Taking $\tilde{\sigma} = \sigma$ *(the matching condition)*, this medium becomes dispersionless and the spurious transient waves are avoided [88, 155]. The interface between the domain and the layer is unreflective, but only for waves with a normal incidence. Similarly to the radiation condition, this layer can be exact for normal waves, and not for oblique waves.

With the *perfectly matched layers* (PMLs), first introduced by Bérenger in 1994 [21], outgoing waves of any incidence are transmitted from the truncated domain

to the layer without any reflection, and are then damped. The layer is then both *perfectly matched* and *dissipative*. Furthermore, by using adequate parameters, outgoing waves are completely damped. The artificial boundary of the truncated domain is then theoretically perfectly modeled. Unfortunately, some instabilities can be brought on in Bérenger's PMLs and an extension for wave-like equations involving additional dynamics is not trivial. Moreover, the design and the implementation of a PML is far more elaborate than with former layers, especially for time-dependent problems. Key ideas about PMLs and a short review are developed in section 2.3.

The above absorbing layers can be employed together with an absorbing boundary condition at the external boundary Γ. This ABC can improve the global performance of the layer by absorbing spurious transients and the outgoing waves that are not sufficiently damped by the medium. However, the PML gives already excellent results without requiring any kind of ABC if its parameters are well-chosen [140]. In this case, the use of an ABC is then not necessary.

2.2.3 Discussion and remarks

Due to both the quantity and the diversity of boundary conditions and layer techniques, a comparison between these boundary treatments is a complicated task. Various points of view could be considered [73]: the mathematical properties (well-posedness and stability), the accuracy (in both continuous and discrete levels), the ease of implementation, the computational cost, the generality, ... Moreover, depending on the application context, other points may appear.

Nevertheless, two general comments can be made on these two families of boundary treatments:

- In a discrete context, a higher accuracy implies a higher computation cost for both families. Indeed, it requires an increase of the order of the hierarchical condition (and thus the order of involved derivatives or the number of additional equations) or the size of the layer (and thus the number of discrete unknowns). Similarly, an accurate numerical approximation of nonlocal boundary conditions will be expensive.
- The extension of boundary treatments to increasingly complex physics and general geometries is not straighforward. Adaptations are required to guarantee a certain level of accuracy and, sometimes, the well-posedness and the stability. Specific reviews of boundary treatments are made in oceanography [27], in electromagnetism [79], in compressible fluid dynamics [41] and for the Schrödinger equation [9].

For the sake of completeness, let us mention the infinite and boundary elements, proposed in the framework of finite element methods. They directly deal with unbounded domains, and do not require the redefinition of problems on truncated domains. Further details can be found in [73] and references herein.

2.3 Key ideas about the PMLs

Bérenger's original strategy to design PMLs [21] is based on a purely mathematical transformation of the Cartesian equations. This transformation, devoid of any physical justification, has found convenient interpretations in the time-harmonic context. Thanks to these interpretations, alternative PMLs have been developed in various wave-like problems for both time and frequency domains.

In this section, key ideas about the PMLs are reviewed. First, Bérenger's strategy is explained and applied to the system (1.3). Fundamental properties of the obtained PML equations are then demonstrated by using plane-wave analyses. The interpretations of the PML and methods to design other versions are finally proposed.

2.3.1 Bérenger's technique to design a PML

Bérenger's technique to design a PML consists in splitting each field and equation into the Cartesian directions [21, 22]. Each equation contains partial derivatives in only one direction. For a damping in one Cartesian direction ($\boldsymbol{e}_x$, $\boldsymbol{e}_y$ or $\boldsymbol{e}_z$), a dissipation term is added in each equation with a partial derivative corresponding to the direction. With this formulation, PMLs with different directions of damping can be used together. Overlap of these layers at corners of the domain are naturally taken into account, as illustrated in Figure 2.3.

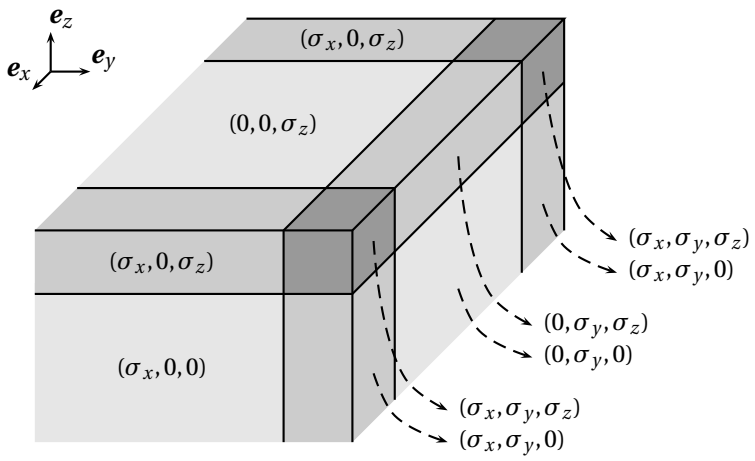


Figure 2.3: Geometry for the problem with the *split PML* of Bérenger for a rectangular cuboid truncated domain (upper-right part). The absorption coefficients are different in each face, each side and each corner, where outgoing waves are damped in respectively one, two and three directions. The figure is inspired from [22].

For the system (1.3), the field p is split into p_x, p_y and p_z, with $p = p_x + p_y + p_z$. At the corners of the domain, where the PML damps in the three directions, the fields are then governed by

$$\begin{cases} \dfrac{\partial p_x}{\partial t} + a\dfrac{\partial u_x}{\partial x} = -\sigma_x p_x, & \dfrac{\partial u_x}{\partial t} + b\dfrac{\partial p}{\partial x} = -\sigma_x u_x, \\ \dfrac{\partial p_y}{\partial t} + a\dfrac{\partial u_y}{\partial y} = -\sigma_y p_y, & \dfrac{\partial u_y}{\partial t} + b\dfrac{\partial p}{\partial y} = -\sigma_y u_y, \\ \dfrac{\partial p_z}{\partial t} + a\dfrac{\partial u_z}{\partial z} = -\sigma_z p_z, & \dfrac{\partial u_z}{\partial t} + b\dfrac{\partial p}{\partial z} = -\sigma_z u_z, \end{cases} \tag{2.5}$$

where u_x, u_y and u_z are the Cartesian components of $\boldsymbol{u}$, and $\sigma_x(x)$, $\sigma_y(y)$ and $\sigma_z(z)$ are the *absorption functions* corresponding to the three Cartesian directions. These functions are non-zero only in the regions with damping in the directions $\boldsymbol{e}_x$, $\boldsymbol{e}_y$ and $\boldsymbol{e}_z$, respectively (see Figure 2.3). Inside the truncated domain, they are equal to zero.

This *split PML*, originally designed for Maxwell's equations in two dimensions [21], has been quickly adapted to three-dimensional Maxwell's equations [22, 100], acoustics [89], geophysical fluid dynamics [45] and elastodynamics [84].

2.3.2 Properties of the PML

The PML is dissipative, dispersionless and perfectly matched for any incidence. In addition, if the absorption function is singular at the external boundary *(perfectly absorption condition)*, outgoing waves are fully absorbed by the layer, without any spurious reflection.

These interesting properties can be demonstrated by means of plane-wave analyses. Here, we consider a PML with damping only in the direction $\boldsymbol{e}_x$ (*i.e.* $\sigma_y = \sigma_z = 0$ everywhere). We use the notation $\sigma(x)$ instead of $\sigma_x(x)$, with $x \in \mathbb{R}^+$ (semi-infinite layer) or $x \in [0, \delta]$ (finite layer). The extension for the corner at the intersection of two or three PMLs is trivial. Here, we study separately the roles of the PML medium, the interface domain/layer and the external boundary.

Plane-wave solution

In the PML, the elementary time-harmonic plane-wave solution of the system (2.5) corresponds to a damped wave of the form

$$\begin{pmatrix} \boldsymbol{p}(\boldsymbol{x}, t) \\ \boldsymbol{u}(\boldsymbol{x}, t) \end{pmatrix} = \begin{pmatrix} \boldsymbol{P} \\ \boldsymbol{U} \end{pmatrix} e^{\imath(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)} e^{-\alpha}, \tag{2.6}$$

where $\boldsymbol{p} = (p_x, p_y, p_z)^T$, $\alpha(x)$ is the attenuation function and $\boldsymbol{P}$ and $\boldsymbol{U}$ are the amplitudes. Injecting this solution in the system (2.5), one obtains the relations

$$\omega = \sqrt{ab}\,\|\boldsymbol{k}\|, \tag{2.7}$$

$$\boldsymbol{P} = \sqrt{\frac{a}{b}}\,\frac{\boldsymbol{k}:\boldsymbol{U}}{\|\boldsymbol{k}\|},$$

$$\alpha(x) = \frac{\cos\theta}{\sqrt{ab}}\int_0^x \sigma(x')\,\mathrm{d}x', \quad \forall x \in \Sigma. \tag{2.8}$$

where $\theta = \arccos(k_x/\|\boldsymbol{k}\|)$ is the angle of incidence, and ':' denote the element-by-element product. Adding the split fields p_x, p_y and p_z, one obtains the solution for the unsplit field p,

$$p(\boldsymbol{x}, t) = P e^{\imath(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} e^{-\alpha}, \tag{2.9}$$

with the amplitude relation

$$P\frac{\boldsymbol{k}}{\|\boldsymbol{k}\|} = \sqrt{\frac{a}{b}}\,\boldsymbol{U}. \tag{2.10}$$

The dispersion relation (2.7) and the amplitude relation (2.10), are identical to those of the original medium (relations (1.6) and (1.7)). The phase velocities and the impedance are also identical. Therefore, the plane waves that propagate in the PML medium have exactly the same properties as those of the original medium, but with an additional factor (second exponential of (2.6) and (2.9)) that diminishes the amplitude of both fields. The shape of the decay is independent of the frequency, but not of the angle of incidence.

Perfect matching of the PML

If continuity conditions are prescribed on the fields at the interface between the truncated domain and the layer, the interface is transparent for waves. This can be shown considering a representative problem with a semi-infinite domain (with $x < 0$) extended by a semi-infinite layer (with $x > 0$). Let us assume that an incident wave that propagates towards the layer is decomposed into a reflected part (in the domain) and a transmitted part (in the layer). The solution is then

$$p(\boldsymbol{x}, t) = \begin{cases} P^i e^{\imath(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} + P^r e^{-2\imath k_x x} e^{\imath(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} & \text{in the domain,} \\ P^t e^{\imath(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} e^{-\alpha} & \text{in the layer,} \end{cases}$$

$$\boldsymbol{u}(\boldsymbol{x}, t) = \begin{cases} \boldsymbol{U}^i e^{\imath(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} + \boldsymbol{U}^r e^{-2\imath k_x x} e^{\imath(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} & \text{in the domain,} \\ \boldsymbol{U}^t e^{\imath(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} e^{-\alpha} & \text{in the layer.} \end{cases}$$

Injecting this solution in the continuity interface conditions and using the amplitude relation (2.10), one immediately obtains that the amplitudes of the reflected

wave, P^r and $\boldsymbol{U}^r$, are zero whatever the value of σ. Therefore, the reflection coefficient associated with the interface is zero, *i.e.*

$$\boxed{r_{\text{interf}} = \left| \frac{P^r}{P^i} \right| = 0.} \tag{2.11}$$

Every incident wave is perfectly transmitted from the truncated domain to the layer, whatever its frequency and angle of incidence. The interface is said *perfectly matched*.

Perfect absorption of the PML

If plane waves are not sufficiently damped, they could be reflected by the external boundary and come back in the domain. To study this case, let us consider a semi-infinite domain (with $x < 0$) extended by a finite layer (with $x \in [0, \delta]$), with the homogeneous Dirichlet boundary condition $\boldsymbol{e}_x \cdot \boldsymbol{u} = 0$ at $x = \delta$. Assuming the outgoing wave is reflected by the external boundary, the general solution becomes

$$p(\boldsymbol{x}, t) = \begin{cases} P^i e^{\imath(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} + P^r e^{-2\imath k_x x} e^{\imath(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} & \text{in the domain,} \\ P^t e^{\imath(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} e^{-\alpha} + P^b e^{-2\imath k_x x} e^{\imath(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} e^{\alpha} & \text{in the layer,} \end{cases}$$

$$\boldsymbol{u}(\boldsymbol{x}, t) = \begin{cases} \boldsymbol{U}^i e^{\imath(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} + \boldsymbol{U}^r e^{-2ik_x x} e^{\imath(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} & \text{in the domain,} \\ \boldsymbol{U}^t e^{\imath(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} e^{-\alpha} + \boldsymbol{U}^b e^{-2\imath k_x x} e^{\imath(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} e^{\alpha} & \text{in the layer.} \end{cases}$$

After some calculations, one obtains the reflection coefficient associated with the layer,

$$\boxed{r_{\text{pml}} = \left| \frac{P^r}{P^i} \right| = \exp\left[-2 \frac{\cos\theta}{\sqrt{ab}} \int_0^{\delta} \sigma(x') \mathrm{d}x' \right],} \tag{2.12}$$

which gives an indication on the reflection of outgoing waves by the layer. As announced before, this reflection coefficient is zero if

$$\boxed{\int_0^{\delta} \sigma(x')\, \mathrm{d}x' = +\infty,} \tag{2.13}$$

e.g. if σ is singular at the external boundary $x = \delta$. The layer is then *perfectly absorbing*.

2.3.3 Interpretations of the PML

While the PML has been introduced by Bérenger without any physical justification, convenient interpretations have been proposed in the time-harmonic context. Assuming a time-dependency in $e^{-\imath\omega t}$, the time-harmonic form of the split

PML equations (2.5) is

$$\begin{cases} -\imath\omega\hat{p}_x + a\dfrac{\partial\hat{u}_x}{\partial x} = -\sigma_x\hat{p}_x, & -\imath\omega\hat{u}_x + b\dfrac{\partial\hat{p}}{\partial x} = -\sigma_x\hat{u}_x, \\ -\imath\omega\hat{p}_y + a\dfrac{\partial\hat{u}_y}{\partial y} = -\sigma_y\hat{p}_y, & -\imath\omega\hat{u}_y + b\dfrac{\partial\hat{p}}{\partial y} = -\sigma_y\hat{u}_y, \\ -\imath\omega\hat{p}_z + a\dfrac{\partial\hat{u}_z}{\partial z} = -\sigma_z\hat{p}_z, & -\imath\omega\hat{u}_z + b\dfrac{\partial\hat{p}}{\partial z} = -\sigma_z\hat{u}_z, \end{cases}$$

where the symbol $\hat{}$ denotes time-harmonic fields. After some manipulations and by adding the three first equations, the system is reduced to the unsplit system

$$\begin{cases} -\imath\omega\hat{p} + a\left(h_x^{-1}\dfrac{\partial\hat{u}_x}{\partial x} + h_y^{-1}\dfrac{\partial\hat{u}_y}{\partial y} + h_z^{-1}\dfrac{\partial\hat{u}_z}{\partial z}\right) = 0, \\ -\imath\omega\hat{u}_x + bh_x^{-1}\dfrac{\partial\hat{p}}{\partial x} = 0, \\ -\imath\omega\hat{u}_y + bh_y^{-1}\dfrac{\partial\hat{p}}{\partial y} = 0, \\ -\imath\omega\hat{u}_z + bh_z^{-1}\dfrac{\partial\hat{p}}{\partial z} = 0, \end{cases} \tag{2.14}$$

with

$$h_x(x) = 1 - \sigma_x(x)/(\imath\omega), \tag{2.15}$$

$$h_y(y) = 1 - \sigma_y(y)/(\imath\omega), \tag{2.16}$$

$$h_z(z) = 1 - \sigma_z(z)/(\imath\omega). \tag{2.17}$$

In the literature, these functions are generally referred as the *stretching functions*.

PML as a stretch of the metric

The PML treatment of the equations can be interpreted as a change of the metric of the space [106, 160, 161]. Indeed, defining the complex operator

$$\tilde{\nabla} \overset{\text{def.}}{=} \operatorname{diag}(h_x^{-1}, h_y^{-1}, h_z^{-1})\,\nabla,$$

the system (2.14) can be written in the more compact form

$$\boxed{\begin{cases} -\imath\omega\hat{p} + a\,\tilde{\nabla}\cdot\hat{\boldsymbol{u}} = 0, \\ -\imath\omega\hat{\boldsymbol{u}} + b\,\tilde{\nabla}\hat{p} = 0. \end{cases}} \tag{2.18}$$

This system is similar to the original wave system (1.3) in its time-harmonic form, but with a modification of the metric of the space. While the Cartesian metric tensor in the original medium is the identity $\operatorname{diag}(1,1,1)$, the PML medium has the complex one $\operatorname{diag}(h_x^2, h_y^2, h_z^2)$.

PML as an anisotropic material absorber

The PML can also be considered as an anisotropic absorbing medium. Indeed, defining the new field

$$\hat{\boldsymbol{u}}^{\text{abs}} = \text{diag}(h_y h_z, h_x h_z, h_x h_y)\, \hat{\boldsymbol{u}},$$

the system (2.14) can be rewritten

$$\boxed{\begin{cases} -\imath\omega\hat{p} + a^{\text{abs}}\, \nabla\cdot \hat{\boldsymbol{u}}^{\text{abs}} = 0, \\ -\imath\omega\hat{\boldsymbol{u}}^{\text{abs}} + \mathbf{B}^{\text{abs}}\, \nabla\hat{p} = 0, \end{cases}} \tag{2.19}$$

with the complex physical parameters

$$a^{\text{abs}} = \frac{a}{h_x h_y h_z} \quad \text{and} \quad \mathbf{B}^{\text{abs}} = \text{diag}\left(\frac{h_y h_z}{h_x}, \frac{h_x h_z}{h_y}, \frac{h_x h_y}{h_z}\right) b.$$

The PML system is again similar to the original wave system (1.3) in his time-harmonic form. Unlike the system (2.18), this version does not require any change of the metric. The PML treatment is made through the complex physical parameters and another definition of the field $\hat{\boldsymbol{u}}$.

In the case of time-harmonic Maxwell's equations, this interpretation can be seen as a consequence of the change of the metric [66, 152, 160]. Indeed, Maxwell's equations are independent of the metric, contrary to the constitutive relations. Therefore, the classical equations are still used in the PML, but with complex material parameters that replace the physical ones [160].

PML as a change of variables in a stretched coordinate system

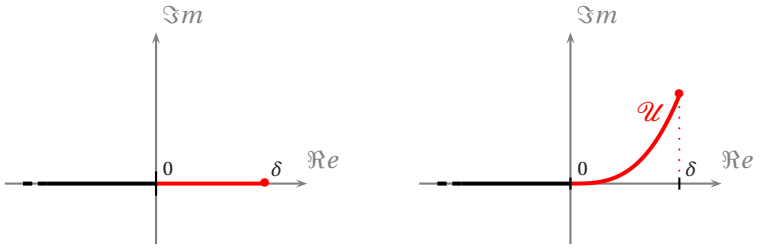
Finally, the PML treatment to design layers can be interpreted as a change of variables in a stretched coordinate system [35, 144]. The real coordinates are simply replaced with complex ones (with a nonzero imaginary part), and a change of variables is then used to come back to a real coordinate system.

For a PML that damps in the $x-$direction, the coordinate $x \in [0,\delta]$ is replaced with $\tilde{x} \in \mathscr{U}$, where $\mathscr{U}$ is a curve in the complex plane (see Figure 2.4). In order to write the PML equations in a real coordinate system, the curve $\mathscr{U}$ is parametrized by the function $\tilde{x}(x_\star)$ with the real argument $x_\star \in [0,\delta]$. Taking the particular function

$$\tilde{x}(x_\star) = \int_0^{x_\star} h_x(x'_\star)\, \mathrm{d}x'_\star = x_\star - \frac{1}{\imath\omega}\int_0^{x_\star} \sigma_x(x'_\star)\, \mathrm{d}x'_\star, \tag{2.20}$$

the complex spatial derivative associated with $\tilde{x}$ is written

$$\frac{\partial}{\partial\tilde{x}} = \frac{1}{1-\sigma_x/\imath\omega}\frac{\partial}{\partial x_\star} = h_x^{-1}\frac{\partial}{\partial x_\star}.$$



(a) Original medium (b) PML medium

Figure 2.4: Complex coordinate stretching of a coordinate. The domain is represented by the black line, while the layer is represented by the red one.

Finally, the PML equations are obtained by replacing the real partial derivative by the complex one, and then by using the change of variables with a parametrization of the curve:

$$\boxed{\frac{\partial}{\partial x} \xrightarrow{\text{rep.}} \frac{\partial}{\partial \tilde{x}} \xrightarrow{\text{c.v.}} h_x^{-1} \frac{\partial}{\partial x_\star}.}$$

Using this procedure for the three Cartesian directions and removing the subscrit $_\star$, the PML system (2.14) is immediately obtained from the original one (1.3) in its time-harmonic form.

2.3.4 Alternative PMLs — Extension to other geometries and physics

In light of the above interpretations, a lot of new PMLs have been proposed for various linear wave-like problems in time-harmonic and time-dependent contexts.

Unsplit PMLs that preserve the original (unsplit) structure of fields are an alternative to Bérenger's *split PMLs* in Cartesian coordinates. The time-harmonic systems (2.18) or (2.19) are unsplit. In the time-dependent context, PML equations are obtained by employing an inverse Fourier transform. The fields can then be governed by the original equations, but with additional terms involving either time-convolutions [143, 170] or additional fields governed by additional differential equations [2, 5, 6, 12, 19, 65].

In other classical orthogonal coordinate systems, PMLs are simply obtained by changing the metric tensor [160] or using the complex coordinate stretching of one of the coordinates [36, 40, 141, 158, 169]. For example, a PML for a spherical truncated domain is provided by stretching the radial coordinate r of the spherical

coordinate system, *i.e.*

$$r \quad \rightarrow \quad \tilde{r}(r) = r - \frac{1}{\imath\omega}\int_{r_0}^{r} \sigma(r')\,dr', \quad \forall r \in [r_0, r_0+\delta],$$

where r_0 is the radius of the spherical domain and δ is the thickness of the layer.

A perfect extension of PMLs for increasing complex problems brings up new difficulties. In advective acoustics, for example, the advection terms need specific attention [3, 51, 90–92, 126, 156]. Evanescent waves are overdamped by PMLs designed with the complex coordinate stretching (2.20). Therefore, spurious errors appear when numerical methods are used [23, 24, 48]. A solution is to slightly change the formula (2.20) or, equivalently, the stretching functions (2.15)–(2.17) [43, 65, 104, 148]. For example, Kuzuoglu and Mittra [104] propose the stretching function

$$h = \kappa + \frac{\sigma}{\alpha - \imath\omega},$$

where κ and α are new positive parameters. The last is introduced to better absorb the evanescent waves.

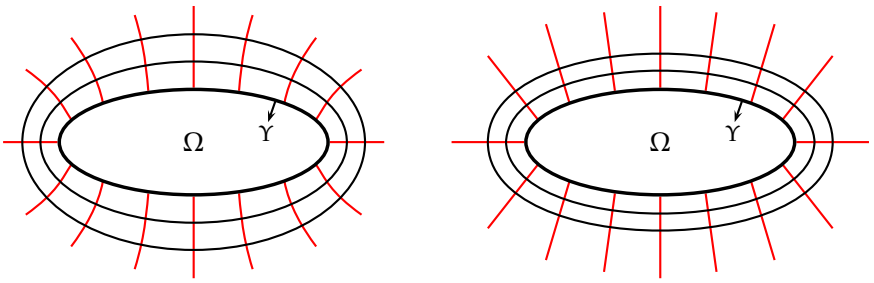
Even though Bérenger's PMLs have demonstrated an overall excellent performance for applications [21, 22, 65], Abarbanel and Gottlieb [1] and later Bécache and Joly [16] have shown that the PML equations designed with Bérenger's technique are not strongly well-posed: spurious modes can be generated in the layer. In addition, some problems have been reported for long-term simulations [4, 18]. Different improved layers exhibiting better properties have been proposed [2, 4, 6, 12, 16, 18, 141, 143].

2.4 A novel formulation of PML for convex truncated domains

In this section, we propose a complete strategy to design PMLs for convex truncated domains with a regular boundary[1] in the time domain. The strategy is applied to derive the equations of a PML for the scalar wave system (1.3). Let us note that PMLs could be derived in a similar manner for other linear wave-like systems.

The proposed strategy is based on a coordinate stretch in a specific curvilinear coordinate system that is associated with the artificial boundary Υ of the truncated domain Ω. This particular stretch has been first proposed by Teixeira and Chew [159]. It has been studied in time-harmonic contexts [105, 106, 161] and first applied in a time-dependent context in the paper [52] with a specific numerical scheme. The novelty here is that we provide a convenient form of the PML

[1] At each point of the boundary, there is a unique tangent line (for a regular curve) or a unique tangent plane (for a regular surface). As a consequence, the proposed strategy cannot be applied to boundaries with corners.



(a) Elliptical coordinate system

(b) Curvilinear coordinate system based on parallel surfaces

Figure 2.5: Coordinate systems associated with an elliptical truncated domain. Radial lines of the elliptical system are curved (a), while those of the curvilinear system based on parallel lines are straight (b).

equations at the continuous level in the time domain. These equations can be easily written in the Cartesian coordinate system, even if the coordinate stretch is performed in the curvilinear one.

This section is organized as follows. First, the specific curvilinear coordinate system is presented. Then, the *nabla* operator in the stretched metric is explicitly derived in the (real) Cartesian coordinate system. Finally, both time-harmonic and time-dependent PML equations are obtained for the scalar wave system.

2.4.1 Curvilinear coordinate system associated with the boundary

Different coordinate systems can be associated with the boundary Υ of a convex truncated domain Ω. For instance, in the case of a two-dimensional elliptical truncated domain, two choices are natural: the elliptical coordinate system, and the curvilinear coordinate system based on lines parallel to Υ (Figure 2.5). For both systems, the radial coordinate is stretched to derive a layer. The two obtained layers are perfectly matched, but with a different geometry [161]. Indeed, the external boundary of the first is elliptical and the radial lines are curved. For the second, the external boundary is not elliptical, but parallel to Υ, and the radial lines are straight.

PMLs used in this thesis are built using curvilinear systems based on parallel surfaces (in three dimensions) and parallel lines (in two dimensions). In the literature, such PMLs are called *conformal PMLs* [52, 159, 161]. Since they can be derived for all convex truncated domains with regular boundary, this approach appears very general. Hereafter, the local curvilinear system is derived for both two-dimensional and three-dimensional cases.

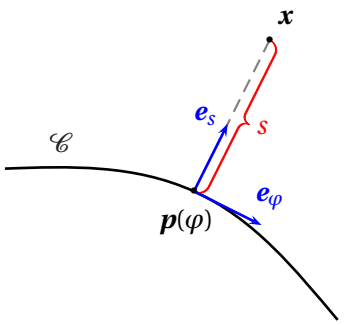


Figure 2.6: Orthonormal curvilinear coordinate system (s,φ) associated with the curve $\mathscr{C}$.

In two dimensions, the boundary of Ω is a curve, denoted $\mathscr{C}$. Let us consider a natural parametrization for this curve in $\mathbb{R}^2$, $\boldsymbol{p}(\varphi)$ with $\left\|\mathrm{d}\boldsymbol{p}/\mathrm{d}\varphi\right\| = 1$ and $\varphi \in [0, 2\pi]$. Since Ω is convex, a point of the layer has a unique closest point on $\mathscr{C}$. Its Cartesian coordinates, respectively $\boldsymbol{x}(s,\varphi)$ and $\boldsymbol{p}(\varphi)$, are then linked by the relation

$$\boldsymbol{x}(s,\varphi) = \boldsymbol{p}(\varphi) + s\,\boldsymbol{e}_s(\varphi), \tag{2.21}$$

where φ is the arc length corresponding to the closest point on $\mathscr{C}$, s is the distance between the two points, and $\boldsymbol{e}_s(\varphi)$ is the unit outward normal of Ω at $\boldsymbol{p}(\varphi)$, as illustrated in Figure 2.6. The pair (s,φ) defines the curvilinear coordinate system.

In three dimensions, the curvilinear coordinate system is an extension of the one proposed above. The boundary of Ω is a surface, denoted $\mathscr{S}$ and parametrized in $\mathbb{R}^3$ by $\boldsymbol{p}(\varphi,\theta)$, with $\varphi \in [0, 2\pi]$ and $\theta \in [0, \pi]$. Since Ω is convex, each point of the layer can be written

$$\boldsymbol{x}(s,\varphi,\theta) = \boldsymbol{p}(\varphi,\theta) + s\,\boldsymbol{e}_s(\varphi,\theta), \tag{2.22}$$

where the parameters φ and θ correspond to the point of $\mathscr{S}$ the closest to $\boldsymbol{x}(s,\varphi,\theta)$, s is again the distance between $\boldsymbol{x}(s,\varphi,\theta)$ and $\boldsymbol{p}(\varphi,\theta)$, and $\boldsymbol{e}_s(\varphi,\theta)$ is the unit outward normal of Ω at $\boldsymbol{p}(\varphi,\theta)$. Let us consider the specific case where the parameters φ and θ are also the arc lengths of the curves $\mathscr{C}_\varphi$ and $\mathscr{C}_\theta$ obtained from the intersection of $\mathscr{S}$ with the planes generated by $\boldsymbol{e}_s$ and one of the principal vectors of the surface, denoted $\boldsymbol{e}_\varphi$ and $\boldsymbol{e}_\theta$. The coordinates (s,φ,θ) define an orthogonal curvilinear system. The vectors $\boldsymbol{e}_s$, $\boldsymbol{e}_\varphi$ and $\boldsymbol{e}_\theta$ are orthonormal and define the so-called Darboux frame [76]. The Frenet relations for $\mathscr{C}_\varphi$ and $\mathscr{C}_\theta$ read

$$\frac{\partial \boldsymbol{p}}{\partial \varphi} = \boldsymbol{e}_\varphi, \qquad \frac{\partial \boldsymbol{e}_\varphi}{\partial \varphi} = -\kappa_\varphi \boldsymbol{e}_s, \qquad \frac{\partial \boldsymbol{e}_s}{\partial \varphi} = \kappa_\varphi \boldsymbol{e}_\varphi,$$
$$\frac{\partial \boldsymbol{p}}{\partial \theta} = \boldsymbol{e}_\theta, \qquad \frac{\partial \boldsymbol{e}_\theta}{\partial \theta} = -\kappa_\theta \boldsymbol{e}_s, \qquad \frac{\partial \boldsymbol{e}_s}{\partial \theta} = \kappa_\theta \boldsymbol{e}_\theta.$$

2.4.2 Coordinate stretch

The PML is built by stretching the radial coordinate s in the time-harmonic context. For both two and three dimensions, the real coordinate $s \in [0,\delta]$ is then replaced with the complex one $\tilde{s} \in \mathscr{U}$, where $\mathscr{U}$ is a curve in the complex plane. Let us consider the parametrization of this curve

$$\tilde{s}(s^\star) = s^\star - \frac{1}{\imath\omega}\int_0^{s^\star} \sigma(s')\,\mathrm{d}s', \quad \text{with } s^\star \in [0,\delta]. \tag{2.23}$$

The time-harmonic PML equations are obtained in the Cartesian coordinate system using the following four-step procedure:

1) change of variables to write the original time-harmonic equations from the Cartesian coordinate system (x,y,z) to the curvilinear one (s,φ,θ),
2) coordinate stretch: the coordinates (s,φ,θ) are replaced with $(\tilde{s},\varphi,\theta)$,
3) change of variables to write the PML equations from the complex curvilinear coordinate system $(\tilde{s},\varphi,\theta)$ to the real one (s,φ,θ) using the relation (2.23),
4) change of variables to write the PML equations from the curvilinear coordinate system (s,φ,θ) to the Cartesian one (x,y,z).

With this procedure, the *nabla* operator ∇ becomes successively

$$\underset{(x,y,z)}{\nabla} \xrightarrow{\text{c.v.}} \underset{(s,\varphi,\theta)}{\nabla} \xrightarrow{\text{rep.}} \underset{(\tilde{s},\varphi,\theta)}{\tilde{\nabla}} \xrightarrow{\text{c.v.}} \underset{(s,\varphi,\theta)}{\tilde{\nabla}} \xrightarrow{\text{c.v.}} \underset{(x,y,z)}{\tilde{\nabla}},$$

where $\tilde{\nabla}$ is the *nabla* operator with complex stretched coordinates. At each step, the operator is written using the coordinates indicated under the symbol.

1) Change of variables $(x,y,z) \xrightarrow{\text{c.v.}} (s,\varphi,\theta)$

The transformation from Cartesian to curvilinear coordinates is made using the Jacobian matrix

$$\begin{pmatrix} \partial_s x & \partial_\varphi x & \partial_\theta x \\ \partial_s y & \partial_\varphi y & \partial_\theta y \\ \partial_s z & \partial_\varphi z & \partial_\theta z \end{pmatrix} = \begin{pmatrix} e_{s,x} & e_{\varphi,x} & e_{\theta,x} \\ e_{s,y} & e_{\varphi,y} & e_{\theta,y} \\ e_{s,z} & e_{\varphi,z} & e_{\theta,z} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1+\kappa_\varphi s & 0 \\ 0 & 0 & 1+\kappa_\theta s \end{pmatrix},$$

where $\kappa_\varphi(\varphi,\theta)$ and $\kappa_\theta(\varphi,\theta)$ are the principal curvatures of $\mathscr{S}$ at $\boldsymbol{p}(\varphi,\theta)$. The first matrix of the right-hand side of this equation contains the Cartesian components of vectors $\boldsymbol{e}_s$, $\boldsymbol{e}_\varphi$ and $\boldsymbol{e}_\theta$. The Jacobian matrix is obtained using the relation (2.22)

and the Frenet relations for $\mathscr{C}_\varphi$ and $\mathscr{C}_\theta$. Therefore, one has

$$\begin{pmatrix} \partial_s \\ \partial_\varphi \\ \partial_\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1+\kappa_\varphi s & 0 \\ 0 & 0 & 1+\kappa_\theta s \end{pmatrix} \begin{pmatrix} e_{s,x} & e_{s,y} & e_{s,z} \\ e_{\varphi,x} & e_{\varphi,y} & e_{\varphi,z} \\ e_{\theta,x} & e_{\theta,y} & e_{\theta,z} \end{pmatrix} \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix}, \tag{2.24}$$

$$\Rightarrow \quad \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} = \begin{pmatrix} e_{s,x} & e_{\varphi,x} & e_{\theta,x} \\ e_{s,y} & e_{\varphi,y} & e_{\theta,y} \\ e_{s,z} & e_{\varphi,z} & e_{\theta,z} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{1+\kappa_\varphi s} & 0 \\ 0 & 0 & \frac{1}{1+\kappa_\theta s} \end{pmatrix} \begin{pmatrix} \partial_s \\ \partial_\varphi \\ \partial_\theta \end{pmatrix}, \tag{2.25}$$

and then in the compact form

$$\nabla = \boldsymbol{e}_s \frac{\partial}{\partial s} + \frac{1}{1+\kappa_\varphi s} \boldsymbol{e}_\varphi \frac{\partial}{\partial \varphi} + \frac{1}{1+\kappa_\theta s} \boldsymbol{e}_\theta \frac{\partial}{\partial \theta}.$$

2) Coordinate stretch $(s,\varphi,\theta) \xrightarrow{\text{rep.}} (\tilde{s},\varphi,\theta)$

In this step, the real coordinate s is simply replaced with the complex one $\tilde{s}$. As a consequence, the real operator ∇ is replaced with the complex one $\tilde{\nabla}$, defined by

$$\tilde{\nabla} = \boldsymbol{e}_s \frac{\partial}{\partial \tilde{s}} + \frac{1}{1+\kappa_\varphi \tilde{s}} \boldsymbol{e}_\varphi \frac{\partial}{\partial \varphi} + \frac{1}{1+\kappa_\theta \tilde{s}} \boldsymbol{e}_\theta \frac{\partial}{\partial \theta}.$$

3) Change of variables $(\tilde{s},\varphi,\theta) \xrightarrow{\text{c.v.}} (s,\varphi,\theta)$

Using the parametrisation (2.23), the change of variables from $(\tilde{s},\varphi,\theta)$ to (s,φ,θ) gives

$$\begin{pmatrix} \partial_s \\ \partial_\varphi \\ \partial_\theta \end{pmatrix} = \begin{pmatrix} \partial_s \tilde{s} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \partial_{\tilde{s}} \\ \partial_\varphi \\ \partial_\theta \end{pmatrix}$$

with $\partial_s \tilde{s} = 1 - \sigma/(\imath\omega)$. Using this relation, the complex operator $\tilde{\nabla}$ thus reads

$$\begin{pmatrix} \partial_{\tilde{x}} \\ \partial_{\tilde{y}} \\ \partial_{\tilde{z}} \end{pmatrix} = \begin{pmatrix} e_{s,x} & e_{\varphi,x} & e_{\theta,x} \\ e_{s,y} & e_{\varphi,y} & e_{\theta,y} \\ e_{s,z} & e_{\varphi,z} & e_{\theta,z} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{1+\kappa_\varphi \tilde{s}} & 0 \\ 0 & 0 & \frac{1}{1+\kappa_\theta \tilde{s}} \end{pmatrix} \begin{pmatrix} \frac{\imath\omega}{\imath\omega-\sigma} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \partial_s \\ \partial_\varphi \\ \partial_\theta \end{pmatrix},$$

and then

$$\tilde{\nabla} = \frac{\imath\omega}{\imath\omega-\sigma} \boldsymbol{e}_s \frac{\partial}{\partial s} + \frac{1}{1+\kappa_\varphi \tilde{s}} \boldsymbol{e}_\varphi \frac{\partial}{\partial \varphi} + \frac{1}{1+\kappa_\theta \tilde{s}} \boldsymbol{e}_\theta \frac{\partial}{\partial \theta}.$$

4) Change of variables $(s,\varphi,\theta) \xrightarrow{\text{c.v.}} (x,y,z)$

Using the relation (2.24), the complex operator $\tilde{\nabla}$ reads

$$\begin{pmatrix} \partial_{\tilde{x}} \\ \partial_{\tilde{y}} \\ \partial_{\tilde{z}} \end{pmatrix} = \begin{pmatrix} e_{s,x} & e_{\varphi,x} & e_{\theta,x} \\ e_{s,y} & e_{\varphi,y} & e_{\theta,y} \\ e_{s,z} & e_{\varphi,z} & e_{\theta,z} \end{pmatrix} \begin{pmatrix} \frac{\imath\omega}{\imath\omega-\sigma} & 0 & 0 \\ 0 & \frac{1+\kappa_\varphi s}{1+\kappa_\varphi \tilde{s}} & 0 \\ 0 & 0 & \frac{1+\kappa_\theta s}{1+\kappa_\theta \tilde{s}} \end{pmatrix} \begin{pmatrix} e_{s,x} & e_{s,y} & e_{s,z} \\ e_{\varphi,x} & e_{\varphi,y} & e_{\varphi,z} \\ e_{\theta,x} & e_{\theta,y} & e_{\theta,z} \end{pmatrix} \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix}.$$

It can be successively written

$$\begin{aligned}\tilde{\nabla} &= \frac{\imath\omega}{\imath\omega-\sigma}\nabla_{\boldsymbol{e}_s} + \frac{1+\kappa_\varphi s}{1+\kappa_\varphi \tilde{s}}\nabla_{\boldsymbol{e}_\varphi} + \frac{1+\kappa_\theta s}{1+\kappa_\theta \tilde{s}}\nabla_{\boldsymbol{e}_\theta},\\ &= \frac{\imath\omega}{\imath\omega-\sigma}\nabla_{\boldsymbol{e}_s} + \frac{1+\kappa_\varphi s}{1+\kappa_\varphi s-\kappa_\varphi\frac{\bar{\sigma}}{\imath\omega}}\nabla_{\boldsymbol{e}_\varphi} + \frac{1+\kappa_\theta s}{1+\kappa_\theta s-\kappa_\theta\frac{\bar{\sigma}}{\imath\omega}}\nabla_{\boldsymbol{e}_\theta},\end{aligned}$$

with

$$\begin{aligned}&\nabla_{\boldsymbol{e}_s} = [\boldsymbol{e}_s(\boldsymbol{e}_s\cdot\nabla)], && \nabla_{\boldsymbol{e}_\varphi} = [\boldsymbol{e}_\varphi(\boldsymbol{e}_\varphi\cdot\nabla)], && \nabla_{\boldsymbol{e}_\theta} = [\boldsymbol{e}_\theta(\boldsymbol{e}_\theta\cdot\nabla)],\\ &\bar{\sigma} = \int_0^r \sigma(s)\,ds, && \bar{\kappa}_\varphi = (\kappa_\varphi^{-1}+s)^{-1}, && \bar{\kappa}_\theta = (\kappa_\theta^{-1}+s)^{-1}.\end{aligned}$$

Finally, the complex operator $\tilde{\nabla}$ reads

$$\boxed{\tilde{\nabla} = \nabla - \frac{\sigma}{\sigma-\imath\omega}\nabla_{\boldsymbol{e}_s} - \frac{\bar{\kappa}_\varphi\bar{\sigma}}{\bar{\kappa}_\varphi\bar{\sigma}-\imath\omega}\nabla_{\boldsymbol{e}_\varphi} - \frac{\bar{\kappa}_\theta\bar{\sigma}}{\bar{\kappa}_\theta\bar{\sigma}-\imath\omega}\nabla_{\boldsymbol{e}_\theta}.} \tag{2.26}$$

2.4.3 PML equations

In the time-harmonic context, the equations of the PML are obtained by remplacing the operator ∇ by the complex one $\tilde{\nabla}$. Using the expression (2.26), the time-harmonic scalar wave system becomes

$$\begin{cases} -\imath\omega\hat{p} + a\nabla\cdot\hat{\boldsymbol{u}} = a\dfrac{\sigma}{\sigma-\imath\omega}\nabla_{\boldsymbol{e}_s}\cdot\hat{\boldsymbol{u}} + a\dfrac{\bar{\kappa}_\varphi\bar{\sigma}}{\bar{\kappa}_\varphi\bar{\sigma}-\imath\omega}\nabla_{\boldsymbol{e}_\varphi}\cdot\hat{\boldsymbol{u}} + a\dfrac{\bar{\kappa}_\theta\bar{\sigma}}{\bar{\kappa}_\theta\bar{\sigma}-\imath\omega}\nabla_{\boldsymbol{e}_\theta}\cdot\hat{\boldsymbol{u}},\\ -\imath\omega\hat{\boldsymbol{u}} + b\nabla\hat{p} = b\dfrac{\sigma}{\sigma-\imath\omega}\nabla_{\boldsymbol{e}_s}\hat{p} + b\dfrac{\bar{\kappa}_\varphi\bar{\sigma}}{\bar{\kappa}_\varphi\bar{\sigma}-\imath\omega}\nabla_{\boldsymbol{e}_\varphi}\hat{p} + b\dfrac{\bar{\kappa}_\theta\bar{\sigma}}{\bar{\kappa}_\theta\bar{\sigma}-\imath\omega}\nabla_{\boldsymbol{e}_\theta}\hat{p}. \end{cases} \tag{2.27}$$

The time-dependent PML equations are obtained using a Fourier transform on the time-harmonic ones. In order to perform this transform, additional fields must be defined. The system (2.27) then becomes

$$\begin{cases} -\imath\omega\hat{p} + a\nabla\cdot\hat{\boldsymbol{u}} = -\sigma\hat{p}_s - \bar{\kappa}_\varphi\bar{\sigma}\hat{p}_\varphi - \bar{\kappa}_\theta\bar{\sigma}\hat{p}_\theta,\\ -\imath\omega\hat{\boldsymbol{u}} + b\nabla\hat{p} = -\sigma\hat{\boldsymbol{u}}_s - \bar{\kappa}_\varphi\bar{\sigma}\hat{\boldsymbol{u}}_\varphi - \bar{\kappa}_\theta\bar{\sigma}\hat{\boldsymbol{u}}_\theta,\\ -\imath\omega\hat{p}_s + a\nabla_{\boldsymbol{e}_s}\cdot\hat{\boldsymbol{u}} = -\sigma\hat{p}_s,\\ -\imath\omega\hat{p}_\varphi + a\nabla_{\boldsymbol{e}_\varphi}\cdot\hat{\boldsymbol{u}} = -\bar{\kappa}_\varphi\bar{\sigma}\hat{p}_\varphi,\\ -\imath\omega\hat{p}_\theta + a\nabla_{\boldsymbol{e}_\theta}\cdot\hat{\boldsymbol{u}} = -\bar{\kappa}_\theta\bar{\sigma}\hat{p}_\theta,\\ -\imath\omega\hat{\boldsymbol{u}}_s + b\nabla_{\boldsymbol{e}_s}\hat{p} = -\sigma\hat{\boldsymbol{u}}_s,\\ -\imath\omega\hat{\boldsymbol{u}}_\varphi + b\nabla_{\boldsymbol{e}_\varphi}\hat{p} = -\bar{\kappa}_\varphi\bar{\sigma}\hat{\boldsymbol{u}}_\varphi,\\ -\imath\omega\hat{\boldsymbol{u}}_\theta + b\nabla_{\boldsymbol{e}_\theta}\hat{p} = -\bar{\kappa}_\theta\bar{\sigma}\hat{\boldsymbol{u}}_\theta. \end{cases}$$

This system can be written in the more compact form

$$\begin{cases} -\imath\omega\hat{p} + a\nabla\cdot\hat{\boldsymbol{u}} = -\sigma\hat{p}_s - \bar{\kappa}_\varphi\bar{\sigma}\hat{p}_\varphi - \bar{\kappa}_\theta\bar{\sigma}(\hat{p}-\hat{p}_s-\hat{p}_\varphi), \\ -\imath\omega\hat{\boldsymbol{u}} + b\nabla\hat{p} = -\sigma\boldsymbol{e}_s(\boldsymbol{e}_s\cdot\hat{\boldsymbol{u}}) - \bar{\kappa}_\varphi\bar{\sigma}\boldsymbol{e}_\varphi(\boldsymbol{e}_\varphi\cdot\hat{\boldsymbol{u}}) - \bar{\kappa}_\theta\bar{\sigma}\boldsymbol{e}_\theta(\boldsymbol{e}_\theta\cdot\hat{\boldsymbol{u}}), \\ -\imath\omega\hat{p}_s + a\nabla_{\boldsymbol{e}_s}\cdot\hat{\boldsymbol{u}} = -\sigma\hat{p}_s, \\ -\imath\omega\hat{p}_\varphi + a\nabla_{\boldsymbol{e}_\varphi}\cdot\hat{\boldsymbol{u}} = -\bar{\kappa}_\varphi\bar{\sigma}\hat{p}_\varphi, \end{cases}$$

where $\hat{\boldsymbol{u}}_s = \boldsymbol{e}_s(\boldsymbol{e}_s\cdot\hat{\boldsymbol{u}})$, $\hat{\boldsymbol{u}}_\varphi = \boldsymbol{e}_\varphi(\boldsymbol{e}_\varphi\cdot\hat{\boldsymbol{u}})$, $\hat{\boldsymbol{u}}_\theta = \boldsymbol{e}_\theta(\boldsymbol{e}_\theta\cdot\hat{\boldsymbol{u}})$ and $\hat{p} = \hat{p}_s + \hat{p}_\varphi + \hat{p}_\theta$. Finally, the inverse Fourier transform gives the time-dependent PML equations

$$\boxed{\begin{cases} \dfrac{\partial p}{\partial t} + a\nabla\cdot\boldsymbol{u} = -\sigma p_s - \bar{\kappa}_\varphi\bar{\sigma}p_\varphi - \bar{\kappa}_\theta\bar{\sigma}(p-p_s-p_\varphi), \\ \dfrac{\partial \boldsymbol{u}}{\partial t} + b\nabla p = -\sigma\boldsymbol{e}_s(\boldsymbol{e}_s\cdot\boldsymbol{u}) - \bar{\kappa}_\varphi\bar{\sigma}\boldsymbol{e}_\varphi(\boldsymbol{e}_\varphi\cdot\boldsymbol{u}) - \bar{\kappa}_\theta\bar{\sigma}\boldsymbol{e}_\theta(\boldsymbol{e}_\theta\cdot\boldsymbol{u}), \\ \dfrac{\partial p_s}{\partial t} + a\nabla_{\boldsymbol{e}_s}\cdot\boldsymbol{u} = -\sigma p_s, \\ \dfrac{\partial p_\varphi}{\partial t} + a\nabla_{\boldsymbol{e}_\varphi}\cdot\boldsymbol{u} = -\bar{\kappa}_\varphi\bar{\sigma}p_\varphi. \end{cases}} \tag{2.28}$$

In two dimensions, the system is reduced to

$$\begin{cases} \dfrac{\partial p}{\partial t} + a\nabla\cdot\boldsymbol{u} = -\sigma p_s - \bar{\kappa}\bar{\sigma}(p-p_s), \\ \dfrac{\partial \boldsymbol{u}}{\partial t} + b\nabla p = -\sigma\boldsymbol{e}_s(\boldsymbol{e}_s\cdot\boldsymbol{u}) - \bar{\kappa}\bar{\sigma}\boldsymbol{e}_\varphi(\boldsymbol{e}_\varphi\cdot\boldsymbol{u}), \\ \dfrac{\partial p_s}{\partial t} + a\nabla_{\boldsymbol{e}_s}\cdot\boldsymbol{u} = -\sigma p_s, \end{cases} \tag{2.29}$$

with $\bar{\kappa} = (\kappa^{-1}+r)^{-1}$, where $\kappa(\varphi)$ is the curvature of the curve $\mathscr{C}$. For a straight truncation surface in the direction $\boldsymbol{e}_s$, the system is, for both two and three dimensions,

$$\begin{cases} \dfrac{\partial p}{\partial t} + a\nabla\cdot\boldsymbol{u} = -\sigma p_s, \\ \dfrac{\partial \boldsymbol{u}}{\partial t} + b\nabla p = -\sigma\boldsymbol{e}_s(\boldsymbol{e}_s\cdot\boldsymbol{u}), \\ \dfrac{\partial p_s}{\partial t} + a\nabla_{\boldsymbol{e}_s}\cdot\boldsymbol{u} = -\sigma p_s. \end{cases} \tag{2.30}$$

In one dimension, the system (2.4) is recovered.

2.5 Conclusion

The perfectly matched layers (PMLs) exhibit very interesting properties for truncating an unbounded domain: the perfect matching and the perfect absorption. The artificial boundary of the truncated domain can then be theoretically perfectly modeled. This explains why many PMLs have been and continue to be developped for a wide range of wave-like problems.

In this chapter, we have reviewed some key elements about the PMLs and their design. In particular, some well-known interpretations used to build PMLs are explained: the change of the metric, the change of physical parameters or the coordinate stretch in the frequency domain. These viewpoints are usefull to develop novel efficient absorbing layers.

We have proposed a procedure to design PMLs for generally-shaped convex truncated domains with regular boundary for time-dependent problems. This permits a great flexibility when choosing the shape of the computational domain. While involving a stretch in a specific curvilinear coordinate system, the final PML equations (2.28) are easily written in cartesian coordinates, which facilitate their implementation in computational codes. The complete procedure is the main original result of this chapter. The effectiveness of obtained PMLs will be illustrated by numerical results in chapters 4 and 5.

CHAPTER 3

Optimizing Perfectly Matched Layers in Discrete Contexts

3.1 Introduction

In a discrete context, when a numerical method is used to solve the wave-like problem, the properties of the PML are altered [34, 167]. Indeed, the perfect matching is lost, and the perfect absorption is no longer ensured by the condition (2.13), valid only for the continuous problem. The PMLs are deteriorated by the discretization. In the worst cases, they become highly reflective.

Fortunately, spurious reflection of outgoing waves by a discretized PML can be reduced by adjusting the discretization and the PML parameters. To this aim, very thick layers, adaptive meshes [121], modified discrete schemes [60] and optimized absorption functions [34, 167] can be used. However, the first two approaches imply an increase of the computational cost, and the others are case-dependent. In particular, the absorption function $\sigma(x)$ is generally optimized by means of costly computational procedures or *a priori* error estimates [39, 83, 130, 142], whose result depends on both the equations and the numerical scheme considered.

Smoothly increasing absorption functions $\sigma(x)$ are commonly used with success in the litterature. Indeed, they provide a smooth and efficient damping of the outgoing waves in the layer. *Polynomial functions* are widely used (see *e.g.* [25, 65]), even if they require a case-dependent optimization of additional parameters. Alternatives are offered by the promising *hyperbolic functions* [26]. In a specific time-harmonic acoustic context with a continuous finite element scheme, Bermúdez et al. [26] showed that these functions do not require any tuning to provide PMLs as efficient as those obtained with optimized polynomial functions.

This chapter deals with the optimization of PMLs when numerical methods are used. Since the absorption function $\sigma(x)$ is a key parameter of the optimization problem, a large part of the chapter is dedicated to its study. In particular, the result of Bermúdez et al. [26] about hyperbolic functions is extended to the time domain. In order to draw conclusions as case-independent as possible, all simulations are systematically performed with different classical numerical methods, and interpretations are proposed.

This chapter is organized as follows. In section 3.2, we propose a preliminary analysis of the PML in one dimension. The change of properties of discretized PMLs are highlighted, together with the influence of the problem parameters. The specific role of the absorption function $\sigma(x)$ is shown and interpreted, together with the existence of optimum values in discrete contexts. In section 3.3, both polynomial and hyperbolic absorption functions are optimized and compared by means of one- and two-dimensional benchmarks. Thereby, the result of Bermúdez et al. [26] is extended to the time domain for different numerical methods. Finally, we propose an interpretation of this result.

3.2 Preliminary analysis of the PML in discrete contexts

When using a PML together with a numerical scheme, it is worthwhile to preserve its ability to simulate accurately the artificial boundary of the truncated domain. In discrete contexts, this ability can be studied from two complementary standpoints.

First, the different contributions to the error on the solution can be studied. Let us recall that, inside the truncated domain, the numerical solution must be as close as possible to the solution of the original continuous problem. The difference between these solutions (*i.e.* the *global error*) is due to two kinds of error, as illustrated in Figure 3.1:

- the *modeling error*, caused by the replacement of the original continuous problem (defined on the unbounded domain $\mathbb{R}^n$) by a modified version (defined on the truncated domain Ω surrounded by the layer Σ);
- the *numerical error*, generated by the discretization of the fields and the equations of the modified continuous problem (in both Ω and Σ).

The modeling error can be studied by using plane-wave analyses of continuous equations, as performed in section 2.3.2. In particular, the reflection coefficient r_{pml} associated to a PML of finite thickness, given by the expression (2.12), is a way to quantify this error. For a finite difference scheme, the numerical error can be evaluated using the truncation[1] error of the scheme.

[1] Here, the term 'truncation' refers to the replacement of infinite sums by finite ones. Truncated Taylor series are used to evaluate the error generated by the replacement of partial derivatives with finite differences in the scheme.

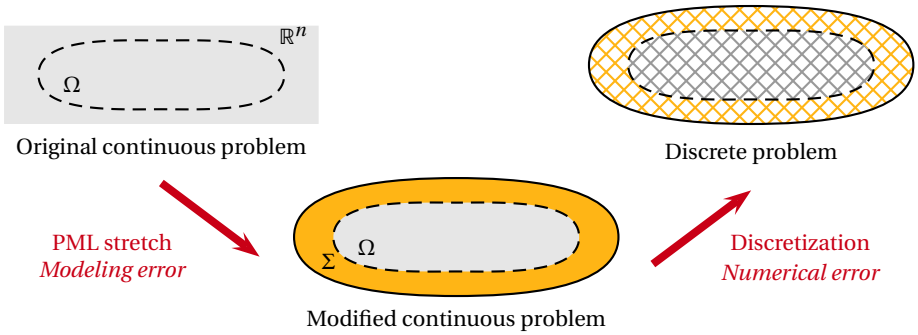


Figure 3.1: When using a PML together with a numerical scheme, the global error can be interpreted as the sum of modeling and numerical errors.

A complementary standpoint is to consider the discrete problem itself as a wave-like problem. Indeed, the discrete equations support discrete plane-wave solutions, which have properties different from their continuous versions. The purpose of a discretized PML is then to absorb as accurately as possible the discrete outgoing waves.

These standpoints lead to two ways to optimize PMLs: minimizing both modeling and numerical errors [130], or minimizing directly a discrete version of the reflection coefficient r_{pml} [39, 83, 142]. The former way has the advantage to highlight the different sources of error. With the latter, it is convenient to have to minimize a single quantity: the discrete reflection coefficient.

In this section, we study the ability of discretized PMLs to simulate an artificial boundary in a simple one-dimensional case. First, we derive and analyse the discrete reflection coefficients associated to discrete plane waves for the finite difference scheme. Then, our approach is validated and extended to other numerical methods by using numerical simulations in the time domain.

The problem considered hereafter is defined on the semi-infinite domain $\Omega = \mathbb{R}^-$, extended with a PML $\Sigma = [0, \delta]$. In both domain Ω and layer Σ, the fields $p(x,t)$ and $u(x,t)$ are governed by

$$\begin{cases} \dfrac{\partial p}{\partial t} + a\dfrac{\partial u}{\partial x} = -\sigma p, \\ \dfrac{\partial u}{\partial t} + b\dfrac{\partial p}{\partial x} = -\sigma u, \end{cases}$$

where the absorption function σ is equal to a constant value $\bar{\sigma}$ in Σ, and to zero in Ω. In this particular case, the reflection coefficient associated to the interface (or

an infinite layer) (2.11) remains

$$\boxed{r_{\text{interface}} = 0,} \tag{3.1}$$

while the reflection coefficient associated to a PML of thickness δ, that is given by the general expression (2.12), becomes

$$\boxed{r_{\text{pml}} = \exp\left[-\frac{2\bar{\sigma}\delta}{\sqrt{ab}}\right].} \tag{3.2}$$

3.2.1 Properties of the PML in the FD context

Let us consider a discretization of the problem defined above with the finite difference (FD) method. The fields are spatially discretized on staggered regular grids. The semi-discrete fields $\tilde{p}_{i+1/2}(t)$ and $\tilde{u}_i(t)$ take their values at the discrete points $x_{i+1/2} = (i+1/2)\Delta x$ and $x_i = i\Delta x$, respectively, where i is the spatial index and Δx is the spatial step. Using central finite differences, the governing semi-discrete equations read

$$\begin{cases} \dfrac{\mathrm{d}\tilde{p}_{i+1/2}}{\mathrm{d}t} + a\dfrac{\tilde{u}_{i+1} - \tilde{u}_i}{\Delta x} = -\sigma_{i+1/2}\,\tilde{p}_{i+1/2}, \\ \dfrac{\mathrm{d}\tilde{u}_i}{\mathrm{d}t} + b\dfrac{\tilde{p}_{i+1/2} - \tilde{p}_{i-1/2}}{\Delta x} = -\sigma_i\,\tilde{u}_i. \end{cases} \tag{3.3}$$

The discrete values of the absorption function, σ_i and $\sigma_{i+1/2}$, are equal to zero in the domain ($i < 0$) and to the constant value $\bar{\sigma}$ in the PML ($i \geq 0$).

Discrete plane-wave solution

As for the continuous case, some properties of the discrete problem can be studied by analysing its plane-wave solution. The complete analyses are in appendix A.

In the domain Ω, the elementary plane-wave solution of the discrete system (3.3) (without the sink terms) reads

$$\begin{aligned} \tilde{p}_{i+1/2}(t) &= P e^{\imath(k(i+1/2)\Delta x - \omega t)}, \\ \tilde{u}_i(t) &= U e^{\imath(ki\Delta x - \omega t)}, \end{aligned}$$

where the amplitudes P and U are linked by $P = \pm\sqrt{a/b}\,U$, while the wave number k (real) and the angular frequency ω (real and positive) are linked by the dispersion relation

$$k = \pm\frac{2}{\Delta x}\arcsin\left(\frac{\Delta x}{2}\frac{\omega}{c}\right), \tag{3.4}$$

with $c = \sqrt{ab}$. Since their phase velocity $\omega/|k|$ depends on the wave number, these plane waves are dispersive, contrary to their continuous versions. To recover the

continuous behavior of waves, and therefore the corresponding dispersion relation, the spatial step must be small in front of the wavelength, *i.e.* $\Delta x \ll c/\omega$. The discrete lattice must be able to reproduce the oscillatory behavior of waves.

In the layer Σ, the oscillatory behavior of discrete plane waves is altered by the absorption function, and the decay of these waves varies with the frequency, in contrast with the continuous case (see section 2.3.2). For a constant absorption function $\bar{\sigma}$, the elementary solution of the system can be written with damped plane waves

$$\tilde{p}_{i+1/2}(t) = Pe^{\imath(\beta(i+1/2)\Delta x - \omega t)},$$
$$\tilde{u}_i(t) = Ue^{\imath(\beta i \Delta x - \omega t)},$$

where β is a complex number. The real part of β is the wave number, while its imaginary part gives the attenuation factor. Injecting these solutions in the governing equations, one obtains the relation

$$\beta = \pm \frac{2}{\Delta x} \arcsin\left(\frac{\Delta x}{2}\frac{\omega + \imath\bar{\sigma}}{c}\right). \tag{3.5}$$

Both real and imaginary parts of β depend on ω and $\bar{\sigma}$. To recover the continuous dispersion relation, the spatial step Δx must be small in comparison with both c/ω and $c/\bar{\sigma}$. The first condition is the same as in the domain Ω (with $\sigma = 0$). The second condition can be interpreted considering the shape of the solution in the layer, illustrated in Figure 3.2. For a constant absorption function, the decay of waves is exponential, with the characteristic length $c/\bar{\sigma}$. Because this decay must be captured by the discrete mesh, the spatial step must be smaller than this characteristic length, *i.e.* $\Delta x \ll c/\bar{\sigma}$.

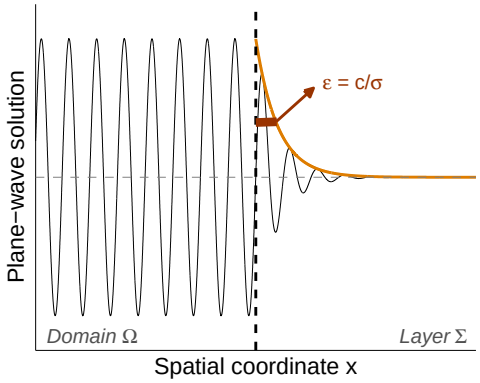


Figure 3.2: Illustration of the exponential decay of a plane wave in a one-dimensional PML for a constant absorption function $\sigma(x) = \bar{\sigma}$. The characteristic length of the decay is $c/\bar{\sigma}$.

Discrete reflection coefficients — Effectiveness of the discretized PML

Discrete versions of the reflection coefficients (3.1) and (3.2) are derived using the procedure already employed for the continuous case in section 2.3.2. They are representative of the effectiveness of the discretized layer for simulating the truncation of the domain. Assuming a semi-infinite domain terminated with a PML, a *discrete reflection coefficient* is the ratio of the amplitudes of an incident and a reflected discrete plane wave. The coefficient associated to the interface (or an infinite PML) is then

$$\boxed{r^\star_{\text{interf}} = \left| \frac{\left(\frac{\sigma\Delta x}{c} - \imath\frac{\omega\Delta x}{c}\right) + e^{\imath\beta\Delta x/2} - e^{-\imath k\Delta x/2}}{\left(\frac{\bar\sigma\Delta x}{c} - \imath\frac{\omega\Delta x}{c}\right) + e^{\imath\beta\Delta x/2} + e^{\imath k\Delta x/2}} \right|,}$$

while the one corresponding to a PML of thickness δ is

$$\boxed{r^\star_{\text{pml}} = \left| \frac{\left(\frac{\bar\sigma\Delta x}{c} - \imath\frac{\omega\Delta x}{c}\right) + \imath\frac{\cos(\beta\delta - \beta\Delta x/2)}{\sin(\beta\delta)} - e^{-\imath k\Delta x/2}}{\left(\frac{\bar\sigma\Delta x}{c} - \imath\frac{\omega\Delta x}{c}\right) + \imath\frac{\cos(\beta\delta - \beta\Delta x/2)}{\sin(\beta\delta)} + e^{\imath k\Delta x/2}} \right|.} \tag{3.6}$$

In these formulas, k and β are linked to ω by the dispersion relations (3.4) and (3.5) with the plus sign in either case. The complete developments are detailed in appendix A. The continuous and discrete versions of the reflection coefficients are plotted for a set of parameters in Figure 3.3.

For a small value of $\bar\sigma$, the behaviour of the continuous solution in the PML is accurately reproduced by the numerical scheme. Indeed, when $\bar\sigma < c/\delta$, the curves of the discrete reflection coefficients are close to those of the corresponding continuous ones (see Figure 3.3). The outgoing waves are then perfectly transmitted from the domain to the PML. However they are not sufficiently damped in both the continuous and the discrete cases. For a finite layer, the waves are indeed reflected by the outer boundary and come back in the domain. Therefore, the poor behavior of the PML is due to the *modeling error* that is already present in the continuous model. A larger $\bar\sigma$ is needed.

If $\bar\sigma$ is too large, the interface domain/layer is reflective, while it is perfectly mached in the continuous case. This spurious reflection of waves is caused by the *discretization error*. The characteristic length of the exponential decay of outgoing waves is lower than the spatial step, *i.e.* if $c/\bar\sigma < \Delta x$. As a consequence, the behavior of the solution cannot be reproduced by the discrete lattice, and numerical errors appear.

Therefore, a constant absorption function $\bar\sigma$ must be chosen in such a way to damp enough outgoing waves ($c/\bar\sigma < \delta$) without inducing a too sharp decrease of the fields in the PML ($c/\bar\sigma > \Delta x$). There exists an optimum value $\bar\sigma_{\text{opt}}$ that corresponds to a compromise. The value $\bar\sigma_{\text{opt}}$ minimizes (3.6). It is such that the char-

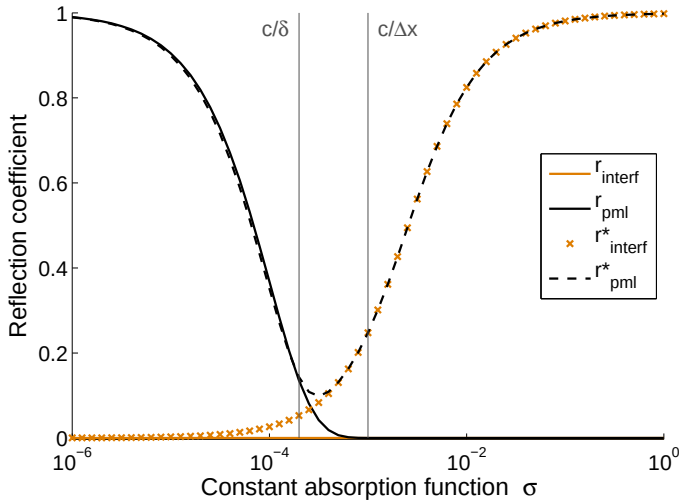


Figure 3.3: Reflection coefficients obtained with plane-wave analyses in both continuous (r_{interf} and r_{pml}) and discrete ($r^{\star}_{\text{interf}}$ and $r^{\star}_{\text{pml}}$) contexts, for the PML thickness $\delta = 5\,\Delta x$, the angular frequency $\omega = 2\pi/(13\,\Delta x)$ and the spatial step $\Delta x = 10^3$. All quantities are normalized to both velocity $\sqrt{ab}$ and impedance $\sqrt{a/b}$, and are then dimensionless.

acteristic length of the exponential decay $c/\bar{\sigma}_{\text{opt}}$ is in the range $[\Delta x, \delta]$, as shown in Figure 3.3.

The impact of the spatial step Δx and the thickness of the layer δ on the errors can be interpreted by analysing the curves in Figure 3.3. An increase of δ moves the curve $r_{\text{pml}}(\bar{\sigma})$ to the left. Similarly, a decrease of Δx moves the curve $r_{\text{interf}}(\bar{\sigma})$ to the right. Therefore, for a given value of $\bar{\sigma}$, they lead to a decrease of the modeling error and the numerical error, respectively. Both ways then decrease the total error, but they do not ensure a significant decrease. Indeed, an increase of δ (resp. decrease of Δx) does not improve significantly the layer if the total error is largely dominated by the numerical error (resp. modeling error).

As a conclusion, $\bar{\sigma}$, δ and Δx are key parameters that influence the effectiveness of the discretized PML. The modeling error is related to the value δ, while the numerical error strongly depends on Δx. The most appropriate approach to improve a PML is to choose δ and Δx, and after to find the corresponding optimum value of $\bar{\sigma}$. If the obtained total error is too large, the procedure must be repeated with a larger δ and/or a smaller Δx, and thus a larger computational cost.

3.2.2 Extension to other numerical methods

The behavior of the discretized PMLs is now studied using the results of time-dependent numerical simulations, with the FD methods as well as the FV, CG and DG ones. For convenience, all quantities are normalized to both velocity $\sqrt{ab}$ and impedance $\sqrt{a/b}$. They are then dimensionless.

One-dimensional benchmark

Let us adapt the previous time-harmonic benchmark to the time domain. A finite domain $\Omega = [-L, 0]$ is extended with the finite PML $\Sigma = [0, \delta]$. An incident Gaussian pulse is used as initial condition with

$$p(x,0) = \exp\left(-\frac{(x+L/4)^2}{R^2}\right),$$
$$u(x,0) = \exp\left(-\frac{(x+L/4)^2}{R^2}\right),$$

where R is a constant parameter, as illustrated in Figure 3.4. This pulse covers a broad range of frequencies, by contrast with the incident sine wave of the time-harmonic benchmark. It is initially centered at the middle of $[-L/2, 0]$. As time goes by, it moves towards the layer and is partly reflected. At the end of the simulation ($t_f = L/2$), the reflected part of the pulse is mainly in $[-L/2, 0]$. However, due to the numerical dispersion of the discrete scheme, some numerical modes can reach $[-L, -L/2]$ and the shape of the reflected signal can be non-Gaussian. Finally, the boundary condition $u = 0$ is used at both $x = -L$ and $x = \delta$.

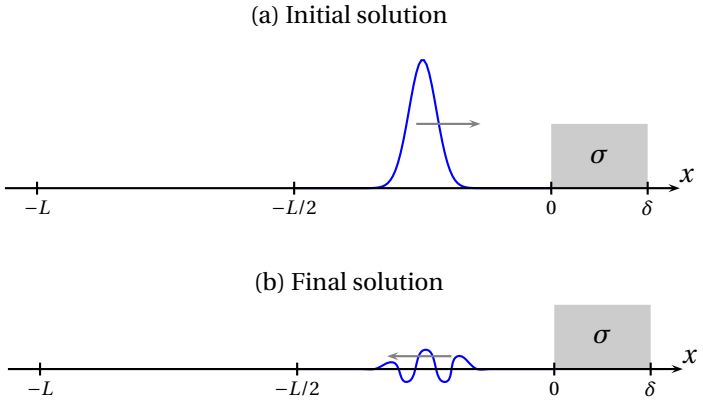


Figure 3.4: One-dimensional benchmark. Field $p(x,t)$ at the initial and final instants. A Gaussian-shaped pulse that moves to the right is prescribed by the initial condition (a). After a reflection by the layer, the reflected part of the pulse moves to the left. At the final instant, the reflected signal (possibly deformed) is mainly in the right part of the domain (b).

The reflection by the PML is quantified with the *relative error* ξ_r of the numerical solution defined by

$$\boxed{\xi_r = \sqrt{\frac{E_{\text{pml}}(t_f)}{E_{\text{wall}}(t_f)}},} \tag{3.7}$$

where $E_{\text{pml}}(t_f)$ is the *total energy* associated to the numerical solution in $[-L,0]$ at the end of the simulation, *i.e.* in the dimensionless case

$$E_{\text{pml}}(t_f) = \int_{-L}^{0} \left(\frac{1}{2} p^2(x,t_f) + \frac{1}{2} u^2(x,t_f) \right) \mathrm{d}x,$$

and $E_{\text{wall}}(t_f)$ is the one obtained by remplacing the PML with the perfectly reflecting boundary condition $u = 0$ at $x = 0$. The value of ξ_r can be then interpreted as the part of the total energy that is reflected by the layer. The value 0 corresponds to a perfectly absorbing layer, while 1 is for a perfectly reflective layer.

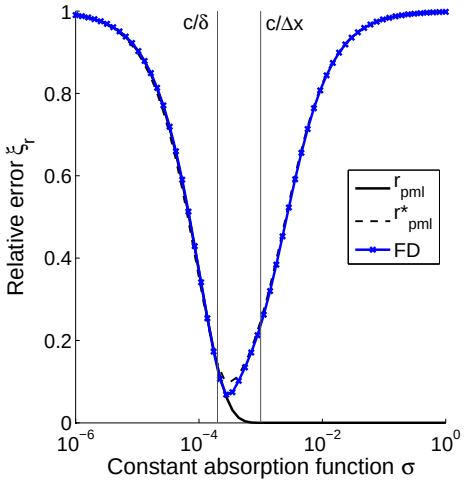
The discrete reflection coefficient $r^\star_{\text{pml}}$ and the relative error ξ_r are two complementary ways to quantify the effectiveness of a PML. The former is for a particular frequency (time-harmonic context), while the latter directly accounts for a range of frequencies through the initial incident pulse (time-dependent context). In the particular case where a harmonic plane wave is used instead of the Gaussian incident pulse, the relative error reduces to the discrete reflection coefficient, *i.e.* $\xi_r = r^\star_{\text{pml}}$.

Numerical results

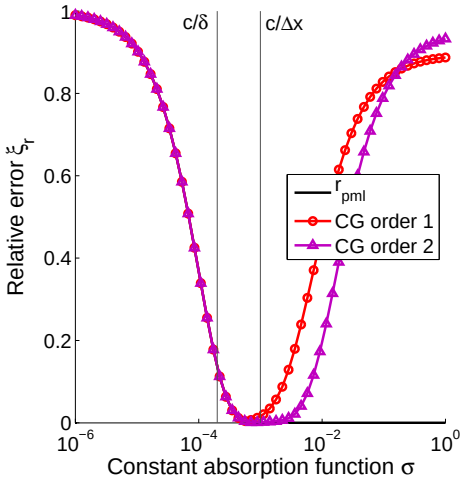
The relative error ξ_r has been computed for values of the parameter $\bar{\sigma}$ in $[10^{-6}, 1]$ with different numerical schemes in a unique setting. The curves of ξ_r as a function of $\bar{\sigma}$ are plotted in Figure 3.5 (decimal scale) and Figure 3.6 (logarithmic scale). We consider centered and upwind fluxes for FV and DG schemes, and first and second-order elements for CG and DG schemes. Let us recall that a DG scheme with zero-order elements reduces to a FV scheme. For all schemes, the spatial step is $\Delta x = 10^3$, the size of the domain is $L = 500\,\Delta x$, the thickness of the layer is $\delta = 5\,\Delta x$, and the parameter of the initial pulse is $R = 10\,\Delta x$. The time-integration is made using the Leapfrog scheme with $\Delta t = 0.5\,10^3$ (FD), or the Crank-Nicolson scheme with $\Delta t = 2.5 10^3$ (FV, CG and DG). To stabilize the CG scheme, PSPG terms are added in the formulation (see section 1.3.3.1) with a numerical diffusion parameter $\kappa = 10^3$.

The numerical results obtained with the FD scheme are consistent with the plane-wave analysis of the previous section. Indeed, in Figure 3.5(a), the curve of the relative error $\xi_r(\bar{\sigma})$ is close to the one of the discrete reflection coefficient $r^\star_{\text{pml}}(\bar{\sigma})$ obtained with the plane-wave analysis. The small difference can be explained by the difference of incident signal: a harmonic one with a single frequency for $r^\star_{\text{pml}}$ and a Gaussian one with a range of frequencies for ξ_r.

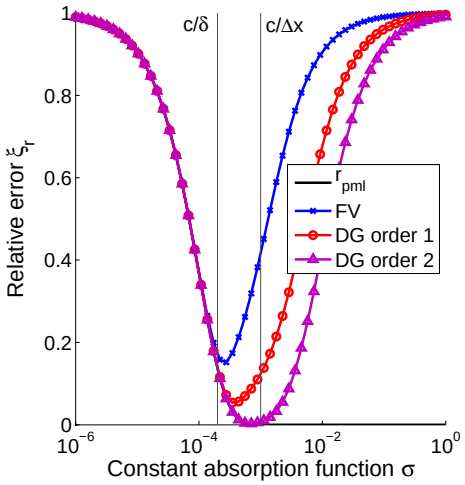
(a) FD scheme



(b) CG scheme



(c) FV/DG schemes with centered fluxes



(d) FV/DG schemes with upwind fluxes

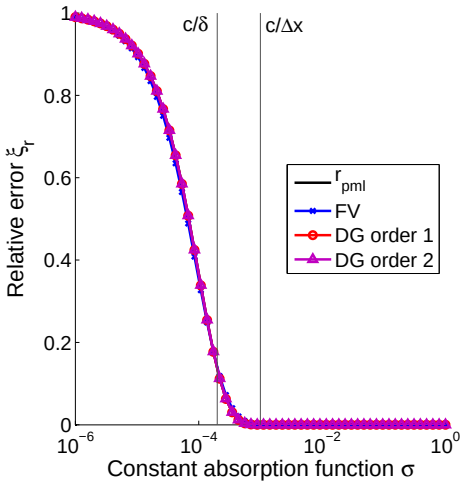


Figure 3.5: One-dimensional benchmark. Relative errors ξ_r of the one-dimensional benchmark as a function of the constant absorption function $\bar{\sigma}$ for different numerical methods.

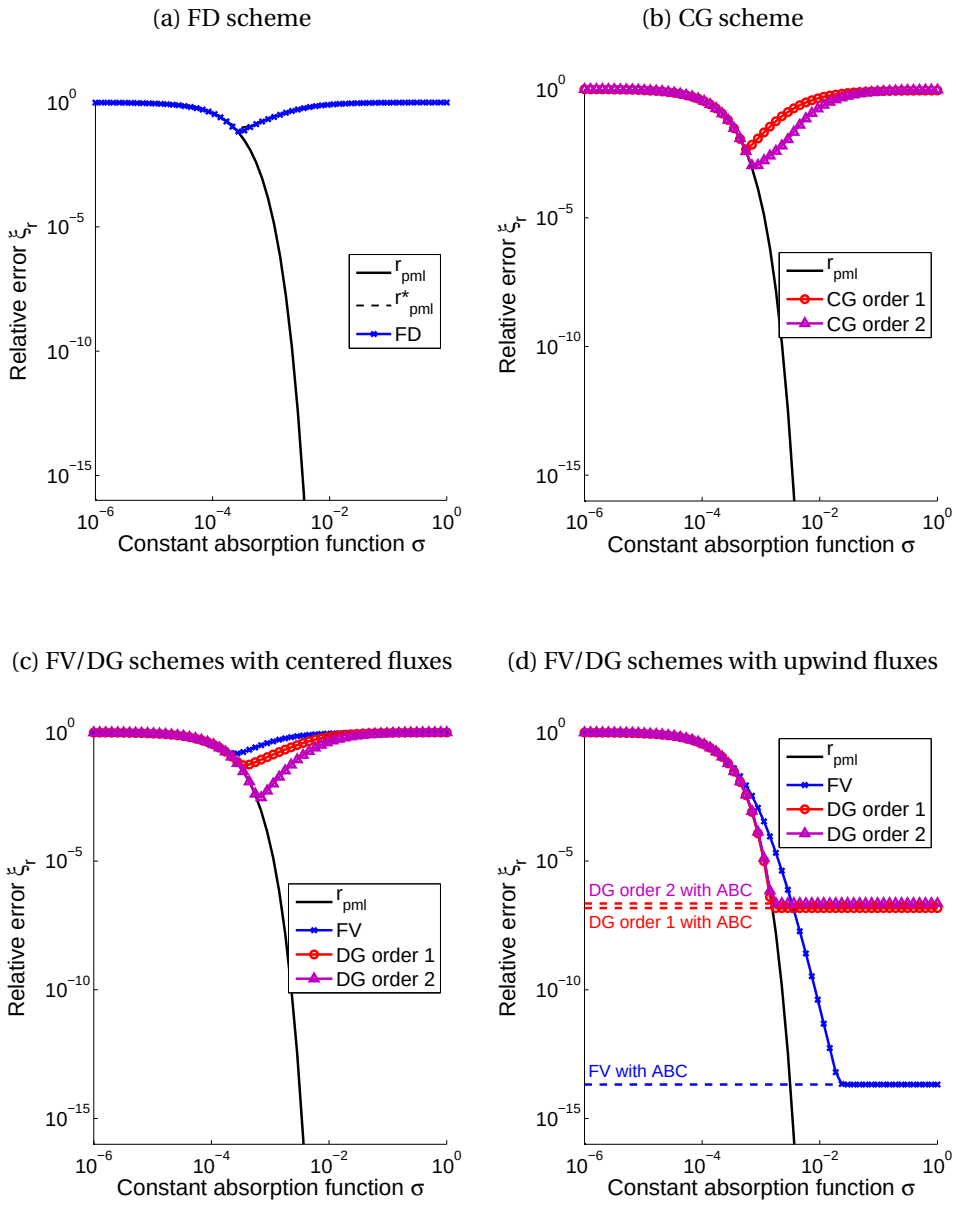


Figure 3.6: One-dimensional benchmark. Same as Figure 3.5, but with a logarithmic scale. In graph (d), horizontal lines are plotted at the positions corresponding to the relative errors obtained with the absorbing boundary condition (ABC) (2.3).

For $\bar{\sigma} \ll c/\delta$, the continuous solution is accurately reproduced by all schemes. Indeed, all curves of relative error are close to the one of the continuous reflection coefficient r_{pml}. Therefore, in each case, the outgoing pulse is not sufficiently damped by the layer and is reflected by the outer boundary. The error is due to the mathematical model.

For $\bar{\sigma} \gg c/\delta$, the exponential decay of outgoing waves is too sharp to be reproduced by the discrete mesh. In this case, the behavior of the discretized PML varies with the numerical scheme.

With the CG scheme and the FV/DG schemes (with centered fluxes), the curves $\xi_r(\bar{\sigma})$ look like the one obtained with the FD scheme (Figures 3.5(b) and 3.5(c)): there is an optimum value of $\bar{\sigma}$ and, beyond this value, the relative error increases with $\bar{\sigma}$. For very high values of $\bar{\sigma}$, the interface domain/layer becomes perfectly reflective. The optimum value of $\bar{\sigma}$ increases when the order of elements increases (see Figures 3.6(b) and 3.6(c)). This can be explained considering that a higher order method can more accurately reproduce the rapid variations of the solution. Therefore, a higher $\bar{\sigma}$, which reduces the characteristic length of the exponential decay, can be used.

While the reflected signal (not shown here) is a Gaussian pulse moving to the right with the CG scheme, it presents rapid oscillations that move faster than the physical velocity ($c = 1$ in this dimensionless context) with the FV/DG schemes, shown in Figure 3.7 for first-order elements. These rapid oscillations can be interpreted by considering that the fields are overdamped in the layer. At the interface, the centered numerical fluxes are then computed with both numerical fields equal to zero at the point of the interface on the side of the layer, *i.e.* at $x = 0^+$. This corresponds to prescribing two homogeneous Dirichlet boundary conditions (one on each field), while only one can be prescribed to have a well-posed problem. Therefore, the well-posedness is lost and instabilities are generated.

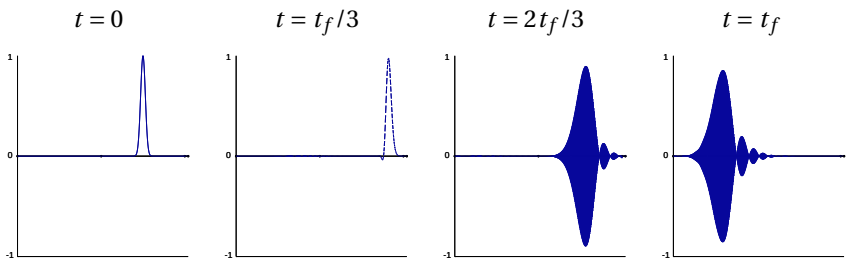


Figure 3.7: One-dimensional benchmark. Discrete field p at different instants of the simulation, obtained with the DG method, centered fluxes and first-order elements. A too large value of the absorption coefficient, $\bar{\sigma} = 10^{-1}$, is used. In the graphs, the horizontal axis represents the spatial coordinate, while the vertical one is for the value of the field.

With the FV/DG schemes and upwind fluxes, unlike with other schemes, the PML remains highly absorbing for large values of $\bar{\sigma}$ (see Figure 3.5(d)). This is due to the upwind fluxes, which are correctly computed at this interface thanks to the Riemann solver. This solver splits the parts of the solution that propagate inward and outward the domain. Outgoing information (*i.e.* outgoing waves) is correctly computed with the interior values. Incoming information (*i.e.* incoming waves), based on the overdamped solution of the layer, are then zero, which is exact. At its best, the PML reaches the accuracy of the absorbing boundary condition (2.3) (see Figure 3.6(d)). This condition is exact for the continuous problem with one spatial dimension (see section 2.2.1). The incoming Riemann invariant is prescribed to zero at the interface, such that no incoming waves enter in the domain. Unfortunately, as shown later, this ideal behavior is lost for multidimensional problems with oblique incident waves. Indeed, the absorbing boundary condition is then approximate and becomes reflective. Similarly, the PML becomes reflective for too large values of $\bar{\sigma}$.

All these numerical results corroborate the conclusion of the plane-wave analysis in the FD context, and extend it to other numerical methods. In nearly all cases, there is an optimum value of $\bar{\sigma}$ to use with the discretized PML. The important exception is a particular case that does not occur in multidimensional problems (for FV and DG schemes with upwind fluxes).

3.3 Spatially varying absorption functions

The effectiveness of a discretized PML is easily improved by using increasing absorption functions $\sigma(x)$ instead of constant ones. Indeed, such spatially varying functions provide decays of outgoing waves that are easier to approximate by the numerical methods than the exponential decay provided by constant functions.

This section deals with the choice of the spatially varying function to use as absorption function $\sigma(x)$. First, we consider this choice as a full optimization problem, where each discrete value of the function is a parameter to optimize. Although this approach is very case-dependent, it provides instructive information. After, polynomial and hyperbolic absorption functions are presented. The additional parameters introduced by these functions are studied. They are optimized by means of one- and two-dimensional benchmarks.

3.3.1 Full optimization of the discrete absorption function

The determination of the spatial distribution of the absorption function can be approached as a full optimization problem: each discrete value of the absorption function is a control variable.

Reusing the benchmark of section 3.2.2 in the one-dimensional FD context, the full optimization procedure consists in minimizing the relative error ξ_r (3.7). This

error is here a function of all the discrete values of the absorption function that are used in the discrete system (3.3), *i.e.* $\sigma_{1/2}$, σ_1, ..., $\sigma_{N_\delta-1}$ and $\sigma_{N_\delta-1/2}$, where N_δ is the number of cells in the layer. The optimization problem then reads

$$\begin{aligned} &\underline{\text{minimize}}_{\{\sigma_i\}} \quad \xi_r(\{\sigma_i\}), \\ &\underline{\text{such that}} \quad 0 \leq \sigma_{1/2} \leq \sigma_1 \leq ... \leq \sigma_{N_\delta-1/2}. \end{aligned}$$

The additional constraint, which enforces the absorption function to be positive and non-decreasing, has been added in order to improve the convergence of the numerical resolution. Here, the relative error ξ_r is computed considering the total energies in $[-L/2, 0]$ instead of $[-L, 0]$. This modification removes the influence of a numerical artefact introduced by the initial condition, which had a little importance until now.

The numerical resolution of the optimization problem using the optimization toolbox of Matlab® suffers from an irregular convergence when the PML is broad or when the model is triggered with a sine wave instead of a Gaussian signal. In either case, a nearly perfect behavior of the PML can be achieved with slightly different distributions of the absorption function, so that the optimum distribution is poorly defined. When the system is triggered with a sine wave, in particular, a nearly perfect behavior can be obtained even for thin absorbing layers, but the optimum distribution (not shown) depends strongly on the wave number and the frequency.

Figure 3.8 shows the optimum spatial distribution $\{\sigma_i\}$ obtained through numerical optimization for a 5-cell PML in the set-up considered in section 3.2.2. The gradual increase of the optimum distribution $\{\sigma_i\}$ in the PML introduces a progressive damping of the incoming wave and avoids therefore the development of large gradients in the solution. This distribution corresponds to the relative error $\xi_r = 3.0256\,10^{-6}$, which is far smaller than what can be obtained with optimized analytical functions (see values in Figure 3.13(a)). Therefore, even if there is no proof that the procedure converges to the global optimum distribution, the present result is already relevant.

3.3.2 Review of analytical absorption functions

Polynomial functions are widely used as absorption functions, *i.e.*

$$\sigma_n(x) = \alpha \left(\frac{x}{\delta}\right)^n,$$

where n and α are additional positive parameters. The parabolic function σ_2 (with $n = 2$) and the cubic one σ_3 (with $n = 3$) are the most frequently used (see *e.g.* [65] and the references of this thesis). As for a constant absorption function, the parameter α of polynomial functions must be large enough to damp the outgoing waves, and small enough to avoid a too sharp variation of the solution that cannot

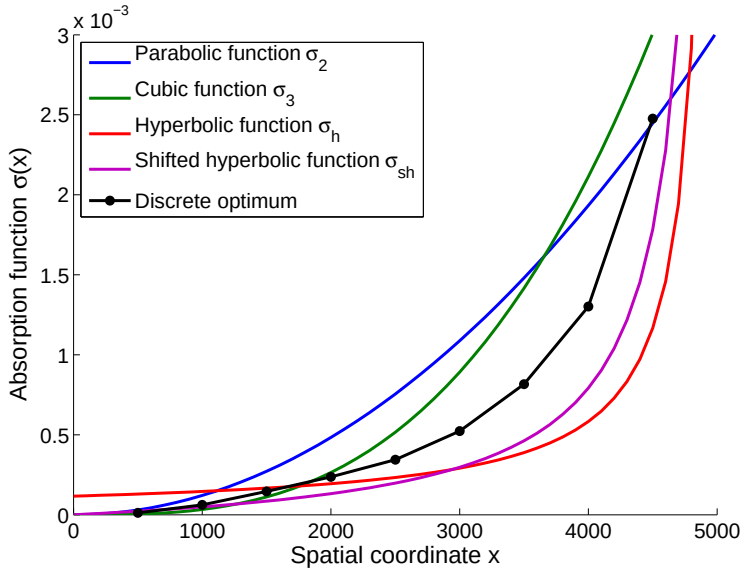


Figure 3.8: Optimized analytical absorption functions and discrete optimum distribution $\{\sigma_{1/2}, \sigma_1, ..., \sigma_{N_\delta-1}, \sigma_{N_\delta-1/2}\}$ for a 5-cell PML with the FD scheme.

be reproduced accurately by the numerical scheme. Nevertheless, the optimum α is not clearly identified. Currently, there exists no general rule to choose it.

As an alternative, *hyperbolic functions* have been proposed by Bermúdez et al. [26],

$$\sigma_{\mathrm{h}}(x) = \frac{\alpha}{\delta - x}, \qquad \sigma_{\mathrm{sh}}(x) = \frac{\alpha}{\delta - x} - \frac{\alpha}{\delta},$$

where α is an additional positive parameter to choose. The second function is called *shifted hyperbolic function*, because it is pushed down to ensure a zero value at the interface domain/layer (*i.e.* at $x = 0$). Both functions are singular at the outer border of the layer (*i.e.* at $x = \delta$). Therefore, since the condition (2.13) is met, the PML is theoretically perfectly absorbing in the continuous context.

In discrete contexts, the parameter α must be optimized for the same reason as with other functions. Fortunately, both functions are optimum with $\alpha \approx c$ (or $\alpha \approx 1$ in the dimensionless context). This was first shown by Bermúdez et al. [26] in a specific time-harmonic acoustic context with the CG method. This result is extended hereafter in time-dependent contexts with different numerical methods.

Let us note that, when a finite element method is used, the numerical integration over an element close to the boundary is tricky because of the large values of σ. This issue can be fixed by adapting the numerical integration [26]. In this work, no particular difficulty has been encountered for both discontinuous Galerkin (DG) and continuous Galerkin (CG) finite element schemes.

3.3.3 Optimization of analytical absorption functions

In this section, polynomial and hyperbolic functions are optimized and compared using one- and two-dimensional benchmarks. The one-dimensional benchmark deals only with waves that are normal to the interface domain/layer, while waves with oblique incidences are considered in the two-dimensional one.

3.3.3.1 Optimization for normal waves

Let us consider again the one-dimensional benchmark of section 3.2.2, with FD, CG and DG schemes. Only first-order elements are employed for the finite element schemes (CG and DG). Figures 3.9, 3.10, 3.11 and 3.12 show the relative error ξ_r as a function of the parameter α for each absorption function and each scheme in different settings. Two sets of numerical parameters are considered for each numerical scheme: those of section 3.2.2 (labelled 'set A') and more accurate ones (labelled 'set B'). In each case, two PML thicknesses δ are used. The values of parameters are given in the caption of figures.

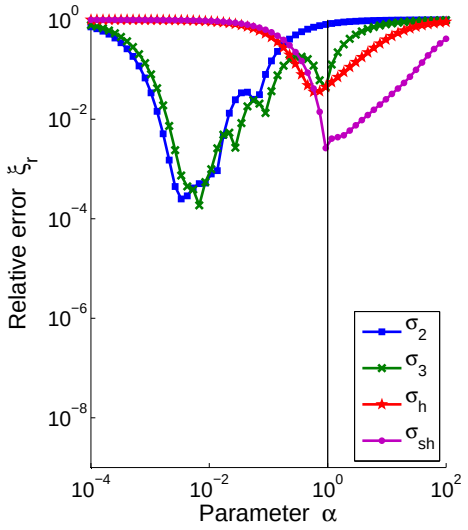
For the PML discretized with the DG scheme and upwind fluxes, the result obtained with a constant absorption function is recovered: the PML remains highly absorbing with very high value of α (see Figure 3.12), by contrast with other numerical schemes. At its best, the PML is as accurate as the absorbing boundary condition (2.3), which is exact in this one-dimensional context.

With other numerical schemes, the curves $\xi_r(\alpha)$ can present several minima. A systematic optimization of absorption functions must therefore be carefully done. For polynomial functions, the best α, which corresponds to the global minimum of each curve, is always in the range $10^{-3}-10^{-2}$, except when the DG scheme with centered fluxes is used. In this case, the best α may have very different values: close to 10^{-3}, 10^{0} or 10^{2} (see Figure 3.11). By contrast, for the hyperbolic functions σ_{h} and σ_{sh}, it is always close to 1.

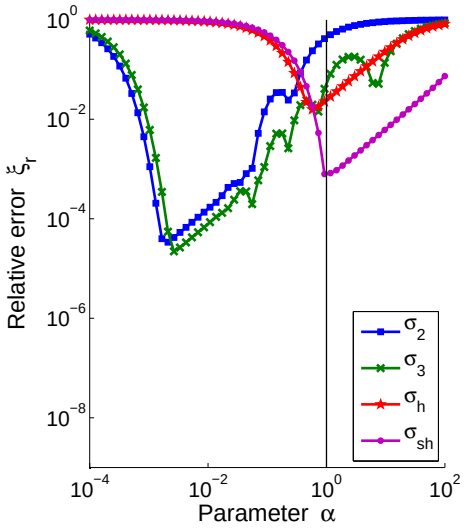
Figure 3.13 shows the minimum ξ_r and the corresponding optimum α as functions of the number of spatial steps in the layer. They are obtained using a systematic procedure with a heuristic minimum search algorithm. The optimum α is searched in the range $10^{-4}-10^{-2}$ for polynomial functions and $10^{-1}-10^{1}$ for hyperbolic ones. Although there is no guarantee that the obtained optimum values correspond to the global minimum of $\xi_r(\alpha)$, the behavior of all curves is coherent: by increasing the thickness of the layer, the minimum relative error ξ_r decreases.

For each absorption function, the optimum value of α is similar in all discrete contexts (Figure 3.13, graphs on the right). It remains in the range $10^{-3}-10^{-2}$ for polynomial functions (σ_2 and σ_3), and close to 1 for hyperbolic ones (σ_{h} and σ_{sh}). In the first case, this optimum value varies slightly following the numerical scheme, and diminishes when the layer thickness increases. By constrast, it is almost always constant with the hyperbolic functions.

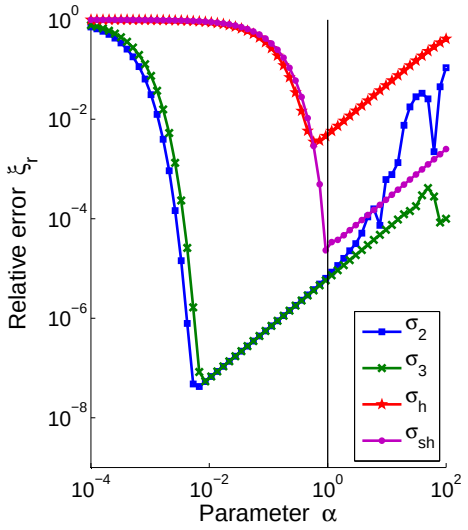
(a) FD scheme with $\delta = 5\,10^3$ and set A



(b) FD scheme with $\delta = 10^4$ and set A



(c) FD scheme with $\delta = 5\,10^3$ and set B



(d) FD scheme with $\delta = 10^4$ and set B

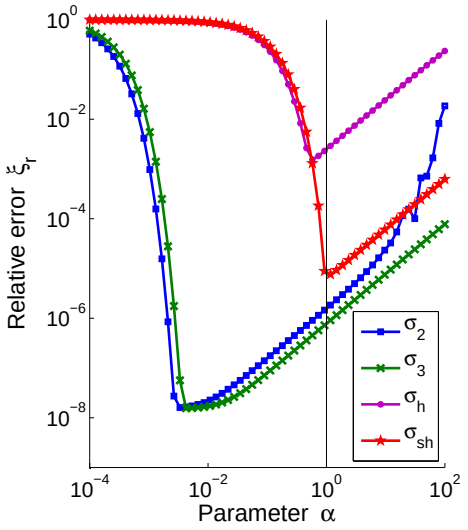
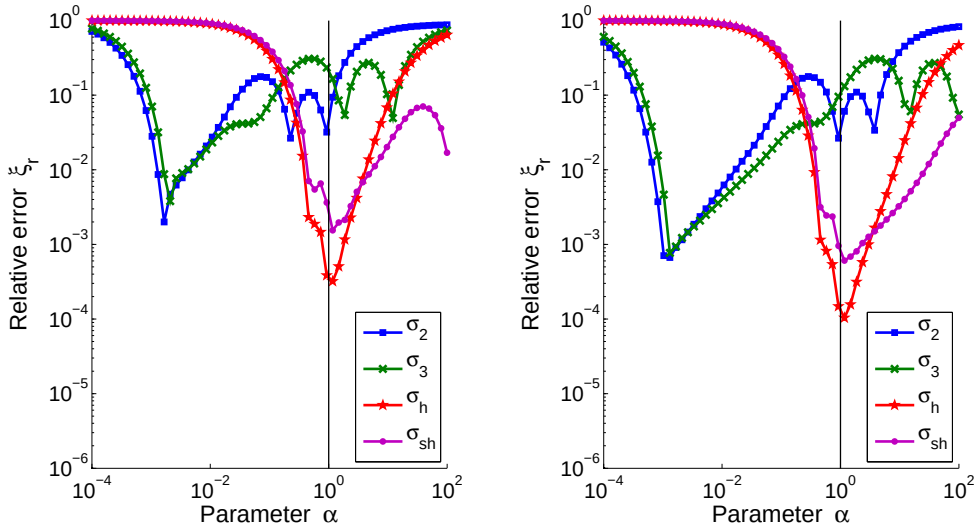


Figure 3.9: One-dimensional benchmark with the finite difference (FD) scheme. Relative error ξ_r as a function of the parameter α for different absorption functions $\sigma(x)$ and two PML thicknesses δ. In each case, two sets of numerical parameters are considered: $\Delta x = 10^3$, $\Delta t = 500$ (set A) and $\Delta x = 100$, $\Delta t = 50$ (set B). All values are dimensionless.

(a) CG scheme with $\delta = 5\,10^3$ and set A

(b) CG scheme with $\delta = 10^4$ and set A



(c) CG scheme with $\delta = 5\,10^3$ and set B

(d) CG scheme with $\delta = 10^4$ and set B

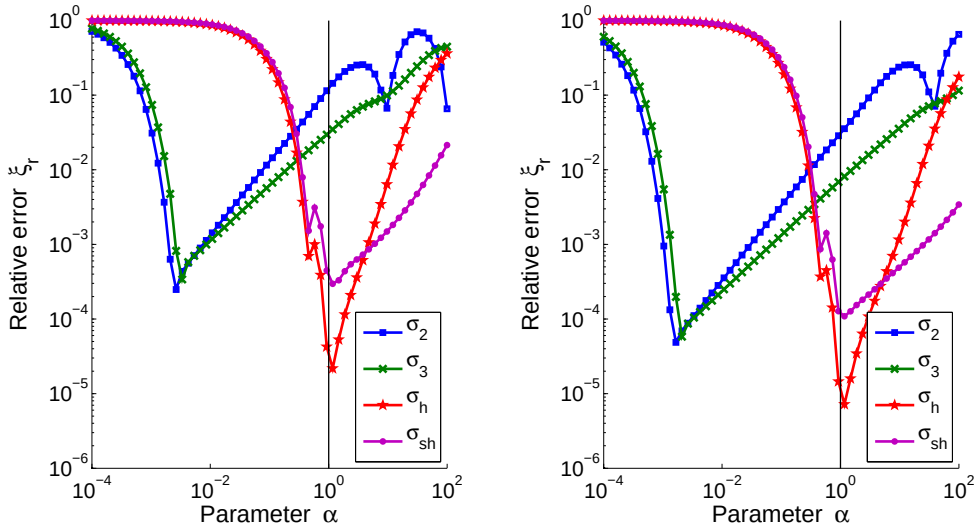
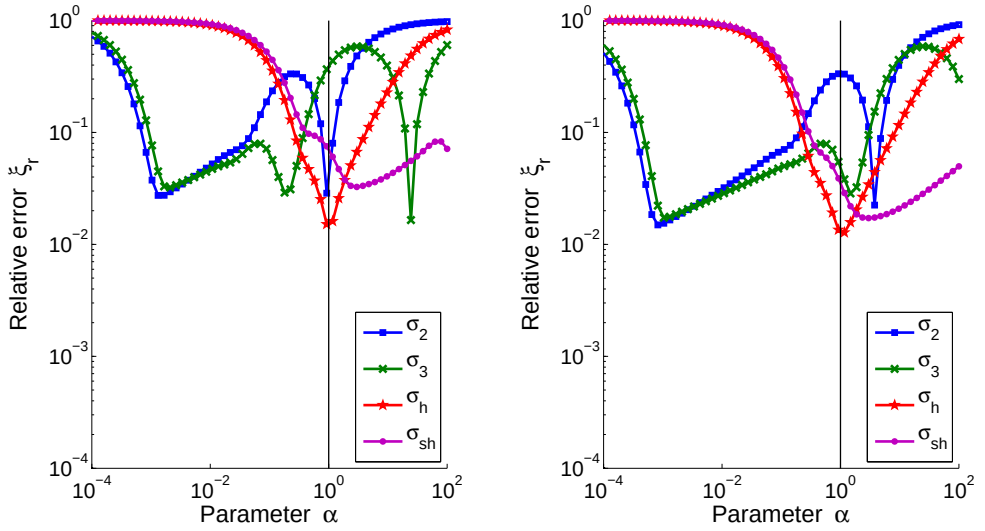


Figure 3.10: One-dimensional benchmark with the continuous Galerkin (CG) scheme. Relative error ξ_r as a function of the parameter α for different absorption functions $\sigma(x)$ and two PML thicknesses δ. In each case, two sets of numerical parameters are considered: $\Delta x = 10^3$, $\Delta t = 2.5\,10^3$ (set A) and $\Delta x = 250$, $\Delta t = 625$ (set B). $\kappa = 1000$ for both settings. All values are dimensionless.

(a) DG scheme with centered fluxes, $\delta = 5\,10^3$ and set A

(b) DG scheme with centered fluxes, $\delta = 10^4$ and set A



(c) DG scheme with centered fluxes, $\delta = 5\,10^3$ and set B

(d) DG scheme with centered fluxes, $\delta = 10^4$ and set B

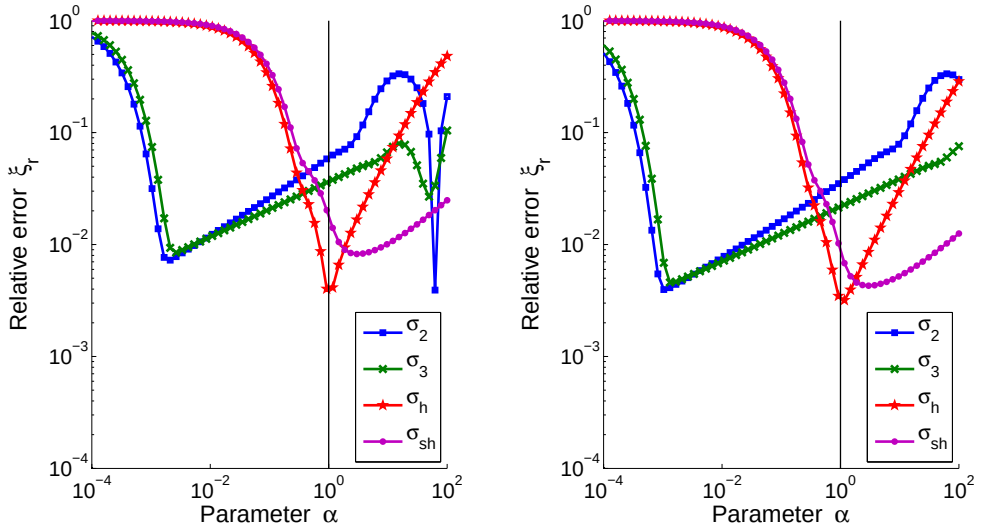
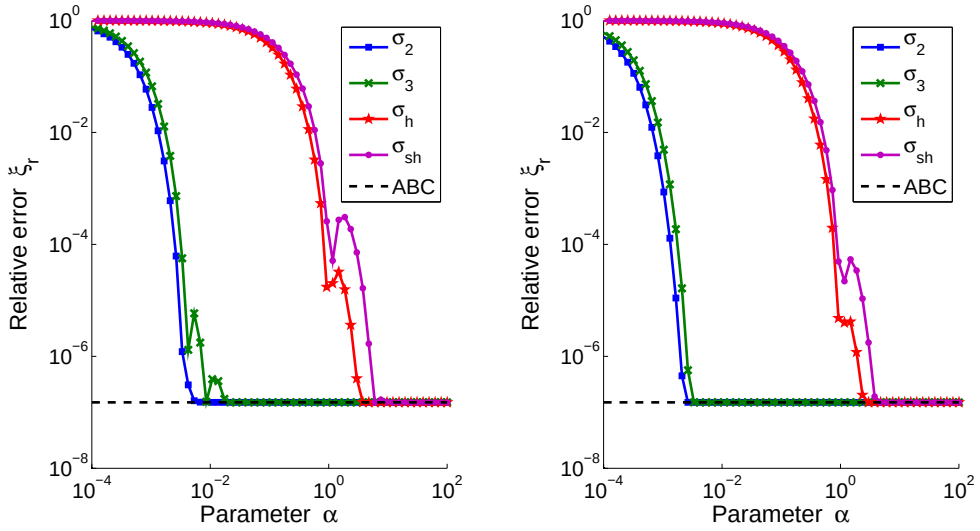


Figure 3.11: One-dimensional benchmark with the discontinuous Galerkin (DG) scheme with centered fluxes. Relative error ξ_r as a function of the parameter α for different absorption functions $\sigma(x)$ and two PML thicknesses δ. In each case, two sets of numerical parameters are considered: $\Delta x = 10^3$, $\Delta t = 2.5\,10^3$ (set A) and $\Delta x = 250$, $\Delta t = 625$ (set B). All values are dimensionless.

(a) DG scheme with upwind fluxes, $\delta = 5\,10^3$ and set A

(b) DG scheme with upwind fluxes, $\delta = 10^4$ and set A



(c) DG scheme with upwind fluxes, $\delta = 5\,10^3$ and set B

(d) DG scheme with upwind fluxes, $\delta = 10^4$ and set B

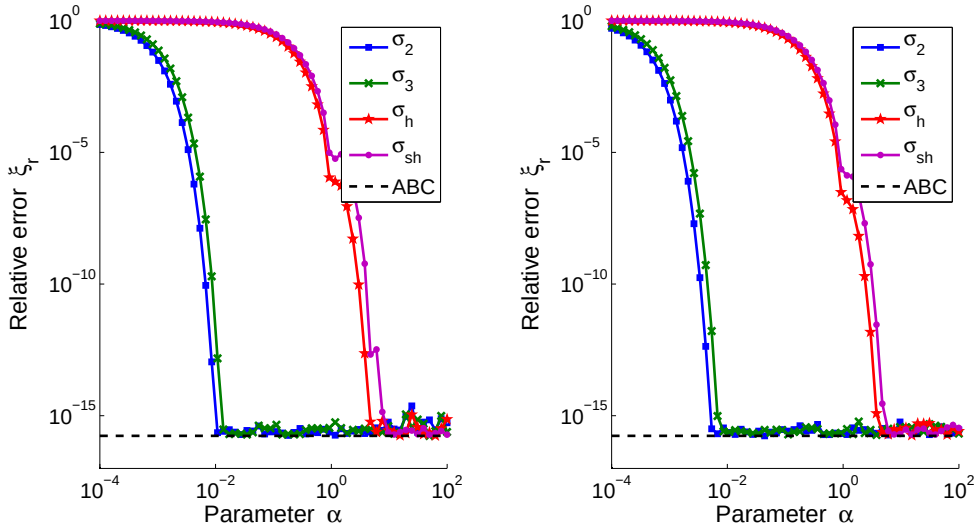
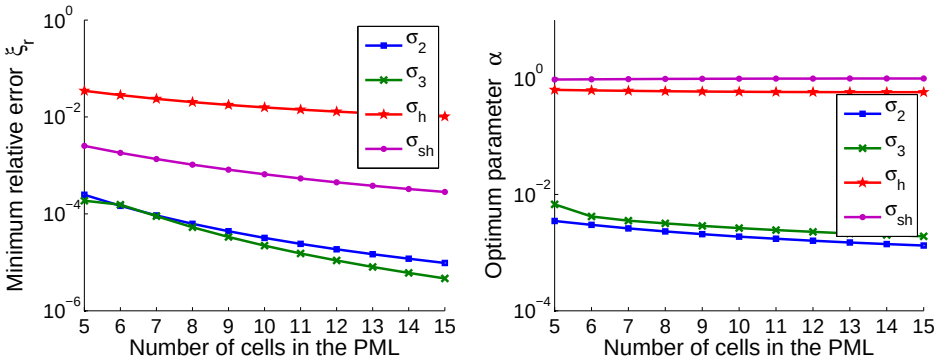
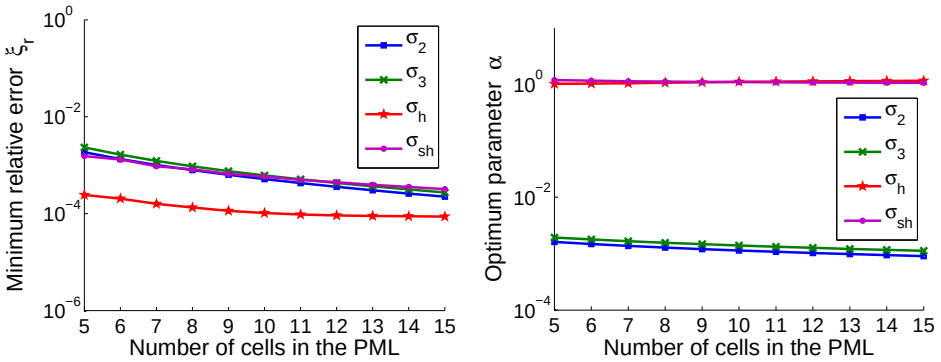


Figure 3.12: One-dimensional benchmark with the discontinuous Galerkin (DG) scheme with upwind fluxes. Relative error ξ_r as a function of the parameter α for different absorption functions $\sigma(x)$ and two PML thicknesses δ. In each case, two sets of numerical parameters are considered: $\Delta x = 10^3$, $\Delta t = 2.5\,10^3$ (set A) and $\Delta x = 250$, $\Delta t = 625$ (set B). The position of the dashed line corresponds to the relative error obtained with the absorbing boundary condition (ABC) (2.3). All values are dimensionless.

(a) FD scheme



(b) CG scheme



(c) DG scheme with centered fluxes

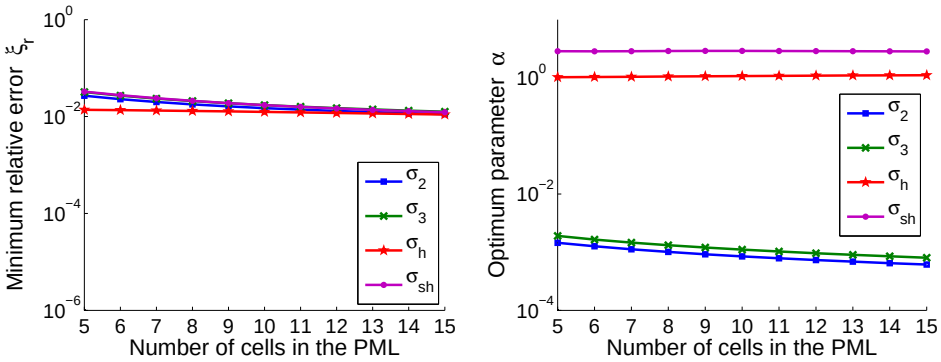


Figure 3.13: One-dimensional benchmark. Minimum relative error ξ_r (left) and corresponding optimum parameter α (right) as a function of the PML thickness δ for the different absorption functions $\sigma(x)$. When δ increases, the spatial step Δx remains constant and the number of cells in the PML increases. For all schemes, $\Delta x = 10^3$. The time step Δt is 500 (FD) or $2.5\,10^3$ (CG and DG). All values are dimensionless.

The effectiveness of the different absorption functions at their best, *i.e.* when the optimum α is used, changes depending on the numerical method (Figure 3.13, graphs on the left). With the FD scheme, the optimized polynomial functions give the smallest relative error. Among the hyperbolic functions, the shifted one σ_{sh} works better than the other. With the CG scheme, σ_{h} outperforms all other optimized functions, which give equivalent results. Finally, with the DG scheme and centered fluxes, all functions give a relative error of the same order of magnitude.

3.3.3.2 Optimization for oblique waves

In order to test the PML with waves with oblique incidences, a two-dimensional benchmark is now considered.

Two-dimensional benchmark

A Gaussian-shaped pulse is initially prescribed in the center of a squared domain using the initial condition on $p(\boldsymbol{x}, t)$:

$$p(\boldsymbol{x}, 0) = \exp\left(-\frac{\|\boldsymbol{x}\|^2}{r^2}\right),$$

where r is a parameter. The field $\boldsymbol{u}(\boldsymbol{x}, t)$ is equal to zero everywhere at $t = 0$. As time goes by, the pulse collapses and circular waves appear. The simulation ends when the main wave front reaches the borders of the square (see the reference solution in Figure 3.14). The homogeneous Dirichlet condition $\boldsymbol{u} \cdot \boldsymbol{n} = 0$ is prescribed at each border of the square, where $\boldsymbol{n}$ is the outward normal.

The PML is tested considering a modified version of the problem with a truncated domain: the upper part of the squared domain is removed and replaced with a PML Σ (see Figure 3.14). The goal is to reproduce the reference solution in the truncated domain Ω. Let us note that, even if the PML is normally used to truncate an unbounded domain, it can replace the removed area of the bounded squared domain in this case. Indeed, during the simulation, the reference solution in the truncated domain is not influenced by the upper boundary of the original squared domain.

The accuracy of the PML is again quantified by using the relative error ξ_r defined as

$$\boxed{\xi_r = \sqrt{\frac{E_{\text{error with pml}}(t_f)}{E_{\text{error with wall}}(t_f)}},}$$

except that the total energies $E_{\text{error with pml}}(t)$ and $E_{\text{error with wall}}(t)$ are now associated to the error on the fields in the truncated domain Ω. The former energy corresponds to the numerical solution obtained with the PML, while the latter is

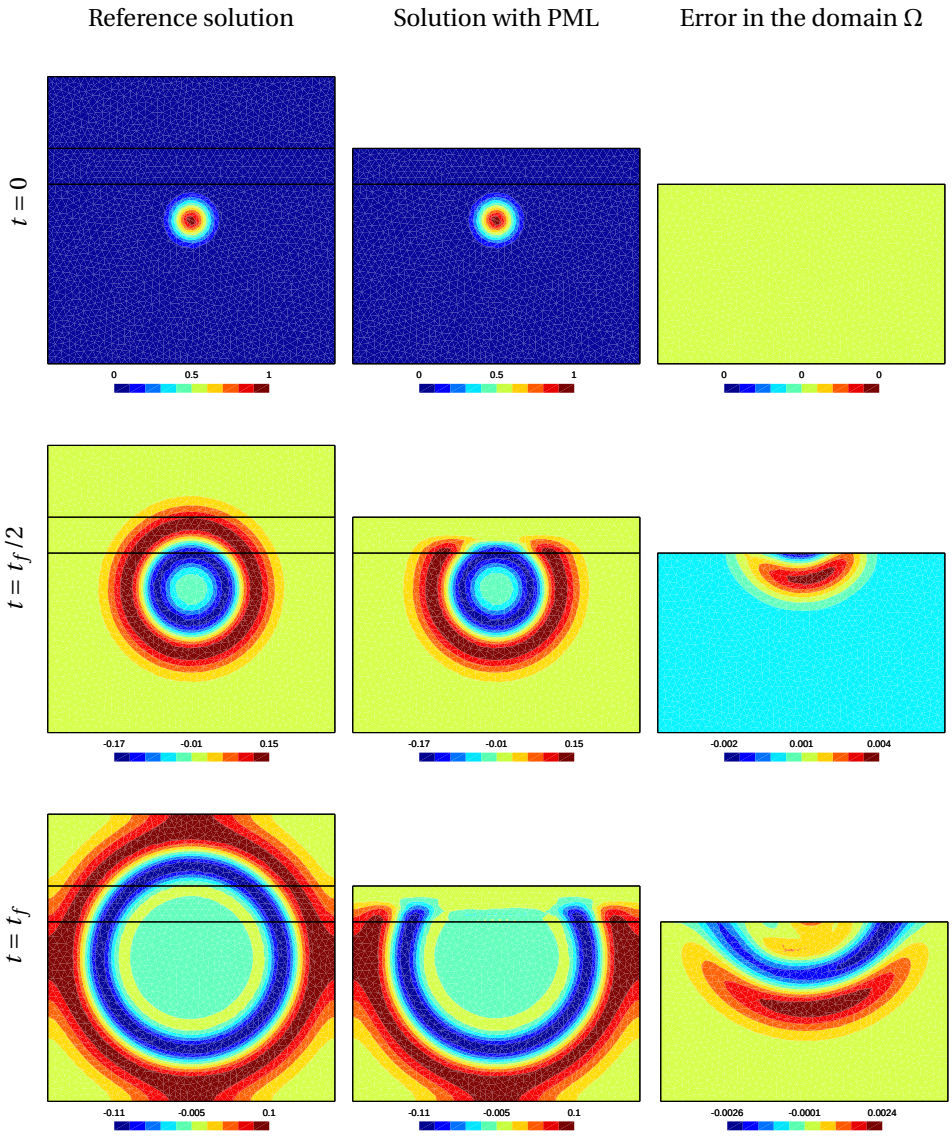


Figure 3.14: Two-dimensional benchmark. Field $p(\boldsymbol{x}, t)$ at different instants of the simulation in the reference domain (left) and the truncated domain Ω extended with a PML Σ (center). The difference between these two solutions in the truncated domain is plotted on the right.

obtained in the case where the interface is perfectly reflective ($\boldsymbol{u}\cdot\boldsymbol{n}=0$ is prescribed at the interface). In this dimensionless context, the first total energy reads

$$E_{\text{error with PML}}(t)=\int_{\Omega}\Big(\big(p_{\text{pml}}(\boldsymbol{x},t)-p_{\text{ref}}(\boldsymbol{x},t)\big)^2+\big\|\boldsymbol{u}_{\text{pml}}(\boldsymbol{x},t)-\boldsymbol{u}_{\text{ref}}(\boldsymbol{x},t)\big\|^2\Big)\,\mathrm{d}\boldsymbol{x},$$

where the numerical fields p_{pml} and $\boldsymbol{u}_{\text{pml}}$ are obtained with the PML, while p_{ref} and $\boldsymbol{u}_{\text{ref}}$ correspond to the reference numerical solution obtained with the original squared domain (see Figure 3.14, left). The layer is then perfectly absorbing for $\xi_r=0$, and perfectly reflective for $\xi_r=1$. In the latter case, the energy reflected by the layer is indeed the same as with a perfectly reflective boundary condition.

As for the one-dimensional benchmark, the relative error ξ_r is computed with numerous values of the parameter α for the different absorption functions and different numerical schemes. For all simulations, the spatial dimensions are $8\,10^4\times 8\,10^4$ for the reference squared domain, and $8\,10^4\times 5\,10^4$ for the truncated domain. We consider the PML thickness $\delta=10^4$, the characteristic length of the Gaussian-shaped pulse $r=5\,10^3$, and the simulation duration is $t_f=3.6\,10^3$. The numerical schemes used are described in section 4.3.2 (DG scheme) and in appendix B.2 (FD and CG schemes). For the FD scheme, the spatial steps are $\Delta x=\Delta y=500$ and the Leapfrog time-stepping is used with $\Delta t=300$. For both CG and DG schemes, the mesh is made of first-order triangular elements with the characteristic size $\ell=1.75\,10^3$. The time-stepping Crank-Nicolson scheme is used with the time step $\Delta t=3\,10^3$. We consider two values of the stabilization parameter κ for the CG scheme, and centered and upwind fluxes for the DG scheme.

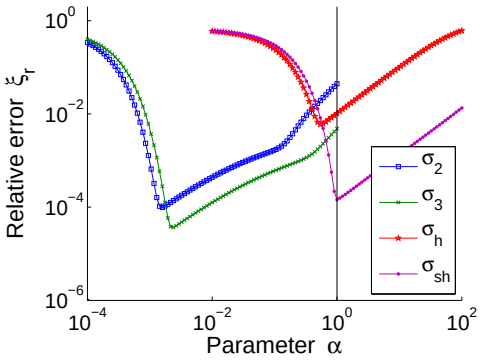
Numerical results

Figure 3.15 shows the relative error ξ_r as a function of the parameter α for the different absorption functions in different numerical contexts.

Contrary to the one-dimensional case, the PML discretized with the DG scheme and upwind fluxes is reflective for large values of α (see Figure 3.15(e)). There is therefore an optimum α (at the minimum of each curve) for each absorption function, like with the other numerical schemes. In each case, the minimum relative error ξ_r is smaller than the one obtained with the ABC (2.3). This is due to the oblique outgoing waves of the two-dimensional benchmark, for which the ABC is only approximate. The PML is more accurate in this context, but it requires to choose a value of α neither large nor small.

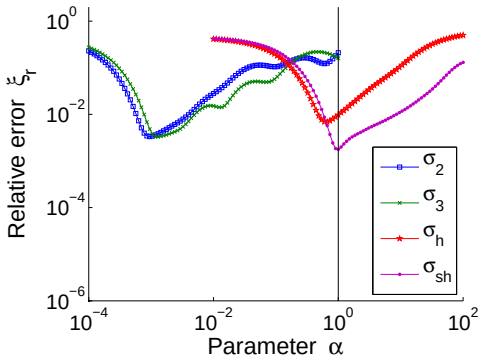
Another difference with the one-dimensional case: the hyperbolic function σ_{h} provides not so good results than other absorption functions. Indeed, when all functions are optimized, the relative error corresponding to σ_{h} is the worst in every discrete context, while this function was efficient with both CG and DG schemes in one dimension. Moreover, when the CG sheme is used, the function σ_{h} gets worse when the stabilization parameter κ increases, while the performance of

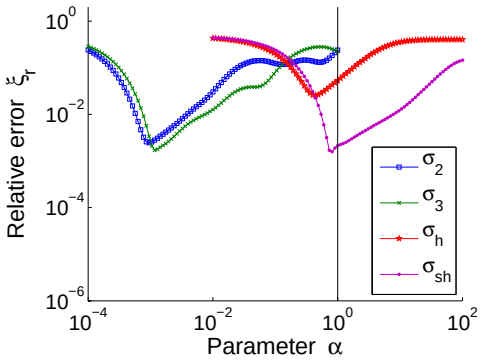
(a) FD scheme



(b) CG scheme with $\kappa = 200$

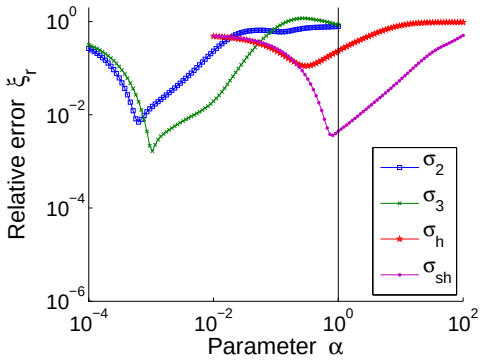
(c) CG scheme with $\kappa = 10^3$





(d) DG scheme with centered fluxes

(e) DG scheme with upwind fluxes



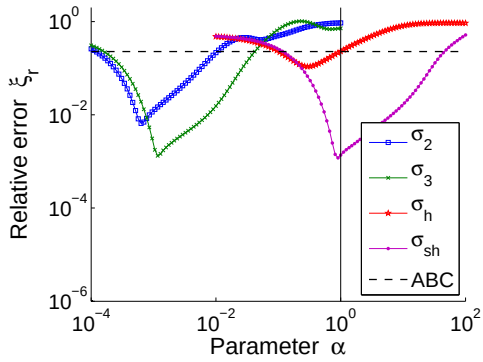


Figure 3.15: Two-dimensional benchmark. Relative error ξ_r as a function of the parameter α for the different absorption functions in different numerical contexts. In graph (e), the position of the horizontal line indicates the relative error obtained with the absorbing boundary condition (ABC) (2.3).

other functions is not affected (see Figure 3.15(c)-(d)). Therefore, σ_{h} is less robust than others.

The cubic function σ_3 and the shifted hyperbolic one σ_{sh} exhibit close relative error ξ_r when they are optimized. The former is slightly better than the latter for the FD scheme and the DG scheme with centered fluxes. The converse holds for the CG scheme with $\kappa = 200$. Between polynomial functions, the cubic one σ_3 is always better than the parabolic one σ_2.

As for the one-dimensional case, the optimum α of the shifted hyperbolic function σ_{sh} is very close to 1. By contrast, for the hyperbolic function σ_{h}, the optimum value is now systematically smaller than 1. Finally, for polynomial functions σ_2 and σ_3, the optimum α is always close to 10^{-3}. This optimum value is coherent with those obtained in one dimension with the same PML thickness (see Figure 3.13 for a PML with 10 cells), while being currently devoided of any interpretation.

3.3.3.3 Interpretation

The performance of hyperbolic functions and the optimum value of α can be interpreted by considering the particular shape of the obtained solution. Indeed, since there is no modeling error with the hyperbolic functions for all $\alpha > 0$ (see section 3.3.2), the total error is entirely due to the numerical error. Since this numerical error depends on the ability of the numerical scheme to reproduce the solution (see section 3.2), the optimum parameter α corresponds to the solution shape that can be described with the best accuracy in a given numerical context.

For a straight PML that damps in the direction $\boldsymbol{e}_x$, the plane-wave solution (2.9) becomes, for the hyperbolic functions σ_{h} with $\alpha = c$,

$$p(x,t) = Pe^{\imath(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}\left[\left(1-\frac{x}{\delta}\right)\right]^{\cos\theta} \tag{3.8}$$

and, for the shifted hyperbolic function σ_{sh} with $\alpha = c$,

$$p(x,t) = Pe^{\imath(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}\left[e^{x/\delta}\left(1-\frac{x}{\delta}\right)\right]^{\cos\theta}.$$

For waves with a normal incidence (*i.e.* $\theta = 0$), the decay of plane waves is then respectively linear and exponential-linear. These plane-wave solutions are illustrated in Figure 3.16.

The linear decay of plane waves provided by the hyperbolic function σ_{h} is ideal. Indeed, linear variations are perfectly represented by numerical schemes built on regular grids. Here, the linear variation modulates the wave oscillation (part $e^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}$ of the solution). For the FD scheme, we show in appendix A.2.2 that the linear decay is captured by the scheme without change of the dispersion proper-

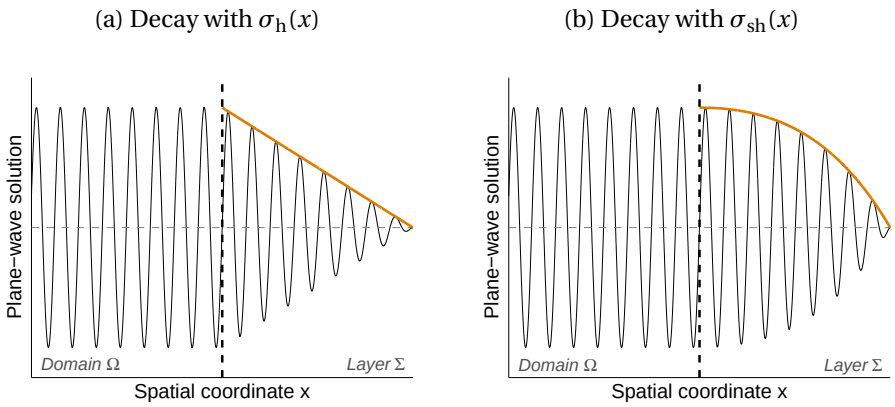


Figure 3.16: Illustration of the decay of a plane wave in a one-dimensional PML for (a) a hyperbolic function $\sigma_h(x)$ and (b) a shifted hyperbolic function $\sigma_{sh}(x)$. The decay is respectively linear and exponential-linear.

ties of the discrete plane waves if α is equal to the group velocity c_g (A.4)[2]. Therefore, both decay and wave oscillation are ideally simulated. The same behavior is expected with other numerical schemes.

However, the spatial derivative of the plane-wave solution is discontinuous at the interface between the domain Ω and the layer Σ (see Figure 3.16(a)). Such solution can be correctly represented with finite element schemes (CG or DG), but not with a FD one. Indeed, the finite difference at the interface is a bad approximation. This explains the observed results for the one-dimensional benchmark: σ_h is good with both CG and DG schemes and bad with the FD one. Unfortunately, in two dimensions, the decay of obliques waves (3.8) is no longer linear, and the discontinuity of the derivative remains. This could explain the bad results observed for all methods with the two-dimensional benchmark.

By contrast, the shifted hyperbolic function σ_{sh} provides a nonlinear decay of plane waves, without discontinuity of the spatial derivative at the interface, as illustrated in Figure 3.16(b). The numerical error is then generated inside the layer, and no longer at the interface. The shape of this solution, with a slow decay, minimizes this error with $\alpha = c$ in all considered discrete contexts.

[2]This assumption is reasonable since c_g differs from the propagation velocity c only for wavelengths close to the spatial step Δx [82]. However, in practical cases, Δx is chosen small enough in comparison to the considered wavelengths.

3.4 Conclusion

The properties of PMLs, shown in the continuous context, are altered when numerical methods are used. While the perfect matching was automatic and the perfect absorption was ensured by an unbounded absorption function $\sigma(x)$, the parameters of a discretized PML must be carefully chosen in order to limit both modeling and numerical errors.

The important elements for the optimization of PMLs are the absorption function σ, the layer thickness δ and the spatial discretization. Tacking a larger δ or a finer mesh improves the PML, but at the price of an increase of the computational cost. Unfortunately, the improvement is not significant if the global error is dominated by respectively the numerical error or the modeling error. It is then crucial to take an absorption function σ that minimize both kinds of error. Since this is done without increasing the computational cost, looking for efficient absorption functions is a very attractive approach.

In discrete contexts, σ must be large enough to sufficiently damp outgoing waves, without introducing large gradients in the solution that cannot be represented by the numerical scheme. In this chapter, we have proposed a systematic comparison between different absorption functions that produce such kind of smooth but efficient damping: commonly used polynomial functions (σ_2 and σ_3) and hyperbolic functions (σ_{h} and σ_{sh}). When they are optimized, σ_2, σ_3 and σ_{sh} provide comparable results, which are better than with σ_{h}, except in specific one-dimensional cases. However, since the optimum value of the free parameter of the polynomial functions does not find any direct interpretation, an optimization procedure remains necessary to use them. By contrast, the function σ_{sh} can be successfully used without any tuning. Indeed, it is nearly optimum when its free parameter α is equal to the propagation velocity c. This function is therefore a convenient choice for a practical use, especially for massive cases where the tuning of the absorption function is not possible.

Since the comparison has been done by means of two representative benchmarks with different classical numerical schemes and different numerical parameters, this conclusion is valid for a broad range of cases involving non-dispersive waves. In chapter 4, the effectiveness of the function σ_{sh} will be assessed for a case of electromagnetism with a more complicated geometry (a convex truncated domain). The study of absorption functions will be extended to cases that account for other physical dynamics in chapter 5.

CHAPTER 4

Perfectly Matched Layers for Acoustic and Electromagnetic Scattering Problems

4.1 Introduction

Scattering problems, which form an important category of problems in both acoustics and electromagnetism, are defined on unbounded domains. Basically, waves are scattered by one or more objects, called *scatterers*. When the equations are linear, the complete mathematical solution can be separated into two contributions, as illustrated in Figure 4.1: the *incident fields*, which corresponds to the solution obtained by removing the scatterers, and the *scattered fields*, which represent the modification of the incident fields due to the presence of the scatterers. When a scattering problem is redefined on a bounded domain Ω, the absorbing boundary condition or the absorbing layer must account for both these fields. In the case where incident waves are coming from the exterior of Ω (case (b) of Figure 4.1), the boundary treatment must be able to prescribe them.

In this chapter, we present numerical acoustic and electromagnetic scattering benchmarks in order to show the effectiveness of the PMLs introduced in chapter 2. Different formulations for scattering problems with PMLs are presented and discussed, including some that easily account for ingoing incident waves. The numerical solutions are obtained with a discontinuous Galerkin scheme.

The chapter is organised as follows. First, equations of acoustics and electromagnetism are introduced (section 4.2). Then, in section 4.3, we derive and study different formulations for scattering problems with PMLs, and propose a DG scheme. Its convergence is studied by means of one- and three-dimensional reference benchmarks. Finally, we present two benchmarks in section 4.4.

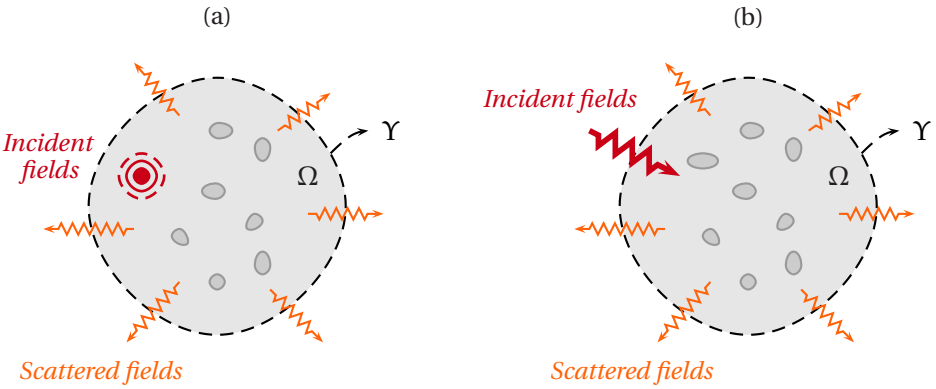


Figure 4.1: Illustration of scattering problems. Incident waves, generated inside (a) or outside (b) the domain of interest Ω, are scattered by one or more objects, called scatterers. These incident waves are described by the *incident fields*, while the modification of the solution due to the presence of scatterers (*i.e.* the scattered waves) is described by the *scattered fields*. The sum of the incident and scattered fields gives the *total fields*.

4.2 Mathematical models

4.2.1 Equations for acoustics

Acoustics is a branch of fluid dynamics concerned with the study of sound waves (see *e.g.* [103, section 15.2] or [147]). This is a linearized theory: variations of fields are small in comparison with steady reference values. In a quiescent fluid, the fluctuations in pressure $p(t,\boldsymbol{x})$ and velocity $\boldsymbol{u}(t,\boldsymbol{x})$ are governed by

$$\boxed{\begin{cases} \dfrac{\partial p}{\partial t} + \rho c^2 \nabla \cdot \boldsymbol{u} = 0, \\ \dfrac{\partial \boldsymbol{u}}{\partial t} + \dfrac{1}{\rho} \nabla p = 0, \end{cases}} \tag{4.1}$$

where ρ is the reference density and c is the propagation velocity of the medium. These parameters are assumed to be constant. In initial-boundary value problems, a solid wall is modeled with the boundary condition $\boldsymbol{n} \cdot \boldsymbol{u} = 0$, where $\boldsymbol{n}$ is the outward normal.

The system (4.1) is obviously related to the scalar wave system (1.3). Their parameters are linked by

$$\begin{aligned} a &\longrightarrow \rho c^2, \\ b &\longrightarrow \rho^{-1}. \end{aligned}$$

Therefore, all results obtained in previous chapters are immediately applicable to acoustic problems defined with the system of equations (4.1).

4.2.2 Equations for electromagnetism

The time-evolution of electromagnetic fields is modeled by Maxwell's equations and constitutive relations (see *e.g.* the reference books [74, 97]). In a linear isotropic nonconducting medium, these equations reduce to

$$\boxed{\begin{cases} \varepsilon \dfrac{\partial \boldsymbol{e}}{\partial t} - \nabla \times \boldsymbol{h} = 0, \\ \mu \dfrac{\partial \boldsymbol{h}}{\partial t} + \nabla \times \boldsymbol{e} = 0, \end{cases}} \tag{4.2}$$

where $\boldsymbol{e}(\boldsymbol{x}, t)$ is the electric field and $\boldsymbol{h}(\boldsymbol{x}, t)$ is the magnetic field. The properties of the medium are taken into account through the permittivity ε and the permeability μ. In this work, these parameters are constant.

In initial-boundary value problems, the boundary condition $\boldsymbol{n} \times \boldsymbol{e} = 0$ is used at the surface of a perfectly electrical conductor (PEC), where $\boldsymbol{n}$ is the outward normal on the boundary. The corresponding condition for a perfectly magnetic conductor is $\boldsymbol{n} \times \boldsymbol{h} = 0$. Finally, the radiation condition for Maxwell's equations, called the Silver-Muller condition, reads

$$\boldsymbol{n} \times \boldsymbol{e} - \sqrt{\frac{\mu}{\varepsilon}}\, \boldsymbol{n} \times (\boldsymbol{n} \times \boldsymbol{h}) = \boldsymbol{n} \times \boldsymbol{e}^{\text{inc}} - \sqrt{\frac{\mu}{\varepsilon}}\, \boldsymbol{n} \times (\boldsymbol{n} \times \boldsymbol{h}^{\text{inc}}),$$

where $\boldsymbol{e}^{\text{inc}}$ and $\boldsymbol{h}^{\text{inc}}$ are the incidents fields. Each of these boundary conditions can be written with only two scalar relations that correspond to the projections in the two directions perpendicular to $\boldsymbol{n}$. Indeed, the scalar relation obtained by projecting a boundary condition on $\boldsymbol{n}$ is trivially ensured.

Reduced systems for TE and TM modes

Wave guide problems with invariance along one direction reduce to problems with only two spatial dimensions. Two particular cases must be considered [74, 97]: the *transverse magnetic* (TM) mode, where $\boldsymbol{e}$ is parallel and $\boldsymbol{h}$ is perpendicular to the invariance direction, and the *transverse electric* (TE) mode, where $\boldsymbol{e}$ is perpendicular and $\boldsymbol{h}$ is parallel to this direction.

For both TM and TE modes, the equations of the system (4.2) reduce to three scalar equations with three scalar unknown fields, that are in correspondence with the scalar wave system. Let us consider a TM mode with invariance of fields in the $z-$direction. Then, $\boldsymbol{e}$ is parallel and $\boldsymbol{h}$ is perpendicular to the $z-$direction, which can be written $\boldsymbol{e} = [0, 0, e_z]$ and $\boldsymbol{h} = [h_x, h_y, 0]$. The Cartesian components of vector

fields are then governed by the system

$$\begin{cases} \varepsilon \dfrac{\partial e_z}{\partial t} - \dfrac{\partial h_y}{\partial x} + \dfrac{\partial h_x}{\partial y} = 0, \\ \mu \dfrac{\partial h_x}{\partial t} + \dfrac{\partial e_z}{\partial y} = 0, \\ \mu \dfrac{\partial h_y}{\partial t} - \dfrac{\partial e_z}{\partial x} = 0. \end{cases} \tag{4.3}$$

This system is immediately recovered from the two-dimensional version of the scalar wave system written with the Cartesian components, *i.e.*

$$\begin{cases} \dfrac{\partial p}{\partial t} + a\left(\dfrac{\partial u_x}{\partial x} + \dfrac{\partial u_y}{\partial y}\right) = 0, \\ \dfrac{\partial u_x}{\partial t} + b\dfrac{\partial p}{\partial x} = 0, \\ \dfrac{\partial u_y}{\partial t} + b\dfrac{\partial p}{\partial y} = 0, \end{cases}$$

by using the substitutions

$$\begin{aligned} p &\longrightarrow e_z, \\ u_x &\longrightarrow -h_y, \\ u_y &\longrightarrow h_x, \\ a &\longrightarrow \varepsilon^{-1}, \\ b &\longrightarrow \mu^{-1}. \end{aligned}$$

Similar substitutions can be written for the TE mode. Therefore, as for the acoustic case, results obtained with the scalar wave system are applicable to two-dimensional electromagnetic problems with TE and TM modes.

4.3 Methods

In this section, we examine different formulations of scattering problems with PMLs. We compare different ways of considering incident fields. Then, the numerical DG scheme used in benchmarks is detailed and studied. All developments are made in the context of the scalar wave system, with a single scatterer.

4.3.1 Formulations for scattering problems with PML

Let us consider a general scattering problem in the time domain. The geometry of the problem is illustrated in Figure 4.2 for the two-dimensional case. The scatterer is denoted Π, and its boundary is $\partial\Pi$. We are interessed in the solution in the bounded domain of interest Ω that contains the scatterer. The domain Ω is such

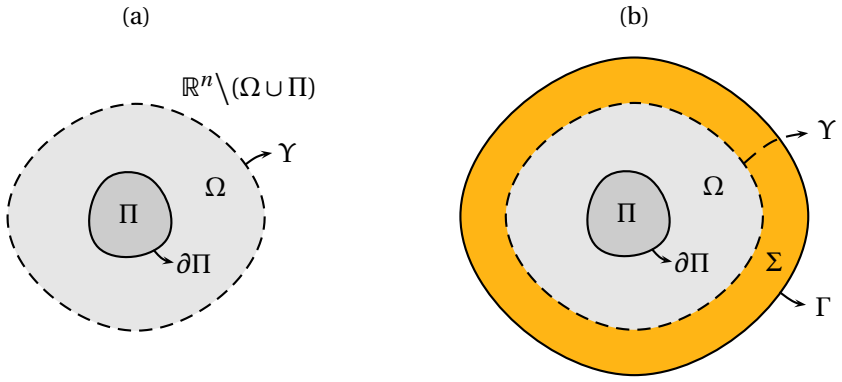


Figure 4.2: Notations for the scattering problem with an unbounded domain (a) and a truncated domain Ω surrounded by a layer Σ (b).

that $\Omega \cap \Pi = \emptyset$, and $\Omega \cup \Pi$ is convex. For the numerical resolution, Ω is used as truncated domain, and is surrounded by a PML Σ (Figure 4.2(b)).

The scattering problem defined in the unbounded domain $\mathbb{R}^n \backslash \Pi$, with $n = 2$ or 3, reads

$$\begin{cases} \dfrac{\partial p}{\partial t} + a\nabla \cdot \boldsymbol{u} = 0, & \forall (\boldsymbol{x}, t) \in \mathbb{R}^n \backslash \Pi \times \mathbb{R}^+, \\ \dfrac{\partial \boldsymbol{u}}{\partial t} + b\nabla p = 0, & \forall (\boldsymbol{x}, t) \in \mathbb{R}^n \backslash \Pi \times \mathbb{R}^+, \\ \boldsymbol{n} \cdot \boldsymbol{u} = 0, & \forall (\boldsymbol{x}, t) \in \partial\Pi \times \mathbb{R}^+, \\ p|_{t=0} = p^{\text{inc}}|_{t=0}, & \forall \boldsymbol{x} \in \mathbb{R}^n \backslash \Pi, \\ \boldsymbol{u}|_{t=0} = \boldsymbol{u}^{\text{inc}}|_{t=0}, & \forall \boldsymbol{x} \in \mathbb{R}^n \backslash \Pi, \end{cases} \tag{4.4}$$

where $p^{\text{inc}}(\boldsymbol{x}, t)$ and $\boldsymbol{u}^{\text{inc}}(\boldsymbol{x}, t)$ are the incident fields. These fields correspond to the solution of the modified problem obtained by removing the scatterer. They are assumed to have a compact support in $\mathbb{R}^n \backslash \Pi$.

The redefinition of the problem (4.4) on the truncated domain Ω with a PML is not immediate. Indeed, if the incident fields have not initially a compact support in Ω, external signals coming from the exterior domain $\mathbb{R}^3 \backslash (\Omega \cup \Pi)$ must be prescribed at the boundary of Ω. However, the PML formulations derived in chapter 2 do not directly account for such external signals.

A solution consists in rewriting the problem with another formulation so that the PML formulations can be used unchanged. Such a formulation is derived thanks to the linearity of the equations. The fields p and $\boldsymbol{u}$ can then be split into incident and scattered parts, *i.e.*

$$p = p^{\text{inc}} + p^{\text{scatt}}, \tag{4.5}$$

$$\boldsymbol{u} = \boldsymbol{u}^{\text{inc}} + \boldsymbol{u}^{\text{scatt}}, \tag{4.6}$$

where p^{scatt} and $\boldsymbol{u}^{\text{scatt}}$ denote the scattered fields. Since the incident fields satisfy the scalar wave system in $\mathbb{R}^n \setminus \Pi$, the problem (4.4) can be rewritten with scattered fields as unknowns instead of the total fields:

$$\begin{cases} \dfrac{\partial p^{\text{scatt}}}{\partial t} + a\nabla\cdot\boldsymbol{u}^{\text{scatt}} = 0, & \forall (\boldsymbol{x},t) \in \mathbb{R}^n\setminus\Pi \times \mathbb{R}^+, \\ \dfrac{\partial \boldsymbol{u}^{\text{scatt}}}{\partial t} + b\nabla p^{\text{scatt}} = 0, & \forall (\boldsymbol{x},t) \in \mathbb{R}^n\setminus\Pi \times \mathbb{R}^+, \\ \boldsymbol{n}\cdot\boldsymbol{u}^{\text{scatt}} = -\boldsymbol{n}\cdot\boldsymbol{u}^{\text{inc}}, & \forall (\boldsymbol{x},t) \in \partial\Pi \times \mathbb{R}^+, \\ p^{\text{scatt}}|_{t=0} = 0, & \forall \boldsymbol{x} \in \mathbb{R}^3\setminus\Pi, \\ \boldsymbol{u}^{\text{scatt}}|_{t=0} = 0, & \forall \boldsymbol{x} \in \mathbb{R}^3\setminus\Pi, \end{cases} \tag{4.7}$$

which is called the *scattered-field formulation* of the problem. Since the initial conditions are zero everywhere, there is no external signal to account for at the boundary of Ω. Incident waves are only prescribed by the boundary condition at the surface of the scatterer.

Thanks to the linearity of the problem, different problem formulations with PMLs can be derived. They exhibit two fundamental differences advantageous or not depending on the application context. First, the unknown fields are not the same: the total fields, the scattered fields or a combination thereof. Second, the incident signals are taken into account differently: through a modification of the interior problem, an adapted PML formulation or a specific interface condition at Υ.

The problem formulations are derived hereafter. Their main properties are highlighted and discussed.

Scattered-field formulation with PML

In the formulation (4.7), all scattered waves are generated by the boundary condition, and are therefore propagated outwards the domain Ω. Since there is no incoming wave, the PML system derived in chapter 2 can be used as it is. In three dimensions, the modified problem reads

$$
\begin{cases}
\text{Equations in } \Omega \text{ and } \Sigma: & \\
\dfrac{\partial p^{\text{scatt}}}{\partial t} + a\nabla\cdot\boldsymbol{u}^{\text{scatt}} = 0, & \forall(\boldsymbol{x},t)\in\Omega\times\mathbb{R}^+, \\
\dfrac{\partial \boldsymbol{u}^{\text{scatt}}}{\partial t} + b\nabla p^{\text{scatt}} = 0, & \forall(\boldsymbol{x},t)\in\Omega\times\mathbb{R}^+, \\
\dfrac{\partial p^{\text{scatt}}}{\partial t} + a\nabla\cdot\boldsymbol{u}^{\text{scatt}} = -\sigma p_s - \bar{\kappa}_\varphi\bar{\sigma}p_\varphi - \bar{\kappa}_\theta\bar{\sigma}(p^{\text{scatt}} - p_s - p_\varphi), & \forall(\boldsymbol{x},t)\in\Sigma\times\mathbb{R}^+, \\
\dfrac{\partial \boldsymbol{u}^{\text{scatt}}}{\partial t} + b\nabla p^{\text{scatt}} = -\sigma\boldsymbol{e}_s(\boldsymbol{e}_s\cdot\boldsymbol{u}^{\text{scatt}}) - \bar{\kappa}_\varphi\bar{\sigma}\boldsymbol{e}_\varphi(\boldsymbol{e}_\varphi\cdot\boldsymbol{u}^{\text{scatt}}) & \\
\qquad\qquad -\bar{\kappa}_\theta\bar{\sigma}\boldsymbol{e}_\theta(\boldsymbol{e}_\theta\cdot\boldsymbol{u}^{\text{scatt}}), & \forall(\boldsymbol{x},t)\in\Sigma\times\mathbb{R}^+, \\
\dfrac{\partial p_s}{\partial t} + a\nabla_{\boldsymbol{e}_s}\cdot\boldsymbol{u}^{\text{scatt}} = -\sigma p_s, & \forall(\boldsymbol{x},t)\in\Sigma\times\mathbb{R}^+, \\
\dfrac{\partial p_\varphi}{\partial t} + a\nabla_{\boldsymbol{e}_\varphi}\cdot\boldsymbol{u}^{\text{scatt}} = -\bar{\kappa}_\varphi\bar{\sigma}p_\varphi, & \forall(\boldsymbol{x},t)\in\Sigma\times\mathbb{R}^+, \\
\text{Boundary conditions:} & \\
\boldsymbol{n}\cdot\boldsymbol{u}^{\text{scatt}} = -\boldsymbol{n}\cdot\boldsymbol{u}^{\text{inc}}, & \forall(\boldsymbol{x},t)\in\partial\Pi\times\mathbb{R}^+, \\
\boldsymbol{n}\cdot\boldsymbol{u}^{\text{scatt}} = 0, & \forall(\boldsymbol{x},t)\in\Gamma\times\mathbb{R}^+, \\
\text{Initial conditions:} & \\
p^{\text{scatt}}\big|_{t=0} = 0, & \forall\boldsymbol{x}\in\Omega\cup\Sigma, \\
\boldsymbol{u}^{\text{scatt}}\big|_{t=0} = 0, & \forall\boldsymbol{x}\in\Omega\cup\Sigma, \\
p_s\big|_{t=0} = 0, & \forall\boldsymbol{x}\in\Sigma, \\
p_\varphi\big|_{t=0} = 0, & \forall\boldsymbol{x}\in\Sigma.
\end{cases}
$$

Incident waves are taken into account through the modified boundary condition on the scatterer.

This formulation is convenient for linear problems with constant parameters a and b without source terms. However, if nonlinear terms are added and/or a and b vary spatially, supplementary terms must be added in the first two equations. Therefore, the computational cost increases because of the numerical evaluation of these terms. This is a weakness of this formulation.

Total-field formulations with PML

If the incident fields have a compact support in Ω, the PML is incorporated in the total-field formulation (4.4) exactly as for scattered-field one: the total fields are governed by the scalar wave system in Ω and by the PML system in Σ. Therefore, both outgoing incident and scattered waves are damped in the PML.

Equations in Ω and Σ:

$$\frac{\partial p}{\partial t} + a\nabla\cdot\boldsymbol{u} = 0, \qquad \forall(\boldsymbol{x},t)\in\Omega\times\mathbb{R}^+,$$

$$\frac{\partial \boldsymbol{u}}{\partial t} + b\nabla p = 0, \qquad \forall(\boldsymbol{x},t)\in\Omega\times\mathbb{R}^+,$$

$$\frac{\partial p}{\partial t} + a\nabla\cdot\boldsymbol{u} = -\sigma p_s - \bar{\kappa}_\varphi\bar{\sigma}p_\varphi - \bar{\kappa}_\theta\bar{\sigma}(p - p_s - p_\varphi), \qquad \forall(\boldsymbol{x},t)\in\Sigma\times\mathbb{R}^+,$$

$$\frac{\partial \boldsymbol{u}}{\partial t} + b\nabla p = -\sigma\boldsymbol{e}_s(\boldsymbol{e}_s\cdot\boldsymbol{u}) - \bar{\kappa}_\varphi\bar{\sigma}\boldsymbol{e}_\varphi(\boldsymbol{e}_\varphi\cdot\boldsymbol{u}) - \bar{\kappa}_\theta\bar{\sigma}\boldsymbol{e}_\theta(\boldsymbol{e}_\theta\cdot\boldsymbol{u}), \qquad \forall(\boldsymbol{x},t)\in\Sigma\times\mathbb{R}^+,$$

$$\frac{\partial p_s}{\partial t} + a\nabla_{\boldsymbol{e}_s}\cdot\boldsymbol{u} = -\sigma p_s, \qquad \forall(\boldsymbol{x},t)\in\Sigma\times\mathbb{R}^+,$$

$$\frac{\partial p_\varphi}{\partial t} + a\nabla_{\boldsymbol{e}_\varphi}\cdot\boldsymbol{u} = -\bar{\kappa}_\varphi\bar{\sigma}p_\varphi, \qquad \forall(\boldsymbol{x},t)\in\Sigma\times\mathbb{R}^+,$$

Boundary conditions:

$$\boldsymbol{n}\cdot\boldsymbol{u} = 0, \qquad \forall(\boldsymbol{x},t)\in\partial\Pi\times\mathbb{R}^+,$$

$$\boldsymbol{n}\cdot\boldsymbol{u} = 0, \qquad \forall(\boldsymbol{x},t)\in\Gamma\times\mathbb{R}^+,$$

Initial conditions:

$$p\big|_{t=0} = p^{\mathrm{inc}}\big|_{t=0}, \qquad \forall\boldsymbol{x}\in\Omega,$$

$$\boldsymbol{u}\big|_{t=0} = \boldsymbol{u}^{\mathrm{inc}}\big|_{t=0}, \qquad \forall\boldsymbol{x}\in\Omega,$$

$$p\big|_{t=0} = 0, \qquad \forall\boldsymbol{x}\in\Sigma,$$

$$\boldsymbol{u}\big|_{t=0} = 0, \qquad \forall\boldsymbol{x}\in\Sigma,$$

$$p_s\big|_{t=0} = 0, \qquad \forall\boldsymbol{x}\in\Sigma,$$

$$p_\varphi\big|_{t=0} = 0, \qquad \forall\boldsymbol{x}\in\Sigma.$$

In order to take into account incident waves coming from the exterior domain $\mathbb{R}^n \setminus (\Omega \cup \Sigma)$, another total-field formulation is obtained from the scattered-field formulation with PML. Using the relations (4.5)–(4.6), the scattered-field formulation is rewritten with the total fields as unknowns:

$$
\begin{cases}
\text{Equations in } \Omega \text{ and } \Sigma: & \\
\quad \dfrac{\partial p}{\partial t} + a\nabla\cdot\boldsymbol{u} = 0, & \forall (\boldsymbol{x},t) \in \Omega \times \mathbb{R}^+, \\
\quad \dfrac{\partial \boldsymbol{u}}{\partial t} + b\nabla p = 0, & \forall (\boldsymbol{x},t) \in \Omega \times \mathbb{R}^+, \\
\quad \dfrac{\partial p}{\partial t} + a\nabla\cdot\boldsymbol{u} = -\sigma p_s - \bar{\kappa}_\varphi\bar{\sigma} p_\varphi - \bar{\kappa}_\theta\bar{\sigma}(p - p^{\text{inc}} - p_s - p_\varphi), & \forall (\boldsymbol{x},t) \in \Sigma \times \mathbb{R}^+, \\
\quad \dfrac{\partial \boldsymbol{u}}{\partial t} + b\nabla p = -\sigma \boldsymbol{e}_s\left(\boldsymbol{e}_s\cdot\boldsymbol{u} - \boldsymbol{e}_s\cdot\boldsymbol{u}^{\text{inc}}\right) & \\
\qquad\qquad -\bar{\kappa}_\varphi\bar{\sigma}\boldsymbol{e}_\varphi\left(\boldsymbol{e}_\varphi\cdot\boldsymbol{u} - \boldsymbol{e}_\varphi\cdot\boldsymbol{u}^{\text{inc}}\right) & \\
\qquad\qquad -\bar{\kappa}_\theta\bar{\sigma}\boldsymbol{e}_\theta\left(\boldsymbol{e}_\theta\cdot\boldsymbol{u} - \boldsymbol{e}_\theta\cdot\boldsymbol{u}^{\text{inc}}\right), & \forall (\boldsymbol{x},t) \in \Sigma \times \mathbb{R}^+, \\
\dfrac{\partial p_s}{\partial t} + a\nabla_{\boldsymbol{e}_s}\cdot\boldsymbol{u} = a\nabla_{\boldsymbol{e}_s}\cdot\boldsymbol{u}^{\text{inc}} - \sigma p_s, & \forall (\boldsymbol{x},t) \in \Sigma \times \mathbb{R}^+, \\
\dfrac{\partial p_\varphi}{\partial t} + a\nabla_{\boldsymbol{e}_\varphi}\cdot\boldsymbol{u} = a\nabla_{\boldsymbol{e}_\varphi}\cdot\boldsymbol{u}^{\text{inc}} - \bar{\kappa}_\varphi\bar{\sigma} p_\varphi, & \forall (\boldsymbol{x},t) \in \Sigma \times \mathbb{R}^+, \\
\text{Boundary conditions:} & \\
\quad \boldsymbol{n}\cdot\boldsymbol{u} = 0, & \forall (\boldsymbol{x},t) \in \partial\Pi \times \mathbb{R}^+, \\
\quad \boldsymbol{n}\cdot\boldsymbol{u} = \boldsymbol{n}\cdot\boldsymbol{u}^{\text{inc}}, & \forall (\boldsymbol{x},t) \in \Gamma \times \mathbb{R}^+, \\
\text{Initial conditions:} & \\
\quad p\big|_{t=0} = p^{\text{inc}}\big|_{t=0}, & \forall \boldsymbol{x} \in \Omega \cup \Sigma, \\
\quad \boldsymbol{u}\big|_{t=0} = \boldsymbol{u}^{\text{inc}}\big|_{t=0}, & \forall \boldsymbol{x} \in \Omega \cup \Sigma, \\
\quad p_s\big|_{t=0} = 0, & \forall \boldsymbol{x} \in \Sigma, \\
\quad p_\varphi\big|_{t=0} = 0, & \forall \boldsymbol{x} \in \Sigma.
\end{cases}
$$

In this formulation, incident waves are prescribed through the boundary condition on Γ, the initial conditions in $\Omega \cup \Sigma$ and additional terms in the PML equations.

By contrast with the scattered-field formulation, the original form of both equations and boundary conditions are preserved in Ω. Additional source terms and spatially varying a and b in Ω are then naturally taken into account. Unfortunately, additional terms appear in all PML equations, increasing the computational cost.

Scattered/total-field formulation with PML

An alternative formulation that combines the advantages of both previous formulations is obtained by writing equations with the total fields p and $\boldsymbol{u}$ in Ω, and with the scattered fields p^{scatt} and $\boldsymbol{u}^{\text{scatt}}$ in Σ. Compatibility relations must be prescribed as interface conditions on Υ. This formulation then reads

$$
\left\{
\begin{array}{l}
\text{Equations in } \Omega \text{ and } \Sigma: \\
\qquad \dfrac{\partial p}{\partial t} + a\nabla\cdot\boldsymbol{u} = 0, \qquad \forall (\boldsymbol{x},t)\in\Omega\times\mathbb{R}^+, \\
\qquad \dfrac{\partial \boldsymbol{u}}{\partial t} + b\nabla p = 0, \qquad \forall (\boldsymbol{x},t)\in\Omega\times\mathbb{R}^+, \\
\dfrac{\partial p^{\text{scatt}}}{\partial t} + a\nabla\cdot\boldsymbol{u}^{\text{scatt}} = -\sigma p_s - \bar{\kappa}_\varphi\bar{\sigma}p_\varphi - \bar{\kappa}_\theta\bar{\sigma}(p^{\text{scatt}} - p_s - p_\varphi), \qquad \forall (\boldsymbol{x},t)\in\Sigma\times\mathbb{R}^+, \\
\dfrac{\partial \boldsymbol{u}^{\text{scatt}}}{\partial t} + b\nabla p^{\text{scatt}} = -\sigma\boldsymbol{e}_s(\boldsymbol{e}_s\cdot\boldsymbol{u}^{\text{scatt}}) - \bar{\kappa}_\varphi\bar{\sigma}\boldsymbol{e}_\varphi(\boldsymbol{e}_\varphi\cdot\boldsymbol{u}^{\text{scatt}}) \\
\qquad\qquad\qquad\qquad -\bar{\kappa}_\theta\bar{\sigma}\boldsymbol{e}_\theta(\boldsymbol{e}_\theta\cdot\boldsymbol{u}^{\text{scatt}}), \qquad \forall (\boldsymbol{x},t)\in\Sigma\times\mathbb{R}^+, \\
\dfrac{\partial p_s}{\partial t} + a\nabla_{\boldsymbol{e}_s}\cdot\boldsymbol{u}^{\text{scatt}} = -\sigma p_s, \qquad \forall (\boldsymbol{x},t)\in\Sigma\times\mathbb{R}^+, \\
\dfrac{\partial p_\varphi}{\partial t} + a\nabla_{\boldsymbol{e}_\varphi}\cdot\boldsymbol{u}^{\text{scatt}} = -\bar{\kappa}_\varphi\bar{\sigma}p_\varphi, \qquad \forall (\boldsymbol{x},t)\in\Sigma\times\mathbb{R}^+, \\
\text{Boundary conditions:} \\
\qquad \boldsymbol{n}\cdot\boldsymbol{u} = 0, \qquad \forall (\boldsymbol{x},t)\in\partial\Pi\times\mathbb{R}^+, \\
\qquad \boldsymbol{n}\cdot\boldsymbol{u}^{\text{scatt}} = 0, \qquad \forall (\boldsymbol{x},t)\in\Gamma\times\mathbb{R}^+, \\
\text{Interface conditions:} \\
\qquad p = p^{\text{scatt}} + p^{\text{inc}}, \qquad \forall (\boldsymbol{x},t)\in\Upsilon\times\mathbb{R}^+, \\
\qquad \boldsymbol{n}\cdot\boldsymbol{u} = \boldsymbol{n}\cdot\boldsymbol{u}^{\text{scatt}} + \boldsymbol{n}\cdot\boldsymbol{u}^{\text{inc}}, \qquad \forall (\boldsymbol{x},t)\in\Upsilon\times\mathbb{R}^+, \\
\text{Initial conditions:} \\
\qquad p\big|_{t=0} = p^{\text{inc}}\big|_{t=0}, \qquad \forall \boldsymbol{x}\in\Omega, \\
\qquad \boldsymbol{u}\big|_{t=0} = p^{\text{inc}}\big|_{t=0}, \qquad \forall \boldsymbol{x}\in\Omega, \\
\qquad p^{\text{scatt}}\big|_{t=0} = 0, \qquad \forall \boldsymbol{x}\in\Sigma, \\
\qquad \boldsymbol{u}^{\text{scatt}}\big|_{t=0} = 0, \qquad \forall \boldsymbol{x}\in\Sigma, \\
\qquad p_s\big|_{t=0} = 0, \qquad \forall \boldsymbol{x}\in\Sigma, \\
\qquad p_\varphi\big|_{t=0} = 0, \qquad \forall \boldsymbol{x}\in\Sigma.
\end{array}
\right.
$$

Incident waves are prescribed through the interface conditions on Υ and the initial conditions in Ω.

This formulation preserves the original equations in Ω, and does not involve additional terms in the PML. It requires specific interface conditions, which are well-known in computational electromagnetism (see *e.g.* [155, section 5.6]).

Let us note that the equations could be nonlinear in the truncated domain Ω. Indeed, they must be linear with constant parameters only in the exterior domain $\mathbb{R}^n \backslash (\Omega \cup \Sigma)$ in order to write the PML equations with scattered fields. This is not necessary in Ω, where the equations remain in a total-field formulation.

4.3.2 Numerical scheme based on the DG method

We derive hereafter a numerical scheme for the PML system (2.28) using the discontinuous Galerkin (DG) method. The DG scheme corresponding to the scalar wave system is given in section 1.3.3.2. These two schemes can be used together for the scattered-field and scattered/total-field formulations of the scattering problem. The former scheme is obviously used in Σ, while the latter is for Ω.

A weak form of the PML equations is obtained from their conservative form, which reads

$$\left\{\begin{aligned}
\frac{\partial p}{\partial t} + a\nabla\cdot\boldsymbol{u} &= -\sigma p_s - \bar{\kappa}_\varphi\bar{\sigma}p_\varphi - \bar{\kappa}_\theta\bar{\sigma}(p - p_s - p_\varphi),\\
\frac{\partial \boldsymbol{u}}{\partial t} + b\nabla p &= -\sigma \boldsymbol{e}_s(\boldsymbol{e}_s\cdot\boldsymbol{u}) - \bar{\kappa}_\varphi\bar{\sigma}\boldsymbol{e}_\varphi(\boldsymbol{e}_\varphi\cdot\boldsymbol{u}) - \bar{\kappa}_\theta\bar{\sigma}\boldsymbol{e}_\theta(\boldsymbol{e}_\theta\cdot\boldsymbol{u}),\\
\frac{\partial p_s}{\partial t} + a\nabla\cdot[\boldsymbol{e}_s(\boldsymbol{e}_s\cdot\boldsymbol{u})] &= -\sigma p_s + \left(\frac{\kappa_\varphi}{1+\kappa_\varphi s} + \frac{\kappa_\theta}{1+\kappa_\theta s}\right)(\boldsymbol{e}_s\cdot\boldsymbol{u}),\\
\frac{\partial p_\varphi}{\partial t} + a\nabla\cdot[\boldsymbol{e}_\phi(\boldsymbol{e}_\phi\cdot\boldsymbol{u})] &= -\bar{\kappa}_\varphi\bar{\sigma}p_\varphi - a\frac{\kappa_\varphi}{1+\kappa_\varphi s}(\boldsymbol{e}_s\cdot\boldsymbol{u}).
\end{aligned}\right. \tag{4.8}$$

The conservative form of the last two equations is derived in appendix B.1. Multiplying the equations of the system (4.8) by the test functions, integrating over an element Ω_e of the layer Σ, and integrating by parts provide the weak form

$$\left\{\begin{aligned}
\left(\frac{\partial p}{\partial t},\hat{p}\right)_{\Omega_e} - \left(a\boldsymbol{u},\nabla\hat{p}\right)_{\Omega_e} + \left\langle f^p,\hat{p}\right\rangle_{\Gamma_e} &= \left(S^p,\hat{p}\right)_{\Omega_e}, \qquad \forall\hat{p},\\
\left(\frac{\partial \boldsymbol{u}}{\partial t},\hat{\boldsymbol{u}}\right)_{\Omega_e} - \left(bp,\nabla\cdot\hat{\boldsymbol{u}}\right)_{\Omega_e} + \left\langle \boldsymbol{f}^{\boldsymbol{u}},\hat{\boldsymbol{u}}\right\rangle_{\Gamma_e} &= \left(\boldsymbol{S}^{\boldsymbol{u}},\hat{\boldsymbol{u}}\right)_{\Omega_e}, \qquad \forall\hat{\boldsymbol{u}},\\
\left(\frac{\partial p_s}{\partial t},\hat{p}_s\right)_{\Omega_e} - \left(a(\boldsymbol{e}_s\cdot\boldsymbol{u})\,\boldsymbol{e}_s,\nabla\hat{p}_s\right)_{\Omega_e} + \left\langle f^{p_s},\hat{p}_s\right\rangle_{\Gamma_e} &= \left(S^{p_s},\hat{p}_s\right)_{\Omega_e}, \qquad \forall\hat{p}_s,\\
\left(\frac{\partial p_\varphi}{\partial t},\hat{p}_\varphi\right)_{\Omega_e} - \left(b\left(\boldsymbol{e}_\varphi\cdot\boldsymbol{u}\right)\boldsymbol{e}_\varphi,\nabla\hat{p}_\varphi\right)_{\Omega_e} + \left\langle f^{p_\varphi},\hat{p}_\varphi\right\rangle_{\Gamma_e} &= \left(S^{p_\varphi},\hat{p}_\varphi\right)_{\Omega_e}, \qquad \forall\hat{p}_\varphi,
\end{aligned}\right. \tag{4.9}$$

where the fluxes f^p, $\boldsymbol{f}^{\boldsymbol{u}}$, f^{p_s} and f^{p_φ} are defined by

$$\begin{aligned}
f^p &= a\,\boldsymbol{n}_e\cdot\boldsymbol{u},\\
\boldsymbol{f}^{\boldsymbol{u}} &= b\,\boldsymbol{n}_e p,\\
f^{p_s} &= a\,(\boldsymbol{n}_e\cdot\boldsymbol{e}_s)(\boldsymbol{e}_s\cdot\boldsymbol{u}),\\
f^{p_\varphi} &= a\,(\boldsymbol{n}_e\cdot\boldsymbol{e}_\varphi)(\boldsymbol{e}_\varphi\cdot\boldsymbol{u}),
\end{aligned}$$

and S^p, $\boldsymbol{S}^{\boldsymbol{u}}$, S^{p_s} and S^{p_φ} contain the source terms written in the right-hand side of the equations (4.8). Γ_e and $\boldsymbol{n}_e$ are respectively the boundary and the unit outward normal of Ω_e. The test functions are denoted with the hat $\hat{}$.

As explained in section 1.3.3.2, the numerical scheme is obtained by approximating each unknown field over each element by a linear combination of local basis functions. These basis functions are then used as test functions in the weak form. Numerical fluxes are finally introduced in boundary terms of each equation.

At the interface between two elements, the numerical fluxes used for the scalar wave system can be reused for the first two equations of the PML system. We consider upwind fluxes, *i.e.*

$$f^{p,\text{num}} = a\,\boldsymbol{n}_e \cdot \{\boldsymbol{u}\} - c\,[\![p]\!], \tag{4.10}$$

$$\boldsymbol{f}^{\boldsymbol{u},\text{num}} = b\,\boldsymbol{n}_e\{p\} - c\,\boldsymbol{n}_e\left(\boldsymbol{n}_e \cdot [\![\boldsymbol{u}]\!]\right). \tag{4.11}$$

Given the similarity between the first equation and last two of the system (4.8), we define the numerical fluxes of the last two equations by

$$f^{p_s,\text{num}} = a\,(\boldsymbol{n}_e \cdot \boldsymbol{e}_s)(\boldsymbol{e}_s \cdot \{\boldsymbol{u}\}) - c\,[\![p_s]\!],$$
$$f^{p_\varphi,\text{num}} = a\,(\boldsymbol{n}_e \cdot \boldsymbol{e}_\varphi)(\boldsymbol{e}_\varphi \cdot \{\boldsymbol{u}\}) - c\,[\![p_\varphi]\!].$$

At the boundary of the computational domain, the numerical fluxes are adapted. The strategy of *ghost cell* is used for $f^{p,\text{num}}$ and $\boldsymbol{f}^{\boldsymbol{u},\text{num}}$ (see end of section 1.3.2). In order to define the last two numerical fluxes, let us study the continuous fluxes. At the boundary Γ, one has $f^{p_s} = f^p$ and $f^{p_\varphi} = 0$, since $\boldsymbol{n}_e$ is parallel to $\boldsymbol{e}_s$ and perpendicular to $\boldsymbol{e}_\varphi$. Therefore, we use the numerical fluxes $f^{p_s,\text{num}} = f^{p,\text{num}}$ and $f^{p_\varphi,\text{num}} = 0$ at Γ. Note that the same strategy can be used for the lateral boundaries of a straight PML

Finally, at the interface Υ, the numerical fluxes $f^{p,\text{num}}$ and $\boldsymbol{f}^{\boldsymbol{u},\text{num}}$ connect the scheme of the scalar wave system (in Ω) with the one of the PML system (in Σ). If these systems are written with the same formulation (scattered-field or total-field), the same numerical fluxes (4.10)–(4.11) are used on both sides of Υ. For a scattered/total-field formulation, these fluxes are adapted on each side by using the relations (4.5)–(4.6). In all cases, the numerical fluxes of the last two PML equations are $f^{p_s,\text{num}} = f^{p,\text{num}}$ and $f^{p_\varphi,\text{num}} = 0$. Since p_s and p_φ are defined only in Σ, these fluxes are defined on Υ only on the side of the layer Σ.

4.3.3 Convergence of the numerical scheme

The convergence of the numerical scheme proposed above is tested by means of one- and three-dimensional benchmarks. For the DG schemes of hyperbolic systems, it is well-known that the L^2-norm of the error on the fields decreases asymptotically accordingly $\mathcal{O}(\ell^{\mathsf{p}+1})$ (see *e.g.* [62, 112]), where ℓ and p denote respectively the characteristic length and the order of elements.

For both benchmarks, the error on the fields is obtained by comparing the numerical solution with the exact solution of the continuous problem. The time-average L^2-norm of the error is computed over the truncated domain Ω by means of the *relative error* ζ_r defined as

$$\boxed{\zeta_r = \sqrt{\frac{E_{\text{error}}}{E_{\text{exact}}}},}$$

with

$$E_{\text{error}} = \frac{1}{t_f}\int_0^{t_f}\int_\Omega \left(\frac{1}{2a}\left|p_{\text{num}} - p_{\text{exact}}\right|^2 + \frac{1}{2b}\left\|\boldsymbol{u}_{\text{num}} - \boldsymbol{u}_{\text{exact}}\right\|^2\right)\mathrm{d}\Omega\,\mathrm{d}t,$$
$$E_{\text{exact}} = \frac{1}{t_f}\int_0^{t_f}\int_\Omega \left(\frac{1}{2a}\left|p_{\text{exact}}\right|^2 + \frac{1}{2b}\left\|\boldsymbol{u}_{\text{exact}}\right\|^2\right)\mathrm{d}\Omega\,\mathrm{d}t,$$

where t_f is the duration of the simulation, p_{exact} and $\boldsymbol{u}_{\text{exact}}$ correspond to the exact analytical solution (of the problem defined on the unbounded domain), while p_{num} and $\boldsymbol{u}_{\text{num}}$ correspond to the numerical solution. E_{error} and E_{exact} then represent the time-average energies associated to the error fields and the exact fields, respectively. The relative error ζ_r, as defined, accounts for the numerical errors, as well as the modeling errors due to the PML (see section 3.2). We use the shifted hyperbolic absorption function

$$\sigma_{\text{sh}}(x) = \frac{\alpha}{\delta}\frac{x}{\delta - x}$$

with $\alpha = c$. Since there is normally no modeling error with this absorption function, the total error is only due to numerical approximations. Therefore, the order of convergence of the DG method should be recovered with both benchmarks.

One-dimensional case

Let us consider again the dimensionless one-dimensional benchmark introduced in section 3.2.2. The truncated domain $\Omega = [-L, 0]$ is extended with a PML $\Sigma = [0, \delta]$. A Gaussian pulse moving to the right is prescribed as initial condition on the second half of Ω. The analytical solution of this problem reads

$$p_{\text{exact}}(x,t) = e^{-(x-L/4-t)^2/R^2},$$
$$u_{\text{exact}}(x,t) = e^{-(x-L/4-t)^2/R^2}.$$

The dimensionless parameters are $L = 1$, $R = 0.05$ and $t_f = L/2$. The fourth-order Runge-Kutta scheme is used for the time stepping with $\Delta t = 0.085\,\ell$ (for first-order elements) and $\Delta t = 0.045\,\ell$ (for second-order elements).

The convergence of the numerical scheme is first verified for the scalar wave system without PML. Simulations are performed in a reference case with an exact condition at $x = 0$. The exact solution is prescribed with an exact Riemann solver. The convergence of the numerical solution with the mesh refinement in shown in Figure 4.3. We see that the convergence $\mathcal{O}(\ell^{p+1})$ is recovered with both first- and second-order elements.

In cases with PML, the asymptotic convergence remains $\mathcal{O}(\ell^{p+1})$. As shown in Figure 4.4 for different PML thicknesses, the error ζ_r is greater or equal to the error obtained in the reference case. The supplementary error is due the approximation of both solution and equations in the PML. We observe that it varies according to ℓ^{p+2}. Since this error decreases faster than ℓ^{p+1} when ℓ decreases, the error curves reach the one of the reference case and the asymptotic convergence rate $\mathcal{O}(\ell^{p+1})$ is recovered.

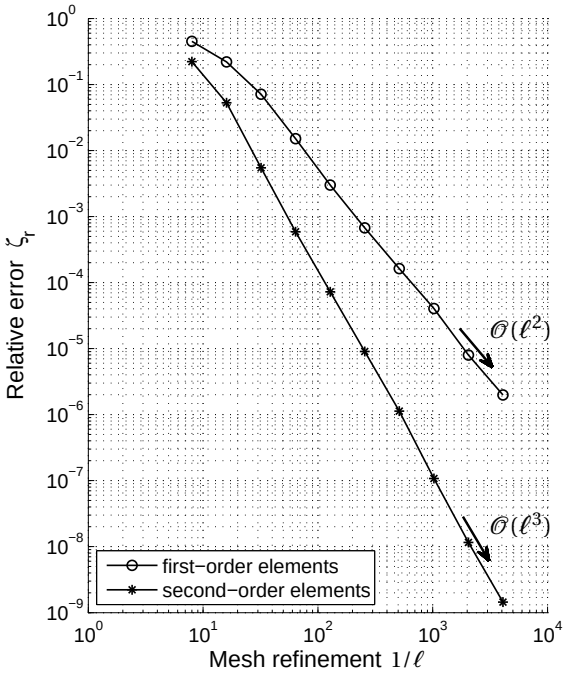
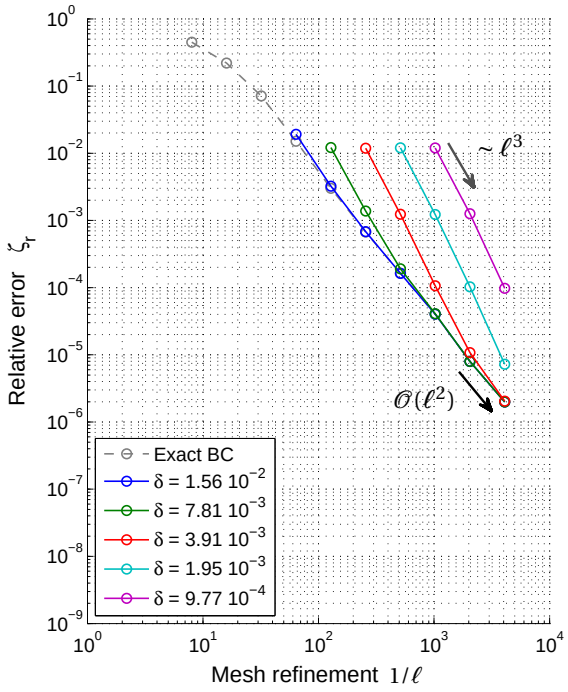


Figure 4.3: Convergence of the numerical solution in the reference case. The one-dimensional benchmark is tested with first- and second-order elements. The exact solution is prescribed at the boundary. The asymptotic convergence $\mathcal{O}(\ell^{p+1})$ is recovered.

(a) with first-order elements



(b) with second-order elements

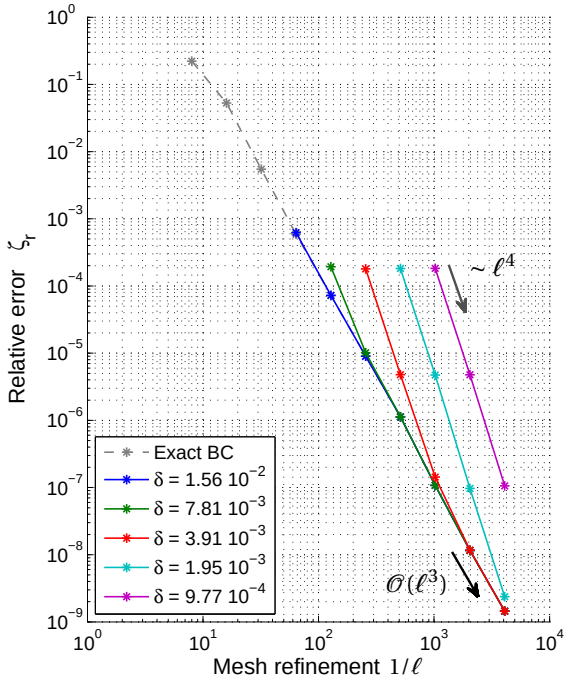


Figure 4.4: Convergence of the numerical solution for different PML thicknesses. The one-dimensional benchmark is tested with first- and second-order elements. The error decreases asymptotically accordingly $\mathcal{O}(\ell^{p+1})$. For thin PML and large ℓ, the error varies as ℓ^{p+2}.

Three-dimensional case

The convergence of the DG scheme is now tested in a three-dimensional acoustic case with the system (4.1). A truncated domain Ω shaped as an ellipsoid of revolution is surrounded by a PML Σ. A Gaussian pulse is prescribed as initial condition on the pressure p,

$$p(\boldsymbol{x},0) = e^{-\|\boldsymbol{x}-\boldsymbol{x}_0\|^2/R^2}, \tag{4.12}$$

while other fields are initially equal to zero. As time goes by, a spherical wave appears and travels towards the exterior of Ω. It reaches the PML with different angles of incidence. The exact solution of this problem reads

$$p_{\text{exact}}(\boldsymbol{x},t) = \frac{1}{2}\left(\frac{r-ct}{r}e^{-(r-ct)^2/R^2} + \frac{r+ct}{r}e^{-(r+ct)^2/R^2}\right),$$
$$\boldsymbol{u}_{\text{exact}}(\boldsymbol{x},t) = \frac{1}{2c\rho}\left[\left(\frac{R^2}{2r^2}+\frac{r-ct}{r}\right)e^{-(r-ct)^2/R^2} - \left(\frac{R^2}{2r^2}+\frac{r+ct}{r}\right)e^{-(r+ct)^2/R^2}\right]\frac{\boldsymbol{x}-\boldsymbol{x}_0}{r},$$

with $r = \|\boldsymbol{x}-\boldsymbol{x}_0\|^{1/2}$. The parameters are $\boldsymbol{x}_0 = (-122.5\,\text{m}, 0, 20\,\text{m})$, $R = 7.5\,\text{m}$, $c = 1.5\,\text{km/s}$, $\rho = 1\,\text{kg/m}^3$ and $t_f = 200\,\text{ms}$. The fourth-order Runge-Kutta scheme is used for the time stepping with $\Delta t = 0.021\,\ell/c$. The axes of the ellipsoid are 330 m ($x-$direction) and 120 m ($y-$ and $z-$directions) in length.

In Figure 4.5, the decay of the relative error ζ_r with the mesh refinement is shown for different cases. We consider a reference case, where the exact solution is prescribed at the artifical boundary, and PMLs with different thicknesses δ. With the thickest PMLs, the error is similar to the reference case, and the decay is close to the asymptotic convergence $\mathcal{O}(\ell^{p+1})$. Unfortunately, with the thinnest

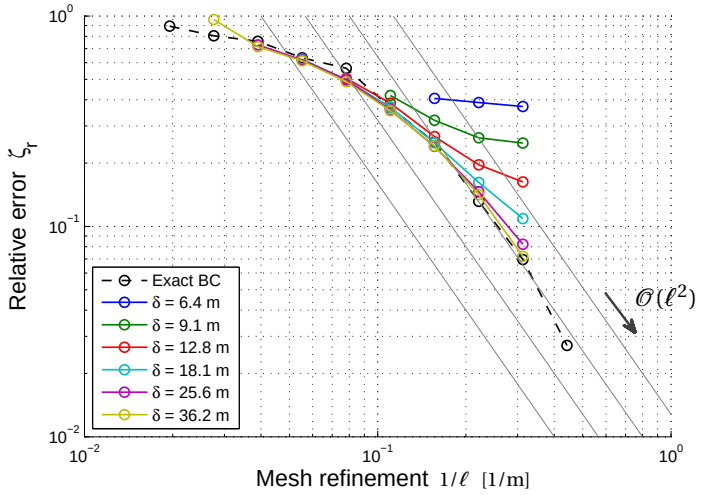


Figure 4.5: Convergence of the numerical scheme with the three-dimensional benchmark for the reference case and the PML case. First-order elements are used.

PMLs, we observe that ζ_r is greater than in the reference case, and its decrease is slower than ℓ^{p+1}. The asymptotic convergence is then lost. This difference with the one-dimensional case could be explained by other errors that can occur in this benchmark (*e.g.* the quality of the mesh and/or the approximate evaluation of s, κ_φ, κ_θ, $\boldsymbol{e}_s$, $\boldsymbol{e}_\varphi$ and $\boldsymbol{e}_\theta$ in the PML equations (2.28)). Further investigations are necessary to study the numerical convergence of the scheme is three dimensions.

4.4 Numerical benchmarks

In this section, two scattering examples are presented in physical contexts in order to illustrate the effectiveness of PMLs: the scattering of electromagnetic waves by a shield in two dimensions, and the scattering of acoustic waves by a submarine in three dimensions.

4.4.1 Electromagnetic scattering by a shield

Let us consider the scattering of transverse magnetic (TM) waves by a perfectly conducting cavity. The geometry and the mesh of the problem are represented in Figure 4.6. The cavity is rectangular (600 mm × 400 mm) of thickness 20 mm with an aperture of 160 mm in the middle of the left side (this cavity is used in [133]). The truncated computational domain Ω is rectangular with rounded corners, and surrounded by a PML Σ of thickness $\delta = 0.2$ m.

The problem is assumed to be invariant in the $z-$direction. The electric field is therefore parallel to the $z-$direction, while the magnetic field is perpendicular to the $z-$direction. They are written $\boldsymbol{e} = [0, 0, e_z]$ and $\boldsymbol{h} = [h_x, h_y, 0]$. The nonzero components e_z, h_x and h_y are governed by the system (4.3) in Ω. Considering a scattered/total-field formulation of the problem, the scattered fields are taken as

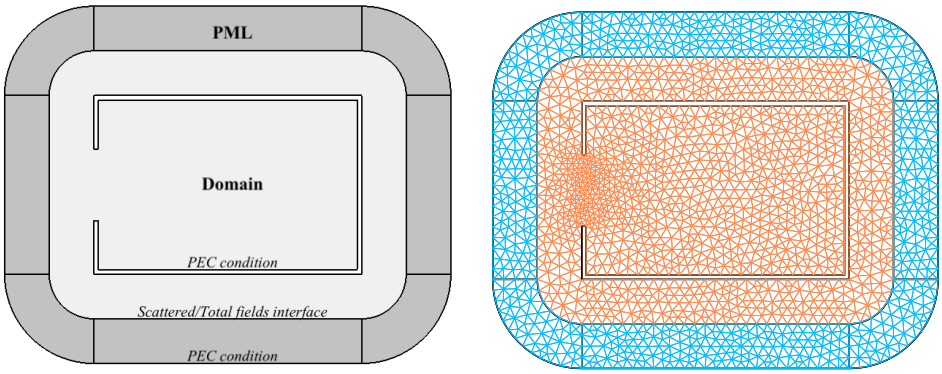


Figure 4.6: Electromagnetic scattering benchmark. Geometry and mesh.

unknowns in Σ, and the incident fields are prescribed through the compatibility relation at the interface Υ.

Evaluation of shielding effectiveness

The present benchmark is an academic problem of electromagnetic compatibility (EMC), already used by Ojeda and Pichon [133] for a study in a time-harmonic context. With the increasing use of electrical, electronic and electromagnetic systems, the study of their undesirable interactions and side effects is an important aspect to be considered by the design engineers. Numerical methods are developed and intensively used for such studies. In order to show the validity of our approach in the time domain, we recover hereafter the resonant modes of the cavity in the same setting as in [133].

Let us consider the modulated incident field

$$e_z^{\text{inc}}(\boldsymbol{x}, t) = E^i \,\text{sinc}\left(2\pi f_r \theta\right) \cos\left(2\pi f_c \theta\right),$$

with the phase

$$\theta = \frac{x - x_0}{c_0} - t,$$

where $f_r = 450\,\text{MHz}$, $f_c = 650\,\text{MHz}$ and $x_0 = 50\,\text{m}$. This *incident signal* contains frequencies in the range $[400\,\text{MHz}, 1000\,\text{MHz}]$. The time series corresponding to the incident signal, and the signal effectively measured during the numerical simulation at the center of the cavity, are plotted in Figure 4.7. The second one is called *transmitted signal*.

Figure 4.8 shows the snapshot of the $z-$component of the incident electric field at the instant $t = 7\,\text{ns}$, and the one that is effectively computed by the numerical scheme. This highlights that, since a total-field formulation is used in domain Ω and a scattered-field one is used in PML Σ, the computed solution is discontinuous at the interface Υ,

In the time-harmonic domain, the ability of the cavity to reduce the transmitted signal is quantified by the *shielding effectiveness* (SE), defined by

$$\text{SE}_{\text{dB}} = 20 \log_{10} \left| \frac{E^{\text{inc}}}{E^{\text{trans}}} \right|,$$

where E^{inc} and E^{trans} and the amplitudes of the incident wave and the transmitted one, respectively. They are obtained in the time-domain by means of a Fourrier transform on the incident signal and the transmitted one.

Figure 4.9 shows the shielding effectiness obtained with our approach for frequencies in the range $[400\,\text{MHz}, 1000\,\text{MHz}]$ and a simulation of $10.5\,\mu\text{s}$. Two minima and one maximum appear in the curve near the resonant frequencies determined by Ojeda and Pichon [133].

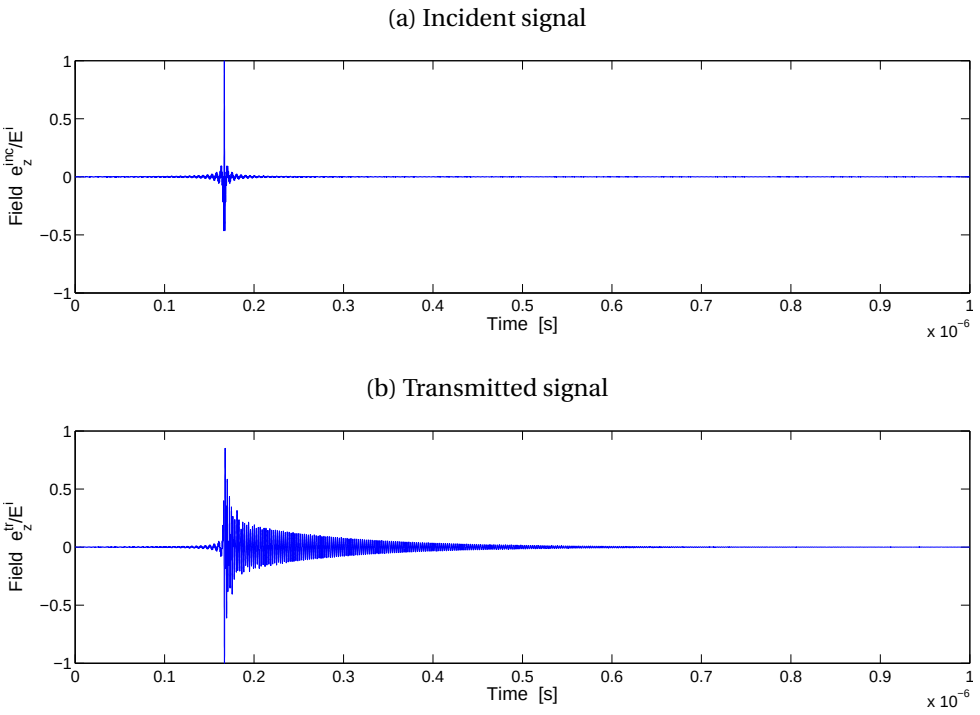


Figure 4.7: Electromagnetic scattering benchmark. Time serie of the component e_z for the incident signal (a) and the transmitted signal (b) at the center of the cavity.

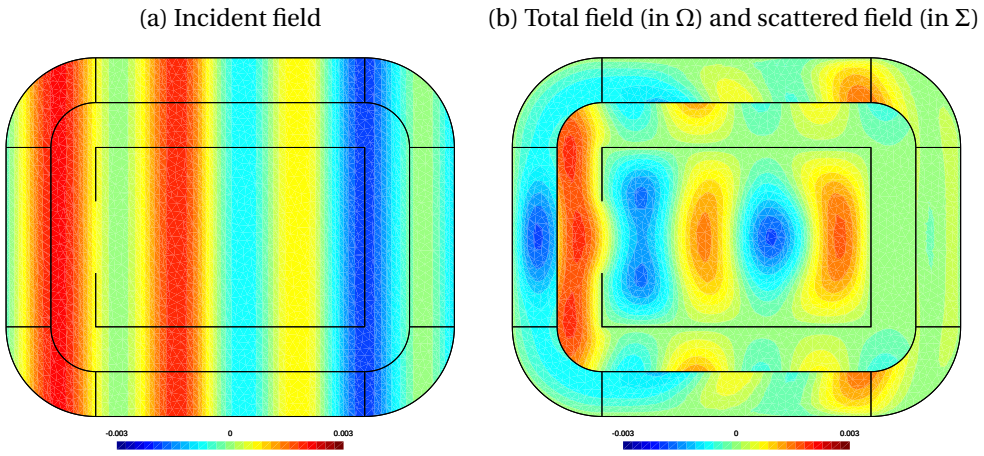


Figure 4.8: Electromagnetic scattering benchmark. Snapshot of the fields at $t = 7$ ns.

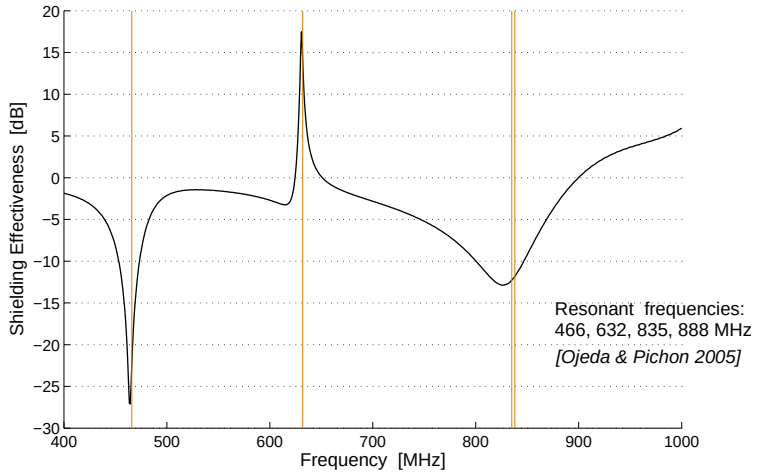


Figure 4.9: Electromagnetic scattering benchmark. Shielding effectiveness as a function of the frequency. The position of vertical lines corresponds to the resonant frequencies obtained in [133].

Optimization of absorption functions

In order to assess the results of section 3.3 in a more complicated case, the different absorption functions are optimized and compared using this benchmark. The absorption functions are

$$\sigma_2(s) = \alpha\left(\frac{s}{\delta}\right)^2, \qquad \sigma_{\mathrm{h}}(s) = \frac{\alpha}{(\delta - s)},$$
$$\sigma_3(s) = \alpha\left(\frac{s}{\delta}\right)^3, \qquad \sigma_{\mathrm{sh}}(s) = \frac{\alpha}{(\delta - s)} - \frac{\alpha}{\delta},$$

where α is the additional parameter to fit, and s is the distance between the position $\boldsymbol{x}$ and the interface Υ.

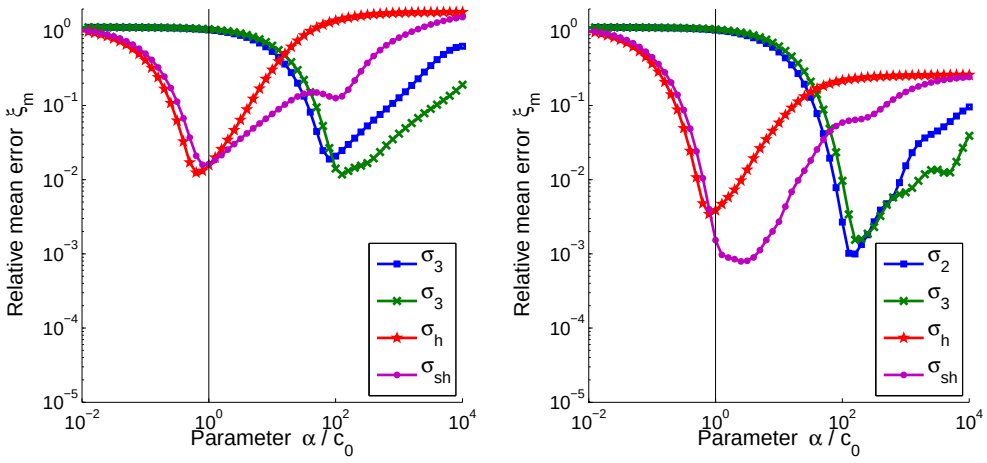
The absorption functions are optimized considering a Gaussian pulse traveling to the right as incident wave, and defining an error associated to the PML to quantify its effectiveness. The Gaussian incident pulse is prescribed through the interface condition. The $z-$component of the incident electric field reads

$$e_z^{\mathrm{inc}}(\boldsymbol{x}, t) = E^i \sqrt{\frac{\mu_0}{\varepsilon_0}} \exp\left(-\left(\frac{(x - x_0) - c_0 t}{R}\right)^2\right),$$

with $x_0 = -1200\,\mathrm{mm}$, $R = 50\,\mathrm{mm}$, $c_0 = 1/\sqrt{\mu_0 \varepsilon_0}$ and $E^i = 0.5\sqrt{\mu_0/\varepsilon_0}$, where μ_0 and ε_0 are the magnetic permeability and the electric permittivity of vacuum. A Crank-Nicolson scheme is used for the time-stepping, with the time $\Delta t = 5\,10^{-11}\,\mathrm{s}$. The duration of the simulation is $t_f = 250\,\Delta t$. Finally, the effectiveness of the PML is

(a) first-order elements and centered fluxes

(b) first-order elements and upwind fluxes



(c) second-order elements and centered fluxes

(d) second-order elements and upwind fluxes

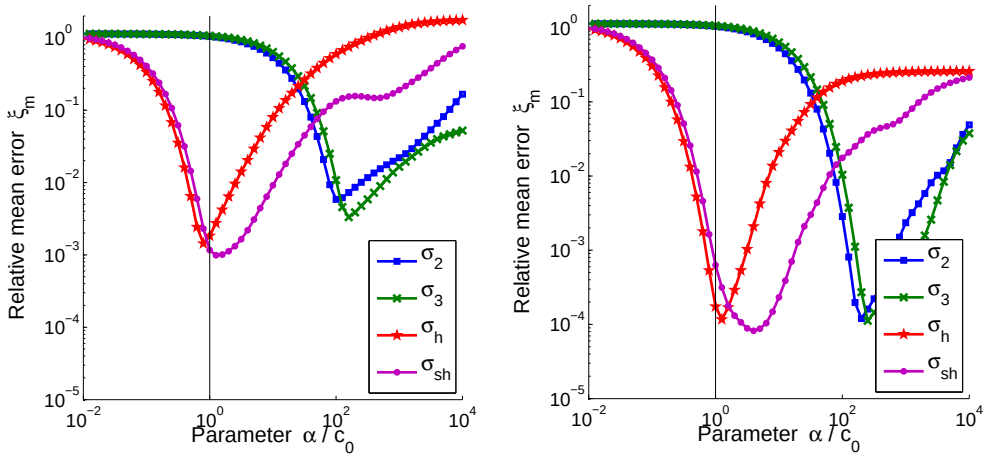


Figure 4.10: Electromagnetic scattering benchmark. Relative mean error ξ_m as a function of α with different absorption functions. First- and second-order elements are used with centered or upwind numerical fluxes. Vertical lines indicate $\alpha = c_0$.

quantified with the relative time-average error ξ_m defined by

$$\boxed{\xi_m = \sqrt{\frac{E_{\text{error}}}{E_{\text{ref}}}}}$$

with

$$E_{\text{error}} = \int_0^{t_f} \int_\Omega \left(\frac{\varepsilon_0}{2} \left\| \boldsymbol{e}_{\text{pml}} - \boldsymbol{e}_{\text{ref}} \right\|^2 + \frac{\mu_0}{2} \left\| \boldsymbol{h}_{\text{pml}} - \boldsymbol{h}_{\text{ref}} \right\|^2 \right) \mathrm{d}\Omega \, \mathrm{d}t,$$
$$E_{\text{ref}} = \int_0^{t_f} \int_\Omega \left(\frac{\varepsilon_0}{2} \left\| \boldsymbol{e}_{\text{ref}} \right\|^2 + \frac{\mu_0}{2} \left\| \boldsymbol{h}_{\text{ref}} \right\|^2 \right) \mathrm{d}\Omega \, \mathrm{d}t,$$

where $\boldsymbol{e}_{\text{ref}}$ and $\boldsymbol{h}_{\text{ref}}$ are reference fields computed with a larger computational domain.

Figure 4.10 shows the error ξ_m as a function of α for different absorption functions in different cases: centered or upwind fluxes, with first- or second-order elements. As in section 3.3, the minimum of each curve corresponds to the optimum parameter α and the minimum ξ_m that can be obtained with the absorption function. In each case, the absorption functions give equivalent results when they are optimized, except the hyperbolic function σ_{h} which does not perform so well with upwind fluxes and first-order elements. The smallest error with σ_{h} and σ_{sh} are always obtained with a value of α close to the propagation velocity c_0. These results are coherent with those obtained in section 3.3.

4.4.2 Acoustic scattering by a submarine

The effectiveness of our PML formulation is now illustrated in a three-dimensional acoustic case: the scattering of a spherical wave by a submarine in an ellipsoidal truncated domain surrounded by a PML. As for the last benchmark of section 4.3.3, the spherical wave is generated by an initial Gaussian pulse of pressure prescribed by the initial condition (4.12). All physical parameters of this benchmark are reused. The length of the submarine is approximately 120 m, the axes of the ellipsoid remain 330 m ($x-$direction) and 120 m ($y-$ and $z-$directions) in length, and the PML thickness is $\delta = 12.5$ m. The mesh is made of 921,761 tetrahedra with the characteristic size ℓ equal to 2.5 m near the PML and 1.25 m in the neighborhood of the submarine. This corresponds to 17,718,896 discrete unknowns, 49.7% of which are inside the PML. The shifted hyperbolic absorption function is used without any tuning. The boundary condition $\boldsymbol{n} \cdot \boldsymbol{u} = 0$ is prescribed at the external boundary.

Figures 4.11 and 4.12 show snapshots of the pressure p at different instants. In the first moments, the spherical wave generated by the initial pulse hits the front of the submarine, and creates perturbations in the pressure field. This is particularly visible at $t = 40$ ms. The perturbations, as well as the primary spherical wave, have correctly left the front zone at $t = 80$ ms. In the remainder of the simulation,

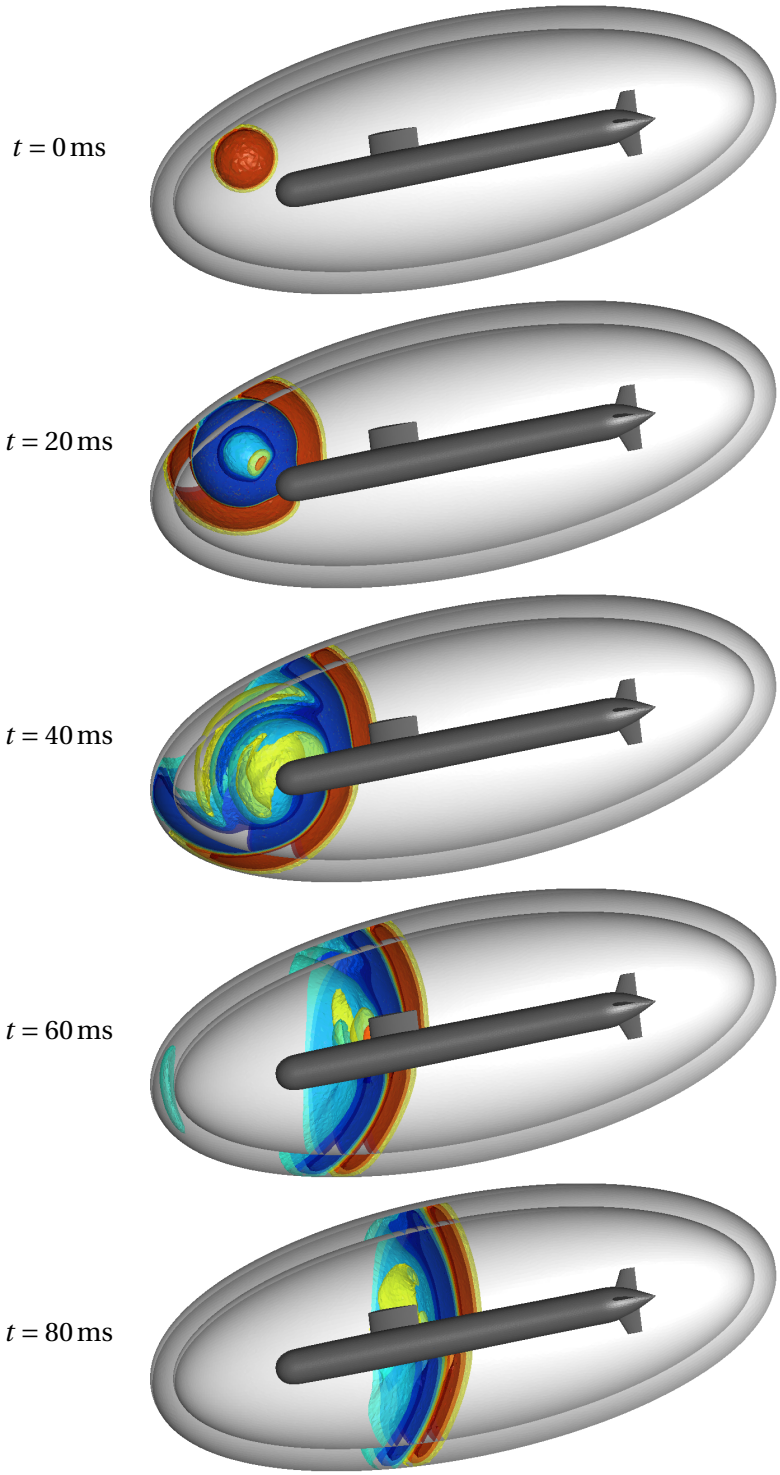


Figure 4.11: Acoustic scattering benchmark. Iso-surfaces of $p(\boldsymbol{x}, t)$ obtained with a thick PML ($\delta = 5\ell$) terminated with the boundary condition $\boldsymbol{n} \cdot \boldsymbol{u} = 0$.

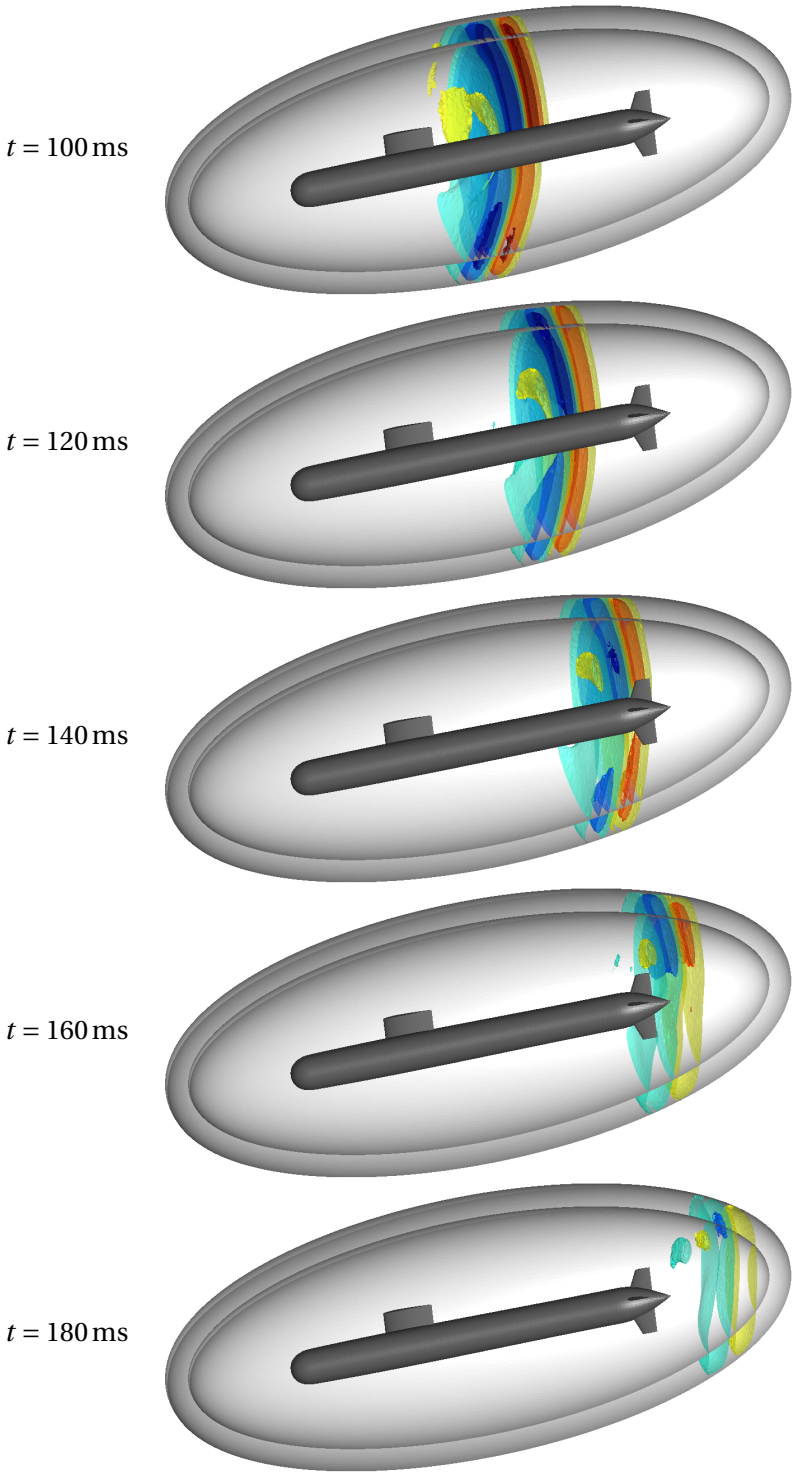


Figure 4.12: Acoustic scattering benchmark. Iso-surfaces of $p(\boldsymbol{x}, t)$ obtained with a thick PML ($\delta = 5\ell$) terminated with the boundary condition $\boldsymbol{n} \cdot \boldsymbol{u} = 0$. *Continued from Figure 4.11.*

Table 4.1: Acoustic scattering benchmark. Mesh statistics.

		ABC	PML ($\delta = \ell$)	PML ($\delta = 2\ell$)	PML ($\delta = 5\ell$)
Number of tetrahedra	in Ω	1,113,462	971,922	898,705	550,421
	in Σ		91,914	145,464	371,340
Number of unknowns	in Ω	17,815,392	15,550,752	14,379,280	8,806,736
	in Σ		2,205,936	3,491,136	8,912,160
Proportion of unknowns	in Ω	100%	87.6%	80.5%	49.7%
	in Σ		12.4%	19.5%	50.3%

the wavefront propagates along the submarine and is nearly grazing at the boundary of the domain. We observe that the spherical wave is not deformed near the interface Υ (the interior ellipsoidal surface), as we wished for, and the pressure is damped inside the layer.

It is instructive to compare this results with what is obtained if the PML is replaced with the basic absorbing boundary condition (ABC) (2.3) with approximately the same computational cost (17,815,392 discrete unknowns here). As shown in Figures 4.13 and 4.14, the spherical wave is partly reflected by the ABC in the first instants of the simulation: a spurious reflected signal appears in the snapshot corresponding to $t = 40$ ms. This signal propagates along the submarine and remains in the domain throughout the simulation. Moreover, a careful look at the primary spherical wave shows that the wavefront is altered near the ellipsoidal surface: the field decays near this boundary. Therefore, for this benchmark, even if the PML is expensive, it is a far better solution than a larger domain with the basic ABC (2.3), which provides visibly unphysical results. Let us note that alternative more effective existing ABCs are to be tested to complete this comparison.

In order to evaluate the limitations of the PML, the simulation is repeated with smaller PML thicknesses: 2.5 m and 5 m, corresponding to ℓ and 2ℓ, respectively. All simulations are performed with nearly the same computational cost: the number of discrete unknowns does not significantly change (see Table 4.1). An increase of the PML thickness is then done at the price of a smaller domain of interest Ω. For each case, a zoom of the solution at $t = 160$ ms is shown in Figure 4.15. We observe that the effectiveness of the PML diminishes when its thickness decreases: spurious signals appears for $\delta = \ell$ and 2ℓ (Figure 4.15, middle). Therefore, the use of a larger PML with a smaller domain of interest is here a good strategy.

The ABC (2.3) is now considered as an alternative boundary condition for the external boundary of the PML, instead of $\boldsymbol{n}\cdot\boldsymbol{u} = 0$. Since this ABC is naturally incorporated in the numerical scheme, there is no supplementary computational cost. The combination of the PML with an ABC termination significantly improves the effectiveness of thin layers. For $\delta = \ell$ and 2ℓ, there are indeed much less spurious signals with the ABC termination than with the Dirichlet boundary condition (Figure 4.15, bottom and middle, respectively). By contrast, for a thicker layer

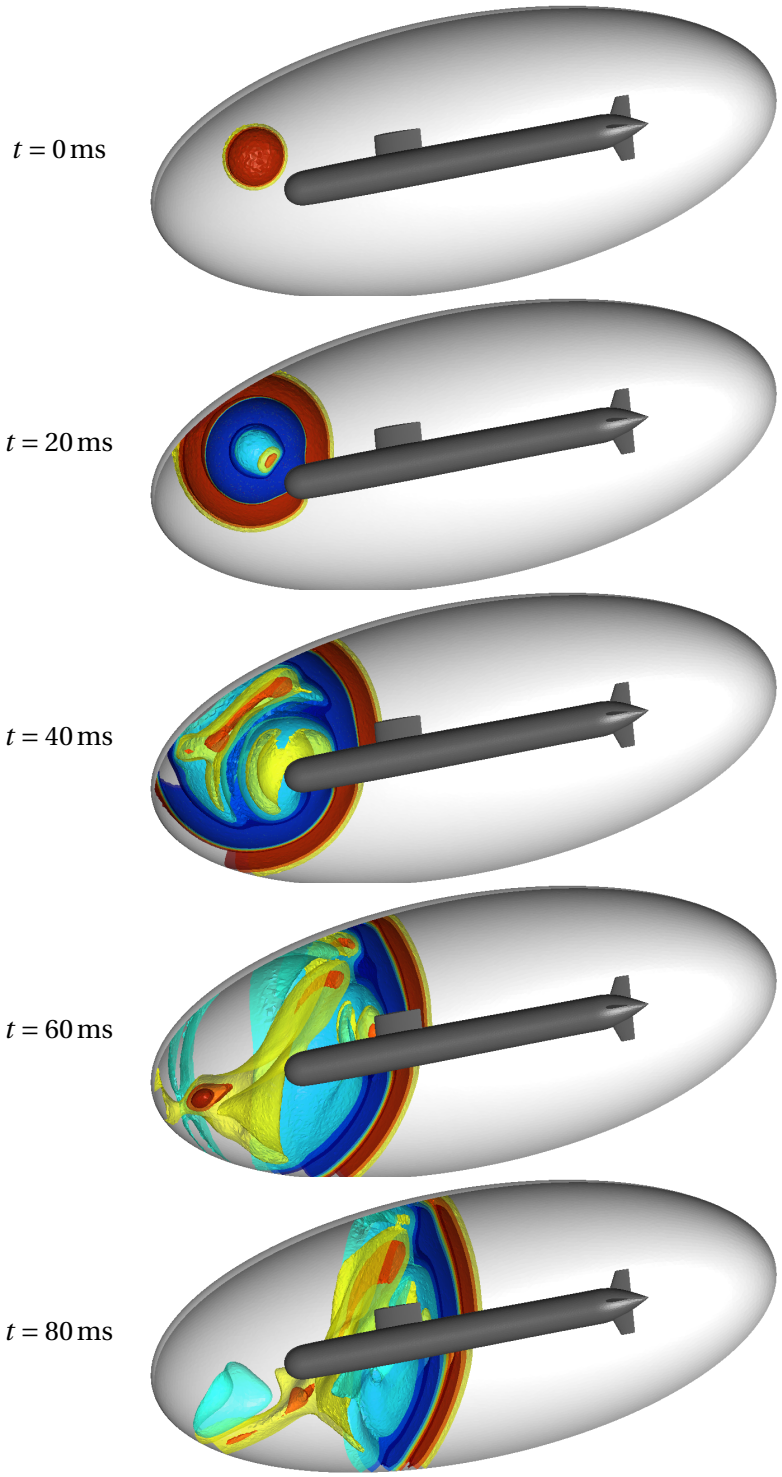


Figure 4.13: Acoustic scattering benchmark. Iso-surfaces of $p(\boldsymbol{x}, t)$ obtained with the ABC.

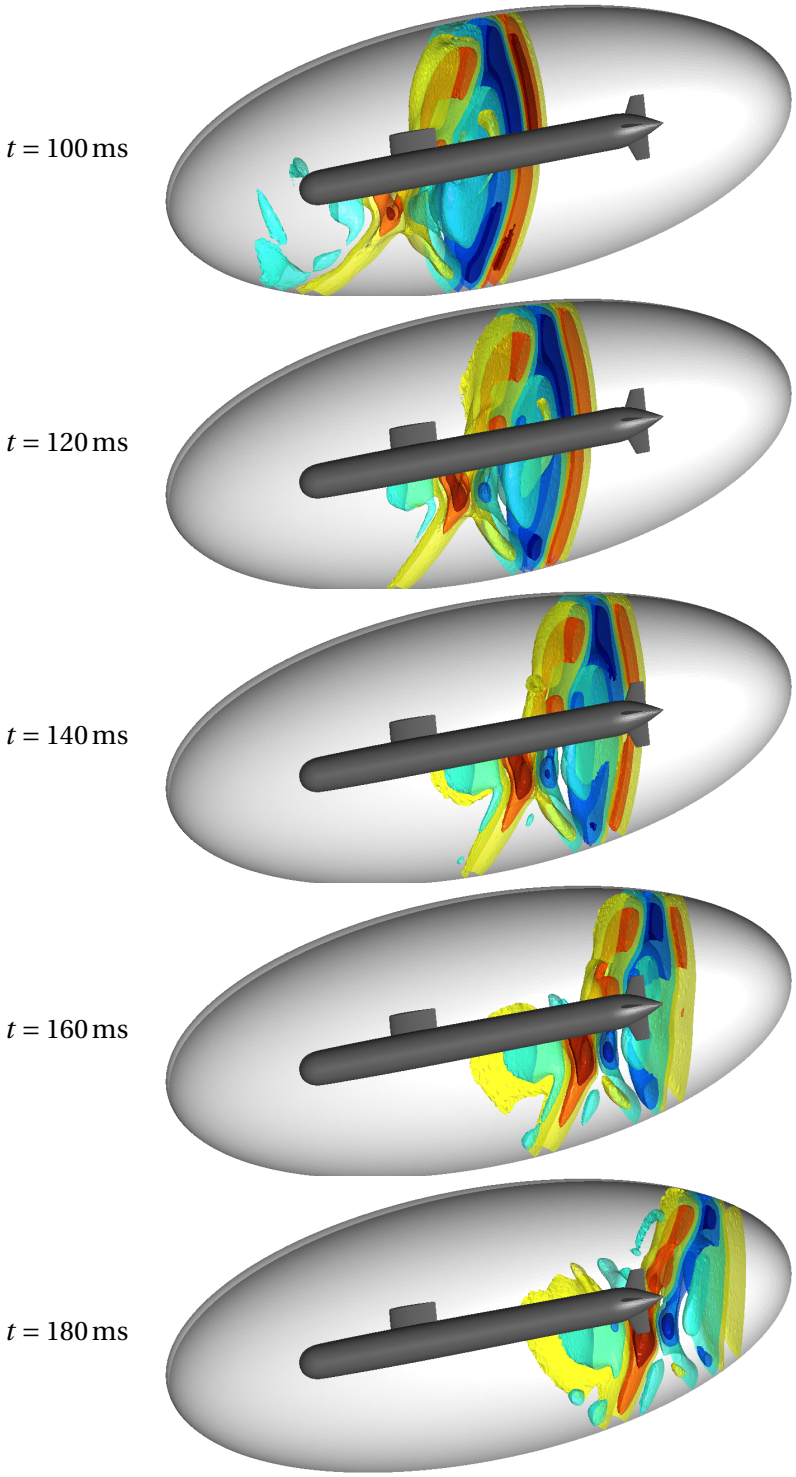


Figure 4.14: Acoustic scattering benchmark. Iso-surfaces of $p(\boldsymbol{x}, t)$ obtained with the ABC. *Continued from Figure 4.13.*

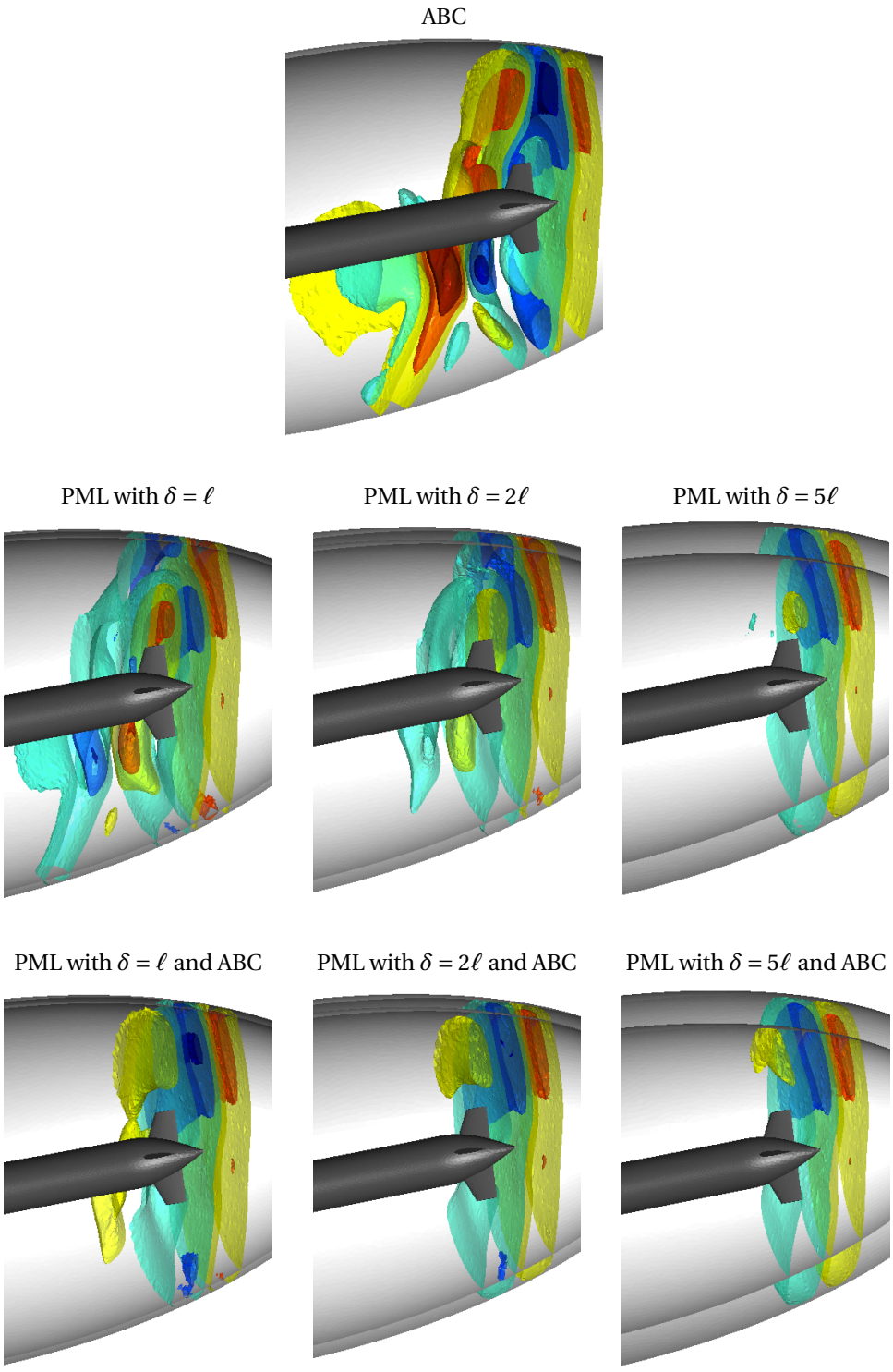


Figure 4.15: Acoustic scattering benchmark. Iso-surfaces of $p(\boldsymbol{x}, t)$ at $t = 160\,\text{ms}$ obtained with the basic ABC *(top)*, PMLs terminated with the boundary condition $\boldsymbol{n} \cdot \boldsymbol{u} = 0$ *(middle)* and PMLs terminated with the basic ABC *(bottom)*. Three PML thicknesses are considered: $\delta = \ell, 2\ell$ and 5ℓ.

($\delta = 5\ell$), the solutions obtained with and without the ABC are very similar: there is no significant improvement. Therefore, an ABC termination can improve poorly efficient PMLs, but is not necessary with thick efficient PMLs. This result has been already highlighted in a finite difference context by Petropoulos [140].

4.5 Conclusion

In this chapter, we have presented numerical results for two scattering benchmarks with PMLs, in two and three dimensions. In both cases, the results show the effectiveness of the PML formulation developed in chapter 2 for truncated domains with non-conventional shape. Along the way, we have confirmed the efficiency of the shifted hyperbolic absorption function σ_{sh} studied in chapter 3.

We have presented a DG scheme for the PML equations, and studied its convergence by means of one- and three-dimensional reference cases. The classical asymptotic convergence $\mathcal{O}(\ell^{n+1})$ of DG schemes is recovered in both cases. Since the numerical solutions are compared to the exact solutions of the unbounded problems, this confirms that there is no modeling error with PMLs.

Finally, we have addressed the issue of taking into account signals generated outside the truncated domain. In particular, we have compared different problem formulations with PMLs that account for such incoming signals. The scattered/total-field formulation is very convenient since it preserves the original model equations inside the domain, and allows one to use the PML equations without any change. The incoming signals are prescribed by only using a specific interface condition. The formulation was tested with the two-dimensional scattering benchmark.

CHAPTER 5

Absorbing Layers for Oceanography

5.1 Introduction

In regional oceanic numerical modeling, the treatment of artificial boundaries at open seas, called *open boundaries*, are often seen as a major source of uncertainty or even error. There are two reasons for this. First, the boundary treatment is supposed to account for what happens outside the model domain, which is usually poorly known: unless the model is embedded in a larger scale model or high resolution observations are available, climatological mean data are usually used to force models along their open boundaries. Second, the treatment is supposed to accurately describe the outward propagation of signals and perturbations of all kinds generated in the model interior as if the model domain was unbounded. While the former problem cannot be solved without appropriate data being available, the latter relies entirely on the mathematical and numerical formulations of the differential problem. It must therefore be solved by the implementation of appropriate boundary treatments.

In this context, a large number of boundary conditions, generally called *open boundary conditions* (OBCs), have been proposed. The simplest one, used for the first time by Elvius and Sundström [58] and popularized by Flather [63], consists in prescribing the value of the ingoing characteristic at the boundary. In this way, an external forcing is naturally taken into account through the value of the characteristic. In comparative studies [113, 131, 135, 137], this OBC appears to be the most robust for a wide range of applications. Many boundary conditions are also derived from the *radiation condition*, first introduced by Sommerfeld [154], written as

$$\frac{\partial \phi}{\partial t} + c\frac{\partial \phi}{\partial x} = 0,$$

where ϕ is any model variable. It has been adapted by various authors in order to consider dynamics encountered in realistic marine applications, such as ingoing

tide waves, large scale currents and dispersive waves generated by the topography, and the stratification or the Earth's rotation [*e.g.* 29, 32, 122, 134, 139, 146, 151, 157]. Different more accurate boundary conditions have been investigated [*e.g.* 30, 59, 115, 117–119, 165]. Particular versions have been designed for models with embedded grids [*e.g.* 28, 31, 50, 116, 132]. Extensive reviews of OBCs have been proposed by Palma and Matano [135, 136] at the end of 90's, and by Blayo and Debreu [27] in 2005.

Artificial layers form a second class of open-sea boundary treatments. The effectiveness of OBCs can be improved when used together with an artificial layer [33, 96, 150]. However, some layers are already efficient without specific boundary termination [108, 114]. Among them, layers based on the *flow relaxation scheme* (FRS) are widely used and particularly appealing by their simplicity, their effectiveness and their robustness [*e.g.* 98, 131, 135, 136]. With this scheme, the numerical solution is progressively relaxed towards an external solution in the layer [46, 47, 114]. It can be interpreted as adding a nudging term to each original model equation

$$\frac{\partial \phi}{\partial t} + F(\phi) = 0,$$

which becomes

$$\frac{\partial \phi}{\partial t} + F(\phi) = -\sigma(\phi - \phi^{\text{ext}}),$$

where ϕ is any model variable, ϕ^{ext} is the corresponding external solution and σ is a positive spatial function. When the external solution ϕ^{ext} is zero, the FRS serves as an absorbing layer, with the single aim of damping waves propagating outwards without spurious reflexion. In this case, σ corresponds to the absorption function encountered in previous chapters.

The FRS has received new attention with the introduction of the concept of *perfectly matched layer* (PML) by Bérenger [21]. In the oceanographic community, Darblade et al. [45] and Navon et al. [127] were the first to design PMLs for shallow water problems. Then, an approximate version of PML, called *pretty good sponge* (PGS), has been proposed by Lavelle and Thacker [108]. This layer exhibits better results than the FRS and requires a lower computational cost than PMLs. Finally, novel stablilized PML versions have been recently proposed in [13, 14].

This chapter is dedicated to the study of absorbing layers in the perspective of marine applications. To this aim, three main topics are addressed in the continuation of the previous chapters: the comparison of different absorbing layers; the choice of the absorption function σ, which is a common issue for the use of all the layers; and the ways of prescribing an external forcing through layers. The entire study is carried out in the context of the shallow water model, which is one of the fundamental models of oceanography. It describes some important physical dynamics encountered in realistic marine applications. Both linear and nonlinear

versions of the model equations are considered in benchmarks. Numerical simulations are performed by using a popular finite volume method with structured grids, or a new generation model based on the discontinuous Galerkin finite element method.

The chapter is structured as follows. In section 5.2, the shallow water model is described, and both classical Flather's OBC and the equations of absorbing layers are derived. These open-sea boundary treatments are compared by using a classical benchmark in section 5.3. Along the way, the comparison of absorption functions, carried out in chapter 3, is extended in this marine context. Section 5.4 deals with the use of absorbing layers in advection-dominated cases. In section 5.5, ways of prescribing an external forcing through layers are discussed. Finally, a limitation of open-sea boundary treatments is highlighted in section 5.6.

5.2 Shallow water model with open boundary

In this section, the equations of the mathematical model considered in this chapter are given. After, different open-sea boundary treatments are derived for the linearized equations.

5.2.1 Shallow water equations

The shallow water model is one of the fundamental models of geophysical fluid dynamics. It provides a mathematical description of barotropic processes that occur in oceans and shelf seas [44]. It enables modeling of tsunamis, tides and dam-breaks, and is a key building block for ocean modeling as well as atmosphere modeling.

With this model, the free-surface elevation of the water $\eta(\boldsymbol{x},t)$ and the depth-averaged velocity $\boldsymbol{u}(\boldsymbol{x},t)$ are governed by the nonlinear equations [see e.g. 44]

$$\begin{cases} \dfrac{\partial \eta}{\partial t} + \nabla\cdot(H\boldsymbol{u}) = 0, \\ \dfrac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u}\cdot(\nabla\boldsymbol{u}) + g\nabla\eta + f\boldsymbol{e}_z\times\boldsymbol{u} = \dfrac{\boldsymbol{\tau}_w-\boldsymbol{\tau}_b}{\rho_0 H} + \dfrac{1}{H}\nabla\cdot(\nu\nabla(H\boldsymbol{u}))\,. \end{cases} \tag{5.1}$$

As illustrated in Figure 5.1, $H(\boldsymbol{x},t)=\eta+h$ denotes the total depth of the water column, $h(\boldsymbol{x})$ is the depth at rest, and the unit vector $\boldsymbol{e}_z$ indicates the vertical direction. The physical parameters are the Coriolis parameter f, the gravitational acceleration g, the surface wind stress $\boldsymbol{\tau}_w$, the bottom stress $\boldsymbol{\tau}_b$, the density ρ_0 and the horizontal eddy viscosity ν.

At the beginning of the study, a simplified version of the model equations is first considered in order to design and compare boundary treatments for open seas. By neglecting the terms corresponding to advection, diffusion and stress,

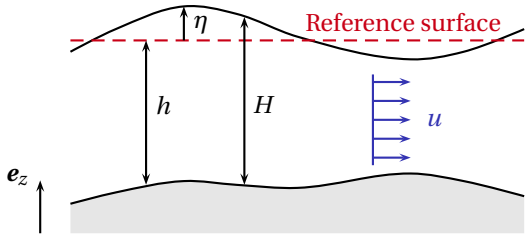


Figure 5.1: Notation for the shallow water model.

the equations (5.1) become

$$\begin{cases} \dfrac{\partial \eta}{\partial t} + h(\nabla \cdot \boldsymbol{u}) = 0, \\ \dfrac{\partial \boldsymbol{u}}{\partial t} + g\nabla\eta + f\boldsymbol{e}_z \times \boldsymbol{u} = 0. \end{cases} \tag{5.2}$$

Although simplified, this linear system includes the description of two kinds of important waves in oceanography: *Kelvin* and *Poincaré waves* [*e.g.* 44]. The former require the support of a lateral boundary such as a coast. They travel along the coast with the phase velocity $\sqrt{gh}$ and are evanescent offshore, *i.e.* they are exponentially decreasing in the direction perpendicular to the coast. The latter are encountered in unbounded domains and are dispersive, *i.e.* the phase velocity depends on the wave number.

In the absence of the Earth's rotation (*i.e.* $f = 0$), the system (5.2) reduces to the scalar wave system (1.3). Both Kelvin and Poincaré waves become classical *pure gravity waves*, which are nonevanescent and nondispersive.

5.2.2 Open-sea boundary treatments

In the context of the shallow water model, classical Flather's *open boundary condition* (OBC) reads

$$\eta - \sqrt{\frac{h}{g}}\boldsymbol{n}\cdot\boldsymbol{u} = \eta^{\text{ext}} - \sqrt{\frac{h}{g}}\boldsymbol{n}\cdot\boldsymbol{u}^{\text{ext}}, \tag{5.3}$$

where $\boldsymbol{n}$ is the outward normal, and the fields η^{ext} and $\boldsymbol{u}^{\text{ext}}$ correspond to the external solution that contains external forcing. This OBC amounts to prescribing the value of the ingoing characteristic variable of the equations at the open boundary, and can then be considered as a particular derivation of the radiation condition (see section 2.2.1). Indeed, assuming the fields vary only along $\boldsymbol{n}$, and considering

$f = 0$, the elevation and the normal velocity are then governed by

$$\begin{cases} \dfrac{\partial \eta}{\partial t} + h\dfrac{\partial (\boldsymbol{n}\cdot\boldsymbol{u})}{\partial x_n} = 0, \\ \dfrac{\partial (\boldsymbol{n}\cdot\boldsymbol{u})}{\partial t} + g\dfrac{\partial \eta}{\partial x_n} = 0, \end{cases}$$

where x_n is the coordinate along $\boldsymbol{n}$. Some algebraic manipulations allow to rewrite this system with the characteristic equations,

$$\begin{cases} \dfrac{\partial w_1}{\partial t} + \sqrt{gh}\,\dfrac{\partial w_1}{\partial x_n} = 0, \\ \dfrac{\partial w_2}{\partial t} - \sqrt{gh}\,\dfrac{\partial w_2}{\partial x_n} = 0, \end{cases}$$

where the characteristic variables, which are Riemann invariants, are

$$w_1 = \eta + \sqrt{\frac{h}{g}}\,(\boldsymbol{n}\cdot\boldsymbol{u}),$$
$$w_2 = \eta - \sqrt{\frac{h}{g}}\,(\boldsymbol{n}\cdot\boldsymbol{u}).$$

The quantity w_1 is transported outward the domain at the velocity $\sqrt{gh}$ and is entirely determined by what happens inside. On the contrary, w_2 is transported inwards and then depends on incoming information. Therefore, the boundary condition (5.3) consists indeed in prescribing the value of the incoming characteristic variable, *i.e.* w_2, with information coming from the exterior.

The idea of the *flow relaxation scheme* (FRS) is to progressively relax the numerical solution towards an external solution in a layer that surrounds the domain. The relaxation is performed by nudging terms that are added to each original model equation [46]. In the layer, the linear system (5.2) then becomes

$$\begin{cases} \dfrac{\partial \eta}{\partial t} + h(\nabla\cdot\boldsymbol{u}) = -\sigma(\eta - \eta^{\text{ext}}), \\ \dfrac{\partial \boldsymbol{u}}{\partial t} + g\nabla\eta + f\,\boldsymbol{e}_z\times\boldsymbol{u} = -\sigma(\boldsymbol{u} - \boldsymbol{u}^{\text{ext}}), \end{cases} \tag{5.4}$$

where σ is a positive spatial function [114]. The full dynamics of the model is then taken into account in a natural way, and the implementation is also straightforward, even in complex geometries. The method is therefore applicable to a wide range of problems. The original equations are easily recovered by using $\sigma = 0$ in the domain of interest.

Using the *pretty good sponge* (PGS) approach developed by Lavelle and Thacker [108], the linearized shallow water equations, written in the Cartesian coordinate

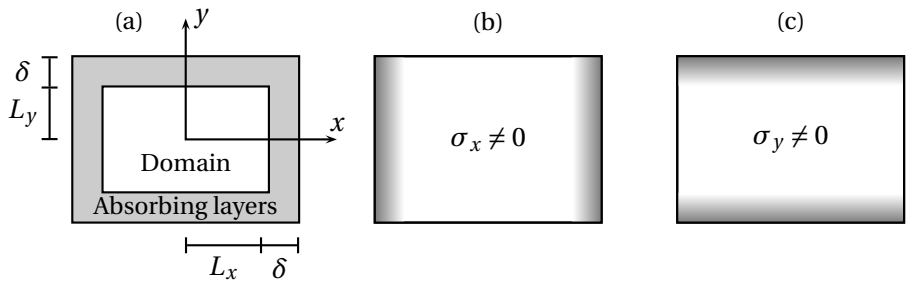


Figure 5.2: (a) Geometry of the two-dimensional problem with the PGS. (b)-(c) Layers where the absorption functions $\sigma_x(x)$ and $\sigma_y(y)$ differ from zero (in gray). *Reproduced from [123] (© 2009 Springer-Verlag).*

system, become

$$\begin{cases} \dfrac{\partial \eta}{\partial t} + h\left(\dfrac{\partial u_x}{\partial x} + \dfrac{\partial u_y}{\partial y}\right) = -(\sigma_x + \sigma_y)(\eta - \eta^{\text{ext}}), \\ \dfrac{\partial u_x}{\partial t} + g\dfrac{\partial \eta}{\partial x} - f u_y = -\sigma_x(u_x - u_x^{\text{ext}}), \\ \dfrac{\partial u_y}{\partial t} + g\dfrac{\partial \eta}{\partial y} + f u_x = -\sigma_y(u_y - u_y^{\text{ext}}), \end{cases} \tag{5.5}$$

where u_x and u_y are the Cartesian components of $\boldsymbol{u}$ in the x- and the y-directions, while $\sigma_x(x)$ and $\sigma_y(y)$ are two positive functions associated to the relaxation of fields in these directions. As shown in Figure 5.2, σ_x (resp. σ_y) is nonzero only in layers where the fields are relaxed along the x-direction (resp. y-direction). These functions vanish in the interior of the domain so that original shallow water equations are recovered.

A set of PML equations can be obtained for the linearized shallow water equations by using the strategy described in chapter 2. In the case of a straight PML, with $\boldsymbol{n}$ as stretch direction, we obtain the PML system

$$\begin{cases} \dfrac{\partial \eta}{\partial t} + h(\nabla \cdot \boldsymbol{u}) = -\sigma \eta_n, \\ \dfrac{\partial \boldsymbol{u}}{\partial t} + g\nabla\eta + f\,\boldsymbol{e}_z \times \boldsymbol{u} = -\sigma \boldsymbol{n}[(\boldsymbol{n}\cdot\boldsymbol{u}) - u_c], \\ \dfrac{\partial \eta_n}{\partial t} + h(\nabla_n \cdot \boldsymbol{u}) = -\sigma \eta_n, \\ \dfrac{\partial u_c}{\partial t} + f\,\boldsymbol{n}\cdot(\boldsymbol{e}_z \times \boldsymbol{u}) = 0. \end{cases} \tag{5.6}$$

This system differs from the one (2.30) obtained for the scalar wave system only by the introduction of the supplementary field u_c. Thanks to this field, the change of the waves properties due to the Earth's rotation is taken into account.

We derive an approximate PML by neglecting the additional field of the PML system (5.6), *i.e.* $u_c \approx 0$, but keeping the Coriolis term in the momentum equation. The system then becomes

$$\begin{cases} \dfrac{\partial \eta}{\partial t} + h(\nabla \cdot \boldsymbol{u}) = -\sigma \eta_n, \\ \dfrac{\partial \boldsymbol{u}}{\partial t} + g\nabla\eta + f\boldsymbol{e}_z \times \boldsymbol{u} = -\sigma \boldsymbol{n}(\boldsymbol{n}\cdot\boldsymbol{u}), \\ \dfrac{\partial \eta_n}{\partial t} + h(\nabla_n \cdot \boldsymbol{u}) = -\sigma \eta_n. \end{cases} \tag{5.7}$$

This alternative layer is a true PML when the Earth's rotation is neglected (*i.e.* $f \approx 0$), and the PGS equations are recovered when the additional field η_n is assumed to be equal to η.

Finally, let us note that all the above layers are perfectly matched for waves with a normal incidence when $f = 0$. Indeed, in this case, the systems (5.4)-(5.7) are equivalent to the system (2.4) with the perfectly matching condition.

5.3 Cases dominated by Poincaré waves

Hereafter, the open-sea boundary treatments presented in the previous section are compared and studied by means of a classical benchmark of the literature. We consider the collapse of a mound of water in an open region [108, 115, 127], which is a case dominated by Poincaré waves. This benchmark is very similar to the one considered in section 3.3.3.2. The novelty here is the influence of Earth's rotation on dynamics, which makes the waves dispersive.

We first present the benchmark. Then, the different open-sea boundary treatments are compared in the basic setting: the flows are governed by the linearized shallow water equations (5.2), which is the simplest system supporting Poincaré waves. The impact of adding supplementary (non-dominant) dynamics such as advection, bottom friction and diffusion is studied by employing the nonlinear model (5.1). After, the optimum values of the absorption functions are compared in order to assess the results of the previous chapters in a modified physical setting.

5.3.1 Description of the benchmark

Let us consider a rectangular domain $\Omega = [-L_x, L_x] \times [-L_y, L_y]$ with $L_x = 1,100\,\text{km}$ and $L_y = 510\text{km}$, where each side is an open boundary. A Gaussian-shaped mound of water is prescribed as initial condition

$$\eta(\boldsymbol{x}, 0) = \eta_0 e^{-\|\boldsymbol{x}\|^2/R^2},$$

with $\eta_0 = 1\,\text{m}$ and $R = 50\,\text{km}$. The fluid is initially at rest, and the reference height $h = 100\,\text{m}$. We use the values $f = 1.028\,10^{-4}\,\text{s}^{-1}$ and $g = 9.85\,\text{m/s}^2$.

Figure 5.3(a) shows some snapshots of the reference solution for the collapse of the initial mound of water and the associated progressive geostrophic adjustment. This reference solution is computed using an extended computational domain in order to avoid any reflection at the open boundaries. During the collapse, circular waves are created and propagate outwards. After about 5h the main wavefront is expected to reach the upper and lower boundaries of the model domain and to hit the absorbing layers at normal incidence. As time goes by, the wavefront propagates along the boundaries and approach the absorbing layers with a decreasing angle of incidence.

In order to highlight the kind of error generated when using an open-sea boundary treatment, snapshots of the error on the elevation are shown in Figure 5.3(b) for the case where PGSs are added along each side of the domain, according to Figure 5.2. The thickness of the layers is $\delta = 130\,\text{km}$ and the shifted hyperbolic absorption function

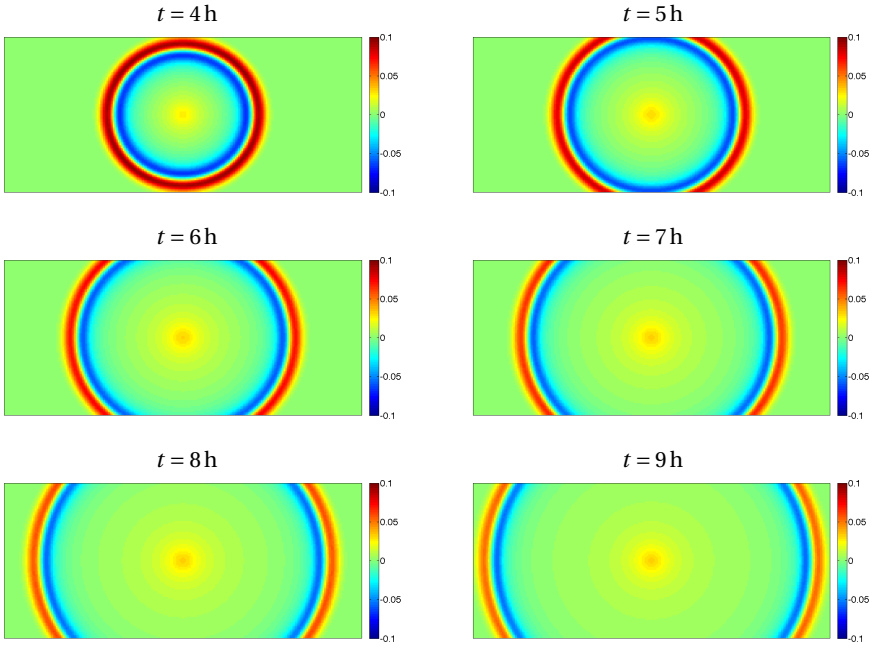
$$\sigma_{\text{sh}}(s) = \frac{\alpha}{\delta - s} - \frac{\alpha}{\delta} \tag{5.8}$$

is used in both x- and y-directions with $\alpha = \sqrt{gh}$, in agreement with the results of chapter 3. The absorption functions σ_x and σ_y of the PGS system (5.5) are explicitly obtained by replacing the variable s of the formula (5.8) by $|x| - L_x$ and $|y| - L_y$, respectively. In this simulation, when the wavefront hits the boundary, the error is small until time $t = 6\text{h}$. This small error is associated with a weak reflection of the waves with quasi-normal incidence. The snapshots taken later show that the error increases as the incidence of the wavefront approaching the boundary decreases and the waves tend to propagate along the boundary. The initial wavefront is partly reflected at the open boundary and the reflected (and damped) wave propagates inside the domain.

5.3.2 Comparison of open-sea boundary treatments

A comparison of boundary treatments is performed in the context of a novel generation ocean model, which is based on the discontinuous Galerkin method. The discretization of shallow water equations is detailed in [42], while additional terms and supplementary equations of PMLs are discretized following the strategy described in chapter 4. The shifted hyperbolic absorbing function (5.8) with $\alpha = \sqrt{gh}$ is used for all the layers. The time-stepping is made with the fourth-order Runge-Kutta method, and the mesh is unstructured and made of triangles. Since our PML formulations do not deal with corners, open-sea boundary treatments are used only on the northern and southern borders of the rectangular domain. A wall condition (*i.e.* $\boldsymbol{n} \cdot \boldsymbol{u} = 0$) is prescribed at other borders — since circular waves generated by the barotropic adjustment do not to reach these borders before the end of the simulation (9h), a specific boundary treatment is not necessary.

(a) Elevation in the domain for the reference solution.



(b) Error on the elevation with the pretty good sponge (PGS).

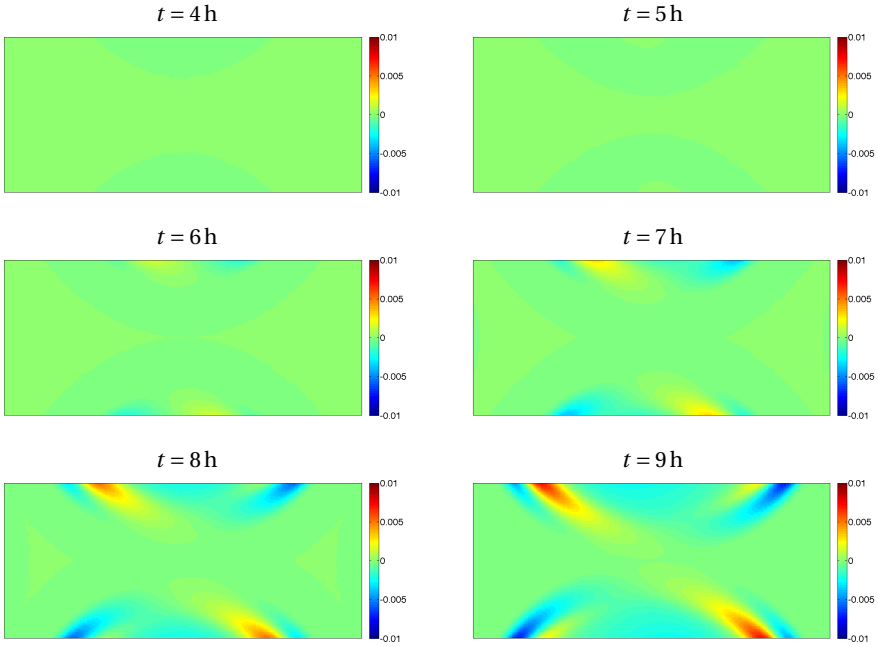


Figure 5.3: Simulation of the collapse of a mound of water. Snapshots of the elevation in the domain in the reference case (a) and of the error on the elevation using the PGS at each boundary (b). The reference solution is computed with an extended domain. The thickness of the layers δ is 130km and the absorption functions are shifted hyperbolic with $\alpha = \sqrt{gh}$. *Adapted from [123] (© 2009 Springer-Verlag).*

Figure 5.4 shows the snapshot of the final error on the elevation for Flather's OBC and the different absorbing layers with the thickness $\delta = 50\,\text{km}$. The OBC, the FRS and the PGS present the same kind of error that has been already observed above: the circular waves generated by the collapse of the mound of water are more and more reflected when the incidence of the wavefront decreases and the waves tend to propagate along the boundary.

The error is much smaller with the PMLs than with the other boundary treatments. With the complete PML, the error looks like noise (Figure 5.5, right). The small spurious oscillations observed are caused by the numerical approximation: numerical errors generated in the layer are propagated towards the domain. For the approximate PML, the aspect of the error is different and its value is one order of magnitude larger (Figure 5.5, left). Because the only difference with the complete PML is the approximation $u_c \approx 0$, the supplementary errors are the result of this approximation. Therefore, the numerical errors are overtaken by new modeling errors.

In order to compare the effectiveness of the different boundary treatments in a quantitative way, the energy associated with the error fields is computed according to

$$E_{\text{error}}(t) = \int_\Omega \left(\frac{g}{2} \left| \eta(\boldsymbol{x}, t) - \eta_{\text{ref}}(\boldsymbol{x}, t) \right|^2 + \frac{h}{2} \left\| \boldsymbol{u}(\boldsymbol{x}, t) - \boldsymbol{u}_{\text{ref}}(\boldsymbol{x}, t) \right\|^2 \right) \mathrm{d}\boldsymbol{x},$$

where the fields $\eta_{\text{ref}}(\boldsymbol{x}, t)$ and $\boldsymbol{u}_{\text{ref}}(\boldsymbol{x}, t)$ correspond to the final solution of a reference run in which reflections are avoided by means of a much larger computational domain. Similarly to our approach of section 3.3.3.2, the relative error

$$\boxed{\xi_r = \sqrt{\frac{E_{\text{error with layer}}(t_f)}{E_{\text{error with wall}}(t_f)}}} \tag{5.9}$$

is used to measure the reflected part of the solution at the final time $t_f = 9\,\text{h}$. Table 5.1 shows the relative error for the different treatments. Different values are used for both the layer thickness δ and the characteristic length of a mesh cell ℓ.

Both Flather's OBC and thin FRS (with $\delta = 50\,\text{km}$) provide similar results: the relative error is close to 0.15 and does not vary with ℓ. Therefore, the only way to improve the results of the FRS is to increase the thickness of the layer δ. For a wider FRS layer, for the PGS and for the approximate PML, a decrease of ℓ reduces the relative error. However, the reduction is small: the decrease of ℓ from 30 km to 20 km and from 20 km to 10 km reduces the relative error by only about $4-5\%$ and 2% in all the cases. By contrast, the reductions are greater with the complete PML: respectively 63% and 89%. Therefore, this corroborates that the results obtained with the complete PML depend strongly on the discretization and relevant *numerical errors*, while the error of other boundary treatments is dominated by *modeling errors*.

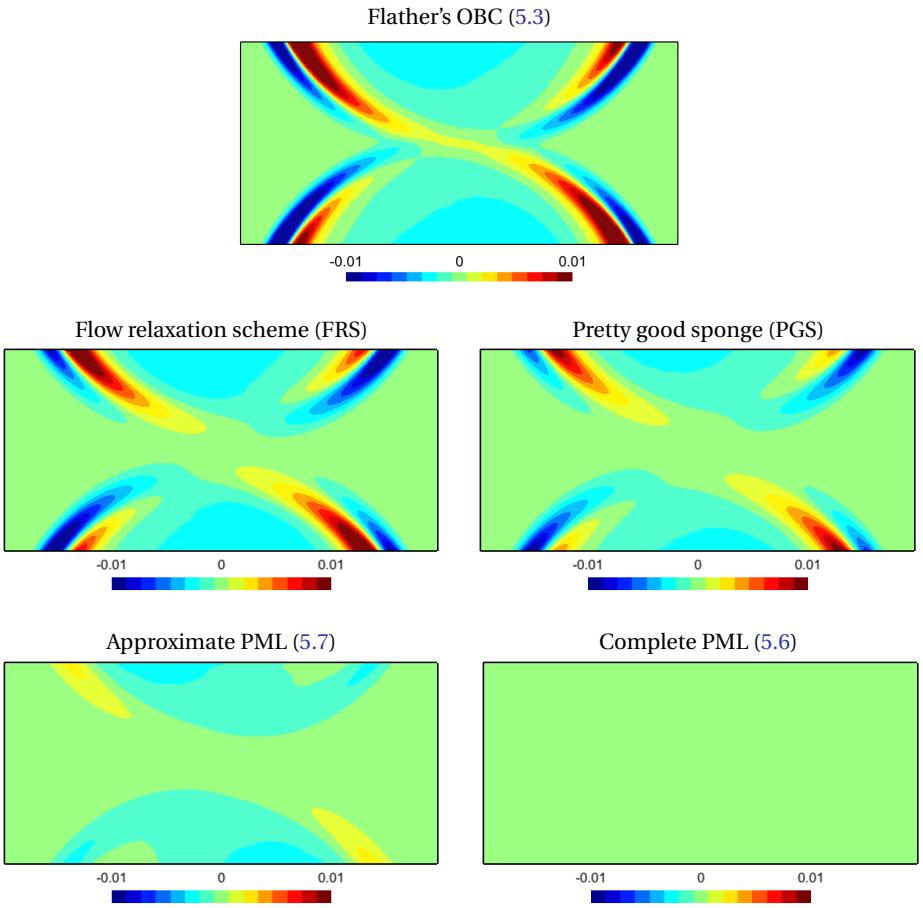


Figure 5.4: Simulation of the collapse of a mound of water with the DG scheme. Error on the elevation at the end of the simulation using Flather's OBC and the different absorbing layers with the thickness $\delta = 100\,\text{km}$ and the shifted hyperbolic absorption function σ_{sh}. The characteristic length size of the mesh cells is $\ell = 20\,\text{km}$ and the time step is $\Delta t = 60\,\text{s}$. The color bar has been fixed from $-0.01\,\text{m}$ to $0.01\,\text{m}$.

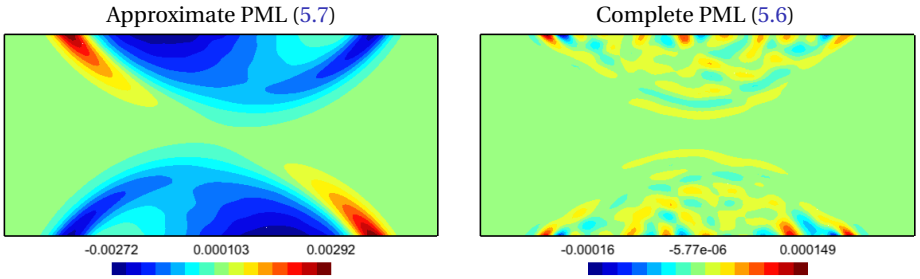


Figure 5.5: Simulation of the collapse of a mound of water with the DG scheme. Error on the elevation at the end of the simulation using the PMLs and the same parameters as Figure 5.4. The color bar goes now from the minimum to the maximum in each case.

Table 5.1: Relative error ξ_r (5.9) for the OBC and different absorbing layers. Three thicknesses of the absorbing layers δ and three characteristic lengths of cell ℓ are considered. For all simulations, the CFL number is $\sqrt{gh}\,(\Delta t)/\ell \approx 0.0942$.

	$\ell = 30\,\text{km}$	$\ell = 20\,\text{km}$	$\ell = 10\,\text{km}$
Flather's OBC	$1.48\ 10^{-1}$	$1.48\ 10^{-1}$	$1.50\ 10^{-1}$
Absorbing layers with $\delta = 50\,\text{km}$			
Flow relaxation scheme (FRS)	$1.49\ 10^{-1}$	$1.49\ 10^{-1}$	$1.49\ 10^{-1}$
Pretty good sponge (PGS)	$1.17\ 10^{-1}$	$1.16\ 10^{-1}$	$1.15\ 10^{-1}$
Approximate PML	$5.37\ 10^{-2}$	$5.04\ 10^{-2}$	$4.97\ 10^{-2}$
Complete PML	$9.18\ 10^{-3}$	$3.49\ 10^{-3}$	$1.55\ 10^{-3}$
Absorbing layers with $\delta = 100\,\text{km}$			
Flow relaxation scheme (FRS)	$1.17\ 10^{-1}$	$1.14\ 10^{-1}$	$1.13\ 10^{-1}$
Pretty good sponge (PGS)	$8.17\ 10^{-2}$	$7.93\ 10^{-2}$	$7.84\ 10^{-2}$
Approximate PML	$4.30\ 10^{-2}$	$4.12\ 10^{-2}$	$4.04\ 10^{-2}$
Complete PML	$1.50\ 10^{-3}$	$5.90\ 10^{-4}$	$1.48\ 10^{-4}$
Absorbing layers with $\delta = 150\,\text{km}$			
Flow relaxation scheme (FRS)	$8.43\ 10^{-2}$	$8.13\ 10^{-2}$	$7.99\ 10^{-2}$
Pretty good sponge (PGS)	$5.43\ 10^{-2}$	$5.22\ 10^{-2}$	$5.12\ 10^{-2}$
Approximate PML	$3.41\ 10^{-2}$	$3.26\ 10^{-2}$	$3.20\ 10^{-2}$
Complete PML	$6.92\ 10^{-4}$	$2.54\ 10^{-4}$	$2.89\ 10^{-5}$

For each δ — ℓ combination, the PGS performs better than the FRS but worse than the approximate PML. The values of relative error ξ_r are close in these cases. By contrast, the values for the PML are lower by one or more orders of magnitude. The difference is larger with thick layers and finer meshes. The thinnest complete PML with a coarse mesh provides a smaller relative error ($9.18\ 10^{-3}$) than the best other layer with a fine mesh ($3.20\ 10^{-2}$). It demonstrates the accuracy of this layer for a benchmark dominated by Poincaré waves.

Impact of additional non-dominant dynamics

The relative error slightly changes when advection, bottom friction and diffusion are included in the simulation. Table 5.2 shows the new values when these dynamics are added separately, and also when they are considered together. Advection is taken into account by using the nonlinear equations (5.1) instead of the linear ones. The bottom friction is modeled through the quadratic bottom stress

$$\boldsymbol{\tau}_b = \frac{\rho_0 g}{C^2}\,\|\boldsymbol{u}\|\,\boldsymbol{u}, \tag{5.10}$$

with the Chezy coefficient $C = 44.4\,\text{m}^{1/2}/\text{s}$. When diffusion is added, the horizontal eddy viscosity is $\nu = 200\,\text{m}^2/\text{s}$.

Two options are considered for the absorbing layers. In the first implementation, labeled '*without dynamics*', the Coriolis term is neglected and the additional dynamics are not considered inside the layers. The full dynamics is taken into account in the variant labeled '*with dynamics*': the Coriolis, nonlinear advection,

Table 5.2: Relative error ξ_r (5.9) for the OBC and the different absorbing layers using different variations of the benchmark with additional dynamics. The parameters are $\delta = 100$km, $\ell = 20$ km and $\Delta t = 60$ s.

Additionnal dynamics:	*None*	*Advection*	*Friction*	*Diffusion*	*All*
Flather's OBC	$1.48\ 10^{-1}$	$1.49\ 10^{-1}$	$1.49\ 10^{-1}$	$1.48\ 10^{-1}$	$1.49\ 10^{-1}$
Absorbing layers '*without dynamics*'					
Flow relaxation scheme (FRS)	$1.18\ 10^{-1}$	$1.18\ 10^{-1}$	$1.19\ 10^{-1}$	$1.18\ 10^{-1}$	$1.19\ 10^{-1}$
Pretty good sponge (PGS)	$9.89\ 10^{-2}$	$9.92\ 10^{-2}$	$9.96\ 10^{-2}$	$9.90\ 10^{-2}$	$1.00\ 10^{-1}$
Approximate PML	$9.03\ 10^{-2}$	$9.07\ 10^{-2}$	$9.11\ 10^{-2}$	$9.05\ 10^{-2}$	$9.16\ 10^{-2}$
Absorbing layers '*with dynamics*'					
Flow relaxation scheme (FRS)	$1.14\ 10^{-1}$	$1.15\ 10^{-1}$	$1.15\ 10^{-1}$	$1.15\ 10^{-1}$	$1.15\ 10^{-1}$
Pretty good sponge (PGS)	$7.93\ 10^{-2}$	$7.96\ 10^{-2}$	$7.97\ 10^{-2}$	$7.95\ 10^{-2}$	$8.01\ 10^{-2}$
Approximate PML	$4.12\ 10^{-2}$	$4.14\ 10^{-2}$	$4.17\ 10^{-2}$	$4.13\ 10^{-2}$	$4.19\ 10^{-2}$
Complete PML	$5.90\ 10^{-4}$	$6.00\ 10^{-4}$	$7.41\ 10^{-4}$	$5.88\ 10^{-4}$	$7.58\ 10^{-4}$
Modified PML	–	$6.38\ 10^{-4}$	$6.24\ 10^{-4}$	$5.88\ 10^{-4}$	–

bottom friction and diffusion are added in the modified continuity and momentum equations used in the layers.

A quick comparison of both variants listed in Table 5.2 shows that the simplification of the dynamics in the absorbing layer can lead to a significant increase of the relative error in all the cases. In order to get the most out of the absorbing layer, it is important to take advantage of the ability of the absorbing layer to accommodate the full dynamics of the system — which is of course also a sensible matter from a physical perspective.

The performances of the different layers with additional dynamics are very similar to those of the original case. The error increases by less than 1% with respect to the linear problem in most of the cases. The only exception is for the complete PML when the friction term is considered: the error increases by about 25%, but stays lower than values obtained with other layers. Therefore, the conclusions of the original case remain valid.

In order to improve the performances of the complete PML, a modified version was tested. Since the governing equation for u_c in the system (5.6) contains the projection of the Coriolis term on $\boldsymbol{n}$, we add the projection of advection, bottom friction and diffusion terms in this equation, which becomes

$$\frac{\partial u_c}{\partial t} + \boldsymbol{n}\cdot\Big[\boldsymbol{u}\cdot(\nabla\boldsymbol{u}) + f\boldsymbol{e}_z\times\boldsymbol{u}\Big] = \boldsymbol{n}\cdot\Big[-\frac{\boldsymbol{\tau}_b}{\rho_0 H} + \frac{1}{H}\nabla\cdot(\nu\nabla(H\boldsymbol{u}))\Big].$$

The obtained values of relative error ξ_r are given in the last line in Table 5.2. They show that this strategy works for the friction term (decrease of the error by 15%), but fails for the others.

5.3.3 Comparison of absorption functions

The choice of the absorption function σ is a common issue for all the absorbing layers. This problem has already been addressed in chapter 3 for the scalar wave system. Since the physical dynamics considered here slightly differ due to Earth's rotation, we aim to check if the shifted hyperbolic function σ_{sh} remains efficient with $\alpha = \sqrt{gh}$. Let us note that the results of chapter 3 directly apply here for the PMLs (5.6) and (5.7) when Earth's rotation is neglected (*i.e.* $f \approx 0$), and for the FRS (5.4) and the PGS (5.5) in one dimension. Indeed, in these cases, the systems reduce to the PML scalar wave system.

The comparison of absorption functions is performed with a layer at each side of the domain. Only the PGS is considered since it is more efficient than the FRS and, as will be explained in section (5.4), it is more robust that PMLs. The equations are discretized on an Arakawa C-grid using a finite volume scheme and the time-stepping is made with a forward-backward strategy [20, 44]. The spatial steps are $\Delta x = \Delta y = 10\,\text{km}$ and the time step is $\Delta t = 150\,\text{s}$. These parameters are similar to those used by Lavelle and Thacker [108] for the same benchmark (but with a square domain).

Figure 5.6 shows the error at the end of the simulations ($t = 9\text{h}$) for two different thicknesses of the PGS and different absorption functions discussed in chapter 3. We consider the parabolic function σ_2 and the hyperbolic function σ_{h}, defined by

$$\sigma_2(s) = \alpha \left(\frac{s}{\delta}\right)^2, \qquad \sigma_{\text{h}}(s) = \frac{\alpha}{\delta - s},$$

where α is a free constant parameter, as well as the shifted hyperbolic function σ_{sh} (5.8) and the constant function $\bar{\sigma}$. The value of the constant function $\bar{\sigma}$ and the free parameter α of σ_2 are numerically optimized in order to reduce the associated error ξ_r for that particular simulation. As in section 3.3.3.2, an optimization of hyperbolic and shifted hyperbolic functions by adjustment of a multiplicative factor is useless: these functions are already optimum when α equals the propagation velocity, which is $\sqrt{gh}$. The results obtained with the parabolic function with $\alpha = 0.9/\Delta t$ advocated by Lavelle and Thacker [108] are also shown. The case where the waves are perfectly reflected is computed using a wall condition at the boundary of the domain Ω, *i.e.* $\boldsymbol{u}\cdot\boldsymbol{n} = 0$.

The error patterns obtained with the parabolic functions are similar to the ones shown in Figure 5.3(b) for the absorption function (5.8). The different solutions are qualitatively similar with very good properties for normal waves and a reduced effectiveness for tangential waves. The optimum constant function $\bar{\sigma}$ gives also better results for normal waves than for tangential waves but they are globally worse than with the spatially varying functions considered.

As expected, the best results are obtained with the thickest absorbing layer ($\delta = 13\,\Delta x$ in Figure 5.6). For a thin absorbing layer ($\delta = 5\,\Delta x$ in Figure 5.6), only

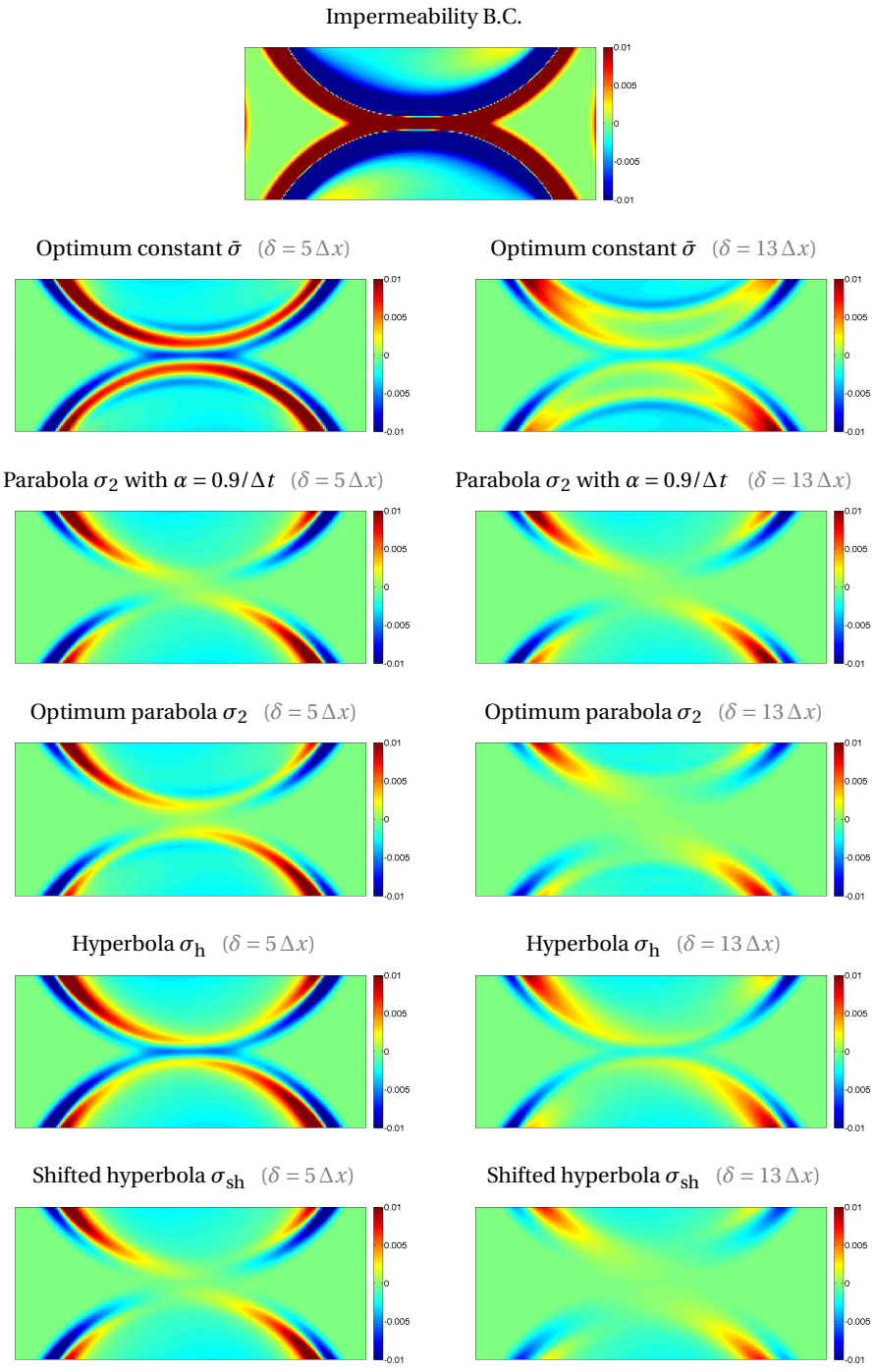


Figure 5.6: Simulation of the collapse of a mound of water. Error on the elevation at the end of the simulation with different absorption functions. The width of the absorbing layer δ is $5\,\Delta x$ (left) or $13\,\Delta x$ (right). The color bar has been fixed from -0.01 m to 0.01 m, as in Figure 5.4. *Adapted from [123] (© 2009 Springer-Verlag).*

Table 5.3: Simulation of the collapse of a mound of water. Relative error ξ_r (5.9) for different absorption functions and for two values of the width of the absorbing layer ($5\,\Delta x$ and $13\,\Delta x$) using the linear shallow water equations.

	$\delta = 5\,\Delta x$	$\delta = 13\,\Delta x$
Optimum constant $\bar{\sigma}$	$1.82\ 10^{-1}$	$1.11\ 10^{-1}$
Parabola σ_2 with $\alpha = 0.9/\Delta t$	$1.21\ 10^{-1}$	$9.38\ 10^{-2}$
Optimum parabola σ_2	$1.15\ 10^{-1}$	$7.06\ 10^{-2}$
Hyperbola σ_h	$1.52\ 10^{-1}$	$8.75\ 10^{-2}$
Shifted hyperbola σ_sh	$1.14\ 10^{-1}$	$5.91\ 10^{-2}$

small differences can be seen between the results obtained with the three spatially varying absorption functions. Slightly larger differences between the different results appear when a thicker absorbing layer is used.

The relative error values listed in Table 5.3 confirm the visual impression from Figure 5.6. The results obtained with the shifted hyperbola are slightly better than those produced with the optimum parabolic function, especially with the $13\,\Delta x$ absorbing layer, where the relative error associated with the shifted hyperbola is smaller than with the optimum parabolic function by about 20%. Increasing the width of the absorbing layer allows for a more gradual damping of outgoing waves obtained using smaller absorption functions.

For the $13\,\Delta x$ absorbing layer, the optimum value of the parameter α of the parabolic function is about $1.5\ 10^{-3}\ \mathrm{s}^{-1}$ while it reaches $3\ 10^{-3}\ \mathrm{s}^{-1}$ for the $5\,\Delta x$ absorbing layer. Relating α to the time step Δt and not to the width of the absorbing layer is therefore not the best solution. The value of $\alpha = 0.9/\Delta t = 6\ 10^{-3}\ \mathrm{s}^{-1}$ estimated using the formula of Lavelle and Thacker [108] is about twice the optimum value for the $5\,\Delta x$ absorbing layer. This affects however little the relative error in this numerical experiment. With the thicker $13\,\Delta x$ absorbing layer, there is however a 25 % difference between the relative errors obtained with the two parabolic functions.

5.4 Cases dominated by advection

The test case considered above is dominated by the propagation of Poincaré waves. While the shifted hyperbolic function σ_sh gives good results in this case, one might wonder if the same function can be used in advection-dominated problems. Therefore, we address the problem of a moving eddy, *i.e.* the advection on an eddy by a mean flow and its own velocity field. This test case is also used, in slightly different forms, by Navon et al. [127], McDonald [115] and Lavelle and Thacker [108].

This section focuses on the comparison of absorption functions for the PGS. It is well-known that the PML formulations are unstable for advection-dominated

cases if they are built without a specific strategy (see *e.g.* [3, 51, 90–92, 126, 156]). Preliminary simulations (not presented here) show that large spurious oscillations are generated inside the PML with our formulations, which do not account for such a strategy. A comparison with these layers is thus not pertinent.

5.4.1 Description of the benchmark

The nonlinear shallow water equations (5.1) are used to describe the movement of the eddy that is initially in geostrophic balance with the same Gaussian-shaped mound of water as in the previous test case. Except the depth at rest h that becomes 500 m, the model parameters and the geometry of the previous benchmark are reused. To move the eddy, a uniform current of 5 m/s is added in both horizontal directions in the initial velocity fields and is supported by a geostrophic tilt of the surface elevation (see [108]). The initial fields are then

$$\eta(\boldsymbol{x},0) = \eta_0 e^{-\|\boldsymbol{x}\|^2/R^2} + \frac{f}{g}\left(xU_y - yU_x\right), \tag{5.11}$$

$$\boldsymbol{u}(\boldsymbol{x},0) = \frac{2g}{fR^2}\eta_0 e^{-\|\boldsymbol{x}\|^2/R^2}\begin{pmatrix} y \\ -x \end{pmatrix} + \begin{pmatrix} U_x \\ U_y \end{pmatrix}, \tag{5.12}$$

where U_x and U_y correspond to the velocities in horizontal directions. In the PGS, the fields are relaxed towards a stationary external solution using nudging terms. This solution includes only the uniform current and the associated tilt of the surface elevation corresponding to the last term in equations (5.11) and (5.12). As time goes by, the eddy is swept out of the open domain by the advecting flow.

The discrete scheme is built considering the conservative form of the nonlinear equations, that is suitable for a numerical discretization using the finite volume approach. As claimed in section 5.2.2, one of the advantages of relaxation methods is that the full nonlinear dynamics can be easily taken into account. In particular, absorption terms can be easily included in the nonlinear shallow water equations. When written in conservative form, the resulting equations for a nonlinear PGS are

$$\begin{cases} \dfrac{\partial H}{\partial t} + \dfrac{\partial (Hu_x)}{\partial x} + \dfrac{\partial (Hu_y)}{\partial y} = -(\sigma_x + \sigma_y)(H - H^{\text{ext}}), \\ \dfrac{\partial (Hu_x)}{\partial t} + \dfrac{\partial (gH^2/2)}{\partial x} + \dfrac{\partial (Hu_x^2)}{\partial x} + \dfrac{\partial (Hu_x u_y)}{\partial y} - fHu_y = -\sigma_x (Hu_x - H^{\text{ext}} u_x^{\text{ext}}), \\ \dfrac{\partial (Hu_y)}{\partial t} + \dfrac{\partial (gH^2/2)}{\partial y} + \dfrac{\partial (Hu_x u_y)}{\partial x} + \dfrac{\partial (Hu_y^2)}{\partial y} + fHu_x = -\sigma_y (Hu_y - H^{\text{ext}} u_y^{\text{ext}}). \end{cases} \tag{5.13}$$

These equations include the advection term for momentum and account for the variability of the water depth. The FBTCS scheme and the Arakawa C-grid are again used for the numerical computation.

The time evolution of the solution is described and illustrated in [108]. During the first simulation hours, the amplitude of the Gaussian eddy decreases through

the generation of adjustment waves, since the initial fields do not achieve a non-linear balance. Afterwards, the eddy is advected towards the open boundary while the height of the rotating mound undergoes a small decrease because of numerical diffusion.

5.4.2 Comparison of absorption functions

The reflection of the outgoing signal on the open boundary can be quantified by computing the relative error (5.9). This is done here for the different absorption functions considered above and for two thicknesses of the absorbing layers (Figure 5.7). The constant function $\bar{\sigma}$ and the multiplicative parameter α of both parabola σ_2 and shifted hyperbola σ_h are tuned to minimize the area under the curves of Figure 5.7. In all the simulations, the error increases during the first hours due to the partial reflection of the eddy on the absorbing layer. After some time, the errors are simply transported by the uniform current and leave the domain, so that the global error decreases.

Increasing the thickness of the absorbing layer from $5\Delta x$ to $13\Delta x$ leads to a significant reduction of the error. In the best cases, the error decreases by a factor of 5. Note that the maximum error in Figure 5.7 appears later when the layer length is increased because the eddy covers then a larger distance in the absorbing layer before being reflected. For both thicknesses of the absorbing layer, the shifted hyperbolic function (5.8) with $\alpha = \sqrt{gh}$ performs significantly better than the parabolic function used by Lavelle and Thacker [108]. The error is however much larger than with numerically optimized absorption functions (uniform, parabolic and shifted hyperbolic).

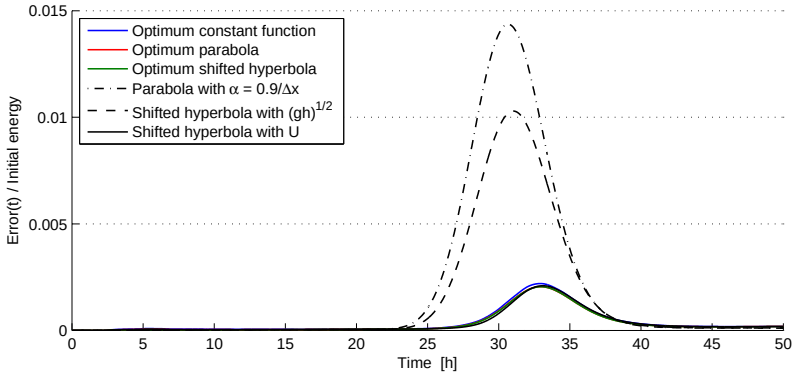
The reason for the bad result obtained with the shifted hyperbolic function is linked to the physics of the problem. With the shifted hyperbolic function, it is implicitly assumed that the signal propagates as surface gravity waves whereas advection is dominant. The speed α is therefore incorrectly estimated and this has obviously an adverse effect on the performance of the absorbing layer. The problem can be easily addressed however by using the normal velocity U in the shifted hyperbolic function instead of $\sqrt{gH}$. As shown in Figure 5.7, the corresponding function leads to optimum results that are comparable with those obtained with numerically tuned functions.

The use of the normal velocity in (5.8) can be justified on theoretical grounds using the same approach as in section 3.3.3.3. Indeed, if advection dominates, the dynamics in the absorbing layer can be approached by

$$\frac{\partial \phi}{\partial t} + U\frac{\partial \phi}{\partial s} = -\sigma(s)\,\phi$$

where $U > 0$ is the advection velocity and $\sigma(s)$ is the absorption function, which vanishes in the domain interior. For a constant absorption function, ϕ undergoes

(a) $\delta = 5\,\mathrm{km}$



(b) $\delta = 13\,\mathrm{km}$

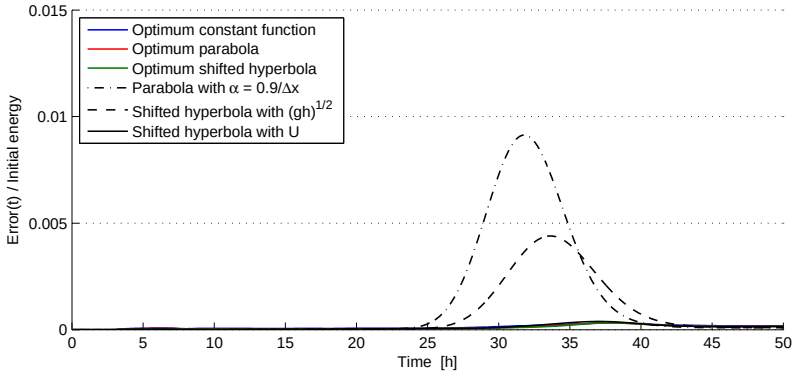


Figure 5.7: Simulation of the moving eddy. Evolution of the error in the domain for different absorption functions and two thicknesses of the absorbing layer. The continuous curves are superimposed. *Reproduced from [123] (© 2009 Springer-Verlag).*

an exponential damping while advected across the absorbing layer. Such an exponential damping can be made very efficient by using large values of σ but is then difficult to describe using a numerical model with limited spatial resolution. However, a linear decrease of the amplitude of outgoing waves can be achieved in the absorbing layer with the hyperbolic function

$$\sigma_{\mathrm{h}}(s) = \frac{U}{\delta - s}.$$

Removing any discontinuity of the function at the interface with the model interior leads then to the modified shifted hyperbolic function

$$\sigma_{\mathrm{sh}}(s) = \sigma_{\mathrm{h}}(s) - \sigma_{\mathrm{h}}(0) = \frac{U}{\delta}\frac{s}{\delta - s},$$

where $s = 0$ at the interface between the domain and the PGS. The results obtained with this modified function are clearly comparable to (and even slightly better than) those obtained with the numerically optimized functions.

5.5 External forcing through an absorbing layer

Forcing an external solution at an open boundary is a critical issue for regional oceanographic models. Tide waves and large scale currents are indeed taken into account through such forcing. Among the absorbing layers presented in section 5.2.2, only the FRS and the PGS permit the prescription of an external solution thanks to the nudging terms. However, they require the knowledge of this solution over the entire layer, which is a problem if it is only poorly known.

For linear cases, we propose to adapt the strategy developed in section 4.3 in the context of scattering problems: the external signal can be forced through an interface condition in a particular formulation of the problem. This formulation can be used together with a FRS or a PGS (instead of forcing through nudging terms), or even a PML. After an illustration of the formulation in a linear case, we discuss the limitation of this approach for nonlinear cases, as well as the impact of poorly-known data.

5.5.1 A strategy for linear cases with complete data

At an open boundary, the fields η and $\boldsymbol{u}$ can be split into two parts: the external fields η^{ext} and $\boldsymbol{u}^{\text{ext}}$ (assumed known), and the residual fields $\tilde{\eta}$ and $\tilde{\boldsymbol{u}}$ (unknown), with

$$\eta = \eta^{\text{ext}} + \tilde{\eta}, \tag{5.14}$$

$$\boldsymbol{u} = \boldsymbol{u}^{\text{ext}} + \tilde{\boldsymbol{u}}. \tag{5.15}$$

The fields $\tilde{\eta}$ and $\tilde{\boldsymbol{u}}$ contain the signals that are not taken into account in the external solution, *i.e.* perturbations and waves of all kinds generated inside the domain. The goal of the absorbing layer is this to allow these signals to leave the domain.

Similarly to the scattered/total-field formulation introduced in section 4.3, a strategy consists in directly computing $\tilde{\eta}$ and $\tilde{\boldsymbol{u}}$ in the layer (instead of η and $\boldsymbol{u}$), and preserving η and $\boldsymbol{u}$ as unknowns in the domain. Since $\tilde{\eta}$ and $\tilde{\boldsymbol{u}}$ must be only damped inside the layer, the PML systems (5.6) and (5.7) can be directly used (η and $\boldsymbol{u}$ must be simply replaced by $\tilde{\eta}$ and $\tilde{\boldsymbol{u}}$ inside the equations). For the FRS and PGS systems, the terms corresponding to external forcing are removed, *i.e.* the FRS system (5.4) becomes

$$\left\{\begin{array}{l} \dfrac{\partial \tilde{\eta}}{\partial t} + h(\nabla \cdot \tilde{\boldsymbol{u}}) = -\sigma \tilde{\eta}, \\ \dfrac{\partial \tilde{\boldsymbol{u}}}{\partial t} + g\nabla\tilde{\eta} + f\,\boldsymbol{e}_z \times \tilde{\boldsymbol{u}} = -\sigma\,\tilde{\boldsymbol{u}}, \end{array}\right. \tag{5.16}$$

and the PGS equations (5.5) become

$$\begin{cases} \dfrac{\partial \tilde{\eta}}{\partial t} + h(\nabla \cdot \tilde{\boldsymbol{u}}) = -(\sigma_x + \sigma_y)\,\tilde{\eta}, \\ \dfrac{\partial \tilde{\boldsymbol{u}}}{\partial t} + g\nabla\tilde{\eta} + f\,\boldsymbol{e}_z \times \tilde{\boldsymbol{u}} = -\operatorname{diag}\{\sigma_x, \sigma_y\}\,\tilde{\boldsymbol{u}}. \end{cases} \tag{5.17}$$

A particular condition must be prescribed at the interface between the domain and the layer, in order to connect the different fields at both sides of the interface. Considering the relations (5.14)-(5.15), one immediately obtains the interface condition

$$\begin{cases} \eta|_{\text{dom}} = \eta^{\text{ext}} + \tilde{\eta}|_{\text{lay}}, \\ \boldsymbol{n}\cdot\boldsymbol{u}|_{\text{dom}} = \boldsymbol{n}\cdot\boldsymbol{u}^{\text{ext}} + \boldsymbol{n}\cdot\tilde{\boldsymbol{u}}|_{\text{lay}}, \end{cases} \tag{5.18}$$

where $\eta|_{\text{dom}}$ and $\boldsymbol{u}|_{\text{dom}}$ correspond to the numerical solution computed in the model domain with the original equations (5.2), while $\tilde{\eta}|_{\text{lay}}$ and $\tilde{\eta}|_{\text{lay}}$ are computed in the layer with the layer equations (5.16) or (5.17).

As announced before, with this formulation, the external forcing is taken into account through the interface condition (5.18). Therefore, only the knowledge of both elevation η^{inc} and normal velocity $\boldsymbol{n}\cdot\boldsymbol{u}^{\text{inc}}$ at the interface are required.

Benchmark

The above strategy is tested with a benchmark involving an external forcing. A circular island with radius 1 km is placed at the center of the rectangular domain $[-L_x, L_x] \times [-L_y, L_y]$, with $L_x = 2.5$ km and $L_y = 5$ km. The northern and southern borders are coasts, while the others are supposed to be open. A Kelvin wave moving eastwards enters in the domain by the western border. If the island is removed, the time evolution of the fields is given by

$$\begin{cases} \eta^{\text{ext}}(t,\boldsymbol{x}) = e^{-(x-x_0-ct)^2/R^2}\, e^{-(f/c)(y-L_y)}, \\ \boldsymbol{u}^{\text{ext}}(t,\boldsymbol{x}) = e^{-(x-x_0-ct)^2/R^2}\, e^{-(f/c)(y-L_y)} \sqrt{\dfrac{g}{h}}\,\boldsymbol{e}_x, \end{cases} \tag{5.19}$$

where the parameters used are $x_0 = -30$km, $R = 12$km, $c = \sqrt{gh}$, $f = 1.02810^{-4}\text{s}^{-1}$, $g = 9.85\,\text{m/s}^2$ and $h = 100\,\text{m}$.

The external solution (5.19) is forced at the western boundary through open-sea boundary treatment. Because this solution does not account for the presence of the island, the fields are modified — the difference is shown in Figure 5.8(c). Using classical terminology of scattering problems, the external solution contains the *incident fields* and the residual fields are the *scattered fields*. Their sum gives the *total fields*.

Simulations are performed with the linear shallow water equations (5.2) and the DG scheme. The characteristic length of mesh cells is $\ell = 200\,\text{m}$. The fourth-order Runge-Kutta time-stepping scheme is used with $\Delta t = 1$s. The computational domain is extended eastward.

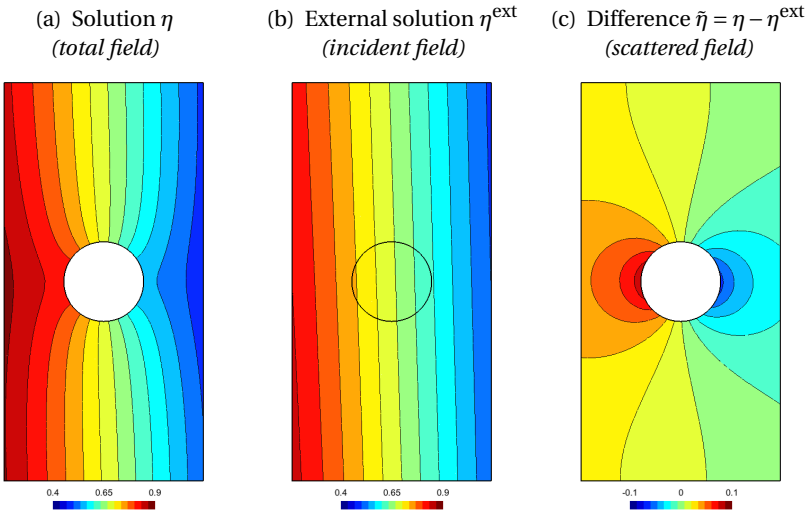


Figure 5.8: Simulation of the scattering of a Kelvin wave by a circular island. Snapshots of the reference elevation in the domain at $t = 12\,\text{min}$ with (a) and without (b) the island. The difference, corresponding to the scattered field, is shown in (c).

Discussion

The snapshots of the field $\tilde{\eta}$ obtained with Flather's OBC and the absorbing layers at $t = 12\,\text{h}$ are shown in Figure 5.9. Let us recall that the fields $\tilde{\eta}$ and $\tilde{\boldsymbol{u}}$ are directly computed (and damped) in the layers, while the original fields are considered in the domain. The field $\tilde{\eta}$ is then obtained by subtracting the external solution (5.19) from the solution effectively computed.

The closeness between reference and layer solutions demonstrates the effectiveness of the interface condition (5.18). Moreover, its use does not change the conclusions of the previous benchmark about boundary treatments. The solution provided by the Flather's OBC in Figure 5.9 is slighly different from the reference solution, especially near the open boundary where the influence of the boundary condition is larger. By contrast, the fields obtained with both complete and approximate PMLs cannot be distinguished from the reference field. Finally, results provided by other layers are coherent with what was obtained before: the FRS is not so good than the PGS, but slighly better than the OBC.

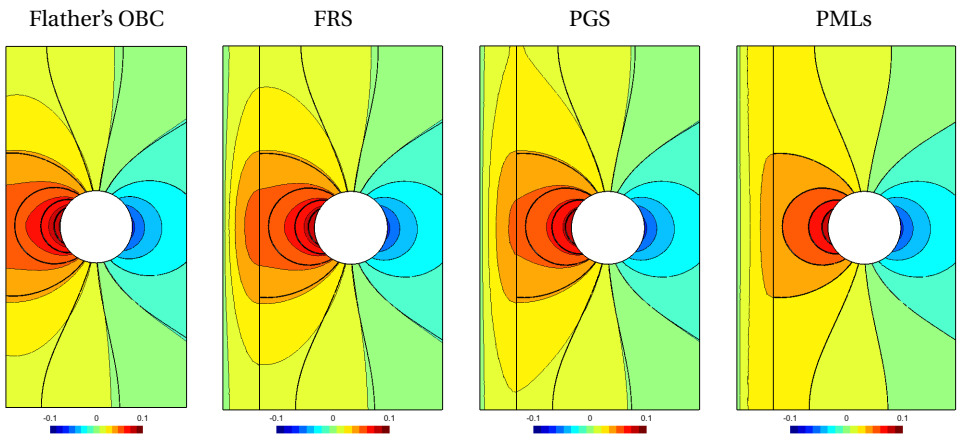


Figure 5.9: Simulation of the scattering of a Kelvin wave by a circular island. Scattered field provided by the OBC and the different layers with the thickness $\delta = 1$ km. Results of both complete and approximate PMLs are identical. In the domain, thick lines correspond to the reference solution shown in Figure 5.8(c).

5.5.2 Extension for nonlinear cases and incomplete external data

The extension of our strategy for more realistic nonlinear cases is not staightforward. Because of nonlinearities, the external (incident) and residual (scattered) fields of the decomposition (5.14)–(5.15) are coupled, and the equations governing residual fields must contain coupling terms that involves the external fields. It is not clear how to account for such coupling terms in the procedure to design an absorbing layer. In order to preserve a formulation with an external forcing only through the interface condition, a first strategy consists in removing these terms from the equations. However, this approximation must be validated, and further investigations are needed.

With the formulations proposed above, the knowledge of both elevation and normal velocity at the open boundary is a requirement. If one of them is unknown, the interface condition (5.18) cannot be normally used. Since the aim of an open-sea boundary treatment is not to rebuild external data, we suggest to address this issue using dedicated existing approaches, such as data assimilation (see *e.g.* [11, 44]) or embedded models (see *e.g.* [7, 28]).

5.6 Large domain *versus* elaborate boundary treatment

Even if more and more accurate open-sea boundary treatments are developed, increasing the size of the domain sometimes remains the best strategy. In practical cases, a good compromise must be reached between the size of the domain and the accuracy (and the complexity) of the boundary treatment. This is illustrated

in this section with a classical benchmark that has been already used to compare OBCs by Røed and Cooper [150] and by Palma and Matano [135, 137].

In this benchmark, the open-sea boundary treatments must deal with waves generated by a real-world wind stress, and influenced by a variable bathymetry. Since this benchmark involves nonlinear dynamics, only Flather's OBC, the FRS and the PGS are considered. The PMLs are excluded for reasons exposed at the beginning of section 5.4.

Description of the benchmark

Eddies and waves are generated in a zonal channel by a cyclone, that moves from North to South, as shown in Figure 5.11. The channel is assumed to be infinitely long and 500 km in width. The bathymetry, represented in Figure 5.10(a), varies along the meridians according to

$$h(y) = 200\,\mathrm{m} + 800\,\mathrm{m}\left(1 + \tanh\left(\frac{y - 150\,\mathrm{km}}{50\,\mathrm{km}}\right)\right),$$

where $y = 0$ and $y = 500\,\mathrm{km}$ correspond to the southern and the northern coast of the channel, respectively. It is similar to the one used by Palma and Matano [135, 137]. Our mesh, shown in Figure 5.10(b), is refined in the shallow part of the domain.

The unknown fields, η and $\boldsymbol{u}$, are governed by the complete nonlinear shallow water equations (5.1). The moving storm induces a wind stress $\boldsymbol{\tau}_w = (\tau_x, \tau_y)$ whose the Cartesian components are

$$\tau_x = -\tau_0 \frac{y_c}{R_c} \exp\left\{\frac{1}{2} - \frac{x_c^2 + y_c^2}{2R_c^2}\right\},$$
$$\tau_y = \tau_0 \frac{x_c}{R_c} \exp\left\{\frac{1}{2} - \frac{x_c^2 + y_c^2}{2R_c^2}\right\},$$

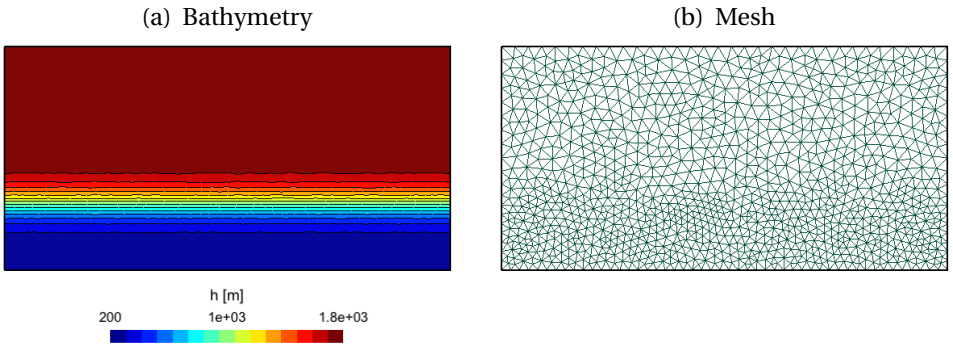


Figure 5.10: Simulation of a traveling storm. Bathymetry and mesh.

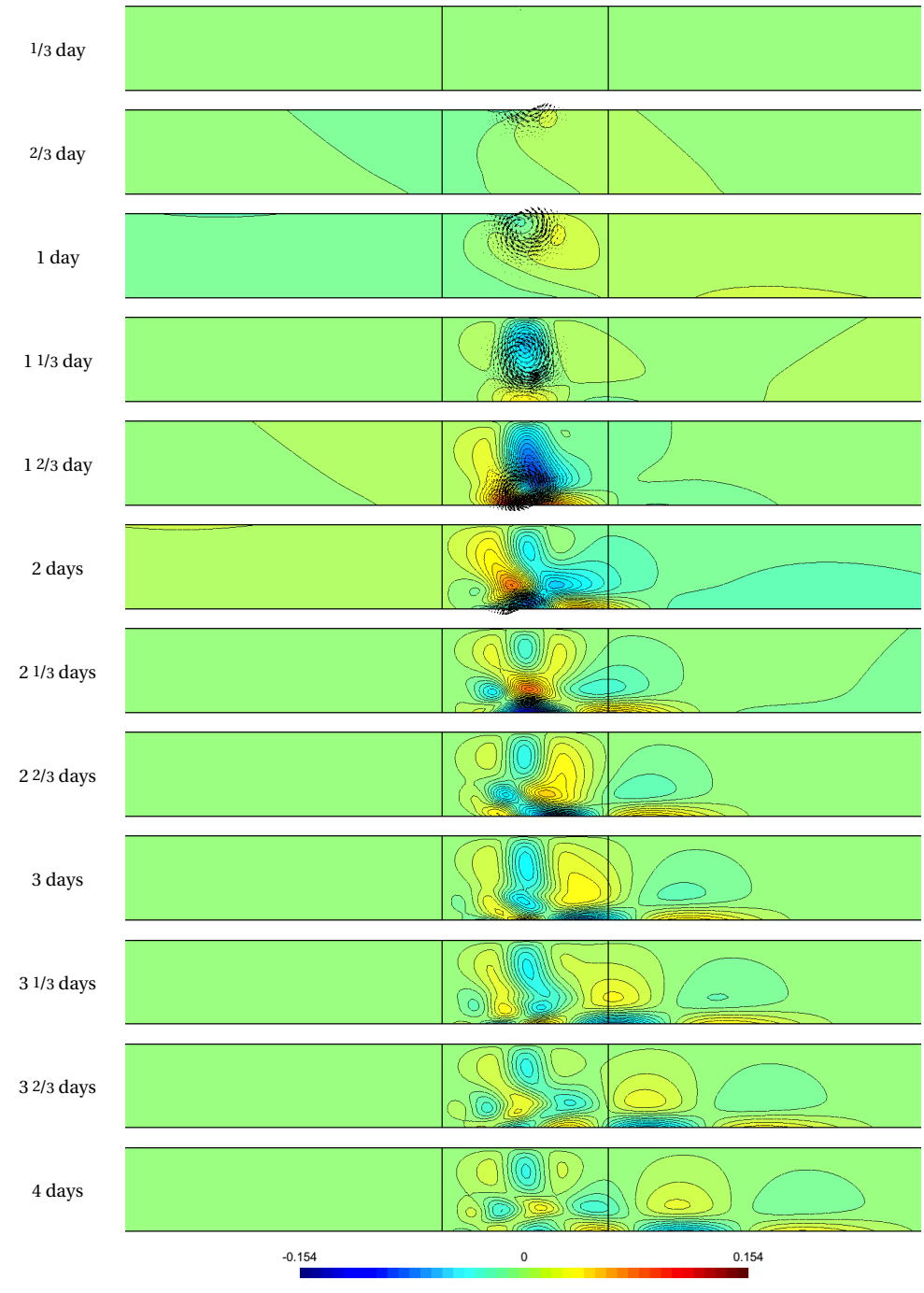


Figure 5.11: Simulation of a traveling storm. Snapshots of the elevation at different instants, obtained with a reference simulation. The reference domain is sufficiently long to remove the influence of its western and eastern borders on the domain of interest (delimited by vertical lines).

where (x_c, y_c) denote the position of the storm center, with

$$x_c = x - x_0,$$
$$y_c = y - y_0 + Ut,$$

(x_0, y_0) is its initial position, and U is its velocity [150]. By contrast with previous works, where the storm moved from the northwest to the southeast, the storm here moves only southwards. The problem parameters are $x_0 = 500\,\text{km}$, $y_0 = 917\,\text{km}$, $U = 500\,\text{km/day}$, $R_c = 100\,\text{km}$ and $\tau_0 = 3\,\text{Pa}$. In addition, we consider the horizontal eddy viscosity $\nu = 200\,\text{m}^2/\text{s}$ and a linear bottom stress $\boldsymbol{\tau}_b = \gamma H \rho_0 \boldsymbol{u}$, with $\gamma = 5\,10^{-5}\,\text{s}^{-1}$. We use the DG scheme described in [42] and a fourth-order Runge-Kutta time-stepping scheme with $\Delta t = 5\,\text{s}$.

During the first two days of the simulation, the storm travels over the channel and generates Kelvin waves and eddies (see Figure 5.11). The Kelvin waves propagate westwards along the northern coast, and eastwards along the southern one. After the storm has gone, topographic waves are generated by the eddies. These waves slowly move eastwards.

Results and discussion

Open-sea boundary treatments are tested by simulating only a part of the channel that is 1,000 km in length, exactly like in the previous works [135, 137, 150]. This area is delimited by the vertical lines in Figure 5.11. Both its western and eastern borders are open boundaries. They correspond to coordinates $x = 0$ and $x = 1{,}000\,\text{km}$, respectively. In this configuration, the direct wind forcing on the open boundaries is negligible.

Snapshots of the elevation obtained with Flather's OBC, the FRS and the PGS are shown in Figure 5.12 at two instants: 2 days (when the storm is just gone) and 4 days (after the generation of topographic waves). In all the cases, the elevations obtained with the OBC and the layers are very close to the reference one (Figure 5.12). By comparing isovalue curves, small differences can be distinguished near the open boundaries. The PGS provides a slightly better result than the FRS, which is itself slightly better than Flather's OBC. The ranking is therefore the same than the one of the previous benchmark. The results are coherent, even if the differences of effectiveness are here smaller.

The small differences between open-sea treatments can be explained by the specificity of this benchmark: the solution is clearly dominated by the eddies. A good modeling of the outgoing waves is less important than a good modeling of these eddies. This is confirmed by two additional results shown in Figure 5.13: the snapshots of the elevation for an extended domain (the extended parts have the same width than the layers), that is ended by a wall condition (*i.e.* $\boldsymbol{n} \cdot \boldsymbol{u} = 0$) or Flather's OBC.

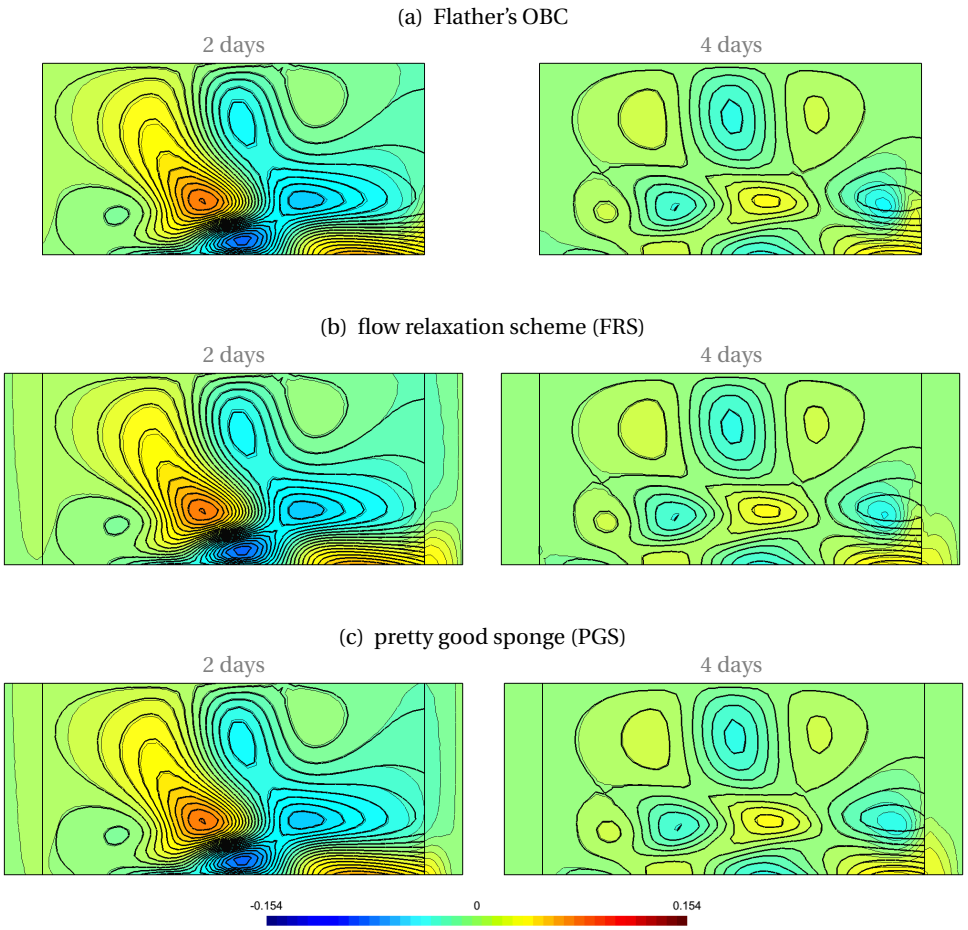


Figure 5.12: Simulation of a traveling storm. Snapshots of the elevation obtained with Flather's OBC (a), the FRS (b) and the PGS (c) at two instants: 2 and 4 days. In the domain of interest, thick lines correspond to the reference solution shown in Figure 5.11.

The solution obtained with the wall condition (Figure 5.13(a)) is clearly different from the reference one, but the shape of the eddies remains. This is surprising, since the wall condition is the worst condition that could be used at an open boundary. Therefore, even with this worst condition, the shape of the solution is coherent. This shows the importance of the storm and the eddies in this benchmark, in comparison with open boundaries.

When Flather's OBC terminates the extended domain, the solution in the domain of interest is slightly better than with the FRS and the PGS above (Figure 5.12). Even if, on the modeling standpoint, both FRS and PGS are better than Flather's OBC, adding nudging terms in a domain extension reduces the quality of the solution. Therefore, a larger domain, which permits a better representation of eddies,

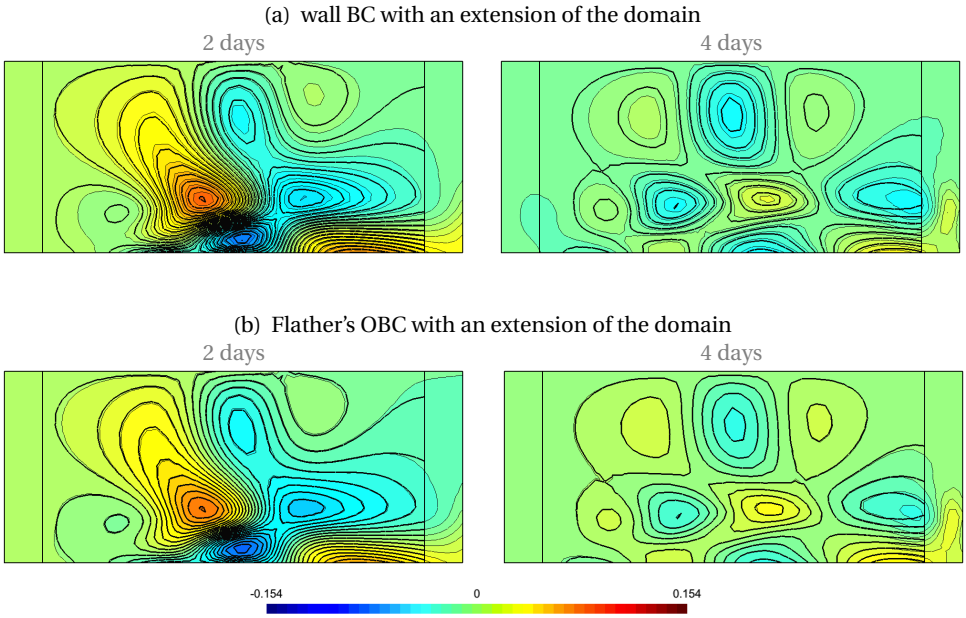


Figure 5.13: Simulation of a traveling storm. Snapshots of the elevation obtained using an extended domain, terminated by a wall (*i.e.* $\boldsymbol{n}\cdot\boldsymbol{u}=0$) (a) and Flather's OBC (b) at two instants: 2 and 4 days. In the domain of interest, thick lines correspond to the reference solution shown in Figure 5.11.

is here more suitable than the elaborate open-sea boundary treatments considered in this chapter.

5.7 Conclusion

Because of their easy implementation and their good properties, absorbing layers are very attractive for numerical ocean models. Introduced at the end of the 80's, the FRS exhibits a good efficiency in comparative studies [98, 131, 135, 136, 149]. The PMLs are promising, even if they are considered in only a few comparative works [108, 127]. Most recently, the so-called PGS has been proposed by Lavelle and Thacker [108], inspired by both FRS and PML formulations. In this chapter, we have compared the FRS, the PGS, classical Flather's boundary condition and two novel absorbing layers designed with the procedure of chapter 2: the PML (5.6) and the approximate PML (5.7).

We have observed that the complete PML (5.6) outperforms other boundary treatments in cases dominated by Poincaré waves. By order of decreasing effectiveness, one then has the approximate PML, the PGS, the FRS and the OBC. However, the considered PML formulations do not deal with advection-dominated

problem due to stability problems. In that case, Lavelle and Thacker [108] have shown the superiority of the PGS on the FRS. As a conclusion, the considered PML is accurate, but for a small range of applications. By contrast, the PGS is more robust, even if it is less accurate than the PML.

The choice of the absorption function σ is a common issue for the use of all the layers. We have extended the discussion of this function, already done in chapter 3 for pure wave propagation problems, in the marine context. To this end, we have considered the PGS in benchmarks dominated by Earth's rotation and advection. While the conclusion about performing functions does not change when Earth's rotation is considered, the results differ when strong advection is present. Indeed, the shifted hyperbolic function (5.8) remains good, but the parameter α must now be choosen close to the mean flow velocity. Therefore, good performances of the PGS can be achieved by adjusting the absorption function according to the physics of the problem.

We have discussed ways of forcing an incoming external signal at the open sea boundary. While an external solution could simply be prescribed throught nudging terms for both FRS and PGS, the issue is not clear for other layers. Inspired by the total/scattered-fields formulation introduced in section 4.3.1, we have proposed to use an interface condition to prescribe the external forcing. This method can be employed together with PMLs, as well as with the FRS and the PGS. External data are then only required at the interface, and not over the whole layer. Since this strategy applies only to linear cases, further developements are required to deal with nonlinear cases.

Finally, we have highlighted two limitations when simulating with open boundaries. First, the open-sea boundary treatment cannot improve the knowledge of external data. If data are incomplete, other strategies must be adressed. Second, the current OBCs and layers cannot simulate all dynamics, and increasing the size of the domain sometimes still remains the best option.

Conclusion

In order to deal with the numerical simulation of wave-like phenomena occuring in large or infinite regions, efficient strategies for building and optimizing absorbing layers have been presented. The main achievements and conclusions of this thesis, already detailed at the end of each chapter, are summarized hereafter. Perspectives for future research are then highlighted.

Main achievements and conclusions

Absorbing layers based on PML techniques exhibit interesting properties for simulating the truncation of unbounded domains. In some cases, PMLs provide an exact treatment of the truncation, while, in other cases, they constitute a good starting point to design alternative approximate layers. Therefore, the major part of this thesis was dedicated to their study.

We have proposed a complete procedure for building PMLs for convex domains with regular boundary, for both time-harmonic and time-dependent problems. This permits a greater flexibility when choosing the shape of the truncated domain. Since the obtained PML equations are written in Cartesian coordinates, they are easily implemented in existing codes. The efficiency of this approach has been shown on time-dependent numerical examples in two and three dimensions. Simulations have been performed with a DG scheme, whose overall numerical convergence $\mathcal{O}(\ell^{n+1})$ has been verified.

The issue of choosing PML parameters that optimize the layer effectiveness has been addressed. By means of novel analytical and numerical results, we have highlighted the roles of the absorption function σ, the layer thickness δ and the discretization density. Since taking a better σ improves the effectiveness of the layer without supplementary computational cost, looking for efficient functions is very appealing. We have performed a systematic comparison of different absorption functions with several classical numerical schemes (finite difference, finite volume and finite elements with both continuous and discontinuous elements). In most cases, the commonly-used polynomial functions and the shifted hyperbolic functions provide equivalent results when they are tuned. The great advantage of the latter is however that neither numerical optimization nor trial and error procedures are required, as its free parameter has a clear physical interpretation and

can be adjusted from the knowledge of the dynamics of the problem. This has been confirmed by numerical simulations in two and three dimensions.

While the PML equations do not *a priori* account for incoming signals generated outside the truncated domain, we have detailed and compared different problem formulations that deal with such forcing. In particular, the hybrid formulation with both scattered and total fields is very appealing. Indeed, the original model equations hold inside the domain without any change, and the PML equations inside the layer can also be used unmodified. The incoming signals are then prescribed by only means of a specific interface condition. This formulation has been tested in a two-dimensional case.

Modeling open-sea boundaries in regional oceanic models brings new challenges. Indeed, such models involve various dynamics (linear and nonlinear), and the external forcing is generally poorly known. In this context, we have compared and studied some boundary treatments considering additional physics. PML formulations are very efficient in linear cases dominated by wave propagation. However, taking into account additional physical dynamics requires further developments. By contrast, the PGS appeared to be more robust and easier to use for a larger range of problems, and is clearly better than the widely-used Flather condition. A comparison of absorption functions has emphasized that the shifted hyperbolic function remains effective, but its free parameter α must be adapted by taking into account each particular physical context. We have also tested the forcing of the external solution by an interface condition. Finally, some limitations of open-sea boundary treatments have been highlighted.

Future prospects

According to Givoli [73], artificial boundary treatments should be built having in mind the following goals: the *well-posedness* of the problem, the *stability* of the numerical scheme, the *accuracy* of the boundary treatment in both continuous and discontinuous levels, the *compatibility* with the numerical scheme, the *ease of implementation* in existing codes and the *computational efficiency*. In addition, the formulation should ideally be *general* and usable in a variety of situations.

In this thesis, we focused on some of the goals listed by Givoli: the accuracy, the compatibility with various numerical schemes, the ease of implementation and the computational efficiency. The other points and several other additional difficulties should be investigated further. In particular, the following aspects would be interesting to investigate:

- *Long-term stability and well-posedness.*

  While the PML formulations presented in this thesis have been successfully applied in numerical simulations, the well-posedness of the PML equations

and the stability of the scheme remain to be investigated for long-term simulations. It is well-known the equations of Bérenger's PML are only well-posed in a weak sense [1, 77]. As a consequence, small errors can linearly grow in the layer [16]. This is not a problem for short-term simulations, but such perturbations are more dangereous with long times. Since the PML equations of this thesis are hyperbolic only in a weak sense (see [77]), we expect the same problem. This should be assessed and overcomed before considering long-term simulations. Some solutions have been proposed in [2, 4, 12, 19, 141], but for conventionally-shaped domains.

- *Extension of PML formulations to other wave-like equations.*

  PML formulations that permit generally-shaped truncated domains are very appealing for various application contexts. Such formulations should be derived with other wave-like equations, such as equations of aeroacoustics, elastodynamics or 3-D electromagnetism. The procedure of section 2.4 could be easily tested with these equations. Further improvements will however be necessary to deal with evanescent waves, advection or anisotropic media, which bring some further difficulties (see section 2.3.4 and references herein).

- *Extensive comparison of boundary treatments.*

  During the last two decades, numerous absorbing layers (perfectly matched or not) and hierarchical boundary conditions have been proposed. Rational comparative studies between available boundary treatments will be helpful in the perspective of applications, where one treatment must be chosen. However, only partial or outdated studies are available. To the best of our knowledge, no comparisons have been proposed between the two above-mentioned families of treatments in the time domain.

- *Efficient absorbing layers with external forcing in oceanography.*

  Due to its effectiveness and its robustness, the PGS is very appealing. This layer should be tested in other cases of oceanography, considering increasingly realistic models (two-layer model, baroclinic model, ...). Since the external solution is generally poorly known, we suggest further investigations on the design of specific interface conditions for prescribing external forcing. Indeed, as highlighted in this work, this approach requires the lowest amount of information about the external solution.

APPENDIX A

Plane-wave analysis in the discrete FD context

When a differential problem is discretized, the properties of the solution change. In the context of wave-like problems, dispersion and dissipation of waves are altered by the discretization. The properties of PMLs (perfect match and perfect absorption) are also altered.

In this appendix, the dispersion and dissipation properties of waves are studied for the scalar wave system (1.2) discretized with the finite difference method. One dimensional cases are considered, with and without PML terms. For a constant absorption function σ, the discrete reflection coefficients associated to both infinite and finite layers are computed.

A.1 Discrete scalar wave system without PML terms

Let us consider the finite difference (FD) scheme for the one-dimensional scalar wave system,

$$\begin{cases} \dfrac{\mathrm{d}\tilde{p}_{i+1/2}}{\mathrm{d}t} + a\dfrac{\tilde{u}_{i+1} - \tilde{u}_i}{\Delta x} = 0, \\ \dfrac{\mathrm{d}\tilde{u}_i}{\mathrm{d}t} + b\dfrac{\tilde{p}_{i+1/2} - \tilde{p}_{i-1/2}}{\Delta x} = 0, \end{cases} \tag{A.1}$$

where $\tilde{p}_{i+1/2}(t)$ and $\tilde{u}_i(t)$ are the semi-discrete fields, and Δx is the spatial step.

The elementary harmonic plane-wave solution of this system reads

$$\tilde{p}_{i+1/2}(t) = Pe^{\imath(k(i+1/2)\Delta x - \omega t)},$$
$$\tilde{u}_i(t) = Ue^{\imath(ki\Delta x - \omega t)},$$

where P and U are the amplitudes of fields, ω is the angular frequency (real and positive) and k is the wave number (real). Injecting this elementary solution in the system (A.1) gives

$$\begin{aligned} &\begin{cases} -\imath\omega P + a\dfrac{e^{\imath k\Delta x/2} - e^{-\imath k\Delta x/2}}{\Delta x} U = 0, \\ -\imath\omega U + b\dfrac{e^{\imath k\Delta x/2} - e^{-\imath k\Delta x/2}}{\Delta x} P = 0, \end{cases} \\ \iff &\begin{cases} -\dfrac{\omega\Delta x}{2} P + a\sin\left(\dfrac{k\Delta x}{2}\right) U = 0, \\ -\dfrac{\omega\Delta x}{2} U + b\sin\left(\dfrac{k\Delta x}{2}\right) P = 0. \end{cases} \end{aligned} \tag{A.2}$$

The last system is true if

$$\begin{aligned} \omega &= \pm c\frac{2}{\Delta x}\sin\left(\frac{k\Delta x}{2}\right), \\ P &= \pm\sqrt{\frac{a}{b}} U, \end{aligned} \tag{A.3}$$

where $c = \sqrt{ab}$ is the propagation velocity of the physical medium. The plus sign corresponds to $k > 0$, while the minus sign is for $k < 0$. One obtains the phase velocity and the group velocity of the discrete medium,

$$\begin{aligned} c_p &\overset{\text{def.}}{=} \frac{\omega}{|k|} = c\frac{2}{k\Delta x}\sin\left(\frac{k\Delta x}{2}\right), \\ c_g &\overset{\text{def.}}{=} \frac{\partial\omega}{\partial k} = \pm c\cos\left(\frac{k\Delta x}{2}\right). \end{aligned} \tag{A.4}$$

Since the phase velocity c_p depends on k, the scheme is dispersive. Since the elementary plane-wave solution is not damped, the scheme is non-dissipative.

A.2 Discrete scalar wave system with PML terms

A.2.1 Constant function $\bar{\sigma}$ case

Let us consider the FD scheme for the one-dimensional scalar wave system with PML terms,

$$\begin{cases} \dfrac{\mathrm{d}\tilde{p}_{i+1/2}}{\mathrm{d}t} + a\dfrac{\tilde{u}_{i+1} - \tilde{u}_i}{\Delta x} = -\bar{\sigma}\,\tilde{p}_{i+1/2}, \\ \dfrac{\mathrm{d}\tilde{u}_i}{\mathrm{d}t} + b\dfrac{\tilde{p}_{i+1/2} - \tilde{p}_{i-1/2}}{\Delta x} = -\bar{\sigma}\,\tilde{u}_i. \end{cases} \tag{A.5}$$

where $\tilde{p}_{i+1/2}(t)$ and $\tilde{u}_i(t)$ are the semi-discrete fields and Δx is the spatial step.

Discrete plane-wave solution

The elementary harmonic damped plane-wave solution of the system (A.5) reads

$$\tilde{p}_{i+1/2}(t) = Pe^{\imath(\beta(i+1/2)\Delta x-\omega t)},$$
$$\tilde{u}_i(t) = Ue^{\imath(\beta i\Delta x-\omega t)},$$

where β is a complexe parameter. The real part of β corresponds to the wave number, while its imaginary part is the damping parameter. Injecting this elementary solution in the system (A.5) gives

$$\begin{cases} -\imath\omega P + a\dfrac{e^{\imath\beta\Delta x/2} - e^{-\imath\beta\Delta x/2}}{\Delta x}U = -\bar{\sigma}P, \\ -\imath\omega U + b\dfrac{e^{\imath\beta\Delta x/2} - e^{-\imath\beta\Delta x/2}}{\Delta x}P = -\bar{\sigma}U. \end{cases}$$
$$\iff \begin{cases} -(\omega+\imath\bar{\sigma})\dfrac{\Delta x}{2}P + a\sin\left(\dfrac{\beta\Delta x}{2}\right)U = 0, \\ -(\omega+\imath\bar{\sigma})\dfrac{\Delta x}{2}U + b\sin\left(\dfrac{\beta\Delta x}{2}\right)P = 0. \end{cases}$$

The last system is true if

$$\omega + \imath\bar{\sigma} = \pm c\frac{2}{\Delta x}\sin\left(\frac{\beta\Delta x}{2}\right),$$
$$P = \pm\sqrt{\frac{a}{b}}U. \tag{A.6}$$

From the first relation, one has an expression for the complex parameter β,

$$\beta = \pm\frac{2}{\Delta x}\arcsin\left(\frac{\Delta x}{2}\frac{\omega+\imath\bar{\sigma}}{c}\right).$$

Reflection coefficient for an infinite layer

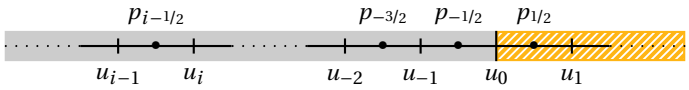
Let us consider a one-dimensional problem with a domain $\Omega = \mathbb{R}^-$ extended on the right side by an infinite PML medium $\Sigma = \mathbb{R}^+$, as shown in Fig. A.1 (a). The semi-discrete fields $\tilde{p}_{i+1/2}(t)$ and $\tilde{u}_i(t)$ are governed by the system (A.1) in domain and the system (A.5) in the layer. At the interface, the field $\tilde{u}_0(t)$ is governed by

$$\frac{\mathrm{d}\tilde{u}_0}{\mathrm{d}t} + b\frac{\tilde{p}_{1/2} - \tilde{p}_{-1/2}}{\Delta x} = -\bar{\sigma}u_0. \tag{A.7}$$

The solution of this problem can be written as the superposition of incident, reflected and transmitted waves. The elementary time-harmonic solution then reads

$$\tilde{p}_{i+1/2}(t) = \begin{cases} P^i e^{\imath(k(i+1/2)\Delta x-\omega t)} + P^r e^{\imath(-k(i+1/2)\Delta x-\omega t)}, & \text{for } i = \ldots,-2,-1 \\ P^t e^{\imath(\beta(i+1/2)\Delta x-\omega t)}, & \text{for } i = 0,1,\ldots \end{cases}$$
$$\tilde{u}_i(t) = \begin{cases} U^i e^{\imath(ki\Delta x-\omega t)} + U^r e^{\imath(-ki\Delta x-\omega t)}, & \text{for } i = \ldots,-1,0 \\ U^t e^{\imath(\beta i\Delta x-\omega t)}, & \text{for } i = 0,1,\ldots \end{cases}$$

(a) Infinite layer



(b) Finite layer

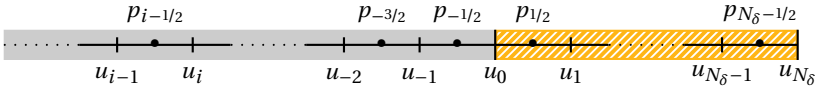


Figure A.1: Discrete grid for the one-dimensional problem with (a) an infinite layer and (b) a finite layer. The orange hatched zone represents the layer.

with $k>0$. Injecting this solution in the governing equation of $\tilde{u}_0(t)$ (A.7) gives

$$(\bar{\sigma}-\imath\omega)(U^i+U^r)+\frac{b}{\Delta x}\left(P^t e^{\imath\beta\Delta x/2}-P^i e^{-\imath k\Delta x/2}-P^r e^{\imath k\Delta x/2}\right)=0.$$

Using the continuity condition of the field $\tilde{u}_i$ at the interface (*i.e.* $U^t=U^i+U^r$) and the impedance relations (A.3) and (A.6), it becomes

$$(\bar{\sigma}-\imath\omega)(P^i-P^r)+\frac{c}{\Delta x}\left((P^i-P^r)e^{\imath\beta\Delta x/2}-P^i e^{-\imath k\Delta x/2}-P^r e^{\imath k\Delta x/2}\right)=0.$$

Finally, one has the discrete reflection coefficient

$$\boxed{r^\star_{\text{interf}}=\left|\frac{P^r}{P^i}\right|=\left|\frac{\left(\frac{\bar{\sigma}\Delta x}{c}-\imath\frac{\omega\Delta x}{c}\right)+e^{\imath\beta\Delta x/2}-e^{-\imath k\Delta x/2}}{\left(\frac{\bar{\sigma}\Delta x}{c}-\imath\frac{\omega\Delta x}{c}\right)+e^{\imath\beta\Delta x/2}+e^{\imath k\Delta x/2}}\right|.}$$

Reflection coefficient for a finite layer

Let us consider now a modified version of the problem, with a finite PML $\Sigma=[0,\delta]$, as shown in Fig. A.1 (b). The solution can be written as the superposition of incident waves, their reflections in the domain Ω, transmitted waves in the PML and their reflections by the outer boundary of the layer ($x=\delta$). Using time-harmonic plane waves, the elementary solution reads

$$\tilde{p}_{i+1/2}(t)=\begin{cases}P^i e^{\imath(k(i+1/2)\Delta x-\omega t)}+P^r e^{\imath(-k(i+1/2)\Delta x-\omega t)}, & \text{for } i=\ldots,-2,-1\\ P^t e^{\imath(\beta(i+1/2)\Delta x-\omega t)}+P^b e^{\imath(-\beta(i+1/2)\Delta x-\omega t)}, & \text{for } i=0,1,\ldots\end{cases}$$

$$\tilde{u}_i(t)=\begin{cases}U^i e^{\imath(ki\Delta x-\omega t)}+U^r e^{\imath(-ki\Delta x-\omega t)}, & \text{for } i=\ldots,-1,0\\ U^t e^{\imath(\beta i\Delta x-\omega t)}+U^b e^{\imath(-\beta i\Delta x-\omega t)}, & \text{for } i=0,1,\ldots\end{cases}$$

with $k>0$ and $\Re e(\beta)>0$.

Relations between amplitudes are obtained with the continuity condition at the interface domain/layer, the boundary condition at the outer boundary of the

layer, *i.e.* $\tilde{u}_{N_\delta} = 0$ and the governing equation of of $u_0(t)$. The continuity condition and the boundary condition give

$$\begin{aligned} U^i + U^r &= U^t + U^b, \\ U^t e^{\imath\beta\delta} + U^b e^{-\imath\beta\delta} &= 0, \end{aligned}$$

and then

$$U^t = \frac{U^i + U^r}{1 - e^{2\imath\beta\delta}}, \tag{A.8}$$

$$U^b = \frac{U^i + U^r}{1 - e^{-2\imath\beta\delta}}. \tag{A.9}$$

The governing equation of $u_0(t)$ (A.7) gives

$$(\bar{\sigma} - \imath\omega)(U^i + U^r) + \frac{b}{\Delta x}\left(P^t e^{\imath\beta\Delta x/2} + P^b e^{-\imath\beta\Delta x/2} - P^i e^{-\imath k\Delta x/2} - P^r e^{\imath k\Delta x/2}\right) = 0.$$

Using the amplitude relations, this relation becomes

$$(\bar{\sigma} - \imath\omega)(U^i + U^r) + \frac{c}{\Delta x}\left(U^t e^{\imath\beta\Delta x/2} - U^b e^{-\imath\beta\Delta x/2} - U^i e^{-\imath k\Delta x/2} + U^r e^{\imath k\Delta x/2}\right) = 0.$$

Using relations (A.8) and (A.9), one has

$$\left(\left(\frac{\bar{\sigma}\Delta x}{c} - \imath\frac{\omega\Delta x}{c}\right) + \frac{e^{\imath\beta\Delta x/2}}{1 - e^{2\imath\beta\delta}} - \frac{e^{-\imath\beta\Delta x/2}}{1 - e^{-2\imath\beta\delta}}\right)(U^i + U^r) - U^i e^{-\imath k\Delta x/2} + U^r e^{\imath k\Delta x/2} = 0.$$

Finally, the discrete reflection coefficient reads

$$\boxed{r^\star_{\text{pml}} = \left|\frac{P^r}{P^i}\right| = \left|\frac{\left(\frac{\bar{\sigma}\Delta x}{c} - \imath\frac{\omega\Delta x}{c}\right) + \imath\frac{\cos(\beta\delta - \beta\Delta x/2)}{\sin(\beta\delta)} - e^{-\imath k\Delta x/2}}{\left(\frac{\bar{\sigma}\Delta x}{c} - \imath\frac{\omega\Delta x}{c}\right) + \imath\frac{\cos(\beta\delta - \beta\Delta x/2)}{\sin(\beta\delta)} + e^{\imath k\Delta x/2}}\right|.}$$

A.2.2 Hyperbolic function $\sigma_{\text{h}}(x)$ case

Let us consider the FD scheme for the one-dimensional scalar wave system with PML terms,

$$\begin{cases} \dfrac{\mathrm{d}\tilde{p}_{i+1/2}}{\mathrm{d}t} + a\dfrac{\tilde{u}_{i+1} - \tilde{u}_i}{\Delta x} = -\sigma_{i+1/2}\,\tilde{p}_{i+1/2}, \\ \dfrac{\mathrm{d}\tilde{u}_i}{\mathrm{d}t} + b\dfrac{\tilde{p}_{i+1/2} - \tilde{p}_{i-1/2}}{\Delta x} = -\sigma_i\,\tilde{u}_i, \end{cases} \tag{A.10}$$

where $\tilde{p}_{i+1/2}(t)$ and $\tilde{u}_i(t)$ are the semi-discrete fields, $\sigma_{i+1/2}$ and σ_i are the discrete values of the absorption function at the same positions as semi-discrete fields, and Δx is the spatial step.

Following the procedure used in the previous sections, the plane-wave analysis of a discrete system provides dispersion and dissipation properties for plane waves. For this purpose, the form of the absorption function must be choosen *a priori* (constant or spatially varying form with a particular shape). However, a complete plane-wave analysis of PMLs is very complicated when σ_i varies spatially. Therefore, the procedure is adapted: the shape of the decay of plane waves is choosen, and both dispersion relation and discrete function σ_i must be found.

Let us assume a plane-wave solution with a linear decay,

$$\tilde{p}_{i+1/2} = P\,\frac{\delta - x_{i+1/2}}{\delta}\,e^{\imath(kx_{i+1/2}-\omega t)},$$
$$\tilde{u}_i = U\,\frac{\delta - x_i}{\delta}\,e^{\imath(kx_i-\omega t)}.$$

By replacing this solution in (A.10), one obtains

$$\left\{\begin{array}{l}(\sigma_{i+1/2}-\imath\omega)P + \dfrac{a}{\Delta x}\left(\dfrac{\delta - x_{i+1}}{\delta - x_{i+1/2}}\,e^{\imath k\Delta x/2} - \dfrac{\delta - x_i}{\delta - x_{i+1/2}}\,e^{-\imath k\Delta x/2}\right)U = 0,\\ (\sigma_i-\imath\omega)U + \dfrac{b}{\Delta x}\left(\dfrac{\delta - x_{i+1/2}}{\delta - x_i}\,e^{\imath k\Delta x/2} - \dfrac{\delta - x_{i-1/2}}{\delta - x_i}\,e^{-\imath k\Delta x/2}\right)P = 0,\end{array}\right.$$
$$\Longleftrightarrow \left\{\begin{array}{l}(\sigma_{i+1/2}-\imath\omega)P + \dfrac{a}{\Delta x}\left(2\imath\sin\left(\dfrac{k\Delta x}{2}\right) - \dfrac{\Delta x}{\delta - x_{i+1/2}}\cos\left(\dfrac{k\Delta x}{2}\right)\right)U = 0,\\ (\sigma_i-\imath\omega)U + \dfrac{b}{\Delta x}\left(2\imath\sin\left(\dfrac{k\Delta x}{2}\right) - \dfrac{\Delta x}{\delta - x_i}\cos\left(\dfrac{k\Delta x}{2}\right)\right)P = 0.\end{array}\right.$$

The imaginary part of the system gives the system (A.2). Therefore, the dispersion relation, the amplitude relation, the phase velocity and the group velocity are exactly the same as for the discrete wave system without PML terms. The real part of the system gives

$$\left\{\begin{array}{l}\sigma_{i+1/2} = \dfrac{a}{\Delta x}\dfrac{\Delta x}{\delta - x_{i+1/2}}\cos\left(\dfrac{k\Delta x}{2}\right)\dfrac{U}{P},\\ \sigma_i = \dfrac{b}{\Delta x}\dfrac{\Delta x}{\delta - x_i}\cos\left(\dfrac{k\Delta x}{2}\right)\dfrac{P}{U}.\end{array}\right.$$

Using the amplitude relation, one has immediately

$$\boxed{\sigma_i = \frac{c_g}{\delta - x_i},}$$

where c_g is the group velocity (A.4). A similar expression is obtained for $\sigma_{i+1/2}$.

APPENDIX B

Additional Material for Numerical Schemes

B.1 Conservative form of the PML scalar wave system

Let us consider the PML scalar wave system

$$\begin{cases} \dfrac{\partial p}{\partial t} + a\nabla\cdot\boldsymbol{u} = -\sigma p_s - \bar{\kappa}_\varphi\bar{\sigma}\, p_\varphi - \bar{\kappa}_\theta\bar{\sigma}(p - p_s - p_\varphi), \\ \dfrac{\partial \boldsymbol{u}}{\partial t} + b\nabla p = -\sigma \boldsymbol{e}_s(\boldsymbol{e}_s\cdot\boldsymbol{u}) - \bar{\kappa}_\varphi\bar{\sigma}\,\boldsymbol{e}_\varphi(\boldsymbol{e}_\varphi\cdot\boldsymbol{u}) - \bar{\kappa}_\theta\bar{\sigma}\,\boldsymbol{e}_\theta(\boldsymbol{e}_\theta\cdot\boldsymbol{u}), \\ \dfrac{\partial p_s}{\partial t} + a\nabla_{\boldsymbol{e}_s}\cdot\boldsymbol{u} = -\sigma p_s, \\ \dfrac{\partial p_\varphi}{\partial t} + a\nabla_{\boldsymbol{e}_\varphi}\cdot\boldsymbol{u} = -\bar{\kappa}_\varphi\bar{\sigma}\, p_\varphi. \end{cases}$$

The differential operator of the third equation can be rewritten as

$$\begin{aligned} \nabla_{\boldsymbol{e}_s}\cdot\boldsymbol{u} &= [\boldsymbol{e}_s(\boldsymbol{e}_s\cdot\nabla)]\cdot\boldsymbol{u}, \\ &= (\boldsymbol{e}_s\cdot\nabla)(\boldsymbol{e}_s\cdot\boldsymbol{u}) - [(\boldsymbol{e}_s\cdot\nabla)\boldsymbol{e}_s]\cdot\boldsymbol{u}, \\ &= \nabla\cdot[\boldsymbol{e}_s(\boldsymbol{e}_s\cdot\boldsymbol{u})] - [\boldsymbol{e}_s(\nabla\cdot\boldsymbol{e}_s) + (\boldsymbol{e}_s\cdot\nabla)\boldsymbol{e}_s]\cdot\boldsymbol{u}. \end{aligned}$$

A similar expression is obtained for $\nabla_{\boldsymbol{e}_\varphi}\cdot\boldsymbol{u}$ and $\nabla_{\boldsymbol{e}_\theta}\cdot\boldsymbol{u}$. Using relations between the basis vectors with the expression of the operator ∇ in curvilinear coordinates (2.25), one obtains

$$\begin{aligned} \nabla\cdot\boldsymbol{e}_s &= \frac{1}{1+\kappa_\varphi s}\boldsymbol{e}_\varphi\cdot(\kappa_\varphi\boldsymbol{e}_\varphi) + \frac{1}{1+\kappa_\theta s}\boldsymbol{e}_\theta\cdot(\kappa_\theta\boldsymbol{e}_\theta) = \frac{\kappa_\varphi}{1+\kappa_\varphi s} + \frac{\kappa_\theta}{1+\kappa_\theta s}, \\ \nabla\cdot\boldsymbol{e}_\varphi &= \frac{1}{1+\kappa_\varphi s}\boldsymbol{e}_\varphi\cdot(-\kappa_\varphi\boldsymbol{e}_s) = 0, \\ \nabla\cdot\boldsymbol{e}_\rho &= \frac{1}{1+\kappa_\theta s}\boldsymbol{e}_\theta\cdot(-\kappa_\theta\boldsymbol{e}_s) = 0, \end{aligned}$$

and

$$(\boldsymbol{e}_s\cdot\nabla)\boldsymbol{e}_s=0,\qquad (\boldsymbol{e}_\varphi\cdot\nabla)\boldsymbol{e}_\varphi=\frac{-\kappa_\varphi}{1+\kappa_\varphi s}\boldsymbol{e}_s,\qquad (\boldsymbol{e}_\rho\cdot\nabla)\boldsymbol{e}_\rho=\frac{-\kappa_\theta}{1+\kappa_\theta s}\boldsymbol{e}_s.$$

Finally, the differential operators of additional equations can be written

$$\begin{aligned}
\nabla_{\boldsymbol{e}_s}\cdot\boldsymbol{u}&=\nabla\cdot[\boldsymbol{e}_s(\boldsymbol{e}_s\cdot\boldsymbol{u})]-\left(\frac{\kappa_\varphi}{1+\kappa_\varphi s}+\frac{\kappa_\theta}{1+\kappa_\theta s}\right)(\boldsymbol{e}_s\cdot\boldsymbol{u}),\\
\nabla_{\boldsymbol{e}_\varphi}\cdot\boldsymbol{u}&=\nabla\cdot[\boldsymbol{e}_\varphi(\boldsymbol{e}_\varphi\cdot\boldsymbol{u})]+\frac{\kappa_\varphi}{1+\kappa_\varphi s}(\boldsymbol{e}_s\cdot\boldsymbol{u}),\\
\nabla_{\boldsymbol{e}_\rho}\cdot\boldsymbol{u}&=\nabla\cdot[\boldsymbol{e}_\rho(\boldsymbol{e}_\rho\cdot\boldsymbol{u})]+\frac{\kappa_\theta}{1+\kappa_\theta s}(\boldsymbol{e}_s\cdot\boldsymbol{u}).
\end{aligned}$$

Therefore, the conservative form of the PML scalar wave system reads

$$\left\{\begin{aligned}
\frac{\partial p}{\partial t}+a\nabla\cdot\boldsymbol{u}&=-\sigma p_s-\bar\kappa_\varphi\bar\sigma p_\varphi-\bar\kappa_\theta\bar\sigma(p-p_s-p_\varphi),\\
\frac{\partial\boldsymbol{u}}{\partial t}+b\nabla p&=-\sigma\boldsymbol{e}_s(\boldsymbol{e}_s\cdot\boldsymbol{u})-\bar\kappa_\varphi\bar\sigma\boldsymbol{e}_\varphi(\boldsymbol{e}_\varphi\cdot\boldsymbol{u})-\bar\kappa_\theta\bar\sigma\boldsymbol{e}_\theta(\boldsymbol{e}_\theta\cdot\boldsymbol{u}),\\
\frac{\partial p_s}{\partial t}+a\nabla\cdot[\boldsymbol{e}_s(\boldsymbol{e}_s\cdot\boldsymbol{u})]&=-\sigma p_s+\left(\frac{\kappa_\varphi}{1+\kappa_\varphi s}+\frac{\kappa_\theta}{1+\kappa_\theta s}\right)(\boldsymbol{e}_s\cdot\boldsymbol{u}),\\
\frac{\partial p_\varphi}{\partial t}+a\nabla\cdot[\boldsymbol{e}_\varphi(\boldsymbol{e}_\varphi\cdot\boldsymbol{u})]&=-\bar\kappa_\varphi\bar\sigma p_\varphi-a\frac{\kappa_\varphi}{1+\kappa_\varphi s}(\boldsymbol{e}_s\cdot\boldsymbol{u}).
\end{aligned}\right.$$

In two dimensions, this system reduces to

$$\left\{\begin{aligned}
\frac{\partial p}{\partial t}+a\nabla\cdot\boldsymbol{u}&=-\sigma p_s-\bar\kappa\bar\sigma(p-p_s),\\
\frac{\partial\boldsymbol{u}}{\partial t}+b\nabla p&=-\sigma\boldsymbol{e}_s(\boldsymbol{e}_s\cdot\boldsymbol{u})-\bar\kappa\bar\sigma\boldsymbol{e}(\boldsymbol{e}\cdot\boldsymbol{u}),\\
\frac{\partial p_s}{\partial t}+a\nabla\cdot[\boldsymbol{e}_s(\boldsymbol{e}_s\cdot\boldsymbol{u})]&=-\sigma p_s+\frac{\kappa_\varphi}{1+\kappa s}(\boldsymbol{e}_s\cdot\boldsymbol{u}).
\end{aligned}\right.\tag{B.1}$$

B.2 Numerical schemes for the PML scalar wave system

In this section, numerical schemes based on finite difference (FD) and continuous Galerkin (CG) methods are derived for the two-dimensional PML scalar wave system. They are used in chapter 3.

B.2.1 Finite difference scheme

The FD scheme is derived for a straight boundary with the outward normal $\boldsymbol{e}_x$. In this case, the system reduces to the system (2.30), with a single additional unknown field $p_n(\boldsymbol{x},t)$.

The additional field $p_n(\boldsymbol{x},t)$ is discretized exactly as the original one $p(\boldsymbol{x},t)$, *i.e.*

$$\tilde{p}_s\big|_{i+1/2,j+1/2}(t) \approx p_n(\boldsymbol{x}_{i+1/2,j+1/2},t).$$

Using central differences to approximate partial derivatives, the scheme then reads

$$\begin{cases} \dfrac{\mathrm{d}\tilde{p}_{i+1/2,j+1/2}}{\mathrm{d}t} + a\left(\dfrac{\tilde{u}_{i+1,j+1/2}-\tilde{u}_{i-1,j+1/2}}{\Delta x} + \dfrac{\tilde{v}_{i+1/2,j+1}-\tilde{v}_{i+1/2,j-1}}{\Delta y}\right) = -\sigma_{i+1/2}\,\tilde{p}_s\big|_{i+1/2,j+1/2}, \\ \dfrac{\mathrm{d}\tilde{u}_{i,j+1/2}}{\mathrm{d}t} + b\dfrac{\tilde{p}_{i+1/2,j+1/2}-\tilde{p}_{i-1/2,j+1/2}}{\Delta x} = -\sigma_i\,\tilde{u}_{i,j+1/2}, \\ \dfrac{\mathrm{d}\tilde{v}_{i+1/2,j}}{\mathrm{d}t} + b\dfrac{\tilde{p}_{i+1/2,j+1/2}-\tilde{p}_{i+1/2,j-1/2}}{\Delta y} = 0, \\ \dfrac{\mathrm{d}\,\tilde{p}_s\big|_{i+1/2,j+1/2}}{\mathrm{d}t} + a\dfrac{\tilde{u}_{i+1,j+1/2}-\tilde{u}_{i-1,j+1/2}}{\Delta x} = -\sigma_{i+1/2}\,\tilde{p}_s\big|_{i+1/2,j+1/2}. \end{cases}$$

where $\sigma_i = \sigma(x_i)$ and $\sigma_{i+1/2} = \sigma(x_{i+1/2})$ are the discrete values of the absorption function.

B.2.2 Continuous finite element scheme

Multiplying the conservative form (B.1) of the equations by test functions, integrating them over $\Omega\cup\Sigma$ and using integration by parts, one gets the weak form:

$$\begin{cases} \left(\dfrac{\partial p}{\partial t},\psi_i\right)_{\Omega\cup\Sigma} + \left\langle a\,(\boldsymbol{u}\cdot\boldsymbol{n}_{\Omega\cup\Sigma}),\psi_i\right\rangle_{\partial(\Omega\cup\Sigma)} - \left(a\boldsymbol{u},\nabla\psi_i\right)_{\Omega\cup\Sigma} = -\left(S^p,\psi_i\right)_\Sigma, \\ \left(\dfrac{\partial \boldsymbol{u}}{\partial t},\psi_i\right)_{\Omega\cup\Sigma} + \left(b\nabla p,\psi_i\right)_{\Omega\cup\Sigma} = -\left(\boldsymbol{S}^{\boldsymbol{u}},\psi_i\right)_\Sigma, \\ \left(\dfrac{\partial p_n}{\partial t},\psi_i\right)_\Sigma + \left\langle a\,(\boldsymbol{e}_s\cdot\boldsymbol{u})\,(\boldsymbol{e}_s\cdot\boldsymbol{n}_\Sigma),\psi_i\right\rangle_{\partial\Sigma} - \left(a(\boldsymbol{e}_s\cdot\boldsymbol{u})\boldsymbol{e}_s,\nabla\psi_i\right)_\Sigma = -\left(S^{p_n},\psi_i\right)_\Sigma, \end{cases} \tag{B.2}$$

where $\boldsymbol{n}_{\Omega\cup\Sigma}$ and $\boldsymbol{n}_\Sigma$ are the outward unit normals of $\Omega\cup\Sigma$ and Σ, respectively. S^p, S^{p_n} and $\boldsymbol{S}^{\boldsymbol{u}}$ are the source terms written in the right-hand side of the equations of the system (B.1).

Unfortunately, the formulation (B.2) generates spurious numerical oscillations. The pressure-stabilization Petrov-Galerkin method (PSPG method) consists in adding a stabilization term in the left-hand side of the first equation of (B.2),

$$-\left(\kappa\boldsymbol{R}^{\boldsymbol{u}},\nabla\psi_i\right)_{\Omega\cup\Sigma},$$

where κ is the numerical stabilization parameter and $\boldsymbol{R}^{\boldsymbol{u}}$ is the residue associated to the momentum equation

$$\boldsymbol{R}^{\boldsymbol{u}} = \frac{\partial \boldsymbol{u}}{\partial t} + b\nabla p + \boldsymbol{S}_{\boldsymbol{u}}.$$

Considering the similarities between the governing equations of p and p_n, a corresponding stabilization term is added in the last equations,

$$-\left(\kappa(\boldsymbol{e}_n\cdot\boldsymbol{R}^{\boldsymbol{u}})\boldsymbol{e}_n,\nabla\psi_i\right)_\Sigma.$$

Bibliography

[1] S. Abarbanel and D. Gottlieb. A mathematical analysis of the PML method. *Journal of Computational Physics*, 134(2):357–363, 1997.

[2] S. Abarbanel and D. Gottlieb. On the construction and analysis of absorbing layers in CEM. *Applied Numerical Mathematics*, 27:331–340, 1998.

[3] S. Abarbanel, D. Gottlieb, and J. S. Hesthaven. Well-posed perfectly matched layers for advective acoustics. *Journal of Computational Physics*, 154(2): 266–283, 1999.

[4] S. Abarbanel, D. Gottlieb, and J. S. Hesthaven. Long time behavior of the perfectly matched layer equations in computational electromagnetics. *Journal of Scientific Computing*, 17(1-4):405–22, 2002.

[5] S. Abarbanel, D. Stanescu, and M. Y. Hussaini. Unsplit variables perfectly matched layers for the shallow water equations with Coriolis forces. *Computational Geosciences*, 7(4):275–294, 2003.

[6] S. Abarbanel, H. Qasimov, and S. Tsynkov. Long-time performance of unsplit PMLs with explicit second order schemes. *Journal of Scientific Computing*, 41(1):1–12, 2009.

[7] A. Alvera-Azcárate, A. Barth, and R. H. Weisberg. A nested model of the cariaco basin (venezuela): description of the basin's interior hydrography and interactions with the open ocean. *Ocean Dynamics*, 59(1):97–120, 2009.

[8] X. Antoine, H. Barucq, and A. Bendali. Bayliss–turkel-like radiation conditions on surfaces of arbitrary shape. *Journal of Mathematical Analysis and Applications*, 229(1):184–211, 1999.

[9] X. Antoine, A. Arnold, C. Besse, M. Ehrhardt, and A. Schädle. A review of transparent and artificial boundary conditions techniques for linear and nonlinear Schrödinger equations. *Communications in Computational Physics*, 4(4):729–796, 2008.

[10] A. Arakawa and V. R. Lamb. Computational design of the basic dynamical processes of the UCLA general circulation model. *Methods in Computational Physics*, 17:173–265, 1977.

[11] A. Barth, A. Alvera-Azcarate, J.-M. Beckers, R. H. Weisberg, L. Vandenbulcke, F. Lenartz, and M. Rixen. Dynamically constrained ensemble perturbations: Application to tides on the West Florida shelf. *Ocean Science*, 5(3):259, 2009.

[12] H. Barucq and M. Fontes. Well-posedness and exponential stability of Maxwell-like systems coupled with strongly absorbing layers. *Journal de Mathématiques Pures et Appliquées*, 87(3):253–273, 2007.

[13] H. Barucq, J. Diaz, and M. Tlemcani. New absorbing layers conditions for short water waves. *Journal of Computational Physics*, 229(1):58–72, 2010.

[14] H. Barucq, J. Diaz, and M. Tlemcani. Perfectly matched layers for linearized and non linear shallow water equation. In *Proceedings of the 11th Waves conference, Tunisia*, June 2013.

[15] A. Bayliss and E. Turkel. Radiation boundary conditions for wave-like equations. *Communications on Pure and Applied Mathematics*, 33:707–725, 1980. doi: 10.1002/cpa.3160330603.

[16] E. Bécache and P. Joly. On the analysis of Bérenger's perfectly matched layers for Maxwell's equations. *ESAIM - Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique*, 36(1):87–119, 2002.

[17] E. Bécache and A. Prieto. Remarks on the stability of cartesian pmls in corners. *Applied Numerical Mathematics*, 62(11):1639–1653, 2012.

[18] E. Bécache, A.-S. Bonnet-BenDhia, and G. Legendre. Perfectly matched layers for the convected Helmholtz equation. *SIAM Journal on Numerical Analysis*, 42(1):409–433, 2004.

[19] E. Bécache, P. G. Petropoulos, and S. D. Gedney. On the long-time behavior of unsplit perfectly matched layers. *IEEE Transactions on Antennas and Propagation*, 52(5):1335–1342, 2004.

[20] J.-M. Beckers and E. Deleersnijder. Stability of a FBTCS scheme applied to the propagation of shallow-water inertia-gravity waves on various space grids. *Journal of Computational Physics*, 108:95–104, 1993.

[21] J.-P. Bérenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, 114(2):185–200, 1994.

[22] J.-P. Bérenger. Three-dimensional perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, 127(2): 363–379, 1996.

[23] J.-P. Bérenger. Evanescent waves in PML's: origin of the numerical reflection in wave-structure interaction problems. *IEEE Transactions on Antennas and Propagation*, 47(10):1497–1503, 1999.

[24] J.-P. Bérenger. Numerical reflection of evanescent waves by PMLs: origin and interpretation in the FDTD case. Expected consequences to other finite methods. *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, 13(2-3):103–114, 2000.

[25] J.-P. Bérenger. *Perfectly Matched Layer (PML) for Computational Electromagnetics*. Morgan & Claypool, 2007.

[26] A. Bermúdez, L. Hervella-Nieto, A. Prieto, and R. Rodríguez. An exact

bounded perfectly matched layer for time-harmonic scattering problems. *SIAM Journal on Scientific Computing*, 30(1):312–338, 2007.

[27] E. Blayo and L. Debreu. Revisiting open boundary conditions from the point of view of characteristic variables. *Ocean Modelling*, 9(3):231–252, 2005.

[28] E. Blayo and V. Martin. Building generalized open boundary conditions for fluid dynamics problems. *International Journal for Numerical Methods in Fluids*, 2012.

[29] A. Blumberg and L. Kantha. Open boundary condition for circulation models. *Journal of Hydraulic Engineering*, 111:237–255, 1985.

[30] A. Bousquet, M. Petcu, M.-C. Shiue, R. Temam, and J. Tribbia. Boundary conditions for limited area models based on the shallow water equations. *Communications in Computational Physics*, 14:664–702, 2013.

[31] S. Cailleau, V. Fedorenko, B. Barnier, E. Blayo, and L. Debreu. Comparison of different numerical methods used to handle the open boundary of a regional ocean circulation model of the bay of biscay. *Ocean Modelling*, 25 (1-2):1–16, 2008.

[32] A. L. Camerlengo and J. J. O'Brien. Open boundary conditions in rotating fluids. *Journal of Computational Physics*, 35:12–35, 1980.

[33] D. C. Chapman. Numerical treatment of cross-shelf open boundaries in a barotropic coastal ocean model. *Journal of Physical Oceanography*, 15: 1060–1075, 1985.

[34] W. C. Chew and J. M. Jin. Perfectly matched layers in the discretized space: An analysis and optimization. *Electromagnetics*, 16(4):325–340, 1996.

[35] W. C. Chew and W. H. Weedon. A 3D perfectly matched medium from modified Maxwell's equations with stretched coordinates. *Microwave and Optical Technology Letters*, 7(13):599–604, 1994.

[36] W. C. Chew, J. M. Jin, and E. Michielssen. Complex coordinate stretching as a generalized absorbing boundary condition. *Microwave and Optical Technology Letters*, 15(6):363–369, 1997.

[37] B. Cockburn, G. E. Karniadakis, and C.-W. Shu. The development of discontinuous Galerkin methods. In B. Cockburn, G. E. Karniadakis, and C.-W. Shu, editors, *Discontinuous Galerkin Methods*, volume 11 of *Lecture Notes in Computational Science and Engineering*, pages 3–50. Springer, 2000.

[38] F. Collino. High order absorbing boundary conditions for wave propagation models: straight line boundary and corner cases. In *Proceedings of the* 2^{nd} *Wave conference, Newark (DE)*, pages 161–171, 1993.

[39] F. Collino and P. B. Monk. Optimizing the perfectly matched layer. *Computer Methods in Applied Mechanics and Engineering*, 164(1-2):157–171, 1998.

[40] F. Collino and P. B. Monk. The perfectly matched layer in curvilinear coordinates. *SIAM Journal on Scientific Computing*, 19:2061–2090, 1998.

[41] T. Colonius. Modeling artificial boundary conditions for compressible flow. *Annual Review of Fluid Mechanics*, 36:315–345, 2004.

[42] R. Comblen, J. Lambrechts, J.-F. Remacle, and V. Legat. Practical evaluation of five partly discontinuous finite element pairs for the non-conservative shallow water equations. *International Journal for Numerical Methods in Fluids*, 63(6):701–724, 2010.

[43] D. Correia and J.-M. Jin. On the development of a higher-order PML. *IEEE Transactions on Antennas and Propagation*, 53(12):4157–4163, 2005.

[44] B. Cushman-Roisin and J.-M. Beckers. *Introduction to Geophysical Fluid Dynamics: Physical and Numerical Aspects*. Academic Press, 2011.

[45] G. Darblade, R. Baraille, A. Y. LeRoux, X. Carton, and D. Pinchon. Non reflecting boundary conditions for a two dimensional linearized barotropic shallow-water model. *Comptes Rendus de l'Academie des Sciences Series I Mathematics*, 324(4):485–490, 1997.

[46] H. C. Davies. A lateral boundary formulation for multi-level prediction models. *Quarterly Journal of the Royal Meteorological Society*, 102:405–418, 1976.

[47] H. C. Davies. Limitations of some common lateral boundary schemes used in regional NWP models. *Monthly Weather Review*, 11:1002–1012, 1983.

[48] J. De Moerloose and M. Stuchly. Behavior of Berenger's ABC for evanescent waves. *IEEE Microwave and Guided Wave Letters*, 5(10):344–346, 1995.

[49] T. De Mulder. *Stabilized finite element methods (SUPG, GLS,...) for incompressible flows*. Lecture Notes. von Karman Institute for Fluid Dynamics, Rhode-Saint-Genese, Belgium, 1997.

[50] L. Debreu and E. Blayo. Two-way embedding algorithms: a review. *Ocean Dynamics*, 58:415–428, 2008.

[51] J. Diaz and P. Joly. A time domain analysis of PML models in acoustics. *Computer Methods in Applied Mechanics and Engineering*, 195(29-32):3820–3853, 2006.

[52] B. Donderici and F. L. Teixeira. Conformal perfectly matched layer for the mixed finite element time-domain method. *IEEE Transactions on Antennas and Propagation*, 56(4):1017–1026, 2008.

[53] P. Dular. *Modélisation du champ magnétique et des courants induits dans des systèmes tridimensionnels non linéaires*. PhD thesis, Université de Liège, November 1994.

[54] P. Dular and C. Geuzaine. *GetDP: a general environment for the treatment of discrete problems*. http://www.geuz.org/getdp/.

[55] P. Dular, C. Geuzaine, F. Henrotte, and W. Legros. A general environment for the treatment of discrete problems and its application to the finite element method. *IEEE Transactions on Magnetics*, 34(5):3395–3398, 1998.

[56] D. R. Durran. *Numerical Methods for Wave Equations in Geophysical Fluid*

Dynamics. Springer, 1999.

[57] K. Duru and G. Kreiss. A well-posed and discretely stable perfectly matched layer for elastic wave equations in second order formulation. *Communications in Computational Physics*, 11(5):1643, 2012.

[58] T. Elvius and A. Sundström. Computationally efficient schemes and boundary conditions for a fine-mesh barotropic model based on the shallow-water equations. *Tellus*, 25:132–156, 1973.

[59] B. Engquist and A. Majda. Absorbing boundary conditions for numerical simulation of waves. *Mathematics of Computation*, 139:629–651, 1977.

[60] S. Ervedoza and E. Zuazua. Perfectly matched layers in 1-d: Energy decay for continuous and semi-discrete waves. *Numerische Mathematik*, 109(4): 597–634, 2008.

[61] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, second edition, 2010.

[62] J. E. Flaherty, L. Krivodonova, J.-F. Remacle, and M. S. Shephard. Aspects of discontinuous Galerkin methods for hyperbolic conservation laws. *Finite Elements in Analysis and Design*, 38(10):889–908, 2002.

[63] R. Flather. A tidal model of the north-west european continental shelf. *Mémoires Société Royale des Sciences de Liège*, 10:141–164, 1976.

[64] V. Fontaine, M. Aubourg, and P. Guillon. Curved perfectly matched layer in frequency domain. *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, 13(2-3):155–165, 2000.

[65] S. Gedney. Perfectly matched layer absorbing boundary conditions. In A. Taflove, editor, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, pages 273–328. Artech House, 2005.

[66] S. D. Gedney. An anisotropic perfectly matched layer-absorbing medium for the truncation of FDTD lattices. *IEEE Transactions on Antennas and Propagation*, 44(12):1630–1639, 1996.

[67] C. Geuzaine. *High order hybrid finite element schemes for Maxwell's equations taking thin structures and global quantities into account*. PhD thesis, Université de Liège, December 2001.

[68] C. Geuzaine and J.-F. Remacle. *Gmsh: a finite element mesh generator with built-in pre- and post-processing facilities*. http://www.geuz.org/gmsh/.

[69] C. Geuzaine and J.-F. Remacle. Gmsh: A 3-d finite element mesh generator with built-in pre-and post-processing facilities. *International Journal for Numerical Methods in Engineering*, 79(11):1309–1331, 2009.

[70] D. Givoli. Non-reflecting boundary-conditions. *Journal of Computational Physics*, 94(1):1–29, 1991.

[71] D. Givoli. *Numerical Methods for Problems in Infinite Domains*. Elsevier, 1992.

[72] D. Givoli. High-order local non-reflecting boundary conditions: a review. *Wave Motion*, 39(4):319–326, 2004.

[73] D. Givoli. Computational absorbing boundaries. In *Computational Acoustics of Noise Propagation in Fluids*, chapter 5, pages 145–166. Springer, Berlin, 2008.

[74] D. J. Griffiths. *Introduction to Electrodynamics*. Addison-Wesley, third edition, 1999.

[75] C. E. Grosch and S. A. Orszag. Numerical solution of problems in unbounded regions: Coordinate transforms. *Journal of Computational Physics*, 25(3):273–295, 1977.

[76] H. W. Guggenheimer. *Differential geometry*. Dover, 1977.

[77] B. Gustafsson, H.-O. Kreiss, and J. Oliger. *Time Dependent Problems and Difference Methods*. John Wiley & Sons, 1995.

[78] T. Hagstrom. Radiation boundary conditions for the numerical simulation of waves. *Acta Numerica*, 8:47–106, 1999.

[79] T. Hagstrom. Radiation boundary conditions for Maxwell's equations: a review of accurate time-domain formulations. *Journal of Computational Mathematics*, 25:305–336, 2007.

[80] L. Halpern, S. Petit-Bergez, and J. Rauch. The analysis of matched layers. *Confluentes Mathematici*, 3(02):159–236, 2011.

[81] E. Hanert, V. Legat, and E. Deleersnijder. A comparison of three finite elements to solve the linear shallow water equations. *Ocean Modelling*, 5: 17–35, 2002.

[82] I. Harari. Reducing spurious dispersion, anisotropy and reflection in finite element analysis of time-harmonic acoustics. *Computer Methods in Applied Mechanics and Engineering*, 140(1-2):39–58, 1997.

[83] I. Harari and U. Albocher. Studies of FE/PML for exterior problems of time-harmonic elastic waves. *Computer methods in applied mechanics and engineering*, 195(29):3854–3879, 2006.

[84] F. D. Hastings, J. B. Schneider, and S. L. Broschat. Application of the perfectly matched layer (PML) absorbing boundary condition to elastic wave propagation. *The Journal of the Acoustical Society of America*, 100:3061, 1996.

[85] E. Hecht. *Optics*. Addison-Wesley, fourth edition, 2002.

[86] J. S. Hesthaven and T. Warburton. *Nodal discontinuous Galerkin methods: algorithms, analysis, and applications*, volume 54. Springerverlag New York, 2008.

[87] C. Hirsch. *Numerical Computation of Internal and External Flows: The Fundamentals of Computational Fluid Dynamics, Volume 1*. John Wiley & Sons, second edition, 2007.

[88] R. Holland and J. W. Williams. Total-field versus scattered-field finite-

difference codes: A comparative assessment. *IEEE Transactions on Nuclear Science*, 30(6):4583–4588, 1983.

[89] F. Q. Hu. On absorbing boundary conditions for linearized euler equations by a perfectly matched layer. *Journal of Computational Physics*, 129:201–219, 1996.

[90] 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.

[91] F. Q. Hu. A perfectly matched layer absorbing boundary condition for linearized euler equations with a non-uniform mean flow. *Journal of Computational Physics*, 208(2):469–492, 2005.

[92] F. Q. Hu. Development of PML absorbing boundary conditions for computational aeroacoustics: A progress review. *Computers & Fluids*, 37(4):336–348, 2008.

[93] F. Q. Hu, M. Y. Hussaini, and P. Rasetarinera. An analysis of the discontinuous galerkin method for wave propagation problems. *Journal of Computational Physics*, 151(2):921–946, 1999.

[94] T. J. R. Hughes, L. P. Franca, and M. Balestra. A new finite element formulation for computational fluid dynamics: V. Circumventing the Babuska-Brezzi condition: a stable Petrov-Galerkin formulation of the Stokes problem accommodating equal-order interpolations. *Computer Methods in Applied Mechanics and Engineering*, 59(1):85–99, 1986.

[95] J.-F. Imhoff, G. Meunier, X. Brunotte, and J. Sabonnadiere. An original solution for unbounded electromagnetic 2D- and 3D-problems throughout the finite element method. *IEEE Transactions on Magnetics*, 26(5):1659–61, 1990.

[96] M. Israeli and S. A. Orszag. Approximation of radiation boundary conditions. *Journal of Computational Physics*, 41(1):115–135, 1981.

[97] J. D. Jackson. *Classical Electrodynamics*. Wiley, third edition, 1998.

[98] T. G. Jensen. Open boundary conditions in stratified ocean models. *Journal of Marine Systems*, 16(3-4):297–322, 1998.

[99] J. Jin. *The Finite Element Method in Electromagnetics (2nd edition)*. John Wiley & Sons, 2002.

[100] D. Katz, E. Thiele, and A. Taflove. Validation and extension to three dimensions of the Bérenger PML absorbing boundary condition for FD-TD meshes. *IEEE Microwave and Guided Wave Letters*, 4(8):268–270, 1994.

[101] J. B. Keller and D. Givoli. Exact non-reflecting boundary-conditions. *Journal of Computational Physics*, 82(1):172–192, 1989.

[102] G. Kreiss and K. Duru. Discrete stability of perfectly matched layers for anisotropic wave equations in first and second order formulation. *BIT Numerical Mathematics*, pages 1–23, 2013.

[103] P. K. Kundu, I. M. Cohen, and D. R. Dowling. *Fluid Mechanics*. Academic Press, fifth edition, 2011.

[104] M. Kuzuoglu and R. Mittra. Frequency dependence of the constitutive parameters of causal perfectly matched anisotropic absorbers. *IEEE Microwave and Guided Wave Letters*, 6(12):447–449, 1996.

[105] M. Lassas and E. Somersalo. Analysis of the PML equations in general convex geometry. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics*, 131:1183–1207, 2001.

[106] M. Lassas, J. Liukkonen, and E. Somersalo. Complex Riemannian metric and absorbing boundary condition. *Journal de Mathématique Pures et Appliquées*, 80(7):739–768, 2001.

[107] S. Laurens. *Approximation de Haute Précision des Problèmes de diffraction*. PhD thesis, Université Toulouse III - Paul Sabatier, March 2010.

[108] J. W. Lavelle and W. C. Thacker. A pretty good sponge: Dealing with open boundaries in limited-area ocean models. *Ocean Modelling*, 20(3):270–292, 2008.

[109] R. J. LeVeque. *Finite volume methods for hyperbolic problems*, volume 31. Cambridge University Press, 2002.

[110] J. Li and Y. Huang. Perfectly matched layers. In *Time-Domain Finite Element Methods for Maxwell's Equations in Metamaterials*, pages 215–240. Springer, 2013.

[111] E. Marchandise. *Simulation of three-dimensional two-phase flows: coupling of a stabilized finite element method with a discontinuous level set approach*. PhD thesis, Université catholique de Louvain, December 2006.

[112] E. Marchandise, N. Chevaugeon, and J.-F. Remacle. Spatial and spectral superconvergence of discontinuous Galerkin method for hyperbolic problems. *Journal of Computational and Applied Mathematics*, 215(2):484–494, 2008.

[113] P. Marchesiello, J. C. McWilliams, and A. Shchepetkin. Open boundary conditions for long-term integration of regional oceanic models. *Ocean Modelling*, 3:1–20, 2001.

[114] E. A. Martinsen and H. Engedahl. Implementation and testing of a lateral boundary scheme as an open boundary condition in a barotropic ocean model. *Coastal Engineering*, 11:603–627, 1987.

[115] A. McDonald. A step toward transparent boundary conditions for meteorological models. *Monthly Weather Review*, 130:140–151, 2002.

[116] A. McDonald. Transparent boundary conditions for the shallow-water equations: testing in a nested environment. *Monthly weather review*, 131 (4):698–705, 2003.

[117] A. McDonald. Transparent lateral boundary conditions for baroclinic waves: a study of two elementary systems of equations. *Tellus*, 57A:171–182, 2005.

[118] A. McDonald. Transparent lateral boundary conditions for baroclinic waves II. Introduicing potential vorticity waves. *Tellus*, 58A:210–220, 2006.

[119] A. McDonald. Transparent lateral boundary conditions for baroclinic waves III. Including vertical shear. *Tellus*, 61A:227–231, 2009.

[120] C. Mercier. *Numerical schemes for tracer studies in shallow shelf seas*. PhD thesis, Université de Liège, October 2012.

[121] C. Michler, L. Demkowicz, J. Kurtz, and D. Pardo. Improving the performance of perfectly matched layers by means of hp-adaptivity. *Numerical Methods for Partial Differential Equations*, 23(4):832–858, 2007.

[122] M. J. Miller and A. J. Thorpe. Radiation conditions for the lateral boundaries of limited-area numerical models. *Quarterly Journal of the Royal Meteorological Society*, 107:615–628, 1981.

[123] A. Modave, E. Deleersnijder, and E. J. M. Delhez. On the parameters of absorbing layers for shallow water models. *Ocean Dynamics*, 60(1):65–79, 2010.

[124] N. Moës and T. Belytschko. X-FEM, de nouvelles frontières pour les éléments finis. *Revue Européenne des Eléments*, 11(2-4):305–318, 2002.

[125] V. Mouysset. *Une méthode de sous-domaines pour la résolution des équations de Maxwell instationnaires en présence d'un ensemble non-connexe d'objets diffractant*. PhD thesis, Université Toulouse III - Paul Sabatier, December 2006.

[126] F. Nataf. A new approach to perfectly matched layers for the linearized Euler system. *Journal of Computational Physics*, 214(2):757–772, 2006.

[127] I. M. Navon, B. Neta, and M. Y. Hussaini. A perfectly matched layer approach to the linearized shallow water equations models. *Monthly Weather Review*, 132(6):1369–1378, 2004.

[128] J.-C. Nédélec. Mixed finite elements in $\mathbb{R}^3$. *Numerische Mathematik*, 35(3): 315–341, 1980.

[129] A. Nicolet, J.-F. Remacle, B. Meys, A. Genon, and W. Legros. Transformation methods in computational electromagnetism. *Journal of Applied Physics*, 75 (10):6036–6038, 1994.

[130] A. Nissen and G. Kreiss. An optimized perfectly matched layer for the Schrödinger equation. *Communication in Computational Physics*, 9:147–179, 2011.

[131] J. Nycander and K. Döös. Open boundary conditions for barotropic waves. *Journal of Geophysical Research-Oceans*, 108(C5):3168–3187, 2003.

[132] P. Oddo and N. Pinardi. Lateral open boundary conditions for nested limited area models: A scale selective approach. *Ocean Modelling*, 20(2):134–156, 2008.

[133] X. Ojeda and T. Pichon. Combining the finite element method and a Padé

approximation for scattering analysis application to radiated electromagnetic compatibility problems. *Journal of Electromagnetic Waves and Applications*, 19(10):1375–1390, 2005.

[134] I. Orlanski. A simple boundary condition for unbounded hyperbolic flows. *Journal of Computational Physics*, 21:251–269, 1976.

[135] E. D. Palma and R. P. Matano. On the implementation of passive open boundary conditions for a general circulation model: The barotropic mode. *Journal of Geophysical Research-Oceans*, 103(C1):1319–1341, 1998.

[136] E. D. Palma and R. P. Matano. On the implementation of open boundary conditions for a general circulation model: The three-dimensional case. *Journal of Geophysical Research-Oceans*, 105(C4):8605–8627, 2000.

[137] E. D. Palma and R. P. Matano. Dynamical impacts associated with radiation boundary conditions. *Journal of Sea Research*, 46(2):117–132, 2001.

[138] S. A. Parrish and F. Q. Hu. PML absorbing boundary conditions for the linearized and nonlinear euler equations in the case of oblique mean flow. *International Journal for Numerical Methods in Fluids*, 60(5):565–589, 2009.

[139] R. A. Pearson. Consistent boundary conditions for numerical models of systems that admit dispersive waves. *Journal of the Atmospheric Sciences*, 31: 1481–1489, 1974.

[140] P. G. Petropoulos. On the termination of the perfectly matched layer with local absorbing boundary conditions. *Journal of Computational Physics*, 143 (2):665–673, 1998.

[141] P. G. Petropoulos. Reflectionless sponge layers as absorbing boundary conditions for the numerical solution of Maxwell equations in rectangular, cylindrical, and spherical coordinates. *SIAM Journal on Applied Mathematics*, 60(3):1037–1058, 2000.

[142] P. G. Petropoulos. An analytical study of the discrete perfectly matched layer for the time-domain Maxwell equations in cylindrical coordinates. *IEEE Transactions on Antennas and Propagation*, 51(7):1671–1675, 2003.

[143] P. G. Petropoulos, L. Zhao, and A. C. Cangellaris. A reflectionless sponge layer absorbing boundary condition for the solution of Maxwell's equations with high-order staggered finite difference schemes. *Journal of Computational Physics*, 139(1):184–208, 1998.

[144] C. Rappaport. Perfectly matched absorbing boundary conditions based on anisotropic lossy mapping of space. *IEEE Transactions on Microwave Theory and Techniques*, 5(3):90–92, 1995.

[145] P. Raviart and J. Thomas. A mixed finite element method for second order elliptic problems. In I. Galligani and E. Magenes, editors, *Mathematical Aspects of Finite Element Methods*, volume 606 of *Lecture Notes in Mathematics*, pages 292–315. Springer Berlin Heidelberg, 1977.

[146] W. H. Raymond and H. L. Kuo. A radiation boundary condition for multi-

dimensional flows. *Quarterly Journal of the Royal Meteorological Society*, 110:535–551, 1984.

[147] S. W. Rienstra and A. Hirschberg. *An introduction to acoustics*. Lecture notes. Eindhoven University of Technology. Version June 30, 2013. Online ressource: http://www.win.tue.nl/~sjoerdr/papers/boek.pdf.

[148] J. Roden and S. Gedney. Convolution PML (CPML): An efficient FDTD implementation of the CFS–PML for arbitrary media. *Microwave and Optical Technology Letters*, 27(5):334–339, 2000.

[149] L. P. Røed and C. Cooper. Open boundary conditions in numerical ocean models. In J. O'Brien, editor, *Advanced Physical Oceanography Numerical Modelling*, volume 186, pages 411–436. NATO ASI Series C, 1986.

[150] L. P. Røed and C. Cooper. A study of various open boundary conditions for wind-forced barotropic numerical ocean models. In J. Nihoul and B. Jamart, editors, *Three-Dimensional Models of Marine and Estuarine Dynamics*, volume 45, pages 305–335. Elsevier, 1987.

[151] L. P. Røed and O. M. Smedstad. Open boundary conditions for forced waves in a rotating fluid. *SIAM Journal on Scientific and Statistical Computing*, 5 (2):414–426, 1984.

[152] Z. Sacks. A perfectly matched anisotropic absorber for use as an absorbing boundary condition. *IEEE Transactions on Antennas and Propagation*, 43 (12):1460–1463, 1995.

[153] G. D. Smith. *Numerical Solution of Partial Differential Equations: Finite Difference Methods*. Oxford University Press, third edition, 1985.

[154] A. Sommerfeld. *Partial differential equations in physics*. Academic Press, 1949.

[155] A. Taflove and S. C. Hagness. *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Artech House, third edition, 2005.

[156] C. K. W. Tam, L. Auriault, and F. Cambuli. Perfectly matched layer as an absorbing boundary condition for the linearized Euler equations in open and ducted domains. *Journal of Computational Physics*, 144(1):213–234, 1998.

[157] Y. Tang and R. Grimshaw. Radiation boundary conditions in barotropic coastal ocean numerical models. *Journal of Computational Physics*, 123(1): 96–110, 1996.

[158] F. L. Teixeira and W. C. Chew. Systematic derivation of anisotropic PML absorbing media in cylindrical and spherical coordinates. *IEEE Microwave and Guided Wave Letters*, 7(11):371–373, 1997.

[159] F. L. Teixeira and W. C. Chew. Analytical derivation of a conformal perfectly matched absorber for electromagnetic waves. *Microwave and Optical Technology Letters*, 17(4):231–236, 1998.

[160] F. L. Teixeira and W. C. Chew. Differential forms, metrics, and the reflectionless absorption of electromagnetic waves. *Journal of Electromagnetic Waves*

and Applications, 13(5):665–686, 1999.

[161] F. L. Teixeira and W. C. Chew. Complex space approach to perfectly matched layers: a review and some new developments. *International Journal of Numerical Modelling-Electronic Networks Devices and Fields*, 13(5):441–455, 2000.

[162] R. Tezaur, A. Macedo, C. Farhat, and R. Djellouli. Three-dimensional finite element calculations in acoustic scattering using arbitrarily shaped convex artificial boundaries. *International journal for numerical methods in engineering*, 53(6):1461–1476, 2002.

[163] E. Toro. *Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction*. Springer-Verlag Berlin Heidelberg, 2009.

[164] S. V. Tsynkov. Numerical solution of problems on unbounded domains. A review. *Applied Numerical Mathematics*, 27(4):465–532, 1998.

[165] G. K. Verboom and A. Slob. Weakly-reflective boundary conditions for two-dimensional shallow water flow problems. *Advances in Water Resources*, 7 (4):192–197, 1984.

[166] G. B. Whitham. *Linear and Nonlinear Waves*. John Wiley & Sons, 1974.

[167] D. C. Wittwer and R. W. Ziolkowski. How to design the imperfect Bérenger PML. *Electromagnetics*, 16(4):465–485, 1996.

[168] K. Yee. Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, 14(3):302–307, 1966.

[169] L. Zhao. The generalized theory of perfectly matched layers (GT-PML) in curvilinear coordinates. *International Journal of Numerical Modelling-Electronic Networks Devices and Fields*, 13(5):457–469, 2000.

[170] L. Zhao and A. C. Cangellaris. GT-PML: Generalized theory of perfectly matched layers and its application to the reflectionless truncation of finite-difference time-domain grids. *IEEE Transactions on Microwave Theory and Techniques*, 44(12):2555–2563, 1996.

List of Figures

List of Tables

List of Symbols and Abbreviations

Common symbols

$\mathbb{R}$	:	Set of real numbers
$\mathbb{C}$	:	Set of complex numbers
Ω	:	Truncated domain
Σ	:	Absorbing layer
Υ	:	Truncation interface, interface between Ω and Σ
Γ	:	External boundary of the absorbing layer Σ
t	:	Time instant [s]
$\boldsymbol{x}$	:	Spatial position [m]
$\boldsymbol{e}_x$, $\boldsymbol{e}_y$, $\boldsymbol{e}_z$	:	Cartesian directions (unit vectors)
x, y, z	:	Cartesian coordinates [m]
$\boldsymbol{e}_s$, $\boldsymbol{e}_\varphi$, $\boldsymbol{e}_\theta$	:	Curvilinear directions associated to Υ (unit vectors)
s, φ, θ	:	Curvilinear coordinates associated to Υ
$\mathbf{I}$	:	Identity tensor
$\imath$	:	Imaginary unit ($=\sqrt{-1}$)
ω	:	Angular frequency [rad/s]
$\boldsymbol{k}$	:	Wave vector [m^{-1}]
k	:	Wave number ($=\|\boldsymbol{k}\|$)
c	:	Phase velocity ($=\omega/k$)
σ	:	Absorption function [s^{-1}]
σ_x, σ_y, σ_z	:	Absorption functions in the cartesian directions [s^{-1}]
ℓ	:	Characteristic size of a mesh cell (edge length) [m]

Operators

^{T}	:	Transpose
$\cdot$	:	Scalar product
$\times$	:	Vector product
$\|\boldsymbol{a}\|$	:	2-norm of a vector, equals to $\sqrt{\boldsymbol{a}\cdot\boldsymbol{a}}$
d_x	:	First-order derivative
$d^n_{x^n}$	:	n^{th}-order derivative
∂_x	:	First-order partial derivative
$\partial^n_{x^n}$	:	n^{th}-order partial derivative
∇	:	Operator *nabla*, equal to $(\partial_x,\partial_y,\partial_z)^T$ in three dimensions
$\nabla_{\boldsymbol{a}}$	:	Operator $[\boldsymbol{a}(\boldsymbol{a}\cdot\nabla)]$, where $\boldsymbol{a}$ is a vector

Specific symbols for acoustics

p	: Pressure [A/m]
$\boldsymbol{u}$	: Velocity [m/s]
ρ	: Density [g/m^3]

Specific symbols for electromagnetism

$\boldsymbol{e}$	: Electric field [V/m]
$\boldsymbol{h}$	: Magnetic field [A/m]
ε	: Electric permittivity [F/m]
ε_0	: Electric permittivity of vacuum ($\approx 8.854187817\,10^{-12}$ F/m)
μ	: Magnetic permeability [H/m]
μ_0	: Magnetic permeability of vacuum ($= 4\pi\,10^{-7}$ H/m)
Z	: Impedance ($= \sqrt{\mu/\varepsilon}$)
Y	: Admittance ($= \sqrt{\varepsilon/\mu}$)

Specific symbols for oceanography

η	: Free-surface elevation [m]
$\boldsymbol{u}$	: Depth-averaged velocity [m/s]
f	: Coriolis coefficient [s^{-1}]
g	: Gravitational acceleration [m/s^2]
h	: Depth at rest [m]
H	: Total depth [m]
ν	: Horizontal eddy viscosity [m^2/s]
$\boldsymbol{\tau}_b, \boldsymbol{\tau}_w$	: Surface wind stress, bottom stress [Pa]

Abbreviations

CG	: Continuous Galerkin
DG	: Discontinuous Galerkin
FD	: Finite difference
FE	: Finite element
FV	: Finite volume
ABC	: Absorbing boundary condition
OBC	: Open boundary condition
PML	: Perfectly matched layer
FRS	: Flow relaxation scheme
PGS	: Pretty good sponge