AN ABSTRACT OF THE THESIS OF

Dojin Kim for the degree of Doctor of Philosophy in Mathematics presented on June 4, 2015.

Title: The Variable Speed Wave Equation and Perfectly Matched Layers

Abstract approved: _____

David V. Finch

A perfectly matched layer (PML) is widely used to model many different types of wave propagation in different media. It has been found that a PML is often very effective and also easy to set, but still many questions remain. We introduce a new formulation from regularizing the classical Un-Split PML of the acoustic wave equation and show the well-posedness and numerical efficiency. A PML is designed to absorb incident waves traveling perpendicular to the PML, but there is no effective absorption of waves traveling with large incident angles. We suggest one method to deal with this problem and show well-posedness of the system, and some numerical experiments. For the 1-d wave equation with a constant speed equipped a PML, stability and the exponential decay rate of energy has been proved, but the question for variable sound speed equation remained open. We show that the energy decays exponentially in the 1-d PML wave equation with variable sound speed. Most PML wave equations appear as a first-order hyperbolic system with as a zero-order perturbation. We introduce a general formulation and show well-posedness and stability of the system. Furthermore we develop a discontinuous Galerkin method and analyze both the semi-discrete and fully discretized system and provide *a priori* error estimations.

The Variable Speed Wave Equation and Perfectly Matched Layers

by

Dojin Kim

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented June 4, 2015
Commencement June 2016

Doctor of Philosophy thesis of Dojin Kim presented on June 4, 2015

APPROVED:

Major Professor, representing Mathematics

Chair of the Department of Mathematics

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Dojin Kim, Author

ACKNOWLEDGEMENTS

Academic

I am really indebted to my advisor Dr. David V. Finch for his patience and guidance for many years. I also appreciate Dr. Ralph E. Showalter for good advice in research, and the department for giving me a chance to study at OSU and support for many years. This research is partially supported by NSF grant DMS-1009194.

Personal

I wish to thank to my loving parents, Changgyu Kim and Youngphil Kwon for all their support, patience, and love.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

THE VARIABLE SPEED WAVE EQUATION AND PERFECTLY MATCHED LAYERS

1. INTRODUCTION

1.1. Problem From Thermoacoustic Tomography

In thermoacoustic tomography, a short electro-magnetic pulse is emitted into a body and irradiated tissue generates acoustic waves. Thermoacoustic tomography aims to recover the degree and distribution of energy deposition from measurement of the generated acoustic waves on the surface of the body.

The problem is mathematically modeled in the following way. We assume that $c(x) > 0$ is the acoustic sound speed at location x. Let u solve the problem

$$\begin{cases} \partial_{tt} u - c^2 \Delta u = 0 \quad \text{in } (0,T) \times \mathbb{R}^n, \\ u|_{t=0} = f, \\ \partial_t u|_{t=0} = 0, \end{cases}$$

where $T > 0$ is fixed.

Assume that f is supported in $\bar{\Omega}$, where $\Omega \subset R^n$ is some smooth bounded domain. The mathematical measurements on the boundary of Ω are modeled by the operator

$$\Lambda f := u|_{[0,T]\times\partial\Omega}.$$

The goal is to reconstruct the initial value f at $t = 0$, using the data of Λf.

If $T = \infty$, then there are some well known results for many different cases of sound speed c, geometry, dimension n, and data (see [33] for example). One reconstruction method

when $T < \infty$ is fixed, greater than the length of the longest geodesic in Ω, is introduced in [33]. It is to get an approximate solution of the thermoacoustic problem by the following time reversal method. Given $h = \Lambda f$, let v solve

$$\begin{cases} (\partial_t^2 - c^2\Delta)v = 0 \text{ in } (0,T)\times\Omega, \\ v|_{[0,T]\times\partial\Omega} = h, \\ v|_{t=T} = \phi, \\ \partial_t v|_{t=T} = 0, \end{cases} \tag{1.1}$$

where ϕ solves the elliptic boundary problem

$$-c^2\Delta\phi = 0, \quad \phi|_{\partial\Omega} = h(T,\cdot).$$

Then the following pseudo-inverse operator is defined from the range of Λ to $H^1(\Omega)$:

$$Ah := v(0,\cdot) \quad \text{ in } \bar{\Omega}.$$

Let $(\Omega, c^{-2}g)$ be a non-trapping Riemannian manifold. i.e., $T(\Omega) < \infty$ where $T(\Omega)$ is the supremum of the length of all geodesics of the metric $c^{-2}g$ in $\bar{\Omega}$.

Theorem 1.1 [33] *Let $T > T(\Omega)$. Then $A\Lambda = Id - K$, where K is compact in $H^1(\Omega)$, and $\|K\|_{H^1(\Omega)} < 1$. In particular, $Id-K$ is invertible on $H^1(\Omega)$, and the inverse thermoacoustic problem has an explicit solution of the form*

$$f = \sum_{m=0}^{\infty} K^m Ah, \quad h := \Lambda f.$$

This theorem motivates a line of research, which is to find a good numerical approximation of $\|K\|_{H^1(\Omega)}$. For the numerical implementation, there are three separate procedures:

1. (Forward-Collect data) Implement acoustic wave equation with Cauchy initial data which is compactly supported in Ω.

2. (Elliptic Boundary Problem) Implement elliptic boundary problem with data at $(t, x) \in T \times \partial\Omega$.

3. (Backward-Reconstruct initial value) Implement acoustic wave equation reversely imposing boundary value from the collected data.

The forward problem is set in the unbounded domain in $\mathbb{R}^n$, which must be truncated for numerical experiments. In the first step of the simulation it is required to have a large enough computational domain to ensure that the data collected on the boundary of Ω is not affected by reflected waves from the numerical boundary, but it is expensive to implement such a large additional domain. There has been much attention devoted to numerical simulation of wave equations in reasonably sized computational domains avoiding reflecting waves from the boundaries. Two general methods are the development of non-reflecting boundary conditions or by surrounding the computational domain by absorbing layers.

1.2. Background of Perfectly Matched Layers

One of the most effective and straight forward ways to truncate an unbounded domain numerically is to surround the computational domain with thin artificial absorbing layers. This is called the Perfectly Matched Layers (PML) method. In 1994, the PML method was first introduced by J. P. Berenger [21] who found absorbing boundary layers for Maxwell's equations. The key property of a PML method is that it is originally designed so that waves incident upon the PML from a non-PML region do not reflect at the interface. This property allows the PML to strongly attenuate by the absorption and exponentially decay outgoing waves in the layers. Since its introduction, there have been several modified reformulations of PML for both Maxwell's equations [50] and for other wave-type equations, such as elastodynamics [51], the linearized Euler equations

[17, 20], and Helmholtz equations [16, 49]. Berenger's original formulation is called a split-field PML, or split PML because the electromagnetic field is split into two unphysical fields. A later formulation called uniaxial PML or UPML [44] describes the PML without any splitting as the ordinary wave equation with a combination of artificial anisotropic absorbing materials. Thus it has become more popular because its simplicity and efficiency. Although both Berenger's formulation and UPML were originally derived by manually computing the solutions for a wave incident on the PML at an arbitrary angle, and then finding conditions under which incident plane waves do not reflect from the PML interface in a homogeneous medium, both of these formulations were later rederived by a much more flexible and general way using a method called a complex coordinate stretching [50]. The complex-coordinate approach is essentially based on analytic continuation of the wave equations into complex spatial coordinates in the layer which changes outward propagating waves to exponentially decaying waves. In this viewpoint, a PML is allowed to be derived for many different media as well as for many different wave type equations. The next section, following [47], explains how the solutions of wave equations exponentially decay in the PML.

1.3. Complex Coordinate Stretching

The complex coordinate stretching can be viewed as the complex coordinate change of variables (see chapter 2.1. for e.g.) in the frequency domain of a wave equation. For more detail, a new complex coordinates $\tilde{x}$ is stretched from the real value x in the new media $x > a$,

$$\tilde{x} = \begin{cases} x & \text{if } x \leq a, \\ x + iy(x) & \text{if } x > a, \end{cases} \tag{1.2}$$

where y is a real nonnegative C^1 function. Then the exponential

$$e^{ik\tilde{x}} = e^{ikx+iy} = e^{ikx}e^{-ky},$$

which has exponential decay for $k > 0$ as y increases. More specifically, y is defined as

$$y(x) = \frac{1}{\omega}\int_a^x \sigma(s)ds, \tag{1.3}$$

where a damping function $\sigma := \sigma(x)$ is a positive C^0 function vanishing when $x < a$ and ω is the frequency. That gives a new complex coordinates,

$$\tilde{x} = x + i\frac{\int_a^x \sigma(s)ds}{\omega}. \tag{1.4}$$

Furthermore,

$$\frac{\partial}{\partial\tilde{x}} = \frac{\partial x}{\partial\tilde{x}}\frac{\partial}{\partial x} = \frac{1}{1+i\frac{\sigma(x)}{\omega}}\frac{\partial}{\partial x}.$$

The reason for applying the frequency ω is that the decay is then independent of the wave number, so that it depends only on spatial position and the sound speed c as follows, by the dispersion relation,

$$e^{ik\tilde{x}} = e^{ik\left(x+i\frac{\int_a^x \sigma(s)ds}{\omega}\right)} = e^{ikx-\frac{k}{\omega}\int_a^x \sigma(s)ds} = e^{ikx-\frac{1}{c}\int_a^x \sigma(s)ds}.$$

1.4. Limitations of PML

The PML method has been widely adapted to different types of wave equations in various media, but there are some limitations such as unavoidable reflection or even exponential growth from the interface between the computational domain and the layer. First, the method is designed to be reflection-less to the positive x-direction for the exact, continuous wave equations. Once the equations equipped with a PML are discretized, there is no guarantee of non-reflection of numerical solutions. But this weakness can be dealt with by making the absorption coefficient σ increase gradually from zero over a PML

simultaneously making a layer thicker or increasing the resolution to get acceptable reflections [2]. Next, the basic idea of PML is that the solutions of wave equations are analytic functions in the normal direction to the boundary and can be analytically continued to the complex plane, so that PML is not applicable in some inhomogeneous media (see [47] for more detail). Another problem is when waves propagate tangentially to a PML because it is assumed that waves moves in the direction perpendicular to the PML, but not all waves hit the interface as it is designed. For example, as the radiation approaches glancing incidence, the reflection is getting bigger. Setting a PML far from the domain of interest mitigates this problem but it costs more from the bigger computation domain. Besides the perfect matching of layers it is also desirable that the equations governing the PML be well-posed in a mathematical view. We focus on the investigation of the well-posedness and the stability of PML wave equations in the next section.

1.5. Well-posedness and Stability

Since the time the PML method was introduced, the well-posedness and stability issue has been investigated in many different ways in different media, but there still remain some questions. This is important when a PML derived for a constant coefficient linear problem is to be applied to a non-linear problem or a problem with variable coefficients. If the linearized problem is only *weakly well-posed*, the corresponding non-linear problem or variable coefficient problem can be *ill-posed*. Most PML wave equations have the form of lower-order perturbations of a first order hyperbolic system. Therefore, the general stability theory for first order hyperbolic systems may ensure the well-posedness of the Cauchy problem associated to the PML equations. Let us assume that a first order system of partial differential equations for complex valued function on $\mathbb{R}^{1+d}$ is obtained from the

wave equation with a PML,

$$\mathcal{L}(x,\partial_t,\partial_x)U := \partial_t U + \sum_{l=1}^{d} \mathcal{A}_l \partial_l U + \mathcal{B}(x)U = 0, \tag{1.5}$$

with the principal part of $\mathcal{L}$, denoted $\mathcal{L}_1$,

$$\mathcal{L}_1(\partial_t,\partial_x) := \partial_t + \sum_{l=1}^{d} \mathcal{A}_l \partial_l, \tag{1.6}$$

having constant matrix coefficients $\mathcal{A}_l$. The Cauchy problem for $\mathcal{L}$ is to find a solution U defined on $[0,\infty)\times\mathbb{R}^d$ satisfying (1.5) with prescribed initial data $U(0,\cdot)$. Following [26]

Definition 1.1 *The Cauchy problem for $\mathcal{L}_1$ is weakly well posed if there exist $q>0, K>0$ and $\alpha\in\mathbb{R}$ so that for any initial value in $H^q(\mathbb{R}^d)$, there is a unique solution $U\in C^0([0,\infty);L^2(\mathbb{R}^d))$ with*

$$\forall t\geq 0, \|U(t,\cdot)\|_{L^2(\mathbb{R}^d)} \leq Ke^{\alpha t}\|U(0,\cdot)\|_{H^q(\mathbb{R}^d)}.$$

The Cauchy problem for $\mathcal{L}_1$ is weakly stable if there is a unique solution $U\in C^0([0,\infty);L^2(\mathbb{R}^d))$ with

$$\forall t\geq 0, \|U(t,\cdot)\|_{L^2(\mathbb{R}^d)} \leq K(1+t)^q\|U(0,\cdot)\|_{H^q(\mathbb{R}^d)}.$$

When the conclusion holds with $q=0$, the Cauchy problem is called strongly well posed or strongly stable, respectively.

Theorem 1.2 *1. The Cauchy problem for $\mathcal{L}_1$ is weakly well posed if and only if for each $\xi\in\mathbb{R}^d$, the eigenvalues of $\mathcal{L}_1(0,\xi)$ are real.*

2. The Cauchy problem for $\mathcal{L}_1$ is strongly well posed if and only if for each $\xi\in\mathbb{R}^d$, the eigenvalues of $\mathcal{L}_1(0,\xi)$ are real and $\mathcal{L}_1(0,\xi)$ is uniformly diagonalizable, there is an invertible $S(\xi)$ satisfying,

$$S(\xi)^{-1}\mathcal{L}_1(0,\xi)S(\xi) = \textit{diagonal}, \quad S, S^{-1}\in L^\infty(\mathbb{R}^d_\xi).$$

3. *If $\mathcal{B}$ has constant coefficients, then the Cauchy problem for $\mathcal{L}$ is weakly well posed (or weakly stable, respectively) if and only if there exists $M \geq 0$ (M=0, respectively) such that for any $\xi \in \mathbb{R}^d$, $\det \mathcal{L}(\tau, \xi) = 0 \rightarrow |Im\ \tau| \leq M$.*

The analytical stability of PMLs have already been claimed by several authors in the cases of Maxwell's equations [14, 37, 27], stability for wave equations [46, 14, 30], stability for elastic wave [18], and shown unstable for anisotropic media [13]. Mostly, stability has been shown by investigating eigenvalues of the first order hyperbolic equations obtained from the constant speed wave equations with PMLs. Additionally a general interpretation of a PML is that the restriction of the equations to the PML equation in the computational domain coincides with the original problem [11]. In the view of a PML, damping terms are required to vanish identically in the computational region, so that the condition of constant damping terms generates discontinuity on the interface between the computational and PML regions. There is a restricted stability result for a general first-order hyperbolic system of the acoustic wave propagation in corners [15]. Alternatively energy techniques can be used to answer the question of stability for a PML associated to wave equations with variable sound speed and damping terms as well as in the case of a constant damping. In studying the constant damping case by energy techniques, a property widely used by several authors [14] is

$$(\partial_t + \sigma)\partial_x = \partial_x(\partial_t + \sigma).$$

As the previous comments about the continuation of a PML, a constant damping creates a jump on the interface, so that the equality is not quite true. There is also another published claim [7] of stability for 3-d second order PML wave equation using an energy method, but in our opinion the proof is not clear in choosing test functions. Therefore we haven't seen clear stability proof for the PML wave equation in several dimensions ($d \geq 2$). In 1 dimension, stability is proved and even the decay rate for the constant speed case is proved [46]. The decay rate in dimension one for variable sound speed and

in higher dimensions for all sounds speeds remains an open question. This stability issue motivates us to propose the following questions:

1. Give a clear answer of analytical stability for a PML wave equation with a constant or variable sound speed in a higher dimension if it is, or provide a counter example if it isn't (in both the case of a constant and non constant damping).

2. Construct a PML wave equation which is stable and efficient in a higher dimension.

3. Figure out energy decay rate in a PML wave equation with variable sound speed in 1 dimension.

1.6. Organization of this Thesis

The organization of this dissertation is as follows:
Chapter 2 will introduce new regularized system of acoustic wave equation associated to the Un-Split PML. Regularizing a term in the classical Un-Split PML we can show well-posedness of the system by energy technique without losing the efficiency of PML.
In chapter 3 we introduce additional damping in a classical PML, which derives new formulation. With this damping we introduce two type of system, Split and Un-Split PML, and show well-posedness and numerical efficiency.
Chapter 4 we show the energy decay for the 1-d acoustic wave equation with variable sound speed equipped a PML.
Chapter 5 will introduce a system of first order hyperbolic equation with low order damping. We show that the system is well-posed and that energy decays. Introducing local discontinuous Galerkin (LDG) method, we show an *a priori* error estimate of LDG for the system.
In chpater 6, as a continuation of the chapter 5, we present a fully discretized discontinuous Galerkin method for the system and show an *a priori* L^2-error estimate under

additional regularity assumptions.

2. REGULARIZED SYSTEM OF UN-SPLIT PML ACOUSTIC WAVE EQUATION

In this section, we start with a second order system of acoustic 2-d wave equations with variable sound speed associated to an Un-Split PML as was introduced in [30]. The well-posedness and also stability of the system have not been clearly established yet, since the damping terms appear in the first order term. This was claimed in [30, 7], but the arguments have errors or are incomplete.

We introduce a new regularized system and show the well-posedness using energy techniques and the standard Galerkin method. The idea of regularization of a specific term is suggested by [20], in which the same technique is first applied to the Split PML formulation of the linearized Euler equations.

First, we show how a PML is applied in the acoustic wave equation.

2.1. Coordinate Transform

We consider the second order acoustic wave equation with variable sound speed in a domain $\Omega_0 \subset\subset [-a,a]\times[-b,b] \subset \mathbb{R}^2$, for some $T>0$,

$$u_{tt}(\vec{x},t) - c(\vec{x})^2 \Delta u(\vec{x},t) = 0, \quad \forall (\vec{x},t) \in \mathbb{R}^2 \times (0,T], \tag{2.1}$$

with the initial condition $u(\vec{x},0) = f, u_t(\vec{x},0) = 0$ with $supp(f) \subset \Omega_0$.

We assume that the sound speed $c(\vec{x}) > 0$ is bounded by

$$0 < c_* \leq c(\vec{x}) \leq c^* < \infty. \tag{2.2}$$

After the even extension of the solution in the entire time domain, i.e., $u(\vec{x},-t) = u(\vec{x},t)$ for all $t \in \mathbb{R}$, we take *Fourier* transform of u in time,

$$\hat{u}(\vec{x},\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} u(\vec{x},t) e^{-i\omega t} dt, \quad \text{for any } \omega \in \mathbb{R}.$$

Then $\hat{u}$ satisfies the Helmholtz equation,

$$-\frac{\omega^2}{c^2}\hat{u}(\vec{x},\omega) = \frac{\partial}{\partial x}\left(\frac{\partial \hat{u}}{\partial x}(\vec{x},\omega)\right) + \frac{\partial}{\partial y}\left(\frac{\partial \hat{u}}{\partial y}(\vec{x},\omega)\right). \tag{2.3}$$

Let the domain $\Omega = [-a - L_x, a + L_x] \times [-b - L_y, b + L_y]$ consist of the computational domain $[-a, a] \times [-b, b]$ surrounded by PML region, where $a, b, L_x, L_y > 0$.

Next, we introduce the coordinate transform using the damping as introduced in (1.3) ($-\omega$ will be used instead of ω for convenient notation).

$$\vec{x} := (x, y) \mapsto (\tilde{x}(x), \tilde{y}(y)) := \left(x + \frac{1}{i\omega}\int_a^x \sigma_x(s)ds,\ y + \frac{1}{i\omega}\int_b^y \sigma_y(s)ds\right), \tag{2.4}$$

where the damping terms $\sigma_\alpha, \alpha = x, y$, are non-negative C^0 functions vanishing in $[-a, a] \times [-b, b]$.

We apply the new coordinate system in the Helmholtz equation (2.3),

$$-\frac{\omega^2}{c^2}\hat{u}(\vec{x},\omega) = \frac{\partial}{\partial \tilde{x}}\left(\frac{\partial \hat{u}}{\partial \tilde{x}}(\vec{x},\omega)\right) + \frac{\partial}{\partial \tilde{y}}\left(\frac{\partial \hat{u}}{\partial \tilde{y}}(\vec{x},\omega)\right). \tag{2.5}$$

From (2.4), we have the partial differentiation with respect to $\tilde{x}, \tilde{y}$ related to partial derivatives with respect to x, y,

$$\frac{\partial}{\partial \tilde{x}} = \frac{1}{\eta_x}\frac{\partial}{\partial x}, \quad \frac{\partial}{\partial \tilde{y}} = \frac{1}{\eta_y}\frac{\partial}{\partial y}, \quad \eta_\alpha = 1 + \frac{\sigma_\alpha}{i\omega},\ \alpha = x, y.$$

Then, by replacing the partial derivatives in (2.5) and multiplying η_x, η_y, we rewrite it,

$$-\frac{w^2}{c^2}\hat{u} = \frac{1}{\eta_x}\frac{\partial}{\partial x}\left(\frac{1}{\eta_x}\frac{\partial \hat{u}}{\partial x}\right) + \frac{1}{\eta_y}\frac{\partial}{\partial y}\left(\frac{1}{\eta_y}\frac{\partial \hat{u}}{\partial y}\right),$$

or

$$-\eta_x\eta_y\frac{w^2}{c^2}\hat{u} = \frac{\partial}{\partial x}\left(\frac{\eta_y}{\eta_x}\frac{\partial \hat{u}}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\eta_x}{\eta_y}\frac{\partial \hat{u}}{\partial y}\right). \tag{2.6}$$

Simple computations give that

$$\frac{\eta_y}{\eta_x}\frac{\partial \hat{u}}{\partial x} = \frac{(1 + \frac{\sigma_y}{iw})}{(1 + \frac{\sigma_x}{iw})}\frac{\partial \hat{u}}{\partial x} = \frac{(\sigma_y + iw)}{(\sigma_x + iw)}\frac{\partial \hat{u}}{\partial x} = \frac{(\sigma_y - \sigma_x + \sigma_x + iw)}{(\sigma_x + iw)}\frac{\partial \hat{u}}{\partial x} = \frac{\partial \hat{u}}{\partial x} + (\frac{\sigma_y - \sigma_x}{\sigma_x + iw})\frac{\partial \hat{u}}{\partial x},$$

$$\frac{\eta_x}{\eta_y}\frac{\partial \hat{u}}{\partial y} = \frac{(1 + \frac{\sigma_x}{iw})}{(1 + \frac{\sigma_y}{iw})}\frac{\partial \hat{u}}{\partial y} = \frac{(\sigma_x + iw)}{(\sigma_y + iw)}\frac{\partial \hat{u}}{\partial y} = \frac{(\sigma_x - \sigma_y + \sigma_y + iw)}{(\sigma_y + iw)}\frac{\partial \hat{u}}{\partial y} = \frac{\partial \hat{u}}{\partial y} + (\frac{\sigma_x - \sigma_y}{\sigma_y + iw})\frac{\partial \hat{u}}{\partial y},$$

and

$$\eta_x\eta_y\frac{(iw)^2}{c^2}\hat{u} = \Big(1+\frac{\sigma_x}{iw}\Big)\Big(1+\frac{\sigma_y}{iw}\Big)\frac{(iw)^2}{c^2}\hat{u} = \frac{1}{c^2}((iw)^2 + iw(\sigma_x+\sigma_y)+\sigma_x\sigma_y)\hat{u}.$$

Again we rewrite (2.6) to obtain

$$\frac{1}{c^2}((iw)^2 + iw(\sigma_x+\sigma_y)+\sigma_x\sigma_y)\hat{u} = \frac{\partial^2\hat{u}}{\partial x^2}+\frac{\partial^2\hat{u}}{\partial y^2}+\frac{\partial}{\partial x}\left(\frac{\sigma_y-\sigma_x}{\sigma_x+iw}\frac{\partial\hat{u}}{\partial x}\right)+\frac{\partial}{\partial y}\left(\frac{\sigma_x-\sigma_y}{\sigma_y+iw}\frac{\partial\hat{u}}{\partial y}\right). \tag{2.7}$$

We introduce the auxiliary variable $\hat{\vec{\mathbf{q}}} = (\hat{q}_x, \hat{q}_y)^T$

$$(\sigma_x+iw)\hat{q}_x = (\sigma_y-\sigma_x)\frac{\partial\hat{u}}{\partial x},$$

$$(\sigma_y+iw)\hat{q}_y = (\sigma_x-\sigma_y)\frac{\partial\hat{u}}{\partial y}.$$

Applying the inverse *Fourier* transform to $\hat{u}$ and $\hat{\vec{\mathbf{q}}}$ in (2.7) we have the following system of the PML wave equation: For all $(\vec{x},t)\in\Omega\times(0,T]$, $(u,\vec{\mathbf{q}})$ satisfies

$$\begin{cases} \frac{1}{c^2}u_{tt}(\vec{x},t)+\alpha(\vec{x})u_t(\vec{x},t)+\beta(\vec{x})u(\vec{x},t)-\nabla\cdot\vec{\mathbf{q}}(\vec{x},t)-\Delta u(\vec{x},t) &= 0,\\ \vec{\mathbf{q}}_t(\vec{x},t) \quad + \quad A(\vec{x})\vec{\mathbf{q}}(\vec{x},t) \quad + \quad B(\vec{x})\nabla u(\vec{x},t) &= 0, \end{cases} \tag{2.8}$$

with the initial conditions

$$u(\vec{x},0) := u^0 = f, \qquad u_t(\vec{x},0) := u^1 = 0, \qquad \vec{\mathbf{q}}(\vec{x},0) := \vec{\mathbf{q}}^0 = \vec{\mathbf{0}},$$

and the zero *Dirichlet* boundary condition

$$u(\vec{x},\cdot)|_{\partial\Omega} = 0,$$

where

$$\alpha(\vec{x}) = \frac{\sigma_x+\sigma_y}{c^2}, \quad \beta(\vec{x}) = \frac{\sigma_x\sigma_y}{c^2}, \quad A(\vec{x}) = \begin{bmatrix} \sigma_x & 0 \\ 0 & \sigma_y \end{bmatrix}, \quad B(\vec{x}) = \begin{bmatrix} \sigma_x-\sigma_y & 0 \\ 0 & \sigma_y-\sigma_x \end{bmatrix},$$

and $\sigma_x := \sigma_x(x)$ and $\sigma_y := \sigma_y(y)$ are nonnegative C^0 functions which vanish in the computational domain in the sense of the analytical continuation of the PML.

2.2. Regularized System

We introduce the new formulation which consists in regularizing the term in (2.8)

$$\nabla \cdot \vec{\mathbf{q}}.$$

Let $\delta_\varepsilon : H^{-1}(\Omega) \to H^{-1}(\Omega) \cap L^2(\Omega)$ such that $\delta_\varepsilon u$ is an approximation of u in the sense that

$$\delta_\varepsilon u \to u \ \text{ as } \ \varepsilon \to 0$$

for all $u \in H^{-1}(\Omega)$. Then we replace (2.8) by the new equations

$$\begin{cases} \frac{1}{c^2} u_{tt} + \alpha u_t + \beta u - \delta_\varepsilon \nabla \cdot \vec{\mathbf{q}} - \Delta u \ = \ 0, \\ \vec{\mathbf{q}}_t \ + \ A\vec{\mathbf{q}} \ + \ B\nabla u \ = \ 0, \end{cases} \tag{2.9}$$

where $\delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}$ denotes the regularizing operator of the function $\nabla \cdot \vec{\mathbf{q}}$ using the smooth function ρ_ϵ in the following way.

Let $\rho \in C^\infty(\mathbb{R}^2)$ with $supp(\rho) \subseteq B_1(0)$ and $\int_{\mathbb{R}^2} \rho(x)dx = 1$, which is called a *mollifier*. For $\epsilon > 0$, $\rho_\epsilon(x)$ on $\mathbb{R}^2$ is defined by

$$\rho_\epsilon(x) = \epsilon^{-2} \rho(\frac{|x|}{\epsilon}), \tag{2.10}$$

and satisfies $\int_{\mathbb{R}^2} \rho_\epsilon(x)dx = 1$ and $supp(\rho_\epsilon) \subseteq \overline{B_\epsilon(0)}$.

We recall the definition of the Sobolev space $H^1(\Omega)$:

$$H^1(\Omega) = \{\varphi : \varphi, \partial_{x_1}\varphi, \partial_{x_2}\varphi \in L^2(\Omega)\}.$$

and denote

$$H^{-1}(\Omega) = [H_0^1(\Omega)]'$$

the dual space of $H_0^1(\Omega)$ for a Lipschitz domain Ω.

First we define an approximation operator to the identity over $H_0^1(\Omega)$ in Lemma 2.1. To do that, we introduce the following definition.

Definition 2.1 *We say that Ω satisfies the* segment condition *if for each $x_0 \in \partial\Omega$ there is a neighborhood U of x_0 and a point $y_0 \in \mathbb{R}^n$ such that*

$$\bar{\Omega} \cap U + ty_0 \subseteq \Omega, \text{ for } 0 < t < 1. \tag{2.11}$$

Lemma 2.1 *We define a linear bounded operator $\delta^\varepsilon : H_0^1(\Omega) \to H_0^1(\Omega) \cap H^2(\Omega)$ for any $u \in H_0^1(\Omega)$ such that $\delta^\varepsilon \to \mathbb{1}$ in $H_0^1(\Omega)$ as $\varepsilon \to 0$ in the strong operator topology following Theorem 2.6 in* [5].

Proof. By a partition of unity each function $u \in H_0^1(\Omega)$ is a linear combination of functions in $H_0^1(\Omega)$ with small bounded supports. Assume that $u \in H_0^1(\Omega)$ has compact support and *supp* $u \subseteq \bar{\Omega} \cap U$, where U in (2.11). Let $u_t(x) = u(x - ty_0)$ for some y_0 satisfying (2.11) so that *supp* $u_t \subseteq \Omega$ for $0 < t < 1$. Let $\varepsilon > 0$, then there is t_0 such that $0 < t \leq t_0$ implies $\|u_t - u\|_{H^1(\Omega)} < \varepsilon/2$, since translation in $H^1(\Omega)$ is continuous. We can choose $\epsilon' > 0$ such that $supp(\rho_{\epsilon'} * u_{t_0}) \in C_c^\infty(\Omega)$ and $\|\rho_{\epsilon'} * u_{t_0} - u_{t_0}\|_{H^1(\Omega)} < \varepsilon/2$, since $\partial x^\alpha(\rho_{\epsilon'} * u_{t_0}) = \rho_{\epsilon'} * \partial x^\alpha u_{t_0} \to \partial x^\alpha u_{t_0}$ in $L^2(\mathbb{R})$, as $\epsilon' \to 0$, for $x^\alpha = x, y$. Taking $\delta^\varepsilon(u) = \rho_{\epsilon'(\varepsilon)} * u_{t_0}$ we have $\|\delta^\varepsilon(u) - u\|_{H^1(\Omega)} < \varepsilon$ for an arbitrarily given $\varepsilon > 0$, in which the proof is completed. □

Remark 2.1 *Now we consider the operator $\delta_\varepsilon : H^{-1}(\Omega) \to H^{-1}(\Omega) \cap L^2(\Omega)$ given by*

$$\delta_\varepsilon(f) = \mathcal{R} \circ \delta^\varepsilon \circ \mathcal{R}^{-1}(f) \quad \text{for all } f \in H^{-1}(\Omega), \tag{2.12}$$

where $\mathcal{R} := -\Delta + I$ is the Riesz map from $H_0^1(\Omega) \to H^{-1}(\Omega)$. Then

$$\delta_\varepsilon \to \mathbb{1} \text{ as } \varepsilon \to \infty \text{ in the strong operator topology,}$$

and also satisfies $\|\delta_\varepsilon(f)\|_{L^2(\Omega)} \leq C_{\delta_\varepsilon}\|f\|_{H^{-1}(\Omega)}$ for some $C_{\delta_\varepsilon} > 0$, since, by the isometry of $\mathcal{R}$,

$$\begin{aligned} \|\delta_\varepsilon(f) - f\|_{H^{-1}(\Omega)} &= \|\mathcal{R}\delta^\varepsilon\mathcal{R}^{-1}(f) - f\|_{H^{-1}(\Omega)} \\ &= \|\delta^\varepsilon\mathcal{R}^{-1}(f) - \mathcal{R}^{-1}(f)\|_{H_0^1(\Omega)} \\ &= \|\delta^\varepsilon u - u\|_{H_0^1(\Omega)} \to 0 \text{ as } \varepsilon \to 0, \end{aligned}$$

for $u \in H_0^1(\Omega)$ such that $\mathcal{R}(u) = f$.

Note that δ_ε is a linear and bounded operator from $H^{-1}(\Omega)$ to $H^{-1}(\Omega) \cap L^2(\Omega)$.

2.3. Well-posedness of the System

Now we show that the system (2.9) is well-posed, provided our function spaces are defined properly and provided functions σ_x, σ_y satisfy

$$\sigma_x, \sigma_y \in L^\infty(\Omega),$$

which implies that

$$\|\alpha\|_\infty = \|\sigma_x + \sigma_y\|_\infty < \infty, \qquad \|\beta\|_\infty \le \|\sigma_x \sigma_y\|_\infty < \infty, \tag{2.13}$$

$$\|A\|_2 = \max\{\|\sigma_x\|_\infty, \|\sigma_y\|_\infty\} < \infty, \quad \|B\|_2 \le \sqrt{2}(\|\sigma_x\|_\infty + \|\sigma_y\|_\infty) < \infty, \tag{2.14}$$

from the setting of $c(x, y) = 1$ in the PML region.

Now we look for a weak solution of (2.9) in the sense that

$$u \in L^2(0, T; H_0^1(\Omega)), \ \vec{\mathbf{q}} \in L^2(0, T; \mathbb{L}^2(\Omega)), \tag{2.15}$$

with

$$u' \in L^2(0, T; L^2(\Omega)), \ u'' \in L^2(0, T; H^{-1}(\Omega)), \ \vec{\mathbf{q}}' \in L^2(0, T; \mathbb{L}^2(\Omega)), \tag{2.16}$$

which satisfies

$$\begin{cases} < \frac{1}{c^2} u'', w > + (\alpha u', w) + (\beta u, w) - (\delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}, w) + (\nabla u, \nabla w) \ = \ 0, \\ (\vec{\mathbf{q}}', \vec{\mathbf{v}}) \ + \ (A\vec{\mathbf{q}}, \vec{\mathbf{v}}) \ + \ (B\nabla u, \vec{\mathbf{v}}) \ = \ 0, \end{cases} \tag{2.17}$$

for each $w \in H_0^1(\Omega), \vec{\mathbf{v}} \in \mathbb{L}^2(\Omega)$ and *a.e.* time $0 \le t \le T$, and the initial data in a weak sense, i.e.,

$$(u(0), w) = (u^0, w), \quad < u'(0), w >= (u^1, w), \ \text{ and } (\vec{\mathbf{q}}(0), \vec{\mathbf{v}}) = (\vec{\mathbf{q}}^0, \vec{\mathbf{v}}) \tag{2.18}$$

for each $w \in H_0^1(\Omega), \vec{\mathbf{v}} \in \mathbb{L}^2(\Omega)$. Here, $< \cdot, \cdot >$ denotes the duality pairing between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$, $(\cdot, \cdot)$ is the inner product in $L^2(\Omega)$, and also time derivatives are understood in a distributional sense here.

Remark 2.2 *We see that* $u \in C([0,T]; L^2(\Omega))$, $u' \in C([0,T]; H^{-1}(\Omega))$, *and* $\vec{\mathbf{q}} \in C([0,T]; \mathbb{L}^2(\Omega))$. *For the details, see Theorem 2, Chapter 5.9.2* [25]. *Consequently the equalities in* (2.17), (2.18) *make sense.*

We use the standard Galerkin method constructing a finite approximate solutions and establish bounds on certain terms in order to extend to a solution in the given space.

2.3.1 Galerkin Approximations.

We employ the Galerkin method to construct a weak solution (2.15).

Let $\{w_j | j \in \mathbb{N}\}$ be an c^{-2}-weighted orthonormal basis in $L^2(\Omega)$, i.e., $(c^{-2} w_j, w_k) = \delta_{jk}$, where Kronecker delta is given by $\delta_{jk} = \begin{cases} 0, & \text{if } j \neq k \\ 1, & \text{if } j = k, \end{cases}$ of eigenfunctions of the eigenvalue problem

$$\begin{cases} c^2 \Delta w = \lambda w & \text{in } \Omega, \\ w = 0 & \text{on } \partial\Omega. \end{cases}$$

and denote $\mathcal{U}_k$ the space generated by $\{w_1, w_2, \cdots, w_k\}$ in $L^2(\Omega)$. Then we have that $\mathcal{U}_k$ is also c^{-2}-weighted orthogonal basis of $H_0^1(\Omega)$ i.e.,

$$(c^{-2} w_j, w_k) + (\nabla w_j, \nabla w_k) = 0, \text{ if } j \neq k.$$

Let also denote $\mathcal{Q}_k$ the space generated by smooth functions $\{\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \cdots, \vec{\mathbf{v}}_k\}$, such that $\{\vec{\mathbf{v}}_k, k \in \mathbb{N}\}$ is an orthonormal basis of $\mathbb{L}^2(\Omega)$.

We construct approximate solutions $(u_k, \vec{\mathbf{q}}_k), k = 1, 2, 3, \cdots$, in the form

$$u_k(t) = \sum_{j=1}^{k} g_j^k(t) w_j, \tag{2.19}$$

$$\vec{\mathbf{q}}_k(t) = \sum_{j=1}^{k} h_j^k(t)\vec{\mathbf{v}}_j, \tag{2.20}$$

where the coefficients $g_j^k(t),\ h_j^k(t)$ for $0 \le t \le T, j = 1, 2, \cdots, k$ satisfy

$$g_j^k(0) = (u^0, w_j), \tag{2.21}$$

$$g_j^{k\prime}(0) = (u^1, w_j), \tag{2.22}$$

$$\vec{\mathbf{q}}_j^k(0) = (\vec{\mathbf{q}}^0, \vec{\mathbf{v}}_j), \tag{2.23}$$

and

$$\begin{cases} (\frac{1}{c^2}u_k'', w_j) + (\alpha u_k' + \beta u_k - \delta_\varepsilon \nabla\cdot\vec{\mathbf{q}}_k, w_j) + (\nabla u_k, \nabla w_j) &= 0, \\ (\vec{\mathbf{q}}_k', \vec{\mathbf{v}}_j) \ + \ (A\vec{\mathbf{q}}_k, \vec{\mathbf{v}}_j) \ + \ (B\nabla u_k, \vec{\mathbf{v}}_j) &= 0, \end{cases} \tag{2.24}$$

for all $w_j \in \mathcal{U}_k,\ \vec{\mathbf{v}}_j \in \mathcal{Q}_k, j = 1, \cdots, k$. For each integer $k = 1, 2, \cdots$, the standard theory of ordinary differential equations guarantees that the system (2.24) has a solution $(u_k(t),\ \vec{\mathbf{q}}_k(t))$ for $0 \le t \le T$.

2.3.2 Energy Estimates.

We have some estimates uniform in k, which allows to send $k \to \infty$.

Theorem 2.1 *There exists a constant C_T, depending only on σ_x, σ_y, Ω, and T such that*

$$\max_{0\le t\le T}\left(\|\frac{1}{c}u_k'(t)\|_{L^2(\Omega)} + \|\nabla u_k(t)\|_{L^2(\Omega)} + \|\vec{\mathbf{q}}_k(t)\|_{L^2(\Omega)}\right) + \|u_k''\|_{L^2(0,T;H^{-1}(\Omega))} + \|\vec{\mathbf{q}}_k'\|_{L^2(0,T;\mathbb{L}^2(\Omega))}$$
$$\le C_T\left(\|u^0\|_{H_0^1(\Omega)}^2 + \|u^1\|_{L^2(\Omega)}^2 + \|\vec{\mathbf{q}}^0\|_{\mathbb{L}^2(\Omega)}^2\right), \tag{2.25}$$

for all $k = 1, 2, \cdots$

Proof. 1. Let us define the approximate energy by

$$E_k(t) = \|\frac{1}{c}u_k'(t)\|_{L^2(\Omega)}^2 + \|\nabla u_k(t)\|_{L^2(\Omega)}^2 + \|\vec{\mathbf{q}}_k(t)\|_{L^2(\Omega)}^2.$$

Then we apply $(g_j^k)'(t)$ and $h_j^k(t)$ in the first and second equation in (2.24), respectively, sum $j = 1, \cdots, k$ and recall (2.19), (2.20) to obtain

$$\begin{cases} (\frac{1}{c^2}u_k'', u_k') + (\alpha u_k' + \beta u_k - \delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}_k, u_k') + (\nabla u_k, \nabla u_k') &= 0, \\ (\vec{\mathbf{q}}_k', \vec{\mathbf{q}}_k) \quad + \quad (A\vec{\mathbf{q}}_k, \vec{\mathbf{q}}_k) \quad + \quad (B\nabla u_k, \vec{\mathbf{q}}_k) &= 0, \end{cases} \tag{2.26}$$

for a.e. $0 \le t \le T$. Note that $(\frac{1}{c^2}u_k'', u_k') = \frac{d}{dt}\left(\frac{1}{2}\|\frac{1}{c}u_k'\|^2_{L^2(\Omega)}\right)$. Combining two equations, we obtain

$$\frac{1}{2}\frac{d}{dt}E_k + F_k^1 + F_k^2 = 0,$$

where

$$F_k^1 = (\alpha u_k', u_k') + (\beta u_k, u_k') - (\delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}_k, u_k'),$$
$$F_k^2 = (A\vec{\mathbf{q}}_k, \vec{\mathbf{q}}_k) + (B\nabla u_k, \vec{\mathbf{q}}_k).$$

Since the operator $\varphi \longmapsto \delta_\varepsilon(\varphi)$ is continuous from $H^{-1}(\Omega) \longrightarrow L^2(\Omega)$,

$$(\delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}_k, u_k') \le [\delta_\varepsilon]\, \|\nabla \cdot \vec{\mathbf{q}}_k\|_{H^{-1}(\Omega)} \|u_k'\|_{L^2(\Omega)}, \tag{2.27}$$

where $[\delta_\varepsilon] = C_{\delta_\varepsilon}$ is the norm of δ_ε in $\mathcal{L}(H^{-1}(\Omega); L^2(\Omega))$ in Remark 2.1. With Hölder's inequality and the assumption for σ_x, σ_y we estimate F_k^1 as following,

$$|F_k^1| \le \|\alpha\|_\infty \|u_k'\|^2_{L^2(\Omega)} + \frac{1}{2}\|\beta\|_\infty (\|u_k\|^2_{L^2(\Omega)} + \|u_k'\|^2_{L^2(\Omega)}) + [\delta_\epsilon]\, \|\nabla \cdot \vec{\mathbf{q}}_k\|_{H^{-1}(\Omega)} \|u_k'\|_{L^2(\Omega)}.$$

From the property $\|\nabla \cdot \vec{\mathbf{q}}_k\|_{H^{-1}(\Omega)} \le c_0 \|\vec{\mathbf{q}}_k\|_{L^2(\Omega)}$ for some $c_0 > 0$ by the embedding and the Poincaré inequality $\|u_k\|^2_{L^2(\Omega)} \le c_p \|\nabla u_k\|^2_{L^2(\Omega)}$ for some constant $c_p > 0$ when $u_k \in H_0^1(\Omega)$, it follows that there is a constant $c_1 > 0$ such that

$$|F_k^1| \le c_1 E_k.$$

Clearly $|F_k^2| \le \|A\|_2 \|\vec{\mathbf{q}}_k\|^2_{L^2(\Omega)} + \|B\|_2 \|\nabla u_k\|_{L^2(\Omega)} \|\vec{\mathbf{q}}_k\|_{L^2(\Omega)} \le c_2 E_k$, for some constant $c_2 > 0$ by (2.14).

Using the above estimates, $E_k(t)$ satisfies

$$\frac{dE_k}{dt} \le C_k E_k,$$

for a suitable constant $C_k = \max\{c_1, c_2\} > 0$.

Furthermore, Gronwall's inequality yields the estimate

$$E_k(t) \leq E_k(0)e^{C_{k_T}} \leq C_{k_T}\left(\|u^0\|^2_{H_0^1(\Omega)} + \|u^1\|^2_{L^2(\Omega)} + \|\vec{\mathbf{q}}^0\|^2_{\mathbb{L}^2(\Omega)}\right), \tag{2.28}$$

for all $k \in \mathbb{N}$. Since $0 \leq t \leq T$ is arbitrary, we see from this estimate, the Poincaré inequality, and (2.2), that

$$\begin{aligned}&\max_{0\leq t\leq T}\left(\|u_k(t)\|^2_{H_0^1(\Omega)} + \|u_k'(t)\|^2_{L^2(\Omega)} + \|\vec{\mathbf{q}}_k\|_{\mathbb{L}^2(\Omega)}\right)\\ &\leq C\left(\|u^0\|^2_{H_0^1(\Omega)} + \|u^1\|^2_{L^2(\Omega)} + \|\vec{\mathbf{q}}^0\|^2_{\mathbb{L}^2(\Omega)}\right)\end{aligned}$$

for some $C > 0$.

2. Fix any $w \in H_0^1(\Omega), \|w\|_{H_0^1(\Omega)} \leq 1$, and $\vec{\mathbf{v}} \in \mathbb{L}^2(\Omega), \|\vec{\mathbf{v}}\|_{\mathbb{L}^2(\Omega)} \leq 1$, and write $w = w^1 + w^2$ and $\vec{\mathbf{v}} = \vec{\mathbf{v}}^1 + \vec{\mathbf{v}}^2$, where

$$w^1 \in span\{w_j\}_{j=1}^k, \quad (\frac{1}{c^2}w^2, w_j) = 0 \ \ (j = 1, \cdots, k),$$

and

$$\vec{\mathbf{v}}^1 \in span\{\vec{\mathbf{v}}_j\}_{j=1}^k, \qquad (\vec{\mathbf{v}}^2, \vec{\mathbf{v}}_j) = 0 \ \ (j = 1, \cdots, k).$$

Note that $\|w^1\|_{H_0^1(\Omega)} \leq 1$ and $\|\vec{\mathbf{v}}^1\|_{\mathbb{L}^2(\Omega)} \leq 1$.

From (2.19), (2.20), and (2.24) we have

$$\begin{aligned}< \frac{1}{c^2}u_k'', w > \ &= \ (\frac{1}{c^2}u_k'', w) \ = \ (\frac{1}{c^2}u_k'', w^1)\\ &= \ -(\alpha u_k' + \beta u_k, w^1) - (\delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}_k, w^1) + (\nabla u_k, \nabla w^1),\\ (\vec{\mathbf{q}}_k', \vec{\mathbf{v}}) \ &= \ (\vec{\mathbf{q}}_k', \vec{\mathbf{v}}^1) \ = \ -(A\vec{\mathbf{q}}_k, \vec{\mathbf{v}}^1) - (B\nabla u_k, \vec{\mathbf{v}}^1).\end{aligned}$$

Thus we have that

$$| < u_k'', w > | + |(\vec{\mathbf{q}}_k', \vec{\mathbf{v}})| \leq C\left(\|u_k\|_{H_0^1(\Omega)} + \|u_k'\|_{L^2(\Omega)} + \|\vec{\mathbf{q}}_k\|_{\mathbb{L}^2(\Omega)}\right).$$

Consequently we obtain

$$\int_0^T \left(\|u_k''\|_{H^{-1}(\Omega)} + \|\vec{\mathbf{q}}'\|_{\mathbb{L}^2(\Omega)}\right) dt \le C \int_0^T \left(\|u_k\|^2_{H_0^1(\Omega)} + \|u_k'\|^2_{L^2(\Omega)} + \|\vec{\mathbf{q}}_k\|^2_{\mathbb{L}^2(\Omega)}\right) dt$$
$$\le C_T \left(\|u^0\|^2_{H_0^1(\Omega)} + \|u^1\|^2_{L^2(\Omega)} + \|\vec{\mathbf{q}}^0\|^2_{\mathbb{L}^2(\Omega)}\right).$$

□

2.3.3 Existence and Uniqueness.

Now we pass to limits in the Galerkin approximations.

Theorem 2.2 (Existence of weak solution) *Assume the initial data* $(u^0, u^1, \vec{\mathbf{q}}^0)$ *are in* $H_0^1(\Omega) \times L^2(\Omega) \times \mathbb{L}^2(\Omega)$. *Then the system* (2.17) *has a unique weak solution, provided* $\sigma_x, \sigma_y \in L^\infty(\Omega)$.

Proof. 1. From the energy estimates (2.1), we see that

$$\begin{cases} \{u_k\}_{k=1}^\infty \text{ is bounded in } L^2(0,T;H_0^1(\Omega)), \\ \{u_k'\}_{k=1}^\infty \text{ is bounded in } L^2(0,T;L^2(\Omega)), \\ \{u_k''\}_{k=1}^\infty \text{ is bounded in } L^2(0,T;H^{-1}(\Omega)), \\ \{\vec{\mathbf{q}}_k\}_{k=1}^\infty \text{ is bounded in } L^2(0,T;\mathbb{L}^2(\Omega)), \\ \{\vec{\mathbf{q}}_k'\}_{k=1}^\infty \text{ is bounded in } L^2(0,T;\mathbb{L}^2(\Omega)). \end{cases} \tag{2.29}$$

As a consequence there exist subsequences $\{u_{k_m}\} \subset \{u_k\}_{k=1}^\infty$, $\{\vec{\mathbf{q}}_{k_m}\} \subset \{\vec{\mathbf{q}}_k\}_{k=1}^\infty$ and $u \in L^2(0,T;H_0^1(\Omega))$, $\vec{\mathbf{q}} \in L^2(0,T;\mathbb{L}^2(\Omega))$ with $u' \in L^2(0,T;L^2(\Omega))$, $u'' \in L^2(0,T;H^{-1}(\Omega))$, $\vec{\mathbf{q}}' \in L^2(0,T;\mathbb{L}^2(\Omega))$, such that

$$\begin{cases} u_{k_m} \rightharpoonup u \text{ weakly in } L^2(0,T;H_0^1(\Omega)), \\ u_{k_m}' \rightharpoonup u' \text{ weakly in } L^2(0,T;L^2(\Omega)), \\ u_{k_m}'' \rightharpoonup u'' \text{ weakly in } L^2(0,T;H^{-1}(\Omega)), \\ \vec{\mathbf{q}}_{k_m} \rightharpoonup \vec{\mathbf{q}} \text{ weakly in } L^2(0,T;\mathbb{L}^2(\Omega)), \\ \vec{\mathbf{q}}_{k_m}' \rightharpoonup \vec{\mathbf{q}}' \text{ weakly in } L^2(0,T;\mathbb{L}^2(\Omega)), \end{cases} \tag{2.30}$$

since $\frac{d}{dt}$ is continuous.

2. Next fix an integer N and choose functions $w \in C^1(0,T;H_0^1(\Omega))$ and $\vec{\mathbf{v}} \in C^0(0,T;\mathbb{L}^2(\Omega))$ of the forms

$$w(t) = \sum_{j=1}^{N} g^j(t) w_j, \quad \vec{\mathbf{v}}(t) = \sum_{j=1}^{N} h^j(t) \vec{\mathbf{v}}_j, \tag{2.31}$$

where $\{g^j(t)\}_{j=1}^N \subset C^1([0,T])$ and $\{h^j(t)\}_{j=1}^N \subset C^0([0,T])$. Take $k \geq N$, multiply (2.24) by $g^j(t), h^j(t)$, sum $j = 1, \cdots, N$, respectively, and then integrate with respect to t, to obtain

$$\begin{cases} \int_0^T < \frac{1}{c^2} u_k'', w > dt + \int_0^T (\alpha u_k' + \beta u_k - \delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}_k, w) dt + \int_0^T (\nabla u_k, \nabla w) dt & = 0, \\ \int_0^T (\vec{\mathbf{q}}_k', \vec{\mathbf{v}}) dt \quad + \quad \int_0^T (A\vec{\mathbf{q}}_k, \vec{\mathbf{v}}) dt \quad + \quad \int_0^T (B\nabla u_k, \vec{\mathbf{v}}) dt & = 0. \end{cases} \tag{2.32}$$

Note that $\nabla \cdot : \mathbb{L}^2(\Omega) \to H^{-1}(\Omega)$ is continuous and $\delta_\varepsilon : H^{-1}(\Omega) \to L^2(\Omega)$ is also continuous a.e. t $\in [0,T]$, thus we have $\delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}_k \rightharpoonup \delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}$ in $L^2(0,T;L^2(\Omega))$. Set $k = k_m$ and use (2.30) to find in the limit that

$$\begin{cases} \int_0^T < \frac{1}{c^2} u'', w > dt + \int_0^T (\alpha u' + \beta u - \delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}, w) dt + \int_0^T (\nabla u, \nabla w) dt & = 0, \\ \int_0^T (\vec{\mathbf{q}}', \vec{\mathbf{v}}) dt \quad + \quad \int_0^T (A\vec{\mathbf{q}}, \vec{\mathbf{v}}) dt \quad + \quad \int_0^T (B\nabla u, \vec{\mathbf{v}}) dt & = 0. \end{cases} \tag{2.33}$$

This equalities hold for all functions $w \in L^2(0,T;H_0^1(\Omega))$ and $\vec{\mathbf{v}} \in L^2(0,T;\mathbb{L}^2(\Omega))$, since functions of the form (2.31) are dense in these spaces respectively. Therefore it follows that from (2.33)

$$\begin{cases} < \frac{1}{c^2} u'', w > + (\alpha u' + \beta u - \delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}, w) + (\nabla u, \nabla w) & = 0, \\ (\vec{\mathbf{q}}', \vec{\mathbf{v}}) \quad + \quad (A\vec{\mathbf{q}}, \vec{\mathbf{v}}) \quad + \quad (B\nabla u, \vec{\mathbf{v}}) & = 0, \end{cases} \tag{2.34}$$

for all $w \in H_0^1(\Omega)$ and $\vec{\mathbf{w}} \in \mathbb{L}^2(\Omega)$ and a.e. $0 \leq t \leq T$.

Furthermore, $u \in C(0,T;L^2(\Omega)), u' \in C(0,T;H^{-1}(\Omega))$, and $\vec{\mathbf{q}} \in C(0,T;\mathbb{L}^2(\Omega))$.

3. We verify the initial conditions

$$u(0) = u^0, \quad u'(0) = u^1, \quad \text{and} \quad \vec{\mathbf{q}}(0) = \vec{\mathbf{q}}^0. \tag{2.35}$$

Choose any function $w \in C^2([0,T];H_0^1(\Omega))$ with $w(T) = w'(T) = 0$, and $\vec{\mathbf{v}} \in C^1([0,T];\mathbb{L}^2(\Omega))$. Then integrating by parts twice with respect to t in the first equation and once in the second in (2.32), we have

$$\int_0^T < \frac{1}{c^2}w'', u > dt \ + \ \int_0^T (\alpha u' + \beta u - \delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}, w)dt \ + \ \int_0^T (\nabla u, \nabla w)dt$$
$$= -(\frac{1}{c^2}u(0), w'(0)) + < \frac{1}{c^2}u'(0), w(0) >, \tag{2.36}$$
$$-\int_0^T (\vec{\mathbf{v}}', \vec{\mathbf{q}})dt \ + \ \int_0^T (A\vec{\mathbf{q}}, \vec{\mathbf{v}})dt + \int_0^T (B\nabla u, \vec{\mathbf{v}})dt \ = \ -(\vec{\mathbf{q}}(0), \vec{\mathbf{v}}(0)). \tag{2.37}$$

Similarly from (2.33) we deduce

$$\int_0^T < \frac{1}{c^2}w'', u_k > dt \ + \ \int_0^T (\alpha u_k' + \beta u_k - \delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}_k, w)dt \ + \ \int_0^T (\nabla u_k, \nabla w)dt$$
$$= -(\frac{1}{c^2}u_k(0), w'(0)) + < \frac{1}{c^2}u_k'(0), w(0) >,$$
$$-\int_0^T (\vec{\mathbf{v}}', \vec{\mathbf{q}}_k)dt \ + \ \int_0^T (A\vec{\mathbf{q}}_k, \vec{\mathbf{v}})dt \ + \ \int_0^T (B\nabla u_k, \vec{\mathbf{v}})dt \ = \ -(\vec{\mathbf{q}}_k(0), \vec{\mathbf{v}}(0)).$$

We set $k = k_m$ and recall (2.21), (2.22), (2.23), and (2.30), to obtain

$$\int_0^T < \frac{1}{c^2}w'', u > dt \ + \ \int_0^T (\alpha u' + \beta u - \delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}, w)dt \ + \ \int_0^T (\nabla u, \nabla w)dt$$
$$= -(\frac{1}{c^2}u^0, w'(0)) + < \frac{1}{c^2}u^1, w(0) >, \tag{2.38}$$
$$-\int_0^T (\vec{\mathbf{v}}', \vec{\mathbf{q}})dt \ + \ \int_0^T (A\vec{\mathbf{q}}, \vec{\mathbf{v}})dt \ + \ \int_0^T (B\nabla u, \vec{\mathbf{v}})dt \ = \ -(\vec{\mathbf{q}}^0, \vec{\mathbf{v}}(0)). \tag{2.39}$$

Comparing identities (2.38), (2.39), (2.36), (2.37), we conclude (2.35), since $w(0), w'(0)$, and $\vec{\mathbf{v}}(0)$ are arbitrary. Hence $(u, \vec{\mathbf{q}})$ is a weak solution of (2.9).

□

Theorem 2.3 (Uniqueness of weak solution) *A weak solution of* (2.9) *is unique.*

Proof. The idea of the proof is from [25] with the second order hyperbolic problems, but the auxiliary variable $\vec{\mathbf{q}}$ in the system needs to be handled carefully.

1. It suffices to show that the only weak solution of (2.9) with $u^0 \equiv u^1 \equiv 0, \vec{\mathbf{q}}^0 \equiv \vec{\mathbf{0}}$ is

$$u \equiv 0, \ \vec{\mathbf{q}} \equiv \vec{\mathbf{0}}.$$

To verify this, fix $0 \le s \le T$ and set

$$w(t) := \begin{cases} \int_t^s u(\tau)d\tau & \text{if } 0 \le t \le s, \\ 0 & \text{if } s \le t \le T. \end{cases}$$

Then $w(t) \in H_0^1(\Omega)$ for each $0 \le t \le T$, and we have applying $w(t)$ in the first equation in (2.17)

$$\int_0^s < \frac{1}{c^2}u'', w > dt + \int_0^s (\alpha u' + \beta u - \delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}, w)dt + \int_0^s (\nabla u, \nabla w)dt = 0.$$

Since $u(0) = u'(0) = 0$ and $w(s) = 0$, we obtain after integration by parts in the first and second term in the above equation:

$$-\int_0^s < \frac{1}{c^2}u', w' > dt - \int_0^s (\alpha u, w')dt + \int_0^s (\beta u - \delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}, w)dt + \int_0^s (\nabla u, \nabla w)dt = 0.$$

Now note $w' = -u$ for $0 \le t < s$, and so $\nabla w' = -\nabla u$, thus we have

$$\int_0^s < \frac{1}{c^2}u', u > dt + \int_0^s (\alpha u, u)dt + \int_0^s (\beta u - \delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}, w)dt - \int_0^s (\nabla w', \nabla w)dt = 0. \tag{2.40}$$

Applying $\vec{\mathbf{v}}(t) = \rho\vec{\mathbf{q}}(t)$ with $\rho > 0$ in the second equation in (2.17) and $w' = -u$ we have

$$\rho \int_0^s (\vec{\mathbf{q}}', \vec{\mathbf{q}})dt + \rho \int_0^s (A\vec{\mathbf{q}}, \vec{\mathbf{q}})dt - \rho \int_0^s (B\nabla w', \vec{\mathbf{q}})dt = 0.$$

Since $\vec{\mathbf{q}}(0) = \vec{\mathbf{0}}$ and $\nabla w(s) = \vec{\mathbf{0}}$, we also have, after integration by parts in the third term in the above equation

$$\rho \int_0^s (\vec{\mathbf{q}}', \vec{\mathbf{q}})dt + \rho \int_0^s (A\vec{\mathbf{q}}, \vec{\mathbf{q}})dt + \rho \int_0^s (B\nabla w, \vec{\mathbf{q}}')dt = 0. \tag{2.41}$$

Again we apply $-\rho B\nabla w$ in the second equation in (2.17) and $w' = -u$ we obtain

$$-\rho \int_0^s (\vec{\mathbf{q}}', B\nabla w)dt - \rho \int_0^s (A\vec{\mathbf{q}}, B\nabla w)dt + \rho \int_0^s (B\nabla w', B\nabla w)dt = 0. \tag{2.42}$$

Summation of the equations (2.40), (2.41), (2.42) gives that

$$\int_0^s \frac{d}{dt}\left(\frac{1}{2}\|\frac{1}{c}u\|^2_{L^2(\Omega)} - \frac{1}{2}\|\nabla w\|^2_{L^2(\Omega)} + \frac{\rho}{2}\|B\nabla w\|^2_{L^2(\Omega)} + \frac{\rho}{2}\|\vec{\mathbf{q}}\|^2_{\mathbb{L}^2(\Omega)}\right)dt$$

$$= \int_0^s \left(-(\alpha u, u) - (\beta u, w) + (\delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}, w) - \rho(A\vec{\mathbf{q}}, \vec{\mathbf{q}})dt + \rho(A\vec{\mathbf{q}}, B\nabla w)\right) dt.$$

Hence we have that

$$\frac{1}{2}\|\frac{1}{c}u(s)\|^2_{L^2(\Omega)} + \frac{1}{2}\|\nabla w(0)\|^2_{L^2(\Omega)} - \frac{\rho}{2}\|B\nabla w(0)\|^2_{L^2(\Omega)} + \frac{\rho}{2}\|\vec{\mathbf{q}}(s)\|^2_{\mathbb{L}^2(\Omega)}$$

$$= \int_0^s \left(-(\alpha u, u) - (\beta u, w) + (\delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}, w) - \rho(A\vec{\mathbf{q}}, \vec{\mathbf{q}}) + \rho(A\vec{\mathbf{q}}, B\nabla w)\right) dt.$$

Since $\rho\|B\nabla w(0)\|^2_{L^2(\Omega)} \leq \rho\|B\|_2^2\|\nabla w(0)\|^2_{L^2(\Omega)}$, we can take $\rho > 0$ with $\rho^{-1} \geq 2\|B\|_2^2$ in order to have that $\rho\|B\nabla w(0)\|^2_{L^2(\Omega)} \leq \frac{1}{2}\|\nabla w(0)\|^2_{L^2(\Omega)}$.

Using the bounds of α, β, A, B from (2.13), (2.14) in energy estimates and Poincaré inequality for $w \in H_0^1(\Omega)$ we have that, for some $C > 0$,

$$\begin{aligned} &\|u(s)\|^2_{L^2(\Omega)} + \|\nabla w(0)\|^2_{L^2(\Omega)} + \|\vec{\mathbf{q}}(s)\|^2_{\mathbb{L}^2(\Omega)} \\ &\leq C\int_0^s \left(\|u\|^2_{L^2(\Omega)} + \|\nabla w\|^2_{L^2(\Omega)} + \|\vec{\mathbf{q}}\|^2_{\mathbb{L}^2(\Omega)}\right) dt. \end{aligned} \tag{2.43}$$

2. Now let us write

$$v(t) := \int_0^t u(\tau)d\tau \quad (0 \leq t \leq T),$$

then it becomes, by (2.43)

$$\begin{aligned} &\|u(s)\|^2_{L^2(\Omega)} + \|\nabla v(s)\|^2_{L^2(\Omega)} + \|\vec{\mathbf{q}}(s)\|^2_{\mathbb{L}^2(\Omega)} \\ &\leq C\int_0^s \left(\|u\|^2_{L^2(\Omega)} + \|\nabla v(t) - \nabla v(s)\|^2_{L^2(\Omega)} + \|\vec{\mathbf{q}}\|^2_{\mathbb{L}^2(\Omega)}\right) dt. \end{aligned} \tag{2.44}$$

But $\|\nabla v(t) - \nabla v(s)\|^2_{L^2(\Omega)} \leq 2\|\nabla v(t)\|^2_{L^2(\Omega)} + 2\|\nabla v(s)\|^2_{L^2(\Omega)}$, and thus (2.44) implies

$$\|u(s)\|^2_{L^2(\Omega)} + (1 - 2sC)\|\nabla v(s)\|^2_{L^2(\Omega)} + \|\vec{\mathbf{q}}(s)\|^2_{\mathbb{L}^2(\Omega)}$$

$$\leq C\int_0^s \left(\|u\|^2_{L^2(\Omega)} + 2\|\nabla v\|^2_{L^2(\Omega)} + \|\vec{\mathbf{q}}\|^2_{\mathbb{L}^2(\Omega)}\right) dt.$$

Take T_1 small enough in order to get

$$1 - 2T_1C \geq \frac{1}{2}.$$

Then if $0 \leq s \leq T_1$, we have

$$\|u(s)\|^2_{L^2(\Omega)}+\|\nabla v(s)\|^2_{L^2(\Omega)}+\|\vec{\mathbf{q}}(s)\|^2_{\mathbb{L}^2(\Omega)} \leq C\int_0^s \left(\|u\|^2_{L^2(\Omega)} + \|\nabla v\|^2_{L^2(\Omega)} + \|\vec{\mathbf{q}}\|^2_{\mathbb{L}^2(\Omega)}\right) dt.$$

Finally the integral form of Gronwall's inequality implies $u \equiv 0, \vec{\mathbf{q}} \equiv \vec{\mathbf{0}}$ on $[0, T_1]$. Repeat the same argument on the intervals $[kT_1, (k+1)T_1], k = 1, 2, \cdots$ until $u \equiv 0, \vec{\mathbf{q}} \equiv \vec{\mathbf{0}}$ on $[0, T]$, which gives the proof of the uniqueness of a weak solution.

□

2.4. Numerical Results

In this section, we present some numerical results to illustrate the theory presented above.

2.4.1 Numerical Scheme.

For numerical examples, we use a family of finite difference schemes using the half-step staggered grids in space and time. All spatial derivatives are defined with the centered finite differences over 2 or 3 cells, which guarantees a second order approximation in space [40]. For the time discretization we also use the centered finite differences for the first and second order time derivatives on a uniform mesh which is also second order accurate in time. We denote the time step by $\triangle t > 0$ and the spatial mesh step sizes in the x and y directions by $\triangle x > 0$ and $\triangle y > 0$ respectively. Now we define the time level $t^n = n\triangle t$, and spatial nodes $x_\ell = \ell\triangle x$ and $y_j = j\triangle y$ for $n, \ell, j \in \mathbb{N} \cup \{0\}$.

We also define staggered nodes in the time direction and the x and y direction, respectively, as $t^{n\pm\frac{1}{2}} = t^n \pm \frac{1}{2}\triangle t, x_{\ell\pm\frac{1}{2}} = x_\ell \pm \frac{1}{2}\triangle x$, and $y_{j\pm\frac{1}{2}} = y_j \pm \frac{1}{2}\triangle y$ for $n, \ell, j \in \mathbb{N}$ (Figure 2.1). The components of u are discretized at nodes (t^n, x_ℓ, y_j) as $u^n_{\ell,j}$, whereas the components of $\vec{\mathbf{q}} = (q_x, q_y)$ are discretized at $(t^{n+\frac{1}{2}}, x_{\ell+\frac{1}{2}}, y_{j+\frac{1}{2}})$ as $q^{n+\frac{1}{2}}_{\alpha\ell+\frac{1}{2},j+\frac{1}{2}}$ for $\alpha = x, y$. Let us now introduce new notations

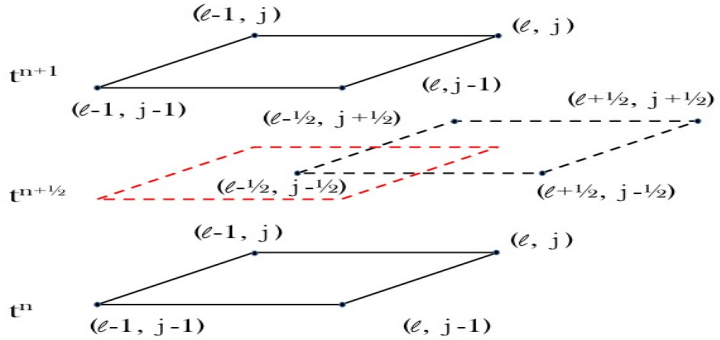

Figure 2.1: Notations for time and spatial grids discretization

$$\mathtt{A}^{x\pm}{}_{\ell+\frac{1}{2}} = 1 \pm \frac{\triangle t}{2}\sigma_{x\ell+\frac{1}{2}}, \quad \mathtt{A}^{y\pm}{}_{j+\frac{1}{2}} = 1 \pm \frac{\triangle t}{2}\sigma_{yj+\frac{1}{2}},$$

and

$$\mathtt{A}^{xy\pm}_{\ell,j} = 1 \pm \frac{\triangle t}{2}(\sigma_{x\ell} + \sigma_{yj}), \quad \sigma_{\alpha_k} = \sigma_\alpha(\alpha_k), \quad \sigma_{\alpha k+\frac{1}{2}} = \sigma_\alpha(\alpha_{k+\frac{1}{2}}), k = \ell, j, \alpha = x, y.$$

Step 1. Compute $\left(q^{n+\frac{1}{2}}_{x\ell+\frac{1}{2},j+\frac{1}{2}}, q^{n+\frac{1}{2}}_{y\ell+\frac{1}{2},j+\frac{1}{2}}\right)$,

$$\mathtt{A}^{x+}_{\ell+\frac{1}{2}} q^{n+\frac{1}{2}}_{x\ell+\frac{1}{2},j+\frac{1}{2}} = \mathtt{A}^{x-}_{\ell+\frac{1}{2}} q^{n-\frac{1}{2}}_{x\ell+\frac{1}{2},j+\frac{1}{2}} - \triangle t(\sigma_{x\ell+\frac{1}{2}} - \sigma_{yj+\frac{1}{2}})\tilde{\partial}_x u^n_{\ell+\frac{1}{2},j+\frac{1}{2}}, \tag{2.45}$$

$$\mathtt{A}^{y+}_{j+\frac{1}{2}} q^{n+\frac{1}{2}}_{y\ell+\frac{1}{2},j+\frac{1}{2}} = \mathtt{A}^{y-}_{j+\frac{1}{2}} q^{n-\frac{1}{2}}_{y\ell+\frac{1}{2},j+\frac{1}{2}} - \triangle t(\sigma_{yj+\frac{1}{2}} - \sigma_{x\ell+\frac{1}{2}})\tilde{\partial}_y u^n_{\ell+\frac{1}{2},j+\frac{1}{2}}, \tag{2.46}$$

where the cell averages of the derivatives of function $u^n_{i+\frac{1}{2},j+\frac{1}{2}}$ are defined as

$$\tilde{\partial}_x u^n_{\ell+\frac{1}{2},j+\frac{1}{2}} = \frac{u^n_{\ell+1,j+1} - u^n_{\ell,j+1} + u^n_{\ell+1,j} - u^n_{\ell,j}}{2\triangle x},$$

$$\tilde{\partial}_y u^n_{\ell+\frac{1}{2},j+\frac{1}{2}} = \frac{u^n_{\ell+1,j+1} - u^n_{\ell+1,j} + u^n_{\ell,j+1} - u^n_{\ell,j}}{2\triangle y}.$$

This allows to compute the regularized term in (2.12)

$$(\delta_\varepsilon \partial_x q_x)^n_{\ell,j}, \quad (\delta_\varepsilon \partial_y q_y)^n_{\ell,j},$$

for $\partial_x q^n_{x\ell,j} = \frac{1}{2}\left(\tilde{\partial}_x q^{n+\frac{1}{2}}_{x\ell,j} + \tilde{\partial}_x q^{n-\frac{1}{2}}_{x\ell,j}\right)$ and $\partial_y q^n_{y\ell,j} = \frac{1}{2}\left(\tilde{\partial}_y q^{n+\frac{1}{2}}_{y\ell,j} + \tilde{\partial}_y q^{n-\frac{1}{2}}_{y\ell,j}\right)$, where the cell averages of the derivatives of function $(q^{n\pm\frac{1}{2}}_{x\ell,j}, q^{n\pm\frac{1}{2}}_{y\ell,j})$ are defined as

$$\tilde{\partial}_x q^{n\pm\frac{1}{2}}_{x\ell,j} = \frac{1}{2\triangle x}\left(q^{n\pm\frac{1}{2}}_{x\ell+\frac{1}{2},j+\frac{1}{2}} - q^{n\pm\frac{1}{2}}_{x\ell-\frac{1}{2},j+\frac{1}{2}} + q^{n\pm\frac{1}{2}}_{x\ell+\frac{1}{2},j-\frac{1}{2}} - q^{n\pm\frac{1}{2}}_{x\ell-\frac{1}{2},j-\frac{1}{2}}\right),$$

$$\tilde{\partial}_y q^{n\pm\frac{1}{2}}_{y\ell,j} = \frac{1}{2\triangle y}\left(q^{n\pm\frac{1}{2}}_{y\ell+\frac{1}{2},j+\frac{1}{2}} - q^{n\pm\frac{1}{2}}_{y\ell+\frac{1}{2},j-\frac{1}{2}} + q^{n\pm\frac{1}{2}}_{y\ell-\frac{1}{2},j+\frac{1}{2}} - q^{n\pm\frac{1}{2}}_{y\ell-\frac{1}{2},j-\frac{1}{2}}\right).$$

Step 2. Compute $u^{n+1}_{\ell,j}$,

$$\mathtt{A}^{xy+}_{\ell,j} u^{n+1}_{\ell,j} = 2u^n_{\ell,j} - \mathtt{A}^{xy-}_{\ell,j} u^{n-1}_{\ell,j} + \triangle t^2\left(-\sigma^{xy}_{\ell,j} u^n_{\ell,j} + c^2_{\ell,j}((\delta_\varepsilon \partial_x q_x)^n_{\ell,j} + (\delta_\varepsilon \partial_y q_y)^n_{\ell,j}) + c^2_{\ell,j}\triangle_n u^n_{\ell,j}\right), \tag{2.47}$$

where

$$\sigma^{xy}_{\ell,j} = \sigma_{x\ell}\sigma_{yj},\ c_{\ell,j} = c(x_\ell, y_j),$$
$$\triangle_n u^n_{\ell,j} = \frac{u^n_{\ell+1,j} - 2u^n_{\ell,j} + u^n_{\ell-1,j}}{\triangle x^2} + \frac{u^n_{\ell,j+1} - 2u^n_{\ell,j} + u^n_{\ell,j-1}}{\triangle y^2}.$$

2.4.2 Layer Parameters.

We now describe the damping and regularization used in the system (2.9) following [20]. In the absorbing layer, the choice of the damping functions can be constant, linear, or quadratic, etc. In our implementations, we use damping functions of the form;

$$\sigma_{x_k}(x_k) = \begin{cases} 0 & \text{for } |x_k| < a_k, \quad k = 1, 2, \\ \bar{\sigma}_0\left(\frac{|x_k - a_k|}{L_k} - \frac{\sin(\frac{2\pi|x_k - a_k|}{L_k})}{2\pi}\right) & \text{for } a_k \le |x_k| \le a_k + L_k, k = 1, 2, \end{cases} \tag{2.48}$$

where $L_k, k = 1, 2$, are thickness of PML layers. The smooth function $\rho_\epsilon(x, y)$ chosen in the following examples is constant on a rectangle centered at zero,

$$\rho_\epsilon(x, y) = \rho_{\epsilon 1}(x)\rho_{\epsilon 2}(y) \text{ with } \rho_{\epsilon_k}(\xi) = \begin{cases} \frac{1}{\epsilon_k} & \text{if } \xi \in [-\frac{\epsilon_k}{2}, \frac{\epsilon_k}{2}], \\ 0 & \text{elsewhere.} \end{cases}$$

Given a 2-D finite difference grid with space steps $\triangle x$ and $\triangle y$, a possible choice is $\epsilon_1 = n_x\triangle x$ and $\epsilon_2 = n_y\triangle y$ with $n_x, n_y \in \mathbb{N}$. For instance, with $n_x = n_y = 1$ and usual integration formulas, we discretize the regularized term $\delta_\varepsilon(v)_{\ell,j} := (\rho_\epsilon * v)_{\ell,j}$ a discretization of the convolution product of ρ_ϵ by a function v given by

$$\begin{aligned}(\rho_\epsilon * v)_{\ell,j} =& \frac{1}{16}\left(4v_{\ell,j} + 2v_{\ell+1,j} + 2v_{\ell-1,j} + 2v_{\ell,j+1} + 2v_{\ell,j-1}\right. \\ &\left. + v_{\ell+1,j+1} + v_{\ell-1,j+1} + v_{\ell+1,j-1} + v_{\ell-1,j-1}\right).\end{aligned} \tag{2.49}$$

Remark 2.3 *We impose the zero Dirichlet boundary condition on* u^n. *The choice of the function* ρ_ϵ *in* (2.49) *can be considered as a way of discretizing the identity operator.*

Now we introduce some stability results of the scheme:

2.4.3 Stability Analysis for the Scheme

In this section, we use standard *von Neumann* stability analysis technique to show the stability of the scheme (2.45), (2.46), (2.47) under additional assumptions. We assume that σ_x and σ_y are constants and $\sigma_x = \sigma_y = \sigma_\alpha \geq 0$ in this section.

First we have the stability of the scheme in the computational domain.

Remark 2.4 *The CFL condition of the scheme* (2.45)-(2.47) *in the computation area* $(i.e., \sigma_x = \sigma_y = 0)$ *is*

$$c\frac{\triangle t}{h} \leq \frac{1}{\sqrt{2}},$$

for $\triangle x = \triangle y = h$ *from the standard von Neumann stability analysis technique.*

Next we have the stability result of the scheme with the assumption.

Theorem 2.4 *Let assume that* $\sigma_x = \sigma_y$ *and* c *are constants. The discrete scheme* (2.45)-(2.47) *is stable if it satisfies the CFL condition*

$$c\triangle t \leq \frac{h}{\sqrt{2}}\frac{1}{(1+\frac{\sigma_\alpha^2 h^2}{8c^2})^{1/2}}. \tag{2.50}$$

We define a simple *von Neumann* polynomial and introduce Theorem 2.6 to show the CFL condition.

Definition 2.2 *A polynomial is a simple von Neumann polynomial if all its roots,* r, *lie on the unit disk* $(|B(0,r)| < 1)$ *and its roots on the unit circle are simple roots.*

There is a sufficient stability condition.

Theorem 2.5 [8] *A sufficient stability condition is that* ϕ *be a simple von Neumann polynomial, where* ϕ *be the characteristic polynomial. (see* [8] *for the proof)*

Theorem 2.6 *Let ϕ be a polynomial of degree p written as*

$$\phi(z) = \mathsf{c}_0 + \mathsf{c}_1 z + \cdots + \mathsf{c}_p z^p,$$

where $\mathsf{c}_0, \mathsf{c}_1, \cdots, \mathsf{c}_p \in \mathbb{C}$ and $\mathsf{c}_p \neq 0$. The polynomial ϕ is a simple von Neumann polynomial if and only if ϕ^0 is a simple von Neumann polynomial and $|\phi(0)| \leq |\bar{\phi}(0)|$, where ϕ^0 is defined as

$$\phi^0(z) = \frac{\bar{\phi}(0)\phi(z) - \phi(0)\bar{\phi}(z)}{z},$$

and the conjugate polynomial $\bar{\phi}$ is defined as

$$\bar{\phi}(z) = \bar{\mathsf{c}}_p + \bar{\mathsf{c}}_{p-1} z + \cdots + \bar{\mathsf{c}}_0 z^p,$$

where $\bar{\mathsf{c}}$ is the complex conjugate of c. The main ingredient in the proof of the theorem is Rouché's theorem, the proof is in [23].

Proof of Theorem 2.4.

Assume that $\sigma_x = \sigma_y = \sigma_\alpha$ in the scheme (2.45)-(2.47) and we rewrite the scheme as the second order central difference scheme of the variable u and $\vec{\mathbf{q}}$.

$$\frac{u_{\ell,j}^{n+1} - 2u_{\ell,j}^{n} + u_{\ell,j}^{n-1}}{\triangle t^2} + 2\sigma_\alpha \frac{u_{\ell,j}^{n+1} - u_{\ell,j}^{n-1}}{2\triangle t} + \sigma_\alpha^2 u_{\ell,j}^{n} \tag{2.51}$$
$$= c^2 \left(\frac{u_{\ell+1,j}^{n} - 2u_{\ell,j}^{n} + u_{\ell-1,j}^{n}}{\triangle x^2} + \frac{u_{\ell,j+1}^{n} - 2u_{\ell,j}^{n} + u_{\ell,j-1}^{n}}{\triangle y^2} \right) + (\rho_\epsilon * \partial_x q_x)_{\ell,j}^{n} + (\rho_\epsilon * \partial_y q_y)_{\ell,j}^{n},$$

$$\frac{\vec{\mathbf{q}}_{\ell+\frac{1}{2},j+\frac{1}{2}}^{n+\frac{1}{2}} - \vec{\mathbf{q}}_{\ell+\frac{1}{2},j+\frac{1}{2}}^{n-\frac{1}{2}}}{\triangle t} + \sigma_\alpha \frac{\vec{\mathbf{q}}_{\ell+\frac{1}{2},j+\frac{1}{2}}^{n+\frac{1}{2}} + \vec{\mathbf{q}}_{\ell+\frac{1}{2},j+\frac{1}{2}}^{n-\frac{1}{2}}}{2} = \vec{\mathbf{0}}. \tag{2.52}$$

We assume a spatial dependence of the following form in the field quantities

$$u_{\ell,j}^{n+1} = \hat{u}^{n+1}(k_x, k_y) e^{ik_x x_\ell + ik_y y_j},$$
$$u_{\ell,j}^{n} = \hat{u}^{n}(k_x, k_y) e^{ik_x x_\ell + ik_y y_j},$$
$$\vec{\mathbf{q}}_{\ell+\frac{1}{2},j+\frac{1}{2}}^{n+\frac{1}{2}} = \hat{\vec{\mathbf{q}}}_{\ell+\frac{1}{2},j+\frac{1}{2}}^{n+\frac{1}{2}}(k_x, k_y) e^{ik_x x_{\ell+\frac{1}{2}} + ik_y y_{j+\frac{1}{2}}},$$

with k_x, k_y, the component of the wave vector $\vec{\mathbf{k}}$, i.e. $\vec{\mathbf{k}} = (k_x, k_y)^T$, and the wave number is $k = \sqrt{k_x^2 + k_y^2}$. Then we have the system $\left[\hat{u}^{n+1}, \hat{u}^n, \hat{\mathbf{q}}_x^{n+\frac{1}{2}}, \hat{\mathbf{q}}_y^{n+\frac{1}{2}}\right]^T = G\left[\hat{u}^n, \hat{u}^{n-1}, \hat{\mathbf{q}}_x^{n-\frac{1}{2}}, \hat{\mathbf{q}}_y^{n-\frac{1}{2}}\right]^T$, where the amplification matrix G of the scheme (2.51), (2.52) is given by

$$G = \begin{bmatrix} -\frac{\mathsf{c}_1}{\mathsf{c}_2} & -\frac{\mathsf{c}_0}{\mathsf{c}_2} & C_{\hat{q}_x} & C_{\hat{q}_y} \\ 1 & 0 & 0 & 0 \\ 0 & 0 & \eta & 0 \\ 0 & 0 & 0 & \eta \end{bmatrix},$$

where $C_{\hat{q}_x}$ and $C_{\hat{q}_y}$ satisfy $\mathsf{c}_2\hat{u}^{n+1} + \mathsf{c}_1\hat{u}^n + \mathsf{c}_0\hat{u}^{n-1} = C_{\hat{q}_x}\hat{\mathbf{q}}_x^{n-\frac{1}{2}} + C_{\hat{q}_y}\hat{\mathbf{q}}_y^{n-\frac{1}{2}}$ with $\mathsf{c}_0 = \frac{1}{\triangle t^2} - \frac{\sigma_\alpha}{\triangle t}$, $\mathsf{c}_1 = -\frac{2}{\triangle t^2} - 2c^2\frac{\cos(k_x\triangle x)-1}{\triangle x^2} - 2c^2\frac{\cos(k_y\triangle y)-1}{\triangle y^2} + \sigma_\alpha^2$, $\mathsf{c}_2 = \frac{1}{\triangle t^2} + \frac{\sigma_\alpha}{\triangle t}$, and $\eta = \frac{1-\frac{\triangle t}{2}\sigma_\alpha}{1+\frac{\triangle t}{2}\sigma_\alpha}$. Then the characteristic function of G is given by

$$\phi(G) = (G^2 + \frac{\mathsf{c}_1}{\mathsf{c}_2}G + \frac{\mathsf{c}_0}{\mathsf{c}_2})(G-\eta)^2.$$

Note that $|\eta| < 1$ by the assumption. From the Theorem 2.6 we have that $\phi(G)$ is a simple von Neumann polynomial if and only if $|\mathsf{c}_1| \le |\mathsf{c}_0 + \mathsf{c}_2|$, i.e.,

$$\left|\frac{2}{\triangle t^2} + 2c^2\frac{\cos(k_x h) + \cos(k_y h) - 2}{h^2} - \sigma_\alpha^2\right| \le \frac{2}{\triangle t^2}, \quad \text{for } h = \triangle x = \triangle y.$$

It is satisfied provided (2.50). □

2.4.4 Efficiency of the System

Here we compare the regularized system (2.9) with the original system (2.8) in the two dimensional space with variable sound speed $c(x, y)$ and the initial condition $u(0) = f, u_t(0) = 0$. The sound speed is taken randomly between $c(x, y) \in [0.5, 1.5]$ and smoothed by convolution, and the initial function is given by

$$f(x,y) = \begin{cases} Ce^{-C_0((x-x_0)^2+(y-y_0)^2} & \text{if } (x,y) \in [-a,a]\times[-a,a], \\ 0 & \text{otherwise.} \end{cases}$$

Here Ω is the square domain with $a = 0.5$, surrounded by a PML with $L = 0.1$ with $\triangle x = \triangle y = 0.01$. We use L^2-*error* in the computation area Ω, given by

$$E(t_n) = \sqrt{\frac{1}{N_p}\sum_{k=1}^{N_p}(u_k^n - \tilde{u}_k^n)^2}, \tag{2.53}$$

for N_p is the number of grid points. We compute a reference solution using the second

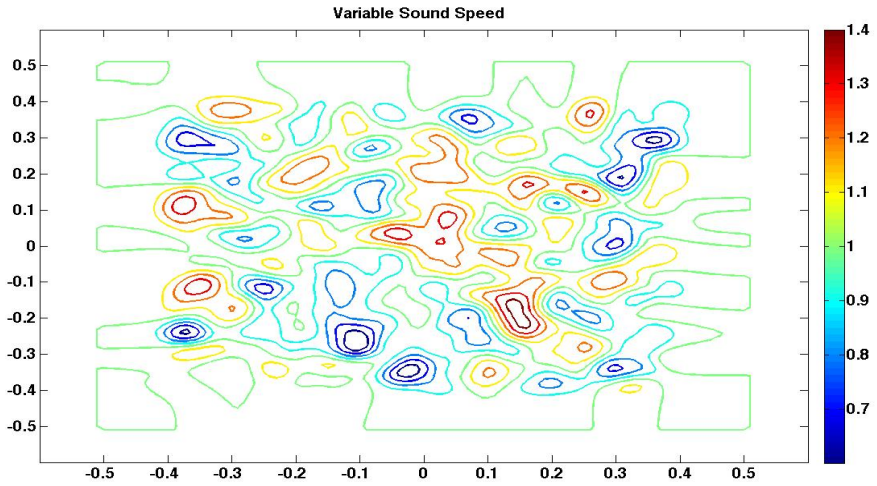

Figure 2.2: Variable Sound speed

order wave equation (2.1) in a much larger domain which doesn't give reflected waves in the chosen time interval. In Figure 2.4, the regularized system leads to slightly larger L^2-error than the one in the classical PML (2.8) when damping is bigger, and there is no large difference in error between the two systems when damping is relatively small. There is one suggestion which express the regularized term (2.12) for better accuracy of the scheme using another smooth function. (see for details; [20]) The maximum error in the computation area between two system is presented in Figure 2.5 with the different damping.

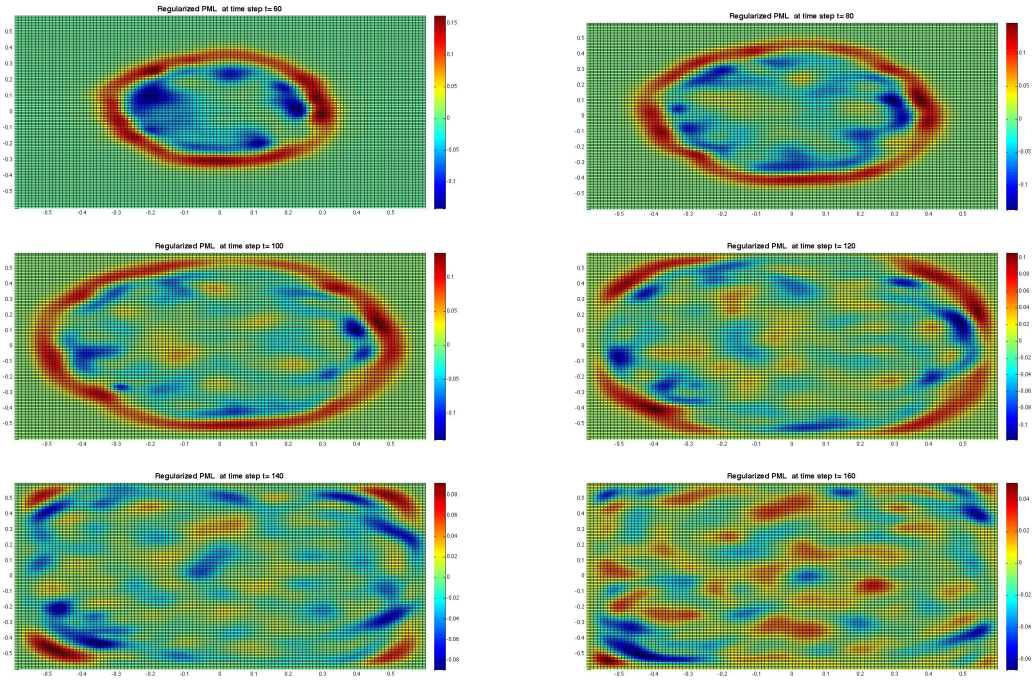

Figure 2.3: Acoustic wave with variable sound speed using regularized PML at time steps 60, 80, 100, 120, 140, 160 (see Appendix for larger figures)

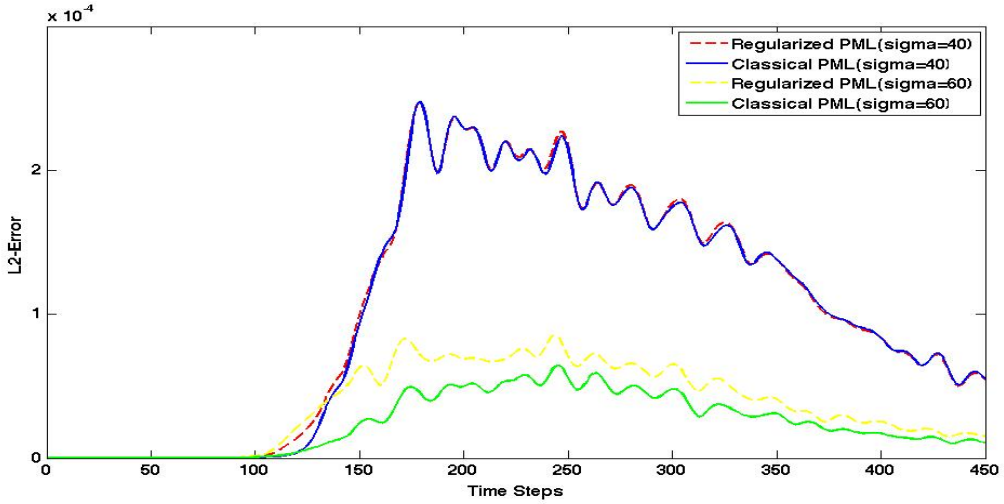

Figure 2.4: L^2-*error* in Computational Domain using $\bar{\sigma}_0 = 40$, $\bar{\sigma}_0 = 60$ in (2.48)

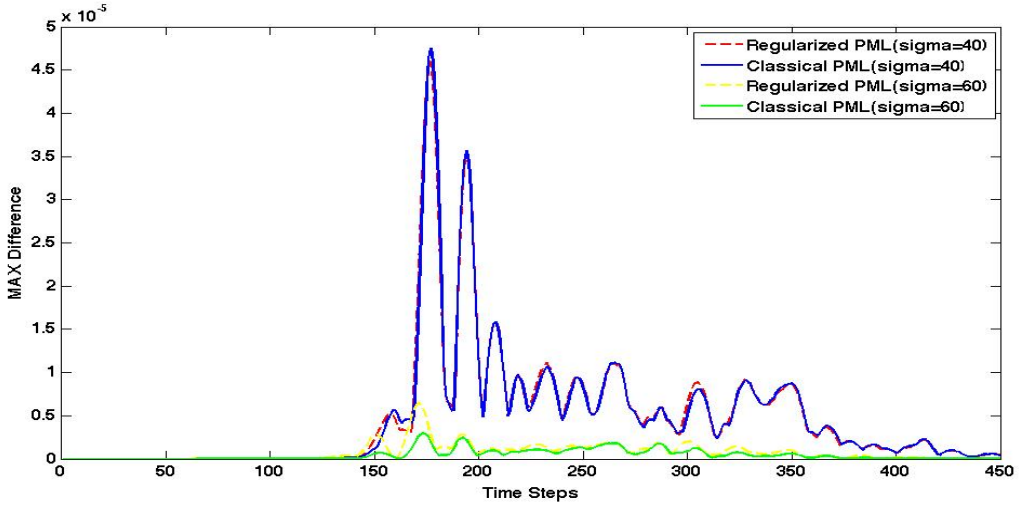

Figure 2.5: Maximum *error* in Computational Domain using $\bar{\sigma}_0 = 40$, $\bar{\sigma}_0 = 60$ in (2.48)

3. MULTI DIRECTIONAL PML

In the PML method, the damping (1.3) is introduced depending on only one variable. That there is no effective absorption for waves with low-grazing incidence angle (high incidence angle) to the interface is one of limitations mentioned in Chapter 1.4.

We introduce additional damping terms, σ_x^y and σ_y^x, which are positive functions of x and y. This is the idea to impose the absorption of incident wave in the parallel to the interface when the wave propagate in a PML. In the coordinate transformation (1.4) we can consider the damping terms which depend on both variables x, y in the PML region. Consider the domain $\Omega = [-a - L_x, a + L_x] \times [-b - L_y, b + L_y]$ as in the previous chapter.

Introduce the coordinate transform of variables in the frequency domain by the following,

$$\tilde{x}(x,y) := x + \frac{1}{iw}\left(\int_a^x \sigma_x(s)ds + \int_0^y \sigma_x^y(s)ds\right) := x + \frac{1}{iw}\sigma^x, \tag{3.1}$$

$$\tilde{y}(x,y) := y + \frac{1}{iw}\left(\int_b^y \sigma_y(s)ds + \int_0^x \sigma_y^x(s)ds\right) := y + \frac{1}{iw}\sigma^y, \tag{3.2}$$

where $\sigma_x^y(x,y)$ and $\sigma_y^x(x,y)$ are non-negative functions in a PML and vanish in the computation area $[-a,a] \times [-b,b]$. We assume that

$$\sigma_x^y(x,y), \sigma_y^x(x,y) \in W^{1,\infty}(\Omega). \tag{3.3}$$

Differentiate $\tilde{x}, \tilde{y}$ with respect to x, y to obtain Jacobian matrix,

$$J = \begin{bmatrix} \frac{\partial \tilde{x}}{\partial x} & \frac{\partial \tilde{x}}{\partial y} \\ \frac{\partial \tilde{y}}{\partial x} & \frac{\partial \tilde{y}}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 + \frac{1}{iw}\partial_x\sigma^x & \frac{1}{iw}\partial_y\sigma^x \\ \frac{1}{iw}\partial_x\sigma^y & 1 + \frac{1}{iw}\partial_y\sigma^y \end{bmatrix} = \begin{bmatrix} \frac{iw+\partial_x\sigma^x}{iw} & \frac{\partial_y\sigma^x}{iw} \\ \frac{\partial_x\sigma^y}{iw} & \frac{iw+\partial_y\sigma^y}{iw} \end{bmatrix},$$

which also gives the inverse of J that is,

$$J^{-1} = \frac{iw}{D}\begin{bmatrix} iw + \partial_y\sigma^y & -\partial_y\sigma^x \\ -\partial_x\sigma^y & iw + \partial_x\sigma^x \end{bmatrix},$$

where $D = (iw)^2 + (\partial_x\sigma^x + \partial_y\sigma^y)iw + \partial_x\sigma^x\partial_y\sigma^y - \partial_x\sigma^y\partial_y\sigma^x$. Then we have the partial derivatives of new coordinate systems,

$$\begin{cases} \frac{\partial}{\partial\tilde{x}} = \frac{iw(iw+\partial_y\sigma^y)}{D}\frac{\partial}{\partial x} - \frac{iw\partial_y\sigma^x}{D}\frac{\partial}{\partial y}, \\ \frac{\partial}{\partial\tilde{y}} = -\frac{iw\partial_x\sigma^y}{D}\frac{\partial}{\partial x} + \frac{iw(iw+\partial_x\sigma^x)}{D}\frac{\partial}{\partial y}. \end{cases} \tag{3.4}$$

We apply this new coordinate systems in the system of first order acoustic wave equation in section 3.1. and in the second order wave equation in section 3.2. to obtain different PML wave equations.

3.1. Multi Directional Un-Split PML

Consider the system of first order acoustic wave equation with variable sound speed,

$$\begin{cases} \frac{1}{c^2}p_t + \nabla\cdot\vec{\mathbf{q}} = 0, & \text{in } \mathbb{R}^2\times(0,T], \\ \vec{\mathbf{q}}_t \ + \ \nabla p \ = \vec{\mathbf{0}}, & \text{in } \mathbb{R}^2\times(0,T], \end{cases} \tag{3.5}$$

with the initial condition $p(x,0) = p_0, \vec{\mathbf{q}}(x,0) = \vec{\mathbf{q}}_0$ and bounds of sound speed $0 < c_* \leq c \leq c^* < \infty$. Then we apply the new coordinates system (3.4) in the frequency domain of the system (3.5), after the even extension of solutions over $\mathbb{R}$ and similar procedure in section 2.1., to obtain,

$$\begin{cases} \frac{D}{c^2}\hat{p} + (iw+\partial_y\sigma^y)\frac{\partial\hat{\mathbf{q}}_x}{\partial x} - \partial_y\sigma^x\frac{\partial\hat{\mathbf{q}}_x}{\partial y} - \partial_x\sigma^y\frac{\partial\hat{\mathbf{q}}_y}{\partial x} + (iw+\partial_x\sigma^x)\frac{\partial\hat{\mathbf{q}}_y}{\partial y} = 0, \\ D\hat{\vec{\mathbf{q}}} \ + \ \begin{bmatrix} iw+\partial_y\sigma^y & -\partial_y\sigma^x \\ -\partial_x\sigma^y & iw+\partial_x\sigma^x \end{bmatrix}\begin{bmatrix} \frac{\partial\hat{p}}{\partial x} \\ \frac{\partial\hat{p}}{\partial y} \end{bmatrix} = 0. \end{cases}$$

Similarly, auxiliary variables $\hat{p}^*$ and $\hat{\vec{\mathbf{q}}}^*$ are introduced by $\hat{p}^* = i\omega\hat{p}$ and $\hat{\vec{\mathbf{q}}}^* = i\omega\hat{\vec{\mathbf{q}}}$. Then the inverse *Fourier Transform* with respect to ω with the direct computations gives the

following formulation,

$$\begin{cases} \frac{1}{c^2}p_t + \frac{1}{c^2}\alpha p + \frac{1}{c^2}\beta p^* + \nabla\cdot\vec{\mathbf{q}} + M_\sigma\vec{\mathbf{q}}^* &= 0, \\ \vec{\mathbf{q}}_t \ + \ \alpha\vec{\mathbf{q}} \ + \ \beta\vec{\mathbf{q}}^* \ + \ \nabla p \ + \ C_\sigma\nabla p^* &= 0, \\ p_t^* &= p, \\ \vec{\mathbf{q}}_t^* &= \vec{\mathbf{q}}, \end{cases} \tag{3.6}$$

with the initial condition $(p, p^*) = (p_0, p_0^*)$, $(\vec{\mathbf{q}}, \vec{\mathbf{q}}^*) = (\vec{\mathbf{q}}_0, \vec{\mathbf{q}}_0^*)$ and the boundary condition $(p, p^*)|_{\partial\Omega} = (0, 0)$, where the coefficients are defined as $\alpha = \partial_x\sigma^x + \partial_y\sigma^y$, $\beta = \partial_x\sigma^x\partial_y\sigma^y - \partial_x\sigma^y\partial_y\sigma^x$,

$$M_\sigma = \begin{bmatrix} \partial_y\sigma^y\frac{\partial}{\partial x} - \partial_y\sigma^x\frac{\partial}{\partial y} & \partial_x\sigma^x\frac{\partial}{\partial y} - \partial_x\sigma^y\frac{\partial}{\partial x} \end{bmatrix},\ C_\sigma = \begin{bmatrix} \partial_y\sigma^y & -\partial_y\sigma^x \\ -\partial_x\sigma^y & \partial_x\sigma^x \end{bmatrix}.$$

We next introduce the regularized formulation of the system (3.6).

3.1.1 The regularized Formulation

We regularize several terms in order to get regularity of weak solutions in (3.6), which derives a new formulation. Recall the linear bounded operator $\delta_\varepsilon : H^{-1}(\Omega) \to L^2(\Omega)$ in (2.12) and the dual operator $\delta'_\varepsilon : L^2(\Omega) \to H_0^1(\Omega)$.

We introduce a new formulation with the regularized term using δ_ε and δ'_ε,

$$\begin{cases} \frac{1}{c^2}p_t + \frac{1}{c^2}\alpha p + \frac{1}{c^2}\beta p^* + \delta_\varepsilon\nabla\cdot\vec{\mathbf{q}} + \delta_\varepsilon\ M_\sigma\vec{\mathbf{q}}^* &= 0, \\ \vec{\mathbf{q}}_t \ + \ \alpha\vec{\mathbf{q}} \ + \ \beta\vec{\mathbf{q}}^* \ + \ \nabla\delta'_\varepsilon p \ + \ C_\sigma\ \nabla\delta'_\varepsilon p^* &= 0, \\ p_t^* &= p, \\ \vec{\mathbf{q}}_t^* &= \vec{\mathbf{q}}, \end{cases} \tag{3.7}$$

with the initial conditions $(p(0), p^*(0)) = (p_0, p_0^*)$ and $(\vec{\mathbf{q}}(0), \vec{\mathbf{q}}^*(0)) = (\vec{\mathbf{q}}_0, \vec{\mathbf{q}}_0^*)$. Note that the zero *Dirichlet* boundary condition $\delta'_\varepsilon p|_{\partial\Omega} = 0$ is imposed in the system (3.7).

We define a weak solution of the system (3.7).

Definition 3.1 *We define*

$$\{p, p^*\} \in L^2(0,T;L^2(\Omega)),\ \{\vec{\mathbf{q}}, \vec{\mathbf{q}}^*\} \in L^2(0,T;\mathbb{L}^2(\Omega)), \tag{3.8}$$

with

$$\{p_t, p_t^*\} \in L^2(0,T;L^2(\Omega)),\ \{\vec{\mathbf{q}}_t, \vec{\mathbf{q}}_t^*\} \in L^2(0,T;\mathbb{L}^2(\Omega)), \tag{3.9}$$

is a weak solution of the initial-value boundary problem (3.7) *provided*

$$\begin{cases} (\frac{1}{c^2}p_t, r) + (\frac{1}{c^2}\alpha p, r) + (\frac{1}{c^2}\beta p^*, r) + (\delta_\varepsilon \nabla\cdot\vec{\mathbf{q}}, r) + (\delta_\varepsilon M_\sigma \vec{\mathbf{q}}^*, r) & = 0, \\ (\vec{\mathbf{q}}_t, \vec{\mathbf{v}}) \ + \ (\alpha\vec{\mathbf{q}}, \vec{\mathbf{v}}) \ + \ (\beta\vec{\mathbf{q}}^*, \vec{\mathbf{v}}) \ + \ (\nabla\delta'_\varepsilon p, \vec{\mathbf{v}}) \ + \ (C_\sigma \nabla\delta'_\varepsilon p^*, \vec{\mathbf{v}}) & = 0, \\ (p_t^*, r^*) \ - \ (p, r^*) & = 0, \\ (\vec{\mathbf{q}}_t^*, \vec{\mathbf{v}}^*) \ - \ (\vec{\mathbf{q}}, \vec{\mathbf{v}}^*) & = 0, \end{cases} \tag{3.10}$$

for all $r, r^* \in L^2(0,T;L^2(\Omega))$, $\vec{\mathbf{v}}, \vec{\mathbf{v}}^* \in L^2(0,T;\mathbb{L}^2(\Omega))$ *which satisfies the* Cauchy *initial data in a weak sense.*

We prove the existence and uniqueness of the weak solution of (3.7).

Theorem 3.1 *We assume that the initial data* $(p_0, p_0^*, \vec{\mathbf{q}}_0, \vec{\mathbf{q}}_0^*) \in [L^2(\Omega)]^2 \times [\mathbb{L}^2(\Omega)]^2$. *The regularized system* (3.7) *admits a unique weak solution satisfying* (3.8), (3.9), *provided* (3.3) *holds true.*

Proof. We define the energy norm by

$$E = \|\frac{1}{c}p\|^2_{L^2(\Omega)} + \|p^*\|^2_{L^2(\Omega)} + \|\vec{\mathbf{q}}\|^2_{\mathbb{L}^2(\Omega)} + \|\vec{\mathbf{q}}^*\|^2_{\mathbb{L}^2(\Omega)}.$$

First we show estimates of the energy. We apply the scalar product of all equations in (3.7) with p, p^* in $L^2(\Omega)$ and $\vec{\mathbf{q}}, \vec{\mathbf{q}}^*$ in $\mathbb{L}^2(\Omega)$ respectively, to obtain the identity,

$$\frac{dE}{dt} + F_1 + F_2 + F_3 + F_4 = 0, \tag{3.11}$$

where

$$F_1 = (\frac{1}{c^2}\alpha p, p) + (\frac{1}{c^2}\beta p^*, p) - (p, p^*),$$

$$F_2 = (\alpha\vec{\mathbf{q}}, \vec{\mathbf{q}}) + (\beta\vec{\mathbf{q}}^*, \vec{\mathbf{q}}) - (\vec{\mathbf{q}}, \vec{\mathbf{q}}^*),$$

$$F_3 = (\delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}, p) + (\nabla\delta'_\varepsilon p, \vec{\mathbf{q}}),$$

$$F_4 = (\delta_\varepsilon M_\sigma \vec{\mathbf{q}}^*, p) + (C_\sigma \nabla\delta'_\varepsilon p^*, \vec{\mathbf{q}}).$$

We have that $|F_1| + |F_2| \leq C_{12}E$ a.e. in t for some $C_{12} > 0$ since $\alpha, \beta \in L^\infty(\tilde{\Omega})$ and the bounds of c in (3.5). It is allowed to have that $F_3 = 0$ by the duality of δ_ε and integration by parts,

$$\begin{aligned}
(\delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}, p) + (\nabla\delta'_\varepsilon p, \vec{\mathbf{q}}) &= p(\delta_\varepsilon \nabla \cdot \vec{\mathbf{q}}) + (\nabla\delta'_\varepsilon p, \vec{\mathbf{q}}), \\
&= (\nabla\cdot)'\delta'_\varepsilon p(\vec{\mathbf{q}}) + (\nabla\delta'_\varepsilon p, \vec{\mathbf{q}}), \\
&= -\nabla\delta'_\varepsilon p(\vec{\mathbf{q}}) + \nabla\delta'_\varepsilon p(\vec{\mathbf{q}}) \\
&= 0,
\end{aligned}$$

since $(\nabla\cdot)' = -\nabla$. The operators

$$\{\vec{\mathbf{q}}^*, p^*\} \to \left\{\delta_\varepsilon M_\sigma \vec{\mathbf{q}}^*, C_\sigma \nabla\delta'_\varepsilon p^*\right\}$$

are continuous from $[\mathbb{L}^2(\Omega)]^2 \times L^2(\Omega) \to [\mathbb{L}^2(\Omega)]^2 \times L^2(\Omega)$ since $\sigma_y^x(x, y)$ and $\sigma_x^y(x, y)$ are in $W^{1,\infty}(\Omega)$, which implies that

$$|F_4| \leq C_4 E \text{ a.e. in } t \in [0, T],$$

for some $C_4 > 0$.

It follows that from (3.11)

$$\frac{dE}{dt} \leq C_T E \text{ a.e. } t \in [0, T] \tag{3.12}$$

for a suitable constant $C_T > 0$.

This is a standard *a priori* estimates, using this estimates we can obtain the existence part of Theorem 3.1 by the standard Galerkin method argument, and also uniqueness can be established by the estimates. We omit details here, since the similar argument was presented in chapter 2. □

Remark 3.1 *We don't present any stability analysis or numerical experiments in this section. Further investigation of the original system and the regularized one remains for further work. But we introduce another simpler formulation with the same technique as in* (3.1) *and* (3.2) *to show that the multi-directional damping PML can be more effective than the classical one.*

3.2. Multi Directional Split PML in the parallel to y-axis

We apply the multi-directional damping (3.1) to the system of first order acoustic wave equation with Split PML techniques parallel to y-axis.

Let the domain $\Omega = [-a - L_x, a + L_x] \times [-b, b]$ consist of the computational domain $[-a, a] \times [-b, b]$ with the PML only parallel to y-axis. The damping $\sigma^x(x, y)$ in (3.1) with $\sigma^y = 0$ in (3.2) is applied as follows:

$$\tilde{x}(x, y(x)) = x + \frac{1}{iw}\sigma^x(x, y) = x + \frac{1}{iw}\left(\int_a^x \sigma_x(s)ds + \int_0^y \sigma_x^y(s)ds\right), \tag{3.13}$$

$$\tilde{y}(y) \quad = y. \tag{3.14}$$

Note that the damping $\sigma_x^y(x)$ depends on both x and y in the PML. The coordinate change with the damping gives Jacobian matrix

$$J = \frac{\partial(\tilde{x}, \tilde{y})}{\partial(x, y)} = \frac{1}{iw}\begin{bmatrix} D & \sigma_x^y \\ 0 & 1 \end{bmatrix},$$

and

$$J^{-1} = \frac{1}{D}\begin{pmatrix} iw & -\sigma_x^y \\ 0 & D \end{pmatrix},$$

where $D = iw + \sigma_x + \frac{\partial}{\partial x}\int_0^y \sigma_x^y(s)ds$.

Now we have

$$\frac{\partial}{\partial \tilde{x}} = \frac{iw}{D}\frac{\partial}{\partial x} - \frac{\sigma_x^y}{D}\frac{\partial}{\partial y}, \tag{3.15}$$

$$\frac{\partial}{\partial \tilde{y}} = \frac{\partial}{\partial y}. \tag{3.16}$$

Following [38], we introduce the split system for the acoustic wave equation in order to apply the coordinate systems (3.15), (3.16). Assume the solution p split into the two fields p^x and p^y satisfying $p = p^x + p^y$ and

$$p_t^x + c^2\frac{\partial}{\partial x}q_x = 0, \qquad p_t^y + c^2\frac{\partial}{\partial y}q_y = 0.$$

Then we have the split system of acoustic wave equation,

$$\begin{cases} p_t^x + c^2\frac{\partial}{\partial x}q_x = 0, \\ p_t^y + c^2\frac{\partial}{\partial y}q_y = 0, \\ q_{xt} + \frac{\partial}{\partial x}(p^x + p^y) = 0, \\ q_{yt} + \frac{\partial}{\partial y}(p^x + p^y) = 0. \end{cases} \tag{3.17}$$

We apply (3.15), (3.16) in the frequency space of (3.17) to obtain

$$\begin{cases} D\frac{1}{c^2}\hat{p}^x + \frac{\partial}{\partial x}\hat{q}_x - \sigma_x^y\frac{1}{iw}\frac{\partial}{\partial y}\hat{q}_x = 0, \\ iw\frac{1}{c^2}\hat{p}^y + \frac{\partial}{\partial y}\hat{q}_y = 0, \\ D\hat{q}_x + \frac{\partial}{\partial x}(\hat{p}^x + \hat{p}^y) - \sigma_x^y\frac{1}{iw}\frac{\partial}{\partial y}(\hat{p}^x + \hat{p}^y) = 0, \\ iw\hat{q}_y + \frac{\partial}{\partial y}(\hat{p}^x + \hat{p}^y) = 0. \end{cases} \tag{3.18}$$

We introduce an auxiliary variable $\hat{q}_x^* = -\frac{1}{iw}\frac{\partial}{\partial y}\hat{q}_x$ to obtain a new formulation after taking the inverse Fourier transform:

$$\begin{cases} \frac{1}{c^2}p_t^x + \frac{\bar{\sigma}}{c^2}p^x + \frac{\partial}{\partial x}q_x + \sigma_x^y q_x^* = 0, \\ \frac{1}{c^2}p_t^y + \frac{\partial}{\partial y}q_y = 0, \\ q_{xt} + \bar{\sigma}_x q_x + \frac{\partial}{\partial x}(p^x + p^y) + \sigma_x^y q_y = 0, \\ q_{y_t} + \frac{\partial}{\partial y}(p^x + p^y) = 0, \\ q_{xt}^* + \frac{\partial}{\partial y}q_x = 0, \end{cases} \tag{3.19}$$

where $\bar{\sigma}_x = \sigma_x + \frac{\partial}{\partial x}\int_0^y \sigma_x^y(s)ds$.

3.2.1 Numerical Results

In this section, we show the system (3.19) is efficient and compare it with the classical Split PML [38]. We use centered differences and the staggered nodes in time and space the same as in section 2.4.1. With the same notation, the components $p^\alpha,\ q_x^*, \alpha = x, y$, are discretized at nodes (t^n, x_i, x_j) as $p^{\alpha n}_{i,j}, q^{*n}_{xi,j}$, and q_α, are discretized at $(t^{n+\frac{1}{2}}, x_{i+\frac{1}{2}}, x_{j+\frac{1}{2}})$ as $q_{\alpha}{}^{n+\frac{1}{2}}_{i+\frac{1}{2},j+\frac{1}{2}}$. This centered time stepping ensures a second order approximation in time. Denote by

$$\mathtt{A}_x^{\pm} = 1 \pm \bar{\sigma}_x \frac{\triangle t}{2}.$$

Step 1. Compute $q_{\alpha}{}^{n+\frac{1}{2}}_{i+\frac{1}{2},j+\frac{1}{2}}, \alpha = x, y,$

$$\mathtt{A}_x^+ q_x{}^{n+\frac{1}{2}}_{i+\frac{1}{2},j+\frac{1}{2}} = \mathtt{A}_x^- q_x{}^{n-\frac{1}{2}}_{i+\frac{1}{2},j+\frac{1}{2}} - \triangle t(\partial_x(u^x+u^y))^n_{i+\frac{1}{2},j+\frac{1}{2}} - \triangle t\sigma_y^x \left(\frac{q_y{}^{n+\frac{1}{2}}_{i+\frac{1}{2},j+\frac{1}{2}} + q_y{}^{n-\frac{1}{2}}_{i+\frac{1}{2},j+\frac{1}{2}}}{2}\right),$$

$$q_y{}^{n+\frac{1}{2}}_{i+\frac{1}{2},j+\frac{1}{2}} = q_y{}^{n-\frac{1}{2}}_{i+\frac{1}{2},j+\frac{1}{2}} - \triangle t(\partial_y(p^x+p^y))^n_{i+\frac{1}{2},j+\frac{1}{2}},$$

where
$$(\partial_x(p^x+p^y))^n_{i+\frac{1}{2},j+\frac{1}{2}} = (\partial_x p^x)^n_{i+\frac{1}{2},j+\frac{1}{2}} + (\partial_x p^y)^n_{i+\frac{1}{2},j+\frac{1}{2}},$$

$$(\partial_x p^x)^n_{i+\frac{1}{2},j+\frac{1}{2}} = \frac{(p^{xn}_{i+1,j+1} - p^{xn}_{i,j+1} + p^{xn}_{i+1,j} - p^{xn}_{i,j})}{2\triangle x},$$

$$(\partial_x p^y)^n_{i+\frac{1}{2},j+\frac{1}{2}} = \frac{(p^{yn}_{i+1,j+1} - p^{yn}_{i,j+1} + p^{yn}_{i+1,j} - p^{yn}_{i,j})}{2\triangle x},$$

and $(\partial_x(p^x+p^y))^n_{i+\frac{1}{2},j+\frac{1}{2}}$ is similarly defined.

Step 2. Compute $q^{*n+1}_{xi,j}$,

$$q^{*n+1}_{xi,j} = q^{*n}_{xi,j} - \frac{\triangle t}{2\triangle y}\left(q_x{}^{n+\frac{1}{2}}_{i+\frac{1}{2},j+\frac{1}{2}} - q_x{}^{n+\frac{1}{2}}_{i+\frac{1}{2},j-\frac{1}{2}} + q_x{}^{n+\frac{1}{2}}_{i-\frac{1}{2},j+\frac{1}{2}} - q_x{}^{n+\frac{1}{2}}_{i-\frac{1}{2},j-\frac{1}{2}}\right).$$

Step 3. Compute $u^{\alpha n}_{i,j}, \alpha = x, y,$

$$\frac{1}{c^2}\mathtt{A}_x^+ p^{xn+1}_{i,j} = \frac{1}{c^2}\mathtt{A}_x^- p^{xn}_{i,j} - \triangle t(\partial_x q_x)^{n+\frac{1}{2}}_{i,j} - \triangle t\sigma_y^x\left(\frac{q_x{}^{n+1}_{i,j} + q_x{}^{n}_{i,j}}{2}\right),$$

$$\frac{1}{c^2}p^{yn+1}_{i,j} = \frac{1}{c^2}p^{yn}_{i,j} - \frac{\triangle t}{2\triangle y}\left(q_y{}^{n+\frac{1}{2}}_{i+\frac{1}{2},j+\frac{1}{2}} - q_y{}^{n+\frac{1}{2}}_{i+\frac{1}{2},j-\frac{1}{2}} + q_y{}^{n+\frac{1}{2}}_{i-\frac{1}{2},j+\frac{1}{2}} - q_y{}^{n+\frac{1}{2}}_{i-\frac{1}{2},j-\frac{1}{2}}\right).$$

We impose smooth variable sound speed $c(x, y) \in [-0.5, 0.5]$ (see FIGURE 3.1), and set

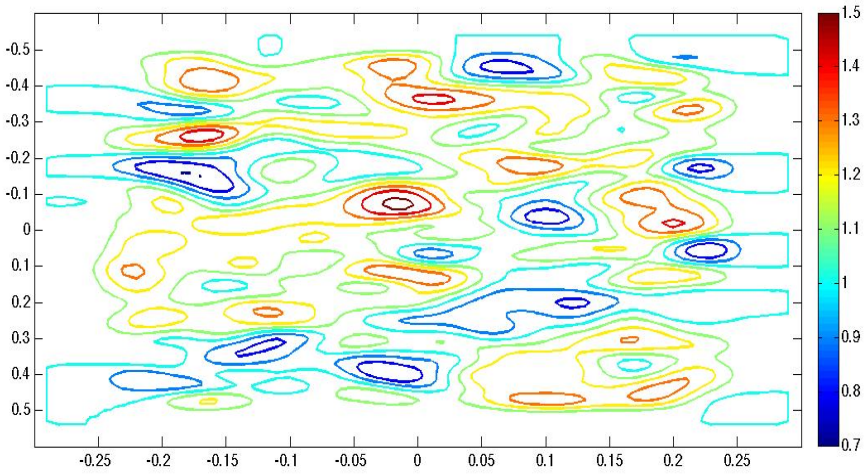

Figure 3.1: Variable sound speed in the computational domain= $[-0.3, 0.3] \times [-0.6, 0.6]$

the damping σ_x^y as

$$\sigma_x^y(x, y) = \sigma_y(y) \int_a^x \sigma_x(s)ds,$$

where σ_x and σ_y are defined as in (2.48) with various maximum damping coefficients $\bar{\sigma}_0$. We consider the computational domain $[-0.3, 0.3] \times [-0.6, 0.6]$ with the PML region $[-0.4, -0.3] \cup [0.3, 0.4]$ parallel to y-axis. We compare the numerical solution obtained with a reference solution, computed with the same numerical scheme on a very large domain $[-0.6, 0.6] \times [-0.6, 0.6]$. In Figure 3.2, it is shown the discrete L^2-*error* in (2.53) of the classical Split PML and Multi-Directional Split PML on the computational domain and the difference of the errors with two different damping. Similarly, Figure 3.3 shows the maximum error of two PMLs in the computational domain. In this case, the multi-directional Split PML leads to a smaller error than that for the classical Split PML for both L^2-*error* and *maximum error*.

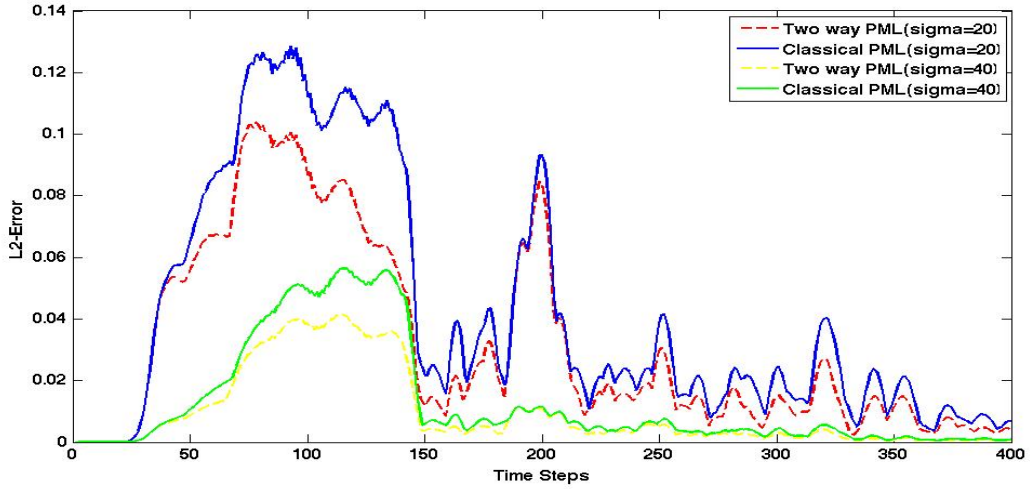

Figure 3.2: L^2-error of classical Split PML and Multi-Directional Split PML in the computational domain

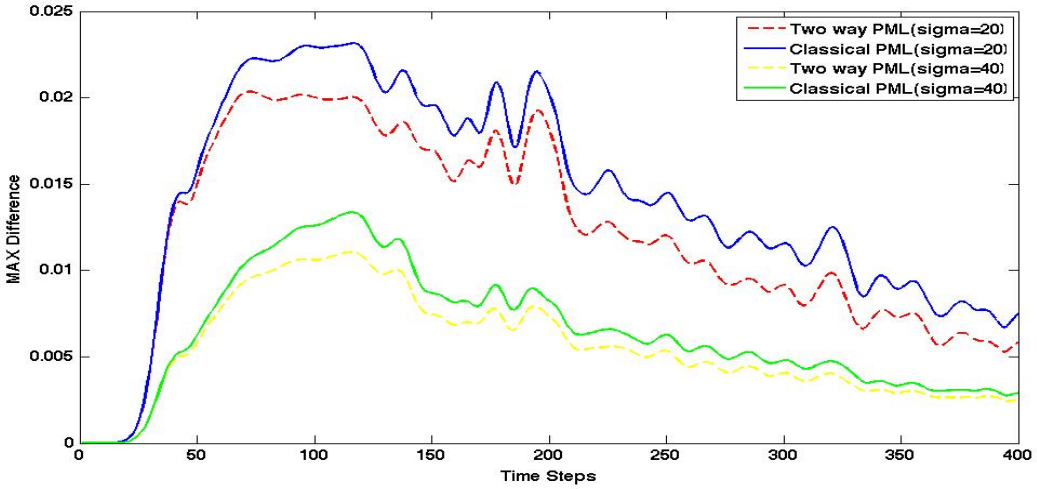

Figure 3.3: Maximum error of classical Split PML and Multi-Directional Split PML in the computational domain

4. PML IN 1-D : ENERGY DECAY FOR THE ACOUSTIC WAVE WITH VARIABLE SOUND SPEED

In this section we investigate the efficiency of the PML method in the acoustic 1-d wave equation with variable sound speed. The energy decay rate is investigated in [46] for a constant speed 1-d continuous and semi-continuous wave equation with one sided PML.

4.1. Energy Decay 1-d PML Wave Equation : Spectrum

First we consider an initial value problem for the acoustic wave equation with variable sound speed $c(x) \in C^1(\mathbb{R})$ in the unbounded domain $\mathbb{R}$,

$$\frac{\partial^2}{\partial t^2}u \;-\; c^2\frac{\partial^2}{\partial x^2}u \;=\; 0, \quad t>0, \tag{4.1}$$

with the initial condition

$$u(x,0) \;=\; f, \quad \text{and} \quad \frac{\partial}{\partial t}u(x,0) \;=\; 0.$$

We introduce the new variables $P = -\frac{\partial}{\partial x}u, Q = \frac{1}{c}\frac{\partial}{\partial t}u$ to obtain the system

$$\begin{cases} \frac{\partial}{\partial t}P \;+\; \frac{\partial}{\partial x}(cQ) \;=\; 0, & t>0, \\ \frac{\partial}{\partial t}Q \;+\; c\frac{\partial}{\partial x}P \;=\; 0, & t>0, \end{cases} \tag{4.2}$$

with the initial conditions $P(x,0) = -\frac{\partial}{\partial x}f,\ Q(x,0) = 0$. Next, we truncate the unbounded domain to the interval $I := [-a-L, a+L]$ with the PML interval $I_\gamma := [-L-a, a]\cup[a, a+L]$ imposing the zero *Dirichlet* boundary condition on ∂I for P.

We assume $c(x) \in C^1(\Omega)$ is bounded below by c_* and above by c^*, i.e.,

$$0 < c_* \le c(x) \le c^* < \infty \text{ in } \Omega, \tag{4.3}$$

and also $c(x) \equiv 1$ in I_γ. Let $\sigma \in L^1(I)$ be a non-trivial and non-negative function vanishing identically in the computational interval $[-a, a]$ and monotone in I_γ, i.e.,

$$\sigma_x(s) \leq \sigma_x(\tau) \text{ if } 0 < s \leq \tau, \text{ or } \tau \leq s < 0. \tag{4.4}$$

Then we have the new following system

$$\begin{cases} \frac{\partial}{\partial t}P(x,t) + \sigma_x(x)P(x,t) + \frac{\partial}{\partial x}(c(x)Q(x,t)) &= 0, \quad t > 0, \\ \frac{\partial}{\partial t}Q(x,t) + \sigma_x(x)Q(x,t) + c(x)\frac{\partial}{\partial x}P(x,t) &= 0, \quad t > 0, \\ P(x,t) &= 0, \quad x \in \partial I, \end{cases} \tag{4.5}$$

with the initial conditions $P(x,0) = -\frac{\partial}{\partial x}f,\ Q(x,0) = 0$.

Let $P = U - V$ and $Q = U + V$, then

$$\begin{cases} \frac{\partial}{\partial t}(U-V) + \sigma_x(U-V) + \frac{\partial}{\partial x}(c(U+V)) = 0, \quad t > 0, \\ \frac{\partial}{\partial t}(U+V) + \sigma_x(U+V) + c\frac{\partial}{\partial x}(U-V) = 0, \quad t > 0. \end{cases} \tag{4.6}$$

From $\frac{\partial}{\partial x}(c(U+V)) = c'(U+V) + c\frac{\partial}{\partial x}(U+V)$ we obtain

$$\begin{cases} \frac{\partial}{\partial t}U + \sigma_x U + c\frac{\partial}{\partial x}U + \frac{1}{2}c'(U+V) = 0, \ t > 0, \\ \frac{\partial}{\partial t}V + \sigma_x V - c\frac{\partial}{\partial x}V - \frac{1}{2}c'(U+V) = 0, \ t > 0. \end{cases} \tag{4.7}$$

Let $M = A + B$, where

$$A\begin{pmatrix} U \\ V \end{pmatrix} = \sigma_x \begin{pmatrix} U \\ V \end{pmatrix} + (c\frac{\partial}{\partial x} + \frac{1}{2}c')\begin{pmatrix} U \\ -V \end{pmatrix}, \tag{4.8}$$

$$B\begin{pmatrix} U \\ V \end{pmatrix} = \frac{1}{2}\, c' \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{pmatrix} U \\ V \end{pmatrix}. \tag{4.9}$$

We define the total energy of the solutions given by

$$E(t) := E(U(t), V(t)) = \frac{1}{2}\int_I \left(|U(x,t)|^2 + |V(x,t)|^2\right) dx. \tag{4.10}$$

Then we have that

$$\frac{d}{dt}E(t) = -\frac{1}{2}\int_I \sigma(x)\left(|U-V|^2 + |U+V|^2\right)dx \leq 0,$$

which shows the well-posedness of the system (4.5) in the space $(P,Q) \in C([0,\infty);[L^2(I)]^2)$. Then we have the following property:

Lemma 4.1 *The operator M has a compact inverse if σ is non-trivial.*

Proof. We construct the inverse of the operator M satisfying

$$M\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} g \\ h \end{pmatrix} \text{ with the boundary condition } U = V \text{ on } \partial I. \tag{4.11}$$

In the PML interval I_γ, $c \equiv 1$ which implies $B = 0$ yielding $\begin{cases} \frac{\partial}{\partial x}U + \sigma_x U &= g, \\ -\frac{\partial}{\partial x}V + \sigma_x V &= h. \end{cases}$

We solve the above equations to obtain

$$\begin{cases} U(a+L) &= e^{-\int_a^{a+L}\sigma_x dx}\left(U(a) + \int_a^{a+L} g e^{\int_a^x \sigma ds}dx\right), \\ V(a+L) &= e^{\int_a^{a+L}\sigma_x dx}\left(V(a) - \int_a^{a+L} h e^{-\int_a^x \sigma ds}dx\right), \end{cases} \tag{4.12}$$

and

$$\begin{cases} U(-a) &= e^{-\int_{-a-L}^{-a}\sigma_x dx}\left(U(-a-L) + \int_{-a-L}^{-a} g e^{\int_{-a-L}^x \sigma ds}dx\right), \\ V(-a) &= e^{\int_{-a-L}^{-a}\sigma_x dx}\left(V(-a-L) - \int_{-a-L}^{-a} h e^{-\int_{-a-L}^x \sigma ds}dx\right). \end{cases} \tag{4.13}$$

In the computation interval $[-a,a]$, from $\sigma_x = 0$, we have

$$\begin{cases} c\frac{\partial}{\partial x}U + \frac{1}{2}c'(U+V) &= g, \\ -c\frac{\partial}{\partial x}V - \frac{1}{2}c'(U+V) &= h. \end{cases}$$

We add and subtract each other to get

$$\begin{cases} c\frac{\partial}{\partial x}(U+V) + c'(U+V) &= g-h, \\ c\frac{\partial}{\partial x}(U-V) &= g+h. \end{cases}$$

Solving the above equations with the property $c(x) \in C^1(I)$ and $c = 1$ at $x = \pm\, a$ we obtain

$$\begin{cases} (U+V)(a) = (U+V)(-a) + \int_{-a}^{a}(g-h)dx, \\ (U-V)(a) = (U-V)(-a) + \int_{-a}^{a}\frac{g+h}{c}dx, \end{cases} \tag{4.14}$$

or

$$\begin{cases} U(a) = U(-a) + \frac{1}{2}\int_{-a}^{a}(g-h)dx + \frac{1}{2}\int_{-a}^{a}\frac{g+h}{c}dx, \\ V(a) = V(-a) + \frac{1}{2}\int_{-a}^{a}(g-h)dx - \frac{1}{2}\int_{-a}^{a}\frac{g+h}{c}dx. \end{cases} \tag{4.15}$$

We combine (4.12), (4.13), and (4.15) taking that $U(x) = V(x)$, at $x = \pm(a+L)$ to have

$$\begin{bmatrix} 1 & -e^{-\int_{-a-L}^{a+L}\sigma_x dx} \\ 1 & -e^{\int_{-a-L}^{a+L}\sigma_x dx} \end{bmatrix} \begin{pmatrix} U(a+L) \\ U(-a-L) \end{pmatrix} = \begin{pmatrix} V_1(g,h) \\ V_2(g,h) \end{pmatrix},$$

where $V_j(g,h), j = 1,2$, are expressions involving g and h and are independent of U or V in $H^1(I)$.

Note that

$$\begin{aligned} det \begin{bmatrix} 1 & -e^{-\int_{-a-L}^{a+L}\sigma_x dx} \\ 1 & -e^{\int_{-a-L}^{a+L}\sigma_x dx} \end{bmatrix} &= -e^{\int_{-a-L}^{a+L}\sigma_x dx} + e^{-\int_{-a-L}^{a+L}\sigma_x dx} \\ &= -e^{-\int_{-a-L}^{a+L}\sigma_x dx}(e^{2\int_{-a-L}^{a+L}\sigma_x dx} - 1) \\ &\neq 0 \end{aligned}$$

if σ_x is non-trivial. Therefore the equation (4.11) is uniquely solvable in $(U,V) \in H^1(I)$ such that $U - V \in H_0^1(I)$ if and only if σ_x is non-trivial.

Furthermore, the inverse of M,

$$M^{-1} : [L^2(I)]^2 \to [H^1(I)]^2$$

is bounded also compact by the compact embedding $H^1(I) \subset\subset L^2(I)$ if σ_x is non-trivial.

□

This Lemma 4.1 implies the spectrum of M is discrete.

Next we investigate the operator A in (4.8).

Definition 4.1 *A collection of functions $\{u_k\}$ in a Hilbert space H is called a Riesz basis for H if $\overline{span\{u_k\}} = H$ and there exist constants $0 < C_A \le C_B < \infty$ such that*

$$C_A \left(\sum_k |a_k|^2 \right) \le \left\| \sum_k a_k u_k \right\|^2 \le C_B \left(\sum_k |a_k|^2 \right)$$

for all sequences of $\{a_k\} \in \ell^2(\mathbb{Z})$.

Lemma 4.2 *Let $\sigma_x \in L^1(I)$ be a non-trivial and non-negative function which vanishes identically in the computational region, $[-a, a]$. Then we have that*

1. *The spectrum of the operator A in (4.8) is identically same as the set of eigenvalues*

$$\lambda_k = \frac{1}{c_L} I_\sigma + \frac{1}{c_L} k\pi i, \text{ where } I_\sigma = \int_I \sigma_x(s) ds, \quad c_L = \int_I \frac{1}{c(s)} ds, \quad k \in \mathbb{Z}, \tag{4.16}$$

   and the eigenfunction corresponding to λ_k is

$$\begin{pmatrix} U_k \\ V_k \end{pmatrix} = \begin{pmatrix} e^{-\int_{-a-L}^{x} \frac{1}{c}(\sigma_x - \lambda_k) dx} \\ e^{\int_{-a-L}^{x} \frac{1}{c}(\sigma_x - \lambda_k) dx} \end{pmatrix}.$$

2. *The eigenfunctions $\{(U_k, V_k)\}$ form a Riesz basis of $[L^2(I)]^2$.*

Proof. Let $A \begin{pmatrix} U \\ V \end{pmatrix} = \lambda \begin{pmatrix} U \\ V \end{pmatrix}$, for $\lambda \in \mathbb{C}$, that is,

$$(c(x)\frac{\partial}{\partial x} + \frac{1}{2}c'(x)) \begin{pmatrix} U \\ -V \end{pmatrix} + \sigma_x \begin{pmatrix} U \\ V \end{pmatrix} = \lambda \begin{pmatrix} U \\ V \end{pmatrix}.$$

Then we have $\frac{\partial}{\partial x} U + \frac{1}{c(x)}(\frac{1}{2}c'(x) + \sigma_x - \lambda)U = 0$ and $\frac{\partial}{\partial x} V - \frac{1}{c(x)}(-\frac{1}{2}c'(x) + \sigma_x - \lambda)V = 0$, which gives

$$\begin{cases} U(x) = U(-a-L) e^{-\int_{-a-L}^{x} \frac{1}{c}(\frac{1}{2}c' + \sigma_x - \lambda) ds}, \\ V(x) = V(-a-L) e^{\int_{-a-L}^{x} \frac{1}{c}(-\frac{1}{2}c' + \sigma_x - \lambda) ds}. \end{cases} \tag{4.17}$$

With the boundary condition $U(x) = V(x)$ at $x = \pm(a + L)$ we have $e^{-\int_I (\sigma_x - \frac{1}{c}\lambda) ds} = e^{\int_I (\sigma_x - \frac{1}{c}\lambda) ds}$ by $c \equiv 1$ on I_γ. Thus $\int_I (\sigma_x(x) - \frac{1}{c}\lambda_k) ds + k\pi i = 0$ for all $k \in \mathbb{Z}$, and we have

the eigenvalues λ_k of A as (4.16). Therefore, the eigenfunction corresponding to λ_k is

$$\begin{pmatrix} U_k \\ V_k \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{c(x)}} e^{-\int_{-a-L}^{x} \frac{1}{c}(\sigma_x - \lambda_k) ds} \\ \frac{1}{\sqrt{c(x)}} e^{\int_{-a-L}^{x} \frac{1}{c}(\sigma_x - \lambda_k) ds} \end{pmatrix}.$$

Define the function θ by

$$\theta(x) = \int_{-a-L}^{x} \frac{1}{c(s)} \left(\sigma_x(s) - \frac{1}{c_L} I_\sigma \right) ds.$$

This function expresses the difference between the damping term σ_x and the average damping $\frac{1}{c_L} I_\sigma$ over $[0, c_L]$. Next we show that the family of (U_k, V_k) forms a Riesz basis in $[L^2(I)]^2$ following the proof (with a constant speed) in [46]. Equivalently that is to show that any pair of functions $(g, h) \in [L^2(I)]^2$ can be written in an unique way in the following sense:

$$(g, h) = \sum a_k (U_k, V_k), \tag{4.18}$$

with

$$\sum |a_k|^2 \simeq \|(g, h)\|^2.$$

We write it, using (4.17), taking the initial $U(-a-L) = 1 = V(-a-L)$ as the following,

$$\begin{cases} g(x) e^{\theta(x)} &= \sum a_k U_k(x) e^{\theta(x)} = \sum a_k e^{\frac{1}{c_L} k\pi i y(x)}, \quad -a-L < x < a+L, \\ h(x) e^{-\theta(x)} &= \sum a_k V_k(x) e^{-\theta(x)} = \sum a_k e^{-\frac{1}{c_L} k\pi i y(x)}, \quad -a-L < x < a+L, \end{cases} \tag{4.19}$$

where

$$y(x) = \phi(x) := \int_{-a-L}^{x} \frac{1}{c(s)} ds.$$

Then, the coefficient $\{a_k\}$ of the decomposition (4.18) of (g, h) in the basis $\{(U_k, V_k)\}$ can be identified as the Fourier coefficients of the function W defined in $(-c_L, c_L)$ by

$$W(y) = \begin{cases} g(\phi^{-1}(y)) e^{\theta(\phi^{-1}(y))}, & 0 < y < c_L, \\ h(\phi^{-1}(y + c_L)) e^{-\theta(\phi^{-1}(y+c_L))}, & -c_L < y < 0. \end{cases} \tag{4.20}$$

Then (4.18) is satisfied if and only if

$$W(y) = \sum_k a_k \exp\left(\frac{1}{c_L} ik\pi y\right), y \in (-c_L, c_L). \tag{4.21}$$

It holds that $W \in L^2(-c_L, c_L)$ since $(g, h) \in [L^2(I)]^2$. This mapping gives an isomorphism $\mathcal{I} : [L^2(I)]^2 \to L^2(-c_L, c_L)$ which sends the eigenvectors (U_k, V_k) to the classical Fourier basis of $L^2(-c_L, c_L)$:

$$\mathcal{I}(g, h) = W, \tag{4.22}$$

where W is the function given in (4.20). This implies that any function $\vec{\mathbf{Q}} \in [L^2(I)]^2$ can be expanded as $\sum_k \vec{\mathbf{Q}}_k$, where the coefficients $\{a_k\}$ satisfy that

$$\|\mathcal{I}\vec{\mathbf{Q}}\|^2_{L^2(-c_L,c_L)} = 2c_L \sum |a_k|^2.$$

The proof is completed. □

We define exponential decay rate of solutions of (4.7) as a function of σ, defined by

$$\omega(\sigma) = \sup\left\{\omega : \exists C, \forall (U_0, V_0) \in [L^2(I)]^2, \forall t, E(t) \le CE(U_0, V_0)\exp(-\omega t)\right\}$$

For each $\omega \le \omega(\sigma)$, we define $C(\omega)$ as the best constant such that

$$\forall (U_0, V_0) \in [L^2(I)]^2, \forall t, E(t) \le C(\omega)E(U_0, V_0)\exp(-\omega t).$$

Definition 4.2 *Let $\sigma(A)$ is the spectrum of the operator A. Then*

$$S(\sigma) := \sup\{Re(\lambda) | \lambda \in \sigma(A)\}$$

is called the spectral abscissa of the operator A.

There is a decay rate of the energy of the operator A in (4.7).

Theorem 4.1 *The energy of the PML system* $\frac{\partial}{\partial t}\begin{pmatrix} U \\ V \end{pmatrix} + A\begin{pmatrix} U \\ V \end{pmatrix} = 0$ *in* (4.7) *is exponentially decaying. In the detail,*

$$\exists C > 0 \text{ such that } \forall t > 0, E(t) \le C\exp(-\omega(\sigma)t),$$

for all solutions. Furthermore, it holds that

$$S(\sigma) = \frac{1}{c_L} I_\sigma = \frac{1}{2}\omega(\sigma), \tag{4.23}$$

and the best constant $C(\omega(\sigma))$ *satisfies*

$$C(\omega(\sigma)) \leq \exp(4\|\theta\|_\infty).$$

Proof. The proof uses the explicit isomorphism $\mathcal{I}$ in (4.22) following by the proof in [46]. Given $(U_0, V_0) \in [L^2(I)]^2$, we expand it in the basis $(U_k, V_k) : (U_0, V_0) = \sum a_k (U_k, V_k)$. Then we have that

$$2E_0 = \|(U_0, V_0)\|^2_{[L^2(I)]^2} \geq \|\mathcal{I}\|^{-2} \|\mathcal{I}(U_0, V_0)\|^2_{L^2(-c_L, c_L)} \geq 2c_L \|\mathcal{I}\|^{-2} \sum |a_k|^2.$$

It easily to check $(U(t), V(t))$ can be expressed by

$$(U(t), V(t)) = \sum a_k \exp(-\lambda_k t)(U_k, V_k),$$

and also obtain that

$$\|\mathcal{I}(U(t), V(t))\|^2_{L^2(-c_L, c_L)} = 2c_L \exp(-2c_L^{-1} t I_\sigma) \sum |a_k|^2.$$

But

$$2E(t) = \|(U(t), V(t))\|^2_{[L^2(I)]^2} \leq \|\mathcal{I}^{-1}\|^2 \|\mathcal{I}(U(t), V(t))\|^2_{L^2(-c_L, c_L)}.$$

We combine the equalities to get

$$E(t) \leq \|\mathcal{I}\|^2 \|\mathcal{I}^{-1}\|^2 \exp(-2c_L^{-1} t I_\sigma) E_0. \tag{4.24}$$

From (4.24) we have $C(\omega(\sigma)) \leq \kappa(\mathcal{I})^2$, where $\kappa(\mathcal{I})$ is the conditioning number $\kappa(\mathcal{I}) = \|\mathcal{I}\| \|\mathcal{I}^{-1}\|$. Applying Parseval's identity to (4.21) we have

$$\|\mathcal{I}(g, h)\|^2_{L^2(-c_L, c_L)} = 2c_L \sum |a_k|^2 = \int_I |g(x)|^2 \exp(2\theta(x)) dx + \int_I |h(x)|^2 \exp(-2\theta(x)) dx.$$

As a consequence, we have that

$$\exp(-2\|\theta\|_\infty) \|(g, h)\|^2_{[L^2(I)]^2} = \exp(-2\|\theta\|_\infty) \int_I \left(|g(x)|^2 + |h(x)|^2 \right) dx$$

$$\leq \|\mathcal{I}((g,h))\|^2_{L^2(-c_L,c_L)}$$

$$\leq \exp(2\|\theta\|_\infty)\|(g,h)\|^2_{[L^2(I)]^2}.$$

Therefore,

$$\|\mathcal{I}\|^2 \leq \exp(2\|\theta\|_\infty), \quad \|\mathcal{I}^{-1}\|^2 \leq \exp(2\|\theta\|_\infty),$$

and

$$C(\omega(\sigma)) \leq \exp(4\|\theta\|_\infty).$$

The equalities (4.23) are shown in the proof. □

Remark 4.1 *Applying the perturbation B to the operator M, i.e., $M = A+B$ the question for the energy decay is unsolved. This remains for further research.*

Next we investigate the energy decay on a computational interval of the 1-d wave equation in unbounded domain following [24]. The result of the following Chapter is presented in [24], and we show detail proof.

4.2. Energy Decay of 1-D Acoustic Wave Equation

We consider the acoustic wave equation with variable sound speed in 1 dimension,

$$u_{tt} = c^2 u_{xx}, \quad \infty < x < \infty. \tag{4.25}$$

Let $[-a, a]$ be a computational interval and $c(x) \equiv 1$ in $\mathbb{R}\backslash[-a+\delta_0, a-\delta_0]$ for small $\delta_0 > 0$. We define the energy of the solution over $[-a, a]$ given by

$$E_{[-a,a]}(t) = \frac{1}{2}\int_{-a}^{a}\left(\frac{1}{c^2}u_t^2(x,t) + u_x^2(x,t)\right)dx.$$

The next Lemma explains that the energy over $[-a, a]$ is exponentially decaying in time.

Lemma 4.3 *The energy $E_{[-a,a]}(t)$ of the solution of the equation* (4.25) *is exponentially decreasing.*

Proof. The constant sound speed $c(x) \equiv 1$ provides the non-reflecting boundary condition at $x = \pm\, a$, by *d'Alembert's solution*,

$$u_t(-a,t) = u_x(-a,t), \quad u_t(a,t) = -u_x(a,t). \tag{4.26}$$

From the definition of $E(t)_{[-a,a]}$, we obtain

$$\begin{aligned}
\frac{d}{dt}E_{[-a,a]}(t) &= \int_{-a}^{a} \left(u_x u_{xt} + \frac{1}{c^2} u_t u_{tt} \right) dx \\
&= \int_{-a}^{a} \left(u_x u_{xt} + u_t u_{xx} \right) dx \\
&= \int_{-a}^{a} \left(u_x u_t \right)_t dx \\
&= -u_t^2(a,t) - u_t^2(-a,t) \\
&\leq 0,
\end{aligned}$$

by the boundary condition (4.26). Thus $E_{[-a,a]}(0) = E_{[-a,a]}(t_0) + \int_0^{t_0} \left(u_t^2(-a,t) + u_t^2(a,t) \right) dt$ for some $t_0 > 0$.

Next, we observe that it is sufficient to show $E_{[-a,a]}(t_0) \leq C_0 \int_0^{t_0} \left(u_t^2(-a,t) + u_t^2(a,t) \right) dt$ for some constant $C_0 > 0$ and some time $t_0 > 0$ to obtain

$$E_{[-a,a]}(t_0) \leq (1 + \frac{1}{C_0})^{-1} E_{[-a,a]}(0),$$

for all solutions of (4.25). It provides $E_{[-a,a]}(kt_0) \leq (1 + \frac{1}{C})^{-k} E_{[-a,a]}(0)$ for any $k \in \mathbb{N}$, so that the energy decays exponentially. To show the above bounds, let $\alpha, \beta : [-a,a] \to \mathbb{R}$ be curves such that

$$\alpha(x) = \frac{1}{c_*}(x+a), \quad \text{and} \quad \beta(x) = t_0 - \frac{1}{c_*}(x+a) \quad \forall x \in [-a,a],$$

for some $t_0 > 4a/c_*$ where c_* is given in (4.3).

We define $F(x)$ given by

$$F(x) = \frac{1}{2} \int_{\alpha(x)}^{\beta(x)} \left(u_x(x,t)^2 + \frac{1}{c^2} u_t^2(x,t) \right) dt.$$

Then we have

$$\frac{d}{dx}F(x) = \frac{1}{2}\left(u_x^2(x,t) + \frac{1}{c^2}u_t^2(x,t)\right)\Big|_{t=\beta(x)} \cdot \beta'(x) - \frac{1}{2}\left(u_x^2(x,t) + \frac{1}{c^2}u_t^2(x,t)\right)\Big|_{t=\alpha(x)} \cdot \alpha'(x)$$
$$+ \int_{\alpha(x)}^{\beta(x)} \left(u_x(x,t)u_{xx}(x,t) + \frac{1}{2}\left(\frac{1}{c^2}\right)' u_t^2(x,t) + \frac{1}{c^2}u_t(x,t)u_{tx}(x,t)\right) dt.$$

But $u_x u_{xx} + \frac{1}{c^2}u_t u_{tx} = \frac{1}{c^2}u_x u_{tt} + \frac{1}{c^2}u_t u_{tx} = \frac{1}{c^2}(u_x u_t)_t$, thus we obtain

$$\begin{aligned}
\frac{d}{dx}F(x) &\leq \frac{1}{2}\left(u_x(x,t) - \frac{1}{c}u_t(x,t)\right)^2\Big|_{t=\beta(x)} \cdot \left(-\frac{1}{c}\right) - \frac{1}{2}\left(u_x(x,t) + \frac{1}{c}u_t(x,t)\right)^2\Big|_{t=\alpha(x)} \cdot \left(\frac{1}{c}\right) \\
&\quad + \frac{1}{2}\int_{\alpha(x)}^{\beta(x)} \left(\frac{1}{c^2}\right)' u_t^2(x,t)dt \\
&\leq \frac{1}{2}\int_{\alpha(x)}^{\beta(x)} \left(\frac{1}{c^2}\right)' u_t^2(x,t)dt \\
&\leq C_F F(x),
\end{aligned}$$

for some $C_F > 0$. By *Gronwall's inequality* we get

$$F(x) \leq F(-a)e^{C_F(x+a)}.$$

Then we have

$$\begin{aligned}
\frac{1}{2}\int_{-a}^{a}\int_{\alpha(x)}^{\beta(x)} \left(u_x(x,t)^2 + \frac{1}{c^2}u_t^2(x,t)\right) dtdx &= \int_{-a}^{a} F(x)dx \\
&\leq \int_{-a}^{a} F(-a)e^{C_F(x+a)}dx \\
&\leq F(-a)\int_{-a}^{a} e^{C_F(x+a)}dx \\
&\leq C'\int_0^{t_0} u_t^2(-a,t)dt,
\end{aligned}$$

using $F(-a) = \frac{1}{2}\int_0^{t_0}\left(u_x^2(-a,t) + \frac{1}{c^2}u_t^2(-a,t)\right)dt = \int_0^{t_0} u_t^2(-a,t)dt$, where $C' = \int_{-a}^{a} e^{C_F(x+a)}dx$. For some $\delta > 0$ such that $\delta E_{[-a,a]}(t_0) \leq \frac{1}{2}\int_{-a}^{a}\int_{\alpha(x)}^{\beta(x)}\left(u_x(x,t)^2 + \frac{1}{c^2}u_t^2(x,t)\right)dtdx$, we obtain the bounds,

$$E_{[-a,a]}(t_0) \leq C_0\int_0^{t_0}\left(u_t^2(-a,t) + u_t^2(a,t)\right)dt \text{ for } C_0 = C'/\delta,$$

which completes the proof. □

We apply Lemma 4.3 to the 1-d PML wave equation to get the energy decay result. A similar argument is claimed in [24], we provide a detailed proof of the claim in the 1-d PML wave equation with variable sound speed.

4.3. Energy Decay 1-d PML Wave Equation

We present similar arguments to obtain the energy decay. First consider the 1-d PML wave equation (4.5) with variable sound speed in $I := [-a-L, a+L]$,

$$\begin{cases} \frac{\partial}{\partial t}P(x,t) + \sigma(x)P(x,t) \ + \frac{\partial}{\partial x}(c(x)Q(x,t)) \ = \ 0, \quad t>0, \\ \frac{\partial}{\partial t}Q(x,t) + \ \sigma(x)Q(x,t) \ + \ c(x)\frac{\partial}{\partial x}P(x,t) \ = \ 0, \quad t>0, \end{cases} \tag{4.27}$$

with the boundary condition $P(x,t) = 0$ at $x \in \partial I$.

Recall that the energy $E_I(t) := E(P(t), Q(t))$ defined in (4.10) over I provides that

$$\frac{d}{dt}E_I(t) = -\int_I \sigma(x)(P^2(x,t) + Q^2(x,t))dx \leq 0.$$

Thus, for some $t_0 > 0$

$$E_I(0) = E_I(t_0) + \int_0^{t_0}\int_I \sigma(x)\left(P^2(x,t) + Q^2(x,t)\right)dxdt. \tag{4.28}$$

Lemma 4.4 *The energy $E_I(t)$ of the solution in (4.27) over I decays exponentially.*

Proof. In a similar way to the proof of Lemma 4.3, it is sufficient to show that

$$E_I(t_0) \leq C\int_0^{t_0}\int_I \sigma(x)\left(P^2(x,t) + Q^2(x,t)\right)dxdt$$

for some $C > 0$. To show this, let us define

$$F(x) = \frac{1}{2}\int_{\alpha(x)}^{\beta(x)}\left(P^2(x,t) + Q^2(x,t)\right)dt,$$

where $\alpha, \beta : I \to \mathbb{R}$ are curves satisfying that

$$\alpha(x) = \frac{1}{c_*}(x+a+L), \quad \text{and} \quad \beta(x) = t_0 - \frac{1}{c_*}(x+a+L) \ \ \forall x \in I,$$

for some $t_0 > \frac{4}{c_*}(a+L)$ where c_* is given in (4.3).

Then we have

$$\begin{aligned}\frac{d}{dx}F(x) &= \frac{1}{2}\left(P^2+Q^2\right)\big|_{t=\beta}\cdot\beta' - \frac{1}{2}\left(P^2+Q^2\right)\big|_{t=\alpha}\cdot\alpha' + \int_\alpha^\beta (PP_x+QQ_x)\,dt.\\ &\le \frac{1}{2}\left(P+Q\right)^2\Big|_{t=\beta}\cdot\left(-\frac{1}{c}\right) - \frac{1}{2}\left(P-Q\right)^2\Big|_{t=\alpha}\cdot\left(\frac{1}{c}\right) - \frac{1}{c}\int_\alpha^\beta \left(2\sigma PQ + c'Q^2\right)dt,\end{aligned}$$

since $PP_x+QQ_x = P\cdot\frac{-Q_t-\sigma Q}{c} + Q\cdot\frac{-P_t-\sigma P-c'Q}{c} = -\frac{1}{c}(PQ)_t - \frac{1}{c}(2\sigma PQ + c'Q^2)$. By the *Cauchy-Schwarz inequality* that $2PQ \le P^2+Q^2$,

$$\frac{d}{dx}F(x) \le \frac{1}{c}\left(\sigma + |c'|\right)F(x),$$

thus *Gronwall's inequality* over $[\xi, x]$ gives that

$$F(x) \le CF(\xi) \quad \text{for all } x,\ \ \xi \le x \le a+L, \tag{4.29}$$

where $C = e^{\int_\xi^x \left(\frac{\sigma}{c}+\frac{|c'|}{c}\right)ds}$.

Let us divide I into $[-a-L, \mu_0]\cup[\mu_0, a+L]$ for some $\mu_0 \in (-a-L, a+L)$ such that $\sigma(\mu_0) = \sigma_\mu$. Note that $\sigma(x) \ge \sigma_\mu$ for $x \le \mu_0$ by the monotonicity of the damping (4.4). From (4.29) we obtain

$$\int_\xi^{a+L} F(x)dx \le C'F(\xi) \text{ for all } \xi,\ \ -a-L \le \xi \le \mu_0,$$

for $C' = \int_\xi^{a+L} e^{\int_\xi^x \left(\frac{\sigma}{c}+\frac{|c'|}{c}\right)ds}dx$. We take $\xi = \mu_0$ to obtain

$$\begin{aligned}\int_{\mu_0}^{a+L} F(x)dx \int_{-a-L}^{\mu_0}\sigma(\xi)d\xi &\le C'F(\mu_0)\int_{-a-L}^{\mu_0}\sigma(\xi)d\xi\\ &\le C'\int_{-a-L}^{\mu_0} F(\mu_0)\sigma(\xi)d\xi\\ &\le C'\int_{-a-L}^{\mu_0}\int_{\alpha(\mu_0)}^{\beta(\mu_0)}\sigma(\xi)\left(P^2(\mu_0,t)+Q^2(\mu_0,t)\right)dtd\xi\\ &\le C'\int_0^{t_0}\int_I \sigma(x)\left(P^2(x,t)+Q^2(x,t)\right)dxdt,\end{aligned}$$

by Fubini's Theorem. Thus $\int_{\mu_0}^{a+L}F(x)dx \le C\int_0^{t_0}\int_I \sigma(x)\left(P^2(x,t)+Q^2(x,t)\right)dxdt$ for $C = \frac{C'}{\int_{-a-L}^{\mu_0}\sigma(\xi)d\xi}$.

Therefore,

$$\begin{aligned}\int_I F(x)dx &= \int_{-a-L}^{\mu_0} F(x)dx + \int_{\mu_0}^{a+L} F(x)dx \\ &\leq \frac{1}{\sigma_\mu}\int_{-a-L}^{\mu_0} \sigma(x)F(x)dx + \int_{\mu_0}^{a+L} F(x)dx \\ &\leq C_0 \int_0^{t_0}\int_I \sigma(x)\left(P^2(x,t)+Q^2(x,t)\right)dxdt,\end{aligned}$$

where $C_0 = \frac{1}{\sigma_\mu} + C'$. For $\delta > 0$ satisfying $\delta E_I(t_0) \leq \frac{1}{2}\int_I \int_{\alpha(x)}^{\beta(x)} \left(P^2(x,t)+Q^2(x,t)\right)dtdx$, we obtain the bounds from (4.28)

$$E_I(t_0) \leq (1+\frac{1}{C})^{-1}E_I(0),$$

for $C = C_0/\delta$, which guarantees the exponential decay of the energy. □

Remark 4.2 *We show the exponential decay of the energy of 1-d PML wave equation with variable sound speed, but the actual decay rate remains a question. question for further investigation.*

5. WAVE EQUATION SYSTEM WITH DAMPING

In this section, we introduce a first order system of acoustic wave equation with zero order damping. We consider the system of the acoustic wave equation in $\mathbb{R}^2 \times I$, where $I = (0, T]$ for some $T > 0$,

$$\begin{aligned} \tfrac{1}{c(x)^2} p_t(x,t) + \nabla \cdot \vec{\mathbf{q}}(x,t) &= 0 \quad \text{in } \mathbb{R}^2 \times I, \\ \mathbf{q}_t(x,t) \ + \ \nabla p(x,t) &= \mathbf{0} \quad \text{in } \mathbb{R}^2 \times I, \end{aligned} \tag{5.1}$$

with $p(\cdot, 0) = f,\ \vec{\mathbf{q}}(\cdot, 0) = \vec{\mathbf{0}}$.

Let $supp(f) \subset \Omega_0 \subset \mathbb{R}^2$ be a bounded *Lipschitz* domain, expand the domain to $\Omega = \Omega_0 \cup \Omega_\gamma$, where Ω_γ is the set of layers surrounding Ω_0. We introduce damping terms σ_p and σ_q in the variables p and $\vec{\mathbf{q}}$, respectively, in the system (5.1), which drives a new damped wave equation in Ω,

$$\begin{cases} \tfrac{1}{c(x)^2} p_t(x,t) + \tfrac{1}{c(x)^2} \sigma_p(x) p(x,t) + \nabla \cdot \vec{\mathbf{q}}(x,t) &= 0 \quad \text{in } \Omega \times I, \\ \vec{\mathbf{q}}_t(x,t) \ + \ \sigma_q(x) \vec{\mathbf{q}}(x,t) \ + \ \nabla p(x,t) &= \vec{\mathbf{0}} \quad \text{in } \Omega \times I, \end{cases} \tag{5.2}$$

with the initial condition $(p, \vec{\mathbf{q}})(x, 0) = (p_0, \vec{\mathbf{q}}_0)$ at $t = 0$, and the zero *Dirichlet* boundary condition $p(x, \cdot) = 0$ on $\partial\Omega$, where the damping terms $\sigma_p, \sigma_q \in L^\infty(\Omega)$ satisfy

$$\begin{aligned} 0 = \sigma_* \le \sigma_p(x), \sigma_q(x) \le \sigma^* \ \text{ if } x \in \Omega, \\ \sigma_p(x) \equiv 0, \quad \sigma_q(x) \equiv 0 \quad \text{if } x \in \Omega_0. \end{aligned} \tag{5.3}$$

Here, the initial conditions are given by $p_0 = f \in H_0^1(\Omega),\ \vec{\mathbf{q}}_0 = \vec{\mathbf{0}}$.

5.1. Well-posedness of the System

In this section we present the well-posedness of the system (5.2). First we assume $c(x) \in C^1(\Omega)$ is bounded below by c_* and above by c^*, i.e.,

$$0 < c_* \le c(x) \le c^* < \infty \text{ in } \Omega, \tag{5.4}$$

$$c(x) \equiv 1 \text{ in } \Omega_\gamma.$$

We can write the two equations in (5.2) in matrix form,

$$\begin{bmatrix} \frac{1}{c^2} & \vec{\mathbf{0}}^T \\ 0 & \mathbf{I} \end{bmatrix} \begin{pmatrix} p_t \\ \vec{\mathbf{q}}_t \end{pmatrix} + \begin{bmatrix} \frac{1}{c^2}\sigma_p & \nabla\cdot \\ \nabla & \sigma_q \mathbf{I} \end{bmatrix} \begin{pmatrix} p \\ \vec{\mathbf{q}} \end{pmatrix} = \begin{pmatrix} 0 \\ \vec{\mathbf{0}} \end{pmatrix}.$$

To show well-posedness of the system we introduce following definitions. Let V_m be a Hilbert space with scalar-product $(\cdot,\cdot)_m$ and denote the corresponding Riesz map from V_m onto the dual V_m' by $\mathcal{M}$. That is,

$$\mathcal{M}u(v) = (u,v)_m, \quad u, v \in V_m$$

Let D be a subspace of V_m, and let $L : D \longrightarrow V_m'$ a linear map.

Definition 5.1 *The linear operator $L : D \longrightarrow V_m'$ is monotone(or non-negative) if*

$$Re\ Lu(u) \geq 0, \quad \forall u \in D.$$

We use the following theorem to show the existence of a solution $(p, \vec{\mathbf{q}})$.

Theorem 5.1 [39] *Assume that L is monotone and $\mathcal{M} + L : D \longrightarrow V_m'$ is surjective. Then, for every $g \in C^1([0,\infty); V_m')$ and $u_0 \in D$, there is a unique $u \in C^1([0,\infty); V_m)$ such that $u(0) = u_0$ and*

$$\mathcal{M}u'(t) + Lu(t) = g(t),\ t \geq 0.$$

We apply Theorem 5.1 to obtain well-posedness of the system (5.2). Let $\mathbb{L}^2_{div}(\Omega) = \{\vec{\mathbf{v}} \in \mathbb{L}^2(\Omega) : \nabla \cdot \vec{\mathbf{v}} \in L^2(\Omega)\}$ and $\mathbb{L}^2(\Omega) = [L^2(\Omega)]^2$.

Theorem 5.2 *For every $(p_0, \vec{\mathbf{q}}_0) \in H_0^1(\Omega) \times \mathbb{L}^2_{div}(\Omega)$ there exists a unique solution $(p, \vec{\mathbf{q}})$ of (5.2) such that $(p, \vec{\mathbf{q}}) \in C^1(\bar{I}; L^2(\Omega) \times \mathbb{L}^2(\Omega)) \cap C(\bar{I}; H_0^1(\Omega) \times \mathbb{L}^2_{div}(\Omega))$ satisfying the initial condition $(p(0), \vec{\mathbf{q}}(0)) = (p_0, \vec{\mathbf{q}}_0)$.*

Proof. Let $V_m := H_m(\Omega) \times \mathbb{L}^2(\Omega)$ where $H_m(\Omega) = \frac{1}{c^2}L^2(\Omega)$ with c^{-2}-weighted L^2-inner product, i.e.,

$$(p, r)_m = (p, r)_{c^{-2}} = \int_\Omega \frac{1}{c^2} p(x) r(x) d\vec{x},$$

and let $D := H_0^1(\Omega) \times \mathbb{L}^2_{div}(\Omega)$.

Then we know that $\frac{1}{c^2}L^2(\Omega) \cong L^2(\Omega)$, and let $\mathcal{M} : V_m \longrightarrow V_m'$ and defined by, for $(p, \vec{\mathbf{q}})^T \in V_m$, and $(r, \vec{\mathbf{v}})^T \in V_m$,

$$\mathcal{M}(p, \vec{\mathbf{q}})^T \left((r, \vec{\mathbf{v}})^T\right) = (p, r)_{c^{-2}} + (\vec{\mathbf{q}}, \vec{\mathbf{v}}),$$

where $(\cdot, \cdot)$ is the L^2-inner product. Let $L : D \longrightarrow V_m'$ is defined by, for any $(p, \vec{\mathbf{q}})^T \in D$ and $(r, \vec{\mathbf{v}})^T \in V_m$,

$$L\,(p, \vec{\mathbf{q}})^T \left((r, \vec{\mathbf{v}})^T\right) = (\sigma_p p, r)_{c^{-2}} + (\sigma_q \vec{\mathbf{q}}, \vec{\mathbf{v}}) + (\nabla p, \vec{\mathbf{v}}) + (\nabla \cdot \vec{\mathbf{q}}, r).$$

From the definition of L,

$$L\,(p, \vec{\mathbf{q}})^T \left((p, \vec{\mathbf{q}})^T\right) = \int_\Omega \left(\frac{\sigma_p}{c^2} p(x)^2 + \sigma_q \vec{\mathbf{q}}^2(x)\right) dx \geq 0,$$

since $\vec{\mathbf{q}} \cdot \vec{\mathbf{n}}(\gamma p) = 0$ where $\gamma : H^1(\Omega) \longrightarrow H^{\frac{1}{2}}(\partial\Omega)$ is the trace map and $\vec{\mathbf{n}}$ is the unit outward norm on $\partial\Omega$, so L is monotone.

To show the surjection, that is, $Rg(\mathcal{M} + L)|_D = V_m'$ define operators

$$\mathcal{A} = \frac{1 + \sigma_p}{c^2} : L^2(\Omega) \longrightarrow L^2(\Omega)',$$

$$\mathcal{B} = \nabla : H_0^1(\Omega) \longrightarrow \mathbb{L}^2(\Omega)',$$

and

$$\mathcal{C} = (1 + \sigma_q)\mathbf{I} : \mathbb{L}^2(\Omega) \longrightarrow \mathbb{L}^2(\Omega)' :$$

it must be shown that for all $(f, \vec{\mathbf{g}})^T \in V_m'$,

$$\exists \begin{pmatrix} p \\ \vec{\mathbf{q}} \end{pmatrix} \in D : \begin{bmatrix} \mathcal{A} & -\mathcal{B}' \\ \mathcal{B} & \mathcal{C} \end{bmatrix} \begin{pmatrix} p \\ \vec{\mathbf{q}} \end{pmatrix} = \begin{pmatrix} f \\ \vec{\mathbf{g}} \end{pmatrix},$$

where $\mathcal{B}' = -\nabla\cdot : \mathbb{L}^2(\Omega) \longrightarrow H^{-1}(\Omega)$ is the dual of $\mathcal{B}$.

We know that it is equivalent to show that there exist $p \in H_0^1(\Omega)$ and $\vec{\mathbf{q}} \in \mathbb{L}^2_{div}(\Omega)$ such that

$$\mathcal{A}p \;-\; \mathcal{B}'\vec{\mathbf{q}} \;=\; f \text{ in } L^2(\Omega)',$$

and

$$\mathcal{B}p \;+\; \mathcal{C}\vec{\mathbf{q}} \;=\; \vec{\mathbf{g}} \text{ in } \mathbb{L}^2(\Omega)'.$$

Equivalently there exists $p \in H_0^1(\Omega)$ such that

$$\mathcal{A}p \;+\; \mathcal{B}'\mathcal{C}^{-1}(\mathcal{B}p \;-\; \vec{\mathbf{g}}) \;=\; f \;\text{ in }\; L^2(\Omega)', \tag{5.5}$$

since $\mathcal{C}$ is bounded below by 1 and $\vec{\mathbf{q}} = \mathcal{C}^{-1}(\vec{\mathbf{g}} - \mathcal{B}p)$. By the definitions of all operators (5.5) is satisfied if we show that there exists $p \in H_0^1(\Omega)$ such that

$$\mathcal{A}p \;+\; \mathcal{B}'\mathcal{C}^{-1}\mathcal{B}p \;=\; \mathcal{B}'\mathcal{C}^{-1}\vec{\mathbf{g}} \;+\; f \text{ in } H^{-1}(\Omega),$$

or equivalently,

$$\mathcal{A}p(r) \;+\; (\mathcal{C}^{-1}\mathcal{B}p, \mathcal{B}r) \;=\; (\mathcal{C}^{-1}\vec{\mathbf{g}}, \mathcal{B}r) \;+\; f(r) \;\; \forall r \in H_0^1(\Omega).$$

The existence of p is guaranteed by coercivity with the constant $\mathsf{C}_\sigma^{-1} = \max\{c^{*2}, 1+\sigma^*\}$ from the following elliptic form. It can be checked that

$$\int_\Omega \left(\frac{1+\sigma_p}{c^2} p(\vec{x})^2 + \nabla p(x)^2 \right) d\vec{x} + \int_\Omega \frac{1}{1+\sigma_q} \nabla p(\vec{x})^2 d\vec{x} \geq \frac{1}{c^{*2}} \int_\Omega p(\vec{x})^2 d\vec{x} + \frac{1}{1+\sigma^*} \int_\Omega \nabla p(\vec{x})^2 d\vec{x}$$
$$\geq \mathsf{C}_\sigma \|p\|^2_{H_0^1(\Omega)},$$

It also follows that $\vec{\mathbf{q}} = \mathcal{C}^{-1}(\vec{\mathbf{g}} - \mathcal{B}p) \in \mathbb{L}^2(\Omega)$ from $p \in H_0^1(\Omega)$. □

Remark 5.1 *1. The solution* $(p, \vec{\mathbf{q}})$ *in* (5.2) *satisfies*

$$\begin{aligned} (\tfrac{1}{c^2} p_t, r) + (\tfrac{1}{c^2}\sigma_p p, r) + (\nabla\cdot\vec{\mathbf{q}}, r) \;&=\; 0 \quad \forall\, r \in H_0^1(\Omega),\ \forall t \text{ in } I, \\ (\vec{\mathbf{q}}_t, \vec{\mathbf{v}}) \;+\; (\sigma_q\vec{\mathbf{q}}, \vec{\mathbf{v}}) \;+\; (\nabla p, \vec{\mathbf{v}}) \;&=\; 0 \quad \forall\, \vec{\mathbf{v}} \in \mathbb{L}^2_{div}(\Omega),\ \forall t \text{ in } I. \end{aligned}$$

2. *The system* (5.2) *with the time-dependent damping terms* $\sigma_p(x,t)$ *and* $\sigma_q(x,t)$ *can be considered, and the well-posedness with the initial condition* $(p_0, \vec{\mathbf{q}}_0)$ *in* $D(L)$ *is obtained from Theorem 4.10, page 245 in* [42].

5.2. Discontinous Galerkin discretization

In this section, we present the Discontinuous Galerkin(DG) method for the system (5.2) of the damped wave equation. The spatial discretization is based on the DG method presented in [4] while the time discretization is based on θ-method with $\theta = \frac{1}{2}$ which is presented in the next section.

The DG methods are locally conservative, stable, and high-order accurate methods which can be easily handled with complex geometric domains, irregular meshes with hanging nodes, and approximations that have polynomials of different degrees in different elements [10].

5.2.1 Spatial Discretization.

We assume that shape-regular meshes $\mathcal{T}_h$ that partition the domain Ω into disjoint elements $\{K\}$ such that $\bar{\Omega} = \cup_{K\in\mathcal{T}_h}\bar{K}$. Thus if $K \in \mathcal{T}_h$, then K is a simplex, i.e., K is a segment if $d = 1$, a triangle or a parallelogram if $d = 2$, and a tetrahedron or a parallelepiped if $d = 3$. The measure of K (length if $d = 1$, area if $d = 2$, and volume if $d = 3$) is denoted by $meas(K)$. It will always be assumed that $meas(K) \neq 0$. The diameters of K and that the largest ball included in K are denoted by h_K and ρ_K, respectively. The ratio of these two quantities is denoted by φ_K. Hence,

$$\varphi_K = \frac{h_K}{\rho_K}, \quad h_K = diam(K), \quad \rho_K = \sup\{r : B_r = \{x : |x-a| \leq r\} \subset K,\ a \in K\}.$$

Note that $\rho_K > 1$. For a family of $\{\mathcal{T}_h\}_{h>0}$, the parameter h refers to

$$h = \max_{K\in\mathcal{T}_h} h_K.$$

We also give the definition for the asymptotic behavior of the family of meshes $\{\mathcal{T}_h\}_{h>0}$.

Definition 5.2 (Shape-regularity) *A family of meshes $\{\mathcal{T}_h\}_{h>0}$ is said to be shape − regular if there exists φ_0 such that*

$$\forall h > 0, \quad \forall K \in \mathcal{T}_h, \quad \varphi_K = \frac{h_K}{\rho_K} \leq \varphi_0.$$

For example, in two dimensions, the triangles in a shape-regular family of triangulations cannot become too flat as $h \to 0$. Generally, it is allowed for irregular meshes with hanging nodes. Here the concept of hanging nodes is that a vertex of an element K^+ belongs to the interior of an edge of another element K^- in the sense that it is a nontrivial convex combination of the end points of K^+. But we don't discuss hanging nodes in detail and avoid the complicated mesh.

However, we assume that the local mesh sizes are of bounded variation; that is, there is a positive constant κ, depending only on the shape-regularity of the mesh, such that

$$\kappa h_K \leq h_{K'} \leq \kappa^{-1} h_K \tag{5.6}$$

for all neighboring elements K and K'. From each adjacent element K^+ and K^- in $\mathcal{T}_h$, we denote the set of all faces by $\mathcal{E}_h$, which consists of both $\mathcal{E}_h^I$ the set of all interior faces of $\partial K^+ \cap \partial K^- \in \mathcal{E}(K^+) \cup \mathcal{E}(K^-)$ and $\mathcal{E}_h^B$ the set of all boundary faces of $\partial K \cap \partial\Omega$, i.e., $\mathcal{E}_h = \mathcal{E}_h^I \cup \mathcal{E}_h^B$, where $\mathcal{E}(K)$ is denoted by the set of all edges of the element K.

For a piecewise smooth scalar-valued function p, define the trace operators on all faces. Let $e \in \mathcal{E}_h^I$ be an interior face shared by elements K^+ and K^-; let $\vec{\mathbf{n}}^{\pm}$ by the unit outward normal vectors on the boundaries $\partial K^{\pm}$ respectively. Denote by $p^{\pm}$ the trace of p taken from within $K^{\pm}$, we define the jump and average of p at $x \in e$ by

$$\{\!\{p\}\!\} := \frac{1}{2}(p^+ + p^-), \qquad [\![p]\!] := p^+ \vec{\mathbf{n}}^+ + p^- \vec{\mathbf{n}}^-. \tag{5.7}$$

Let $\{\!\{p\}\!\} := p$ and $[\![p]\!] := p\vec{\mathbf{n}}$ where $\vec{\mathbf{n}}$ is the unit outward normal vector on $\partial\Omega$ in all boundary faces $e \in \mathcal{E}_h^B$.

If $\vec{\mathbf{q}}$ is a vector-valued function, we set

$$\{\!\{\vec{\mathbf{q}}\}\!\} := \frac{1}{2}(\vec{\mathbf{q}}^+ + \vec{\mathbf{q}}^-), \qquad [\![\vec{\mathbf{q}}]\!] := \vec{\mathbf{q}}^+ \cdot \vec{\mathbf{n}}^+ + \vec{\mathbf{q}}^- \cdot \vec{\mathbf{n}}^-.$$

In a similar way we set $\{\!\{\vec{\mathbf{q}}\}\!\} := \vec{\mathbf{q}}$ and $[\![\vec{\mathbf{q}}]\!] := \vec{\mathbf{q}} \cdot \vec{\mathbf{n}}$ in all boundary faces $e \in \mathcal{E}_h^B$.

Notice that the jump $[\![p]\!]$ of the scalar function p is a vector parallel to $\vec{\mathbf{n}}$ and that $[\![\vec{\mathbf{q}}]\!]$ is the jump of the normal component of the vector function $\vec{\mathbf{q}}$ which is a scalar quantity. Note that there is a trace identity for a vector-valued function $\vec{\mathbf{q}}$ and a scalar-valued function p with continuous normal components across a face $e \in \mathcal{E}_h^I$, by applying the definitions directly one has,

$$p^+(\vec{\mathbf{n}}^+ \cdot \vec{\mathbf{q}}^+) + p^-(\vec{\mathbf{n}}^- \cdot \vec{\mathbf{q}}^-) = [\![p]\!] \cdot \{\!\{\vec{\mathbf{q}}\}\!\} + \{\!\{p\}\!\}[\![\vec{\mathbf{q}}]\!]. \tag{5.8}$$

For a given partition $\mathcal{T}_h$ such as triangulation of Ω and an approximation order $k \geq 1$, we seek an approximate (continuous or possibly discontinuous) solution $(p^h, \vec{\mathbf{q}}^h)$ which is in the finite element space

$$\mathcal{P}^h(\Omega) \times \mathcal{Q}^h(\Omega) = \bigcup_{K \in \mathcal{T}_h} \mathcal{P}^h(K) \times \mathcal{Q}^h(K),$$

where

$$\mathcal{P}^h(K) \times \mathcal{Q}^h(K)$$

$$:= \left\{ (p^h, \vec{\mathbf{q}}^h) \in L^2(K) \times \mathbb{L}^2(K) : (p^h, \vec{\mathbf{q}}^h)|_K \in \mathbb{P}^k(K) \times (\mathbb{P}^k(K))^2 \right\} \quad \forall K \in \mathcal{T}_h,$$

and $\mathbb{P}^k(K)$ is the space of polynomials of total degree at most k on K if K is a simplex. This approximation is said to be non-conformal since $\mathcal{P}^h(\Omega) \not\subset H_0^1(\Omega)$; it is said to be conformal otherwise, e.g. continuous Galerkin methods.

5.2.2 The DG methods

In this section, we define DG methods for the system (5.2) following [4]. We consider only the discretization of this equation in space in this section.

First we assume that $p^h : I \to \mathcal{P}^h(\Omega)$ is absolutely continuous. A DG numerical method is obtained as follows. We discretize the domain Ω, then seek a discontinuous approximate solution $(p^h, \vec{\mathbf{q}}^h)$ on the element K taken in the space $\mathcal{P}^h(K) \times \mathcal{Q}^h(K)$ and determined by requiring that

$$\int_K \frac{1}{c^2} p_t^h r^h \, dx + \int_K \frac{\sigma_p}{c^2} p^h r^h \, dx - \int_K \vec{\mathbf{q}}^h \cdot \nabla_h r^h dx + \int_{\partial K} (\hat{\mathbf{q}}^h \cdot \vec{\mathbf{n}}) r^h ds = 0 \tag{5.9}$$

$$\int_K \vec{\mathbf{q}}_t^h \cdot \mathbf{v}^h dx + \int_K \sigma_q \vec{\mathbf{q}}^h \cdot \vec{\mathbf{v}}^h dx - \int_K p^h \nabla_h \cdot \vec{\mathbf{v}}^h dx + \int_{\partial K} \hat{p}^h (\vec{\mathbf{v}}^h \cdot \vec{\mathbf{n}}) ds = 0 \tag{5.10}$$

for all $(r^h, \vec{\mathbf{v}}^h) \in \mathcal{P}^h(K) \times \mathcal{Q}^h(K)$, where ∇_h and $\nabla_h \cdot$ are the functions whose restriction to each element $K \in \mathcal{T}_h$ are equal to ∇ and $\nabla \cdot$, respectively. To complete the definition of the DG method, it remains to define the two numerical traces, $\hat{p}^h$ and $\hat{\mathbf{q}}^h$. We first begin by finding a stability result for the solution in the original system (5.2). To do that, we multiply the first equation of the system (5.2) by p and integrate over $\Omega \times I,\ I = (0, T)$ to get

$$\frac{1}{2} \int_\Omega \frac{1}{c^2} p^2(\cdot, T) dx + \int_0^T \int_\Omega \frac{\sigma_p}{c^2} p^2 dx dt + \int_0^T \int_\Omega p \nabla \cdot \vec{\mathbf{q}} dx dt = \frac{1}{2} \int_\Omega p^2(\cdot, 0) dx.$$

Then, we multiply the second equation in (5.2) by $\vec{\mathbf{q}}$ and integrate over $\Omega \times I$ to obtain

$$\frac{1}{2} \int_\Omega |\vec{\mathbf{q}}(\cdot, T)|^2 dx + \int_0^T \int_\Omega \sigma_q |\vec{\mathbf{q}}|^2 dx dt + \int_0^T \int_\Omega \nabla p \cdot \vec{\mathbf{q}} dx dt = \frac{1}{2} \int_\Omega |\vec{\mathbf{q}}(\cdot, 0)|^2 dx.$$

Adding these two equations, we have

$$\frac{1}{2} \int_\Omega \left(\frac{1}{c^2} p^2(\cdot, T) + |\vec{\mathbf{q}}(\cdot, T)|^2 \right) dx + \int_0^T \int_\Omega \left(\frac{\sigma_p}{c^2} p^2 + \sigma_q |\vec{\mathbf{q}}|^2 \right) dx dt$$
$$= \frac{1}{2} \int_\Omega \left(p^2(\cdot, 0) + |\vec{\mathbf{q}}(\cdot, 0)|^2 \right) dx.$$

A stability result is immediately followed by this equation. Next, we imitate this procedure for the DG method under consideration.

We begin by taking $r^h = p^h$ in the equation (5.9) defining the DG method and adding over the elements K to get

$$\frac{1}{2} \int_\Omega \frac{1}{c^2} (p^h)^2(\cdot, T) dx + \int_0^T \int_\Omega \frac{\sigma_p}{c^2} (p^h)^2 dx dt$$

$$+\sum_{K\in\mathcal{T}_h}\int_0^T\int_{\partial K}\left(-\vec{\mathbf{q}}^h\cdot\vec{\mathbf{n}}+\hat{\mathbf{q}}^h\cdot\vec{\mathbf{n}}\right)p^h ds dt+\int_0^T\int_\Omega p^h\nabla\cdot\vec{\mathbf{q}}^h dx dt=\frac{1}{2}\int_\Omega (p^h)^2(\cdot,0)dx.$$

Next, we take $\vec{\mathbf{v}}^h=\vec{\mathbf{q}}^h$ in the equation (5.10) and add on the elements to obtain

$$\frac{1}{2}\int_\Omega|\vec{\mathbf{q}}^h(\cdot,T)|^2dx+\int_0^T\int_\Omega\sigma_q|\vec{\mathbf{q}}^h|^2dxdt-\int_0^T\int_\Omega p^h\nabla\cdot\vec{\mathbf{q}}^h dxdt+\sum_K\int_0^T\int_{\partial K}\hat{p}^h(\vec{\mathbf{q}}^h\cdot\vec{\mathbf{n}})\,dsdt$$

$$=\frac{1}{2}\int_\Omega(\vec{\mathbf{q}}^h)^2(\cdot,0)dx.$$

Summing the two equations above, we have that

$$\frac{1}{2}\int_\Omega\left(\frac{1}{c^2}(p^h)^2(\cdot,T)+|\vec{\mathbf{q}}^h(\cdot,T)|^2\right)dx+\int_0^T\int_\Omega\left(\frac{\sigma_p}{c^2}(p^h)^2+\sigma_q|\vec{\mathbf{q}}^h|^2\right)dxdt+\int_0^T\Theta_h dt$$

$$=\frac{1}{2}\int_\Omega\left((p^h)^2(\cdot,0)+(\vec{\mathbf{q}}^h)^2(\cdot,0)\right)dx,$$

where

$$\Theta_h(t)=\sum_{K\in\mathcal{T}_h}\int_{\partial K}\left(p^h\hat{\mathbf{q}}^h\cdot\vec{\mathbf{n}}+(\hat{p}^h-p^h)\vec{\mathbf{q}}^h\cdot\vec{\mathbf{n}}\right)ds.$$

Now we can define consistent numerical traces $\hat{p}^h$ and $\hat{\mathbf{q}}^h$ that provide the quantity $\Theta_h(t)$ non-negative.

Dropping the argument t, we obtain

$$\begin{aligned}\Theta_h&=\sum_{e\in\mathcal{E}_h}\int_e[\![p^h\hat{\mathbf{q}}^h+(\hat{p}^h-p^h)\vec{\mathbf{q}}^h]\!]ds\\&=\sum_{e\in\mathcal{E}_{ih}}\int_e\left([\![p^h]\!]\cdot\hat{\mathbf{q}}^h+\hat{p}^h[\![\vec{\mathbf{q}}^h]\!]-[\![p^h\vec{\mathbf{q}}^h]\!]\right)ds+\int_{\partial\Omega}\left(p^h\hat{\mathbf{q}}^h\cdot\vec{\mathbf{n}}+(\hat{p}^h-p_h)\vec{\mathbf{q}}^h\cdot\vec{\mathbf{n}}\right)ds\\&=\sum_{e\in\mathcal{E}_{ih}}\int_e[\![p^h]\!]\cdot\left(\hat{\mathbf{q}}^h-\{\!\{\vec{\mathbf{q}}^h\}\!\}\right)+[\![\mathbf{q}^h]\!]\left(\hat{p}^h-\{\!\{p^h\}\!\}\right)ds+\int_{\partial\Omega}\left(p^h(\hat{\mathbf{q}}^h-\vec{\mathbf{q}}^h)\cdot\vec{\mathbf{n}}+\hat{p}^h\vec{\mathbf{q}}^h\cdot\vec{\mathbf{n}}\right)ds.\end{aligned}$$

To get non-negative Θ_h, it is enough to take, on $\mathcal{E}_h^I$, i.e., inside the domain Ω,

$$\hat{p}^h=\{\!\{p^h\}\!\}+\mathtt{C}_{22}[\![\vec{\mathbf{q}}^h]\!]-\vec{\mathtt{C}}_{12}\cdot[\![p^h]\!],\quad\hat{\mathbf{q}}^h=\{\!\{\vec{\mathbf{q}}^h\}\!\}+\mathtt{C}_{11}[\![p^h]\!]+\vec{\mathtt{C}}_{12}[\![\vec{\mathbf{q}}^h]\!],$$

for some positive quantities,

$$\mathtt{C}_{11}>0,\mathtt{C}_{22}>0,\mathtt{C}_{11}^1>0,\mathtt{C}_{12}^2>0,\vec{\mathtt{C}}_{12}=\begin{bmatrix}\mathtt{C}_{12}^1 & \mathtt{C}_{12}^2\end{bmatrix}^T,\tag{5.11}$$

and on $\mathcal{E}_h^B$, i.e., its boundary,

$$\hat{p}_h = 0, \quad \hat{\mathbf{q}}^h = \vec{\mathbf{q}}^h + \mathtt{C}_{11} p^h \vec{\mathbf{n}},$$

to finally get

$$\Theta_h = \sum_{e \in \mathcal{E}_h^I} \int_e \left(\mathtt{C}_{11} [\![p^h]\!]^2 + \mathtt{C}_{22} [\![\vec{\mathbf{q}}^h]\!]^2 \right) dx + \sum_{\mathcal{E}_h^B} \int_e \mathtt{C}_{11} (p^h)^2 ds \geq 0.$$

As we can see the vector parameter $\vec{\mathtt{C}}_{12}$ does not have any stabilizing effect; it's not necessary for stability but could be used to enhance the accuracy of the method [6]. In the next section, we will discuss about the above non-negative quantities (5.11) which are necessary quantities to make the system stable. Note that the zero *Dirichlet* boundary condition is imposed weakly through the definition of the numerical trace.

Applying the numerical flux $\hat{p}^h$ and $\hat{\mathbf{q}}^h$ we have the DG system

$$\sum_{K \in \mathcal{T}_h} \int_K \frac{1}{c^2} p_t^h r^h \, dx + \sum_{K \in \mathcal{T}_h} \int_K \frac{\sigma_p}{c^2} p^h r^h dx - \sum_{K \in \mathcal{T}_h} \int_K \vec{\mathbf{q}}^h \cdot \nabla r^h \, dx$$
$$+ \sum_{e \in \mathcal{E}_h^I} \int_e \left(\{\!\{ \vec{\mathbf{q}}^h \}\!\} \cdot [\![r^h]\!] + \mathtt{C}_{11} [\![p^h]\!] \cdot [\![r^h]\!] + \vec{\mathtt{C}}_{12} [\![\mathbf{q}_h]\!] \cdot [\![r^h]\!] \right) ds + \int_{\partial\Omega} r^h \vec{\mathbf{q}}^h \cdot \vec{\mathbf{n}} ds = 0,$$

and

$$\sum_{K \in \mathcal{T}_h} \int_K \vec{\mathbf{q}}_t^h \cdot \vec{\mathbf{v}}^h \, dx + \sum_{K \in \mathcal{T}_h} \int_K \sigma_q \vec{\mathbf{q}}^h \cdot \vec{\mathbf{v}}^h \, dx - \sum_{K \in \mathcal{T}_h} \int_K p^h \nabla \cdot \vec{\mathbf{v}}^h \, dx$$
$$+ \sum_{e \in \mathcal{E}_h^I} \int_e \left(\{\!\{ p^h \}\!\} [\![\vec{\mathbf{v}}^h]\!] \; + \mathtt{C}_{22} [\![\vec{\mathbf{q}}^h]\!] [\![\vec{\mathbf{v}}^h]\!] \; - \vec{\mathtt{C}}_{12} \cdot [\![p^h]\!] [\![\vec{\mathbf{v}}^h]\!] \; \right) ds = 0,$$

for all $(r^h, \vec{\mathbf{v}}^h) \in \mathcal{P}^h(\Omega) \times \mathcal{Q}^h(\Omega)$. This completes the definition of DG method.

5.2.3 Some Properties

We show that the DG method is in fact a mixed formulation. To see this, let us begin by noting that the DG approximate solution $(p^h, \vec{\mathbf{q}}^h)$ can be characterized as the solution of

$$(\frac{1}{c^2} p_t^h, r^h) \; + \; a_h(p^h, r^h) \; - \; b_h'(\vec{\mathbf{q}}^h, r^h) \; = \; 0, \tag{5.12}$$

$$(\vec{\mathbf{q}}_t^h, \vec{\mathbf{v}}^h) \; + \; b_h(p^h, \vec{\mathbf{v}}^h) \; + \; c_h(\vec{\mathbf{q}}^h, \vec{\mathbf{v}}^h) \; = \; 0, \tag{5.13}$$

for all $(r^h, \vec{\mathbf{v}}^h) \in \mathcal{P}^h(\Omega) \times \mathcal{Q}^h(\Omega)$, where

$$a_h(p^h, r^h) = \sum_{K \in \mathcal{T}_h} \int_K \frac{\sigma_p}{c^2} p^h r^h \, dx + \sum_{e \in \mathcal{E}_h^I} \int_e \mathtt{C}_{11} [\![p^h]\!] \cdot [\![r^h]\!] ds, \tag{5.14}$$

$$b_h(p^h, \vec{\mathbf{v}}^h) = - \sum_{K \in \mathcal{T}_h} \int_K p^h \nabla \cdot \vec{\mathbf{v}}^h dx - \sum_{e \in \mathcal{E}_h^I} \int_e \left(\vec{\mathtt{C}}_{12} \cdot [\![p^h]\!] - \{\!\{ p^h \}\!\} \right) [\![\vec{\mathbf{v}}^h]\!] ds, \tag{5.15}$$

$$b'_h(\vec{\mathbf{q}}^h, r^h) = \sum_{K \in \mathcal{T}_h} \int_K \vec{\mathbf{q}}^h \cdot \nabla r^h dx - \sum_{e \in \mathcal{E}_h^I} \int_e \left(\vec{\mathtt{C}}_{12} [\![\vec{\mathbf{q}}^h]\!] + \{\!\{ \vec{\mathbf{q}}^h \}\!\} \right) \cdot [\![r^h]\!] ds, \tag{5.16}$$

$$c_h(\vec{\mathbf{q}}^h, \vec{\mathbf{v}}^h) = \sum_{K \in \mathcal{T}_h} \int_K \sigma_q \vec{\mathbf{q}}^h \cdot \vec{\mathbf{v}}^h dx + \sum_{e \in \mathcal{E}_h^I} \int_e \mathtt{C}_{22} [\![\vec{\mathbf{q}}^h]\!] [\![\vec{\mathbf{v}}^h]\!] ds. \tag{5.17}$$

Remark 5.2 *It holds the equality by the trace identity* (5.8).

$$b_h(p^h, \vec{\mathbf{q}}^h) = b'_h(\vec{\mathbf{q}}^h, p^h) \quad \textit{for all } p^h \in \mathcal{P}^h(\Omega),\ \vec{\mathbf{q}}^h \in \mathcal{Q}^h(\Omega).$$

Note that the second terms in (5.14)-(5.17) correspond to jump and average terms on element boundaries; they vanish when $p, r \in H_0^1(\Omega)$ and $\vec{\mathbf{q}}, \vec{\mathbf{v}} \in \mathbb{L}^2_{div}(\Omega)$. Therefore the above semi-discrete DG formulation (5.12), (5.13) is consistent with the original continuous problem (5.2).

5.3. *A Priori* Error Estimate of DG Method

5.3.1 Preliminaries.

In order to establish an error estimate we introduce the following properties. There is an important inequality in the finite element spaces $\mathcal{P}^h(\Omega)$ and $\mathcal{Q}^h(\Omega)$ which allow that H^1-norm can be bounded above by the L^2-norm. Such an inequality is called an inverse inequality.

Let us introduce the broken Sobolev space of $\mathcal{T}_h$ of the domain Ω,

$$H^s(\mathcal{T}_h) := \{ p \in L^2(\Omega) : p|_K \in H^s(K), \forall K \in \mathcal{T}_h \},$$

with the broken Sobolev norm and seminorm, respectively,

$$\|p\|_{H^s(\mathcal{T}_h)} := \left(\sum_{K \in \mathcal{T}_h} \|p\|_{H^s(K)} \right)^{\frac{1}{2}}, \ |p|_{H^s(\mathcal{T}_h)} := \left(\sum_{K \in \mathcal{T}_h} |p|_{H^s(K)} \right)^{\frac{1}{2}}.$$

The following local inverse inequality can be proved (Appendix A): Let $\{\mathcal{T}_h\}_{h>0}$ be a shape-regular family of meshes in $\mathbb{R}^d$. Then there exists a constant C, independent of h and K, such that, for all $p^h \in \mathbb{P}^k(K)$,

$$\|p^h\|_{1,K} \leq C h_K^{-1} \|p^h\|_{0,K}. \tag{5.18}$$

To obtain a global inverse inequality, that is an inequality not only valid in K but also in the whole domain Ω, the concept of quasi-uniform family of meshes is needed.

Definition 5.3 (Quasi-uniformity) *A family of meshes $\{\mathcal{T}_h\}_{h>0}$ is said to be quasi-uniform if it is shape-regular and there exists $\tau > 0$ such that*

$$\forall h > 0, \forall K \in \mathcal{T}_h, h_K \geq \tau h.$$

Then the following result can be proved (Appendix A): let $\{\mathcal{T}_h\}_{h>0}$ be a quasi-uniform family of affine meshes in $\mathbb{R}^d$, there exists a constant C such that for all $h > 0, K \in \mathcal{T}_h$ and $p^h \in \mathbb{P}^k(K)$

$$\sum_{K \in \mathcal{T}_h} \|p^h\|_{H^1(\mathcal{T}_h)} \leq C h^{-1} \sum_{K \in \mathcal{T}_h} \|p^h\|_{L^2(\mathcal{T}_h)}. \tag{5.19}$$

where the constant C which depends only on the shape regularity of the mesh, the approximation order k, and the dimension d.

Lemma 5.1 (Trace Theorem) *Let $p \in \mathcal{P}^h(\Omega)$ with shape regularity mesh. Then there exists a constant $C_{inv} > 0$ such that*

$$\|p\|_{L^2(\partial K)} \leq C_{inv} (\|p\|_{L^2(K)} (h_K^{-1} \|p\|_{L^2(K)} + \|\nabla p\|_{L^2(K)}))^{\frac{1}{2}}. \tag{5.20}$$

Proof. See Lemma A.3 in [43] for the proof and further details. □

The above bounds will be used in the error estimation later. Now we introduce the function $\mathtt{h}$ in $L^\infty(\mathcal{E}_h)$ related to the local mesh size as

$$\mathtt{h}|_e = \begin{cases} \min\{h_K, h_{K'}\} & \text{if } e \in \mathcal{E}_h^I, e = \partial K \cap \partial K', \\ h_K & \text{if } e \in \mathcal{E}_h^B, e = \partial K \cap \partial\Omega. \end{cases}$$

Then on each $e \in \mathcal{E}_h$, we define the discontinuity stabilization parameters $\alpha_{11} > 0, \alpha_{12} > 0, \alpha_{22} > 0$ in terms of $\mathtt{h}$ by

$$\mathtt{C}_{11} = \alpha_{11}\mathtt{h}^\alpha, \mathtt{C}_{22} = \alpha_{22}\mathtt{h}^\alpha, \vec{\mathtt{C}}_{12} = [\alpha_{12} \quad \alpha_{12}]^T, \tag{5.21}$$

with the parameters $\alpha_{ij}, i \le j, (i, j = 1, 2)$ independent of local mesh sizes. The accuracy of the method relies on the choice of α and we assume $\alpha = 0$. But the specific choice of the stabilization parameter, $\alpha = -1$, i.e., $\mathtt{C}_{11}, \mathtt{C}_{22} = \mathcal{O}(h^{-1})$, makes our DG method an Interior Penalty method (IP;[29]), which provides that the lifting operators (5.26), (5.27) are bounded (see the proof of Remark 5.4).

We now consider the following semi-discrete DG approximation for the spatial discretization of (5.2): Find $(p^h, \vec{\mathbf{q}}^h) : \bar{I} \times \bar{I} \to \mathcal{P}^h(\Omega) \times \mathcal{Q}^h(\Omega)$ such that

$$(\frac{1}{c^2}p_t^h, r^h) + a_h(p^h, r^h) - b_h'(\vec{\mathbf{q}}^h, r^h) \;=\; 0, \quad \forall r^h \in \mathcal{P}^h(\Omega),\ t \in I, \tag{5.22}$$

$$(\vec{\mathbf{q}}_t^h, \vec{\mathbf{v}}^h) + b_h(p^h, \vec{\mathbf{v}}^h) + c_h(\vec{\mathbf{q}}^h, \vec{\mathbf{v}}^h) \;=\; 0, \quad \forall \vec{\mathbf{v}}^h \in \mathcal{Q}^h(\Omega),\ t \in I, \tag{5.23}$$

with

$$p^h(\cdot, 0) \;=\; \Pi_h p_0, \quad \vec{\mathbf{q}}^h(\cdot, 0) \;=\; \mathbf{\Pi}_h \vec{\mathbf{q}}_0, \quad p^h(x, \cdot) \;=\; 0, \quad \forall x \in \partial\Omega.$$

Here Π_h and $\mathbf{\Pi}_h$ denotes the L^2-projections of p and $\vec{\mathbf{q}}$ in $L^2(\Omega)$ and $\mathbb{L}^2(\Omega)$ onto $\mathcal{P}^h(\Omega)$ and $\mathcal{Q}^h(\Omega)$ respectively, that is, for any $p \in L^2(\Omega), \vec{\mathbf{q}} \in \mathbb{L}^2(\Omega)$

$$(\Pi_h p, r^h) = (p, r^h) \text{ and } (\mathbf{\Pi}_h \vec{\mathbf{q}}, \vec{\mathbf{v}}^h) = (\vec{\mathbf{q}}, \vec{\mathbf{v}}^h) \quad \forall r^h \in \mathcal{P}^h(\Omega), \vec{\mathbf{v}}^h \in \mathcal{Q}^h(\Omega), \tag{5.24}$$

and the discrete forms a_h, b_h, and c_h are given by (5.14)-(5.17).

In order to have operator notations in [39], let $\frac{1}{c^2}\mathcal{R}_p + \mathcal{A}_h : \mathcal{P}^h(\Omega) \to [\mathcal{P}^h(\Omega)]'$,

$\mathcal{B}_h : \mathcal{P}^h(\Omega) \to [\mathcal{Q}^h(\Omega)]'$, and $\mathcal{R}_q + \mathcal{C}_h : \mathcal{Q}^h(\Omega) \to [\mathcal{Q}^h(\Omega)]'$ given by

$$\frac{1}{c^2}\mathcal{R}_p p^h(r^h) = (\frac{1}{c^2}p^h, r^h), \qquad \mathcal{R}_q \vec{\mathbf{q}}^h(\vec{\mathbf{v}}^h) = (\vec{\mathbf{q}}^h, \vec{\mathbf{v}}^h),$$
$$\mathcal{A}_h p^h(r^h) = a_h(p^h, r^h), \quad \mathcal{B}_h p^h(\vec{\mathbf{v}}^h) = b_h(p^h, \vec{\mathbf{v}}^h), \quad \mathcal{C}_h \vec{\mathbf{q}}^h(\vec{\mathbf{v}}^h) = c_h(\vec{\mathbf{q}}^h, \vec{\mathbf{v}}^h).$$

Note that the dual operator of $\mathcal{B}_h$, $\mathcal{B}_h' : \mathcal{Q}^h(\Omega) \to [\mathcal{P}^h(\Omega)]'$ satisfies

$$\begin{aligned}
\mathcal{B}_h' \vec{\mathbf{q}}^h(r^h) = &\ \mathcal{B}_h r^h(\vec{\mathbf{q}}^h) \\
= &\ -\sum_{K\in\mathcal{T}_h}\int_K r^h \nabla\cdot\vec{\mathbf{q}}^h dx - \sum_{e\in\mathcal{E}_h^I}\int_e \left(\vec{\mathtt{C}}_{12}\cdot[\![r^h]\!] - \{\!\{r^h\}\!\}\right)[\![\vec{\mathbf{q}}^h]\!] ds \\
= &\ \sum_{K\in\mathcal{T}_h}\int_K \nabla r^h\cdot\vec{\mathbf{q}}^h dx - \sum_{e\in\mathcal{E}_h^I}\int_e \left(\vec{\mathtt{C}}_{12}[\![\vec{\mathbf{q}}^h]\!] + \{\!\{\vec{\mathbf{q}}^h\}\!\}\right)\cdot[\![r^h]\!] ds \\
= &\ b_h'(\vec{\mathbf{q}}^h, r^h),
\end{aligned}$$

which follows from the trace identity (5.8).

Lemma 5.2 *There is a unique semi-discrete solution* $(p^h, \vec{\mathbf{q}}^h)$ *of* (5.22), (5.23) *satisfying*

$$(p^h, \vec{\mathbf{q}}^h) \in C^1\left([0,T]; \mathcal{P}^h(\Omega)\times\mathcal{Q}^h(\Omega)\right).$$

Proof. Theorem 5.1 is used for the proof. We use operator notations of (5.22), (5.23) to get

$$\mathcal{M}_h \begin{pmatrix} p^h \\ \vec{\mathbf{q}}^h \end{pmatrix} + L_h \begin{pmatrix} p^h \\ \vec{\mathbf{q}}^h \end{pmatrix} = 0 \quad \text{in} \quad [\mathcal{P}^h(\Omega)\times\mathcal{Q}^h(\Omega)]',$$

where

$$\mathcal{M}_h = \begin{bmatrix} \frac{1}{c^2}\mathcal{R}_p & 0 \\ 0 & \mathcal{R}_q \end{bmatrix}, \ L_h = \begin{bmatrix} \mathcal{A}_h & -\mathcal{B}_h' \\ \mathcal{B}_h & \mathcal{C}_h \end{bmatrix}.$$

Then we show that L_h is monotone from the definition of L_h,

$$L_h(p^h, \vec{\mathbf{q}}^h)^T((p^h, \vec{\mathbf{q}}^h)^T) = \begin{bmatrix} \mathcal{A}_h p^h - \mathcal{B}_h' \vec{\mathbf{q}}^h \\ \mathcal{B}_h p^h + \mathcal{C}_h \vec{\mathbf{q}}^h \end{bmatrix} \begin{pmatrix} p^h \\ \vec{\mathbf{q}}^h \end{pmatrix}$$

$$= \int_\Omega \left(\frac{\sigma_p}{c^2}(p^h)^2 + \sigma_q \vec{\mathbf{q}}^h \cdot \vec{\mathbf{q}}^h + \nabla p^h \cdot \vec{\mathbf{q}}^h + p^h \nabla \cdot \vec{\mathbf{q}}^h \right) dx + \sum_{e \in \mathcal{E}_{ih}} \int_e \left(C_{11} [\![p^h]\!]^2 + \mathtt{C}_{22} [\![\vec{\mathbf{q}}^h]\!]^2 \right) ds$$

$$+ \sum_{e \in \mathcal{E}_{ih}} \int_e \left(\vec{\mathbf{C}}_{12} [\![\vec{\mathbf{q}}^h]\!] - \{\!\{ \vec{\mathbf{q}}^h \}\!\} \right) \cdot [\![p^h]\!] - \left(\vec{\mathbf{C}}_{12} [\![\vec{\mathbf{q}}^h]\!] + \{\!\{ \vec{\mathbf{q}}^h \}\!\} \right) \cdot [\![p^h]\!] ds$$

$$= \int_\Omega \left(\frac{\sigma_p}{c^2}(p^h)^2 + \sigma_q (\vec{\mathbf{q}}^h)^2 \right) dx + \sum_{e \in \mathcal{E}_{ih}} \int_e \left(C_{11} [\![p^h]\!]^2 + \mathtt{C}_{22} [\![\vec{\mathbf{q}}^h]\!]^2 \right) ds \geq 0,$$

by the trace identity (5.8).

To obtain $Rg(\mathcal{M}_h + L_h) = [\mathcal{P}^h(\Omega) \times \mathcal{Q}^h(\Omega)]'$, it is sufficient to show that $Ker(\mathcal{M}_h + L_h) = \{(0, \vec{\mathbf{0}})\}$. Since

$$\mathcal{M}_h (p^h, \vec{\mathbf{q}}^h)^T \left((p^h, \vec{\mathbf{q}}^h)^T \right) = \int_\Omega \left(\frac{1}{c^2}(p^h)^2 + (\vec{\mathbf{q}}^h)^2 \right) dx$$

$$\geq C \int_\Omega \left((p^h)^2 + (\vec{\mathbf{q}}^h)^2 \right) dx,$$

for some $C = \min\{\frac{1}{c^{*2}}, 1\}$, we can get the surjection, which provides the conclusion. □

To estimate of the difference of the semi-discrete DG solution $(p^h, \vec{\mathbf{q}}^h)$ in (5.22)-(5.23) with analytical solutions $(p, \vec{\mathbf{q}})$ in (5.2) we want to extend to a larger space which contains both solutions. In the next section we show the error estimates.

5.3.2 Extension of DG form

We define the space

$$\mathcal{P}(h) = H_0^1(\Omega) + \mathcal{P}^h(\Omega), \text{ and } \mathcal{Q}(h) = \mathbb{L}^2_{div}(\Omega) + \mathcal{Q}^h(\Omega).$$

with the DG energy norm on $\mathcal{P}(h) \times \mathcal{Q}(h)$,

$$\|(p, \vec{\mathbf{q}})\|_h^2 \;=\; \|p\|^2_{\mathcal{P}(h)} \;+\; \|\vec{\mathbf{q}}\|^2_{\mathcal{Q}(h)},$$

where

$$\|p\|^2_{\mathcal{P}(h)} = \sum_{K \in \mathcal{T}_h} \|p\|^2_{H^1(K)} + \sum_{e \in \mathcal{E}_h} \|\mathtt{C}_{11} [\![p]\!]\|^2_{0,e},$$

$$\|\vec{\mathbf{q}}\|^2_{\mathcal{Q}(h)} = \sum_{K \in \mathcal{T}_h} \|\vec{\mathbf{q}}\|^2_{\mathbb{L}^2_{div}(K)} + \sum_{e \in \mathcal{E}_h} \|\mathtt{C}_{22} [\![\vec{\mathbf{q}}]\!]\|^2_{0,e},$$

and $\mathbb{L}^2_{div}(K) = \{\vec{\mathbf{q}} \in \mathbb{L}^2(K) | \nabla \cdot \vec{\mathbf{q}} \in L^2(K)\}$ with the norm $\|\vec{\mathbf{q}}\|^2_{\mathbb{L}^2_{div}(K)} = \|\vec{\mathbf{q}}\|^2_{\mathbb{L}^2(K)} + \|\nabla \cdot \vec{\mathbf{q}}\|^2_{L^2(K)}$.

For the convenience of notation, let us denote

$$\|\cdot\|_{0,\mathcal{E}_h} := \sum_{e\in\mathcal{E}_h} \|\cdot\|_{0,e}, \quad \|\cdot\|_{0,K} := \|\cdot\|_{L^2(K)} \text{ or } \|\cdot\|_{\mathbb{L}^2(K)}, \quad \|\cdot\|_{0,\Omega} := \|\cdot\|_{L^2(\Omega)} \text{ or } \|\cdot\|_{\mathbb{L}^2(\Omega)}.$$

Furthermore, for $1 \le p \le \infty$ we use the Bochner space $L^p(I; \mathcal{P}(h) \times \mathcal{Q}(h))$,

$$\|(p, \vec{\mathbf{q}})\|_{L^p(I;\mathcal{P}(h)\times\mathcal{Q}(h))} = \begin{cases} (\int_I \|p\|^p_{\mathcal{P}(h)} dt)^{1/p} + (\int_I \|\vec{\mathbf{q}}\|^p_{\mathcal{Q}(h)} dt)^{1/p}, & 1 \le p < \infty, \\ \operatorname{ess\,sup}_{t\in I}(\|p\|_{\mathcal{P}(h)} + \|\vec{\mathbf{q}}\|_{\mathcal{Q}(h)}), & p = \infty. \end{cases}$$

The main result of this section is to establish the $L^2(\Omega)$-*error* estimate. It also gives a bound in the $L^2(\Omega)$-norm of the first time derivative.

Theorem 5.3 *Let the analytical solution* $(p, \vec{\mathbf{q}})$ *of* (5.2) *satisfies*

$$\begin{aligned} (p,\ \vec{\mathbf{q}}) &\in L^\infty(I; H_0^{1+s}(\Omega) \times \mathbb{H}^{1+s}(\Omega)), \\ (p_t,\ \vec{\mathbf{q}}_t) &\in L^1(I; H^s(\Omega) \times \mathbb{H}^s(\Omega)), \end{aligned} \tag{5.25}$$

for a regularity exponent $s > \frac{1}{2}$*, and let* $(p^h, \vec{\mathbf{q}}^h)$ *be the semi-discrete DG approximation obtained by* (5.22), (5.23)*. Then we have the estimate, for the error* $e^p = p - p^h$ *and* $e^{\vec{\mathbf{q}}} = \vec{\mathbf{q}} - \vec{\mathbf{q}}^h$,

$$\begin{aligned} &\sup_{t\in I}\left(\|e^p\|_{0,\Omega} + \|e^{\vec{\mathbf{q}}}\|_{0,\Omega}\right) + \sup_{t\in I}\left(\|[\![e^p]\!]\|_{0,\mathcal{E}_h} + \|[\![e^{\vec{\mathbf{q}}}]\!]\|_{0,\mathcal{E}_h}\right) \le C\left(\|e^p(0)\|_{0,\Omega} + \|e^{\vec{\mathbf{q}}}(0)\|_{0,\Omega}\right) \\ &+ Ch^{\min\{s,k+\frac{1}{2}\}}\left(\|p\|_{L^\infty(I;H^{1+s}(\Omega))} + \|\vec{\mathbf{q}}\|_{L^\infty(I;\mathbb{H}^{1+s}(\Omega))} + \|p_t\|_{L^1(I;H^s(\Omega))} + \|\vec{\mathbf{q}}_t\|_{L^1(I;\mathbb{H}^s(\Omega))}\right), \end{aligned}$$

with a constant C *that is independent of the mesh size* h*.*

Remark 5.3 *The condition* (5.25) *implies that* $(p,\ \vec{\mathbf{q}}) \in C(\bar{I}; H^s(\Omega) \times \mathbb{H}^s(\Omega))$*, thus it is required to have the initial condition* $(p_0,\ \vec{\mathbf{q}}_0) \in H^s(\Omega) \times \mathbb{H}^s(\Omega)$*, and also*

$$\|e^p(0)\|_{0,\Omega} = \|(p - \Pi_h p)(0)\|_{0,\Omega} \le Ch^{\min\{s,k+1\}}\|p\|_{s,\Omega},$$

$$\|e^{\vec{\mathbf{q}}}(0)\|_{0,\Omega} = \|(\vec{\mathbf{q}} - \mathbf{\Pi}_h\vec{\mathbf{q}})(0)\|_{0,\Omega} \leq Ch^{\min\{s,k+1\}}\|\vec{\mathbf{q}}\|_{s,\Omega}.$$

Therefore, Theorem 5.3 *thus implies*

$$\sup_{t\in I}\left(\|e^p\|_{0,\Omega} + \|e^{\vec{\mathbf{q}}}\|_{0,\Omega}\right) + \sup_{t\in I}\left(\|[\![e^p]\!]\|_{0,\mathcal{E}_h} + \|[\![e^{\vec{\mathbf{q}}}]\!]\|_{0,\mathcal{E}_h}\right) \leq Ch^{\min\{s,k+\frac{1}{2}\}},$$

For smooth solutions, Theorem 5.3 *thus yields convergence rates in* L^2*-norm:*

$$\sup_{t\in I}\left(\|e^p\|_{L^2(\Omega)} + \|e^{\vec{\mathbf{q}}}\|_{\mathbb{L}^2(\Omega)}\right) \leq Ch^{k+\frac{1}{2}},$$

where k *is the order of approximation polynomials.*

Following [19] we introduce lifting operators in order to extend the numerical flux to the entire space $\mathcal{P}(h)\times\mathcal{Q}(h)$. We define the lifting operator $\mathcal{L}_h^+ p \in \mathcal{Q}^h(\Omega)$ for $p \in \mathcal{P}(h)$ by

$$\int_\Omega \mathcal{L}_h^+ p\cdot\vec{\mathbf{q}}^h dx \;=\; \sum_{e\in\mathcal{E}_h}\int_e [\![p]\!]\left(\vec{\mathbf{C}}_{12}[\![\vec{\mathbf{q}}^h]\!] + \{\!\{\vec{\mathbf{q}}^h\}\!\}\right) ds, \qquad \forall\vec{\mathbf{q}}^h\in\mathcal{Q}^h(\Omega), \tag{5.26}$$

and also $\mathcal{L}_h^-\vec{\mathbf{q}} \in \mathcal{P}^h(\Omega)$ for $\vec{\mathbf{q}}\in\mathcal{Q}(h)$ by

$$\int_\Omega \mathcal{L}_h^-\vec{\mathbf{q}}\; p^h dx \;=\; \sum_{e\in\mathcal{E}_h}\int_e [\![\vec{\mathbf{q}}]\!]\cdot\left(\vec{\mathbf{C}}_{12}[\![p^h]\!] - \{\!\{p^h\}\!\}\right) ds \qquad \forall p^h\in\mathcal{P}^h(\Omega). \tag{5.27}$$

Note that by the definition of L^2-projection (5.24), we have that

$$\int_\Omega \mathcal{L}_h^+ p\cdot\vec{\mathbf{q}}\; dx \;=\; \int_\Omega \mathcal{L}_h^+ p\cdot\mathbf{\Pi}_h\vec{\mathbf{q}}\; dx \quad \forall p\in\mathcal{P}(h),\vec{\mathbf{q}}\in\mathcal{Q}(h), \tag{5.28}$$

$$\int_\Omega \mathcal{L}_h^-\vec{\mathbf{q}}\; p\; dx \;=\; \int_\Omega \mathcal{L}_h^-\vec{\mathbf{q}}\;\Pi_h p\; dx \quad \forall p\in\mathcal{P}(h),\vec{\mathbf{q}}\in\mathcal{Q}(h). \tag{5.29}$$

Now we extend (5.22), (5.23) using the two lifting functions,

$$(\frac{1}{c^2}p_t, r) \;+\; \tilde{a}_h(p,r) \;-\; \tilde{b}_h'(\vec{\mathbf{q}},r) \;=\; 0 \quad \forall r\in\mathcal{P}(h),\; t\in I, \tag{5.30}$$

$$(\vec{\mathbf{q}}_t,\vec{\mathbf{v}}) \;+\; \tilde{b}_h(p,\vec{\mathbf{v}}) \;+\; \tilde{c}_h(\vec{\mathbf{q}},\vec{\mathbf{v}}) \;=\; 0 \quad \forall\vec{\mathbf{q}}\in\mathcal{Q}(h),\; t\in I, \tag{5.31}$$

where the bilinear forms are given by

$$\tilde{a}_h(p,r) \;=\; \sum_{K\in\mathcal{T}_h}\int_K \frac{\sigma_p}{c^2}pr\;dx \;+\; \sum_{e\in\mathcal{E}_h}\int_e \mathtt{C}_{11}[\![p]\!]\cdot[\![r]\!]ds, \tag{5.32}$$

$$\tilde{b}_h(p,\vec{\mathbf{v}}) \;=\; -\sum_{K\in\mathcal{T}_h}\int_K p\nabla\cdot\vec{\mathbf{v}}dx \;-\; \int_\Omega p\mathcal{L}_h^-\vec{\mathbf{v}}dx, \tag{5.33}$$

$$\tilde{b}'_h(\vec{\mathbf{q}},r) \;=\; \sum_{K\in\mathcal{T}_h}\int_K \vec{\mathbf{q}}\cdot\nabla r dx \;-\; \int_\Omega \vec{\mathbf{q}}\cdot\mathcal{L}_h^+ r dx, \tag{5.34}$$

$$\tilde{c}_h(\vec{\mathbf{q}},\vec{\mathbf{v}}) \;=\; \sum_{K\in\mathcal{T}_h}\int_K \sigma_q\vec{\mathbf{q}}\cdot\vec{\mathbf{v}}dx \;+\; \sum_{e\in\mathcal{E}_h}\int_e \mathtt{C}_{22}[\![\mathbf{q}]\!][\![\mathbf{v}]\!]ds. \tag{5.35}$$

The lifting operators can be bounded provided $\alpha=-1$ (e.g., IP method) as follows:

Remark 5.4 *If the parameter $\alpha=-1$, then there exists a constant C_{inv} which depends only on the shape regularity of the mesh, the approximation order k, and the dimension d, such that*

$$\begin{aligned}\|\mathcal{L}_h^+ p\|_{0,\Omega} &\le \alpha_{11}^{-\frac{1}{2}} C'_{inv}\|\mathtt{C}_{11}[\![p]\!]\|_{0,\mathcal{E}_h},\\ \|\mathcal{L}_h^- \vec{\mathbf{q}}\|_{0,\Omega} &\le \alpha_{22}^{-\frac{1}{2}} C_{inv}\|\mathtt{C}_{22}[\![\vec{\mathbf{q}}]\!]\|_{0,\mathcal{E}_h},\end{aligned}$$

for any $p\in\mathcal{P}(h),\vec{\mathbf{q}}\in\mathcal{Q}(h)$.

Proof. For $p\in\mathcal{P}(\Omega)$ using the definition of $\mathcal{L}_h^+$ and the Riesz representative theorem we have that

$$\begin{aligned}\|\mathcal{L}_h^+ p\|_{0,\Omega} &= \sup_{\vec{\mathbf{q}}\in\mathcal{Q}(h)}\frac{(\mathcal{L}_h^+ p,\vec{\mathbf{q}})}{\|\vec{\mathbf{q}}\|_{0,\Omega}}\\ &= \sup_{\vec{\mathbf{q}}\in\mathcal{Q}(h)}\frac{\sum_{\mathcal{E}_h}\int_e[\![p]\!]\cdot(\vec{\mathbf{C}}_{12}[\![\vec{\mathbf{q}}]\!]+\{\!\{\vec{\mathbf{q}}\}\!\})ds}{\|\vec{\mathbf{q}}\|_{0,\Omega}}\\ &\le \sup_{\vec{\mathbf{q}}\in\mathcal{Q}(h)}\frac{(\sum_{\mathcal{E}_h}\int_e \mathtt{C}_{11}|[\![p]\!]|^2ds)^{\frac{1}{2}}(\sum_{\mathcal{E}_h}\int_e \mathtt{C}_{11}^{-1}|\vec{\mathbf{C}}_{12}[\![\vec{\mathbf{q}}]\!]+\{\!\{\vec{\mathbf{q}}\}\!\}|^2ds)^{\frac{1}{2}}}{\|\vec{\mathbf{q}}\|_{0,\Omega}}\\ &\le \alpha_{11}^{-\frac{1}{2}}\sup_{\vec{\mathbf{q}}\in\mathcal{Q}(h)}\frac{(\sum_{\mathcal{E}_h}\int_e \mathtt{C}_{11}|[\![p]\!]|^2ds)^{\frac{1}{2}}(\sum_{\mathcal{E}_h}\int_e \mathtt{h}|\vec{\mathbf{C}}_{12}[\![\vec{\mathbf{q}}]\!]+\{\!\{\vec{\mathbf{q}}\}\!\}|^2ds)^{\frac{1}{2}}}{\|\vec{\mathbf{q}}\|_{0,\Omega}}\\ &\le \alpha_{11}^{-\frac{1}{2}}(|\vec{\mathbf{C}}_{12}|+1)\sup_{\vec{\mathbf{q}}\in\mathcal{Q}(h)}\frac{(\sum_{\mathcal{E}_h}\int_e \mathtt{C}_{11}|[\![p]\!]|^2ds)^{\frac{1}{2}}(\sum_{\mathcal{T}_h}\int_{\partial K} h_K|\vec{\mathbf{q}}|^2ds)^{\frac{1}{2}}}{\|\vec{\mathbf{q}}\|_{0,\Omega}}\end{aligned}$$

$$\leq \alpha_{11}^{-\frac{1}{2}}(|\vec{\mathbf{C}}_{12}|+1)C_{\text{inv}} \sup_{\vec{\mathbf{q}}\in\mathcal{Q}(h)} \frac{(\sum_{\mathcal{E}_h}\int_e \mathtt{C}_{11}|[\![p]\!]|^2 ds)^{\frac{1}{2}}(\sum_{\mathcal{T}_h}\int_K |\vec{\mathbf{q}}|^2 dx)^{\frac{1}{2}}}{\|\vec{\mathbf{q}}\|_{0,\Omega}}$$

by the Cauchy-Schwarz inequality, the definition (5.21) of α_{11}, and the inverse inequality which is obtained from the combination of (5.18) and (5.20), that is

$$\left(\sum_{\mathcal{T}_h} h_K \int_{\partial K} |\vec{\mathbf{q}}|^2 dx\right)^{\frac{1}{2}} \leq C_{\text{inv}}\|\vec{\mathbf{q}}\|_{0,\Omega},$$

where a constant C_{inv} which depends only on the shape regularity of the mesh, the approximation order k, and the dimension d. Similarly, the second inequality holds, and this completes the proof. □

5.3.3 Error Equations

To derive error equations we define for $r \in \mathcal{P}(h)$, $\vec{\mathbf{v}} \in \mathcal{Q}(h)$ and $p \in H_0^1(\Omega)$, $\vec{\mathbf{q}} \in \mathbb{H}^1(\Omega)$,

$$\mathcal{R}^p(p,\vec{\mathbf{v}}) = \sum_{e\in\mathcal{E}_h}\int_e [\![\vec{\mathbf{v}}]\!]\left(-\vec{\mathtt{C}}_{12}\cdot[\![\Pi_h p - p]\!] + \{\!\{\Pi_h p - p\}\!\}\right) ds, \tag{5.36}$$

$$\mathcal{R}^{\mathbf{q}}(\vec{\mathbf{q}},r) = \sum_{e\in\mathcal{E}_h}\int_e [\![r]\!]\cdot\left(\vec{\mathtt{C}}_{12}[\![\mathbf{\Pi}_h\vec{\mathbf{q}} - \vec{\mathbf{q}}]\!] + \{\!\{\mathbf{\Pi}_h\vec{\mathbf{q}} - \vec{\mathbf{q}}\}\!\}\right) ds. \tag{5.37}$$

The assumption that $p \in H_0^1(\Omega), \vec{\mathbf{q}} \in \mathbb{H}^1(\Omega)$ ensures that $\mathcal{R}^p(p,\vec{\mathbf{v}}), \mathcal{R}^q(\vec{\mathbf{q}},r)$ are well-defined since the trace map of $p, \vec{\mathbf{q}}$ are uniquely defined on all $e \in \mathcal{E}_h$. From the definition (5.7) of jump it directly follows that $\mathcal{R}^p(p,\vec{\mathbf{v}}) = 0, \mathcal{R}^q(\vec{\mathbf{q}},r) = 0$ when $r \in H_0^1(\Omega), \vec{\mathbf{v}} \in \mathbb{H}^1(\Omega)$.

Using the definition of the error equations, we have a following property.

Lemma 5.3 *Let the analytical solution* $(p,\vec{\mathbf{q}})$ *of* (5.2) *satisfy*

$$(p,\vec{\mathbf{q}}) \in L^\infty(I; H_0^1(\Omega)\times\mathbb{H}^1(\Omega)), \quad (p_t,\vec{\mathbf{q}}_t) \in L^1(I; L^2(\Omega)\times\mathbb{L}^2(\Omega)).$$

Let $(p^h,\vec{\mathbf{q}}^h)$ *be the semi-discrete DG approximation obtained by* (5.22), (5.23). *Then the*

error $e^p = p - p^h,\ e^{\vec{\mathbf{q}}} = \vec{\mathbf{q}} - \vec{\mathbf{q}}^h$ *satisfy*

$$(\frac{1}{c^2}e_t^p, r^h) \ + \ \tilde{a}_h(e^p, r^h) \ - \ \tilde{b}'_h(e^{\vec{\mathbf{q}}}, r^h) \ = \ \mathcal{R}^q(\vec{\mathbf{q}}, r^h) \quad \forall r^h \in \mathcal{P}^h(\Omega) \quad a.e.\ in\ I, \tag{5.38}$$

$$(e_t^{\vec{\mathbf{q}}}, \vec{\mathbf{v}}^h) \ + \ \tilde{b}_h(e^p, \vec{\mathbf{v}}^h) \ + \ \tilde{c}_h(e^{\vec{\mathbf{q}}}, \vec{\mathbf{v}}^h) \ = \ \mathcal{R}^p(p, \vec{\mathbf{v}}^h) \quad \forall \vec{\mathbf{v}}^h \in \mathcal{Q}^h(\Omega) \quad a.e.\ in\ I. \tag{5.39}$$

Proof. Let $p^h \in \mathcal{P}^h(\Omega)$ and $\vec{\mathbf{v}}^h \in \mathcal{Q}^h(\Omega)$. Then we obtain that using the discrete formulation in (5.22), (5.23),

$$\begin{aligned}
(\frac{1}{c^2}e_t^p, r^h) + \tilde{a}_h(e^p, r^h) - \tilde{b}'_h(e^{\vec{\mathbf{q}}}, r^h) &= (\frac{1}{c^2}p_t, r^h) + \tilde{a}_h(p, r^h) - \tilde{b}'_h(\vec{\mathbf{q}}, r^h) \quad a.e.\ \text{in}\ I, \\
(e_t^{\vec{\mathbf{q}}}, \vec{\mathbf{v}}^h) + \tilde{b}_h(e^p, \vec{\mathbf{v}}^h) + \tilde{c}_h(e^{\vec{\mathbf{q}}}, \vec{\mathbf{v}}^h) &= (\vec{\mathbf{q}}_t, \vec{\mathbf{v}}^h) + \tilde{b}_h(p, \vec{\mathbf{v}}^h) + \tilde{c}_h(\vec{\mathbf{q}}, \vec{\mathbf{v}}^h) \qquad a.e.\ \text{in}\ I.
\end{aligned}$$

By definitions of $\tilde{b}_h$, the property (5.24) of L^2-projection $\Pi_h, \mathbf{\Pi}_h$, and the definitions (5.26), (5.27) of the lifted element $\mathcal{L}_h^+$, $\mathcal{L}_h^-$, we obtain

$$\begin{aligned}
\tilde{b}_h(p, \vec{\mathbf{v}}^h) &= -\sum_{K\in\mathcal{T}_h}\int_K p\nabla\cdot\vec{\mathbf{v}}^h dx - \sum_{\mathcal{E}_h}\int_e [\![\vec{\mathbf{v}}^h]\!]\left(\vec{\mathsf{C}}_{12}\cdot[\![\Pi_h p]\!] - \{\!\{\Pi_h p\}\!\}\right) ds, \\
\tilde{b}'_h(\vec{\mathbf{q}}, r^h) &= \sum_{K\in\mathcal{T}_h}\int_K \vec{\mathbf{q}}\cdot\nabla r^h dx - \sum_{\mathcal{E}_h}\int_e [\![r^h]\!]\cdot\left(\vec{\mathsf{C}}_{12}[\![\mathbf{\Pi}_h\vec{\mathbf{q}}]\!] + \{\!\{\mathbf{\Pi}_h\vec{\mathbf{q}}\}\!\}\right) ds.
\end{aligned}$$

Since $(p_t, \vec{\mathbf{q}}_t) \in L^1(I; L^2(\Omega)\times\mathbb{L}^2(\Omega))$, we have that $\nabla\cdot\vec{\mathbf{q}} \in L^2(\Omega)$, and $\nabla p \in \mathbb{L}^2(\Omega)$ almost everywhere in I, which implies that p and $\vec{\mathbf{q}}$ have continuous normal components across all interior faces. By integration by parts in element-wise and combination with the trace operators, we get that

$$\begin{aligned}
\tilde{b}_h(p, \vec{\mathbf{v}}^h) &= \sum_{K\in\mathcal{T}_h}\int_K \nabla p\cdot\vec{\mathbf{v}}^h dx - \sum_{\mathcal{E}_h}\int_e [\![\vec{\mathbf{v}}^h]\!]\{\!\{p\}\!\} ds - \sum_{\mathcal{E}_h}\int_e [\![\vec{\mathbf{v}}^h]\!]\left(\vec{\mathsf{C}}_{12}\cdot[\![\Pi_h p]\!] - \{\!\{\Pi_h p\}\!\}\right) ds, \\
\tilde{b}'_h(\vec{\mathbf{q}}, r^h) &= -\sum_{K\in\mathcal{T}_h}\int_K \nabla\cdot\vec{\mathbf{q}} r^h dx + \sum_{\mathcal{E}_h}\int_e [\![r^h]\!]\cdot\{\!\{\vec{\mathbf{q}}\}\!\} ds - \sum_{\mathcal{E}_h}\int_e [\![r^h]\!]\cdot\left(\vec{\mathsf{C}}_{12}[\![\mathbf{\Pi}_h\vec{\mathbf{q}}]\!] + \{\!\{\mathbf{\Pi}_h\vec{\mathbf{q}}\}\!\}\right) ds.
\end{aligned}$$

From the definition of $\mathcal{R}^q(\vec{\mathbf{q}}, r^h)$ and $\mathcal{R}^p(p, \vec{\mathbf{v}}^h)$ in (5.36), (5.37), we have that

$$\begin{aligned}
(\frac{1}{c^2}p_t, r^h) + \tilde{a}_h(p, r^h) - \tilde{b}'_h(\vec{\mathbf{q}}, r^h) &= (\frac{1}{c^2}p_t + \frac{\sigma_p}{c^2}p + \nabla\cdot\vec{\mathbf{q}}, r^h) + \mathcal{R}^q(\vec{\mathbf{q}}, r^h), \\
(\vec{\mathbf{q}}_t, \vec{\mathbf{v}}^h) + \tilde{b}_h(p, \vec{\mathbf{v}}^h) + \tilde{c}_h(\vec{\mathbf{q}}, \vec{\mathbf{v}}^h) &= (\vec{\mathbf{q}}_t + \sigma_q\vec{\mathbf{q}} + \nabla p, \vec{\mathbf{v}}) + \mathcal{R}^p(p, \vec{\mathbf{v}}^h),
\end{aligned}$$

and obtain

$$\begin{aligned}(\frac{1}{c^2}e_t^p, r^h) + \tilde{a}_h(e^p, r^h) - \tilde{b}'_h(e^{\vec{\mathbf{q}}}, r^h) &= (\frac{1}{c^2}p_t + \frac{\sigma_p}{c^2}p + \nabla\cdot\vec{\mathbf{q}}, r^h) + \mathcal{R}^q(\vec{\mathbf{q}}, r^h) &= \mathcal{R}^q(\vec{\mathbf{q}}, r^h),\\ (e_t^{\vec{\mathbf{q}}}, \vec{\mathbf{v}}^h) + \tilde{b}_h(e^p, \vec{\mathbf{v}}^h) + \tilde{c}_h(e^{\vec{\mathbf{q}}}, \vec{\mathbf{v}}^h) &= (\vec{\mathbf{q}}_t + \sigma_q\vec{\mathbf{q}} + \nabla p, \vec{\mathbf{v}}^h) + \mathcal{R}^p(p, \vec{\mathbf{v}}^h) &= \mathcal{R}^p(p, \vec{\mathbf{v}}^h),\end{aligned}$$

where we have used the differential equations in (5.2). □

There is also an important relation between $\tilde{b}_h$ and $\tilde{b}'_h$ from the dual property of ∇ and $-\nabla\cdot$.

Lemma 5.4 *Let the analytical solution* $(p, \vec{\mathbf{q}})$ *of* (5.2) *satisfy*

$$(p, \vec{\mathbf{q}}) \in L^\infty(I; H_0^1(\Omega)\times\mathbb{H}^1(\Omega)), \quad (p_t, \vec{\mathbf{q}}_t) \in L^1(I; L^2(\Omega)\times\mathbb{L}^2(\Omega)).$$

Let $(p^h, \vec{\mathbf{q}}^h)$ *be the semi-discrete DG approximation obtained by* (5.22), (5.23). *Then the following property holds, for all* $r^h \in \mathcal{P}^h(\Omega)$ *and* $\vec{\mathbf{v}}^h \in \mathcal{Q}^h(\Omega)$,

$$-\tilde{b}'_h(e^{\vec{\mathbf{q}}}, \Pi_h p - p^h) + \tilde{b}_h(e^p, \mathbf{\Pi}_h\vec{\mathbf{q}} - \vec{\mathbf{q}}^h) = 0, \tag{5.40}$$

$$-\tilde{b}'_h(\vec{\mathbf{v}}^h, r^h) + \tilde{b}_h(r^h, \vec{\mathbf{v}}^h) = 0. \tag{5.41}$$

Proof. By the definition of $\tilde{b}'_h$, the property (5.26) of lifted element, and the property of L^2-projection, we obtain that

$$\begin{aligned}\tilde{b}'_h(\vec{\mathbf{q}} - \mathbf{\Pi}_h\vec{\mathbf{q}}, r^h) &= \sum_{K\in\mathcal{T}_h}\int_K (\vec{\mathbf{q}} - \mathbf{\Pi}_h\vec{\mathbf{q}})\cdot\nabla r^h dx - \int_\Omega \mathcal{L}_h^+ r^h\cdot(\vec{\mathbf{q}} - \mathbf{\Pi}_h\vec{\mathbf{q}})dx \\ &= -\int_\Omega \mathcal{L}_h^+ r^h\cdot(\vec{\mathbf{q}} - \mathbf{\Pi}_h\vec{\mathbf{q}})dx \\ &= 0.\end{aligned} \tag{5.42}$$

Here, we have used the definition of L^2-projection, $\mathbf{\Pi}_h(\vec{\mathbf{q}} - \mathbf{\Pi}_h\vec{\mathbf{q}}) = \mathbf{\Pi}_h\vec{\mathbf{q}} - \mathbf{\Pi}_h\vec{\mathbf{q}} = 0$. In the similar way it holds that

$$\tilde{b}_h(p - \Pi_h p, \vec{\mathbf{q}}^h) = 0.$$

For $r^h \in \mathcal{P}^h(\Omega)$ and $\vec{\mathbf{v}}^h \in \mathcal{Q}^h(\Omega)$, we use definition of $\tilde{b}'_h$, element-wise integration by parts, and the trace identity (5.8) to obtain that

$$\begin{aligned}
\tilde{b}'_h(\vec{\mathbf{v}}^h, r^h) &= \sum_{K\in\mathcal{T}_h}\int_K \vec{\mathbf{v}}^h\cdot\nabla r^h dx - \int_\Omega \mathcal{L}_h^+ r^h\cdot\vec{\mathbf{v}}^h dx \\
&= -\sum_{K\in\mathcal{T}_h}\int_K \nabla\cdot\vec{\mathbf{v}}^h r^h dx + \sum_{K\in\mathcal{T}_h}\int_K r^h\vec{\mathbf{v}}^h\cdot\vec{\mathbf{n}} ds - \sum_{e\in\mathcal{T}_h}\int_e [\![p^h]\!]\cdot(\vec{\mathtt{C}}_{12}[\![\vec{\mathbf{v}}^h]\!] + \{\!\{\vec{\mathbf{v}}^h\}\!\}) ds \\
&= -\sum_{K\in\mathcal{T}_h}\int_K \nabla\cdot\vec{\mathbf{v}}^h r^h dx - \sum_{e\in\mathcal{T}_h}\int_e \left([\![r^h]\!]\cdot\vec{\mathtt{C}}_{12} - \{\!\{r^h\}\!\}\right)[\![\vec{\mathbf{v}}^h]\!] ds,
\end{aligned}$$

and from the definition $\tilde{b}_h$,

$$\begin{aligned}
\tilde{b}_h(r^h, \vec{\mathbf{v}}^h) &= -\sum_{K\in\mathcal{T}_h}\int_K r^h\nabla\cdot\vec{\mathbf{v}}^h dx - \int_\Omega \mathcal{L}_h^-\vec{\mathbf{v}}^h\ r^h dx \\
&= -\sum_{K\in\mathcal{T}_h}\int_K r^h\nabla\cdot\vec{\mathbf{v}}^h dx - \sum_{e\in\mathcal{T}_h}\int_e \left([\![r^h]\!]\cdot\vec{\mathtt{C}}_{12} - \{\!\{r^h\}\!\}\right)[\![\vec{\mathbf{v}}^h]\!] ds.
\end{aligned}$$

Subtracting $\tilde{b}'_h$ from $\tilde{b}_h$ we have that

$$-\tilde{b}'_h(\vec{\mathbf{v}}^h, r^h) + \tilde{b}_h(r^h, \vec{\mathbf{v}}^h) = 0 \qquad \forall r^h \in \mathcal{P}^h(\Omega), \vec{\mathbf{v}}^h \in \mathcal{Q}^h(\Omega). \tag{5.43}$$

Using the definition of error e^p and $e^{\vec{\mathbf{q}}}$ with the properties (5.42), and (5.43) we obtain

$$\begin{aligned}
\tilde{b}'_h(e^{\vec{\mathbf{q}}}, \Pi_h p - p^h) - \tilde{b}_h(e^p, \mathbf{\Pi}_h\vec{\mathbf{q}} - \vec{\mathbf{q}}^h) &= \tilde{b}'_h(\mathbf{\Pi}_h\vec{\mathbf{q}} - \vec{\mathbf{q}}^h, \Pi_h p - p^h) - \tilde{b}_h(\Pi_h p - p^h, \mathbf{\Pi}_h\vec{\mathbf{q}} - \vec{\mathbf{q}}^h) \\
&= 0,
\end{aligned}$$

which completes the proof. □

5.3.4 Approximation Properties.

We recall the following L^2-projection approximation properties; see [34].

Lemma 5.5 *Let $K \in \mathcal{T}_h$. Then the following properties hold:*

(i) For $p \in H^s(K), s \geq 0$, we have

$$\|p - \Pi_h p\|_{L^2(K)} \leq C h_K^{\min\{s,k+1\}} \|p\|_{H^s(K)},$$

with a constant C that is independent of the local mesh size h_K and depends only on the shape-regularity of the mesh, the approximation order k, the dimension d, and the regularity exponent s.

(ii) For $p \in H^{1+s}(K), s > \frac{1}{2}$, we have

$$\begin{aligned} \|\nabla p - \nabla(\Pi_h p)\|_{L^2(K)} &\leq C h_K^{\min\{s,k\}} \|p\|_{H^{1+s}(K)}, \\ \|p - \Pi_h p\|_{L^2(\partial K)} &\leq C h_K^{\min\{s,k\}+\frac{1}{2}} \|p\|_{H^{1+s}(K)}, \\ \|\nabla p - \mathbf{\Pi}_h \nabla(p)\|_{L^2(\partial K)} &\leq C h_K^{\min\{s,k+1\}-\frac{1}{2}} \|p\|_{H^{1+s}(K)}, \end{aligned} \tag{5.44}$$

with a constant C that is independent of the local mesh size h_K and depends only on the shape-regularity of the mesh, the approximation order k, the dimension d, and the regularity exponent s.

As a consequence of the approximation properties in Lemma 5.5, we have the following results. Let us denote for convenience, $\|\cdot\|_{s,K} := \|\cdot\|_{H^s(K)}$ and $\|\cdot\|_{s,\Omega} := \|\cdot\|_{H^s(\Omega)}$, and the same as $\mathbb{H}^s(K)$ and $\mathbb{H}^s(\Omega)$, respectively.

Lemma 5.6 *Let $p \in H^{1+s}(\Omega), s > \frac{1}{2}$. Then the following hold:*

$$\begin{aligned} \|\{\!\{\Pi_h p - p\}\!\}\|_{0,\mathcal{E}_h} &\leq C h^{\min\{s,k\}+\frac{1}{2}} \|p\|_{1+s,\Omega}, \\ \|[\![\Pi_h p - p]\!]\|_{0,\mathcal{E}_h} &\leq C h^{\min\{s,k\}+\frac{1}{2}} \|p\|_{1+s,\Omega}, \end{aligned}$$

with a constant C that is independent of the local mesh size h_K and depends only on the shape-regularity of the mesh, the approximation order k, the dimension d, and the regularity exponent s.

Proof. It's directly obtained from Lemma 5.5 and definition of jump and average on faces of elements K. □

Lemma 5.7 *Let $(p, \vec{\mathbf{q}}) \in H^{1+s}(\Omega) \times \mathbb{H}^{1+s}(\Omega)$ with $s > \frac{1}{2}$. Then the following hold:*

(i) For $r \in \mathcal{P}(h), \vec{\mathbf{v}} \in \mathcal{Q}(h)$, the forms (5.36) *and* (5.37) *can be bounded by*

$$\begin{aligned} |\mathcal{R}^p(p,\vec{\mathbf{v}})| &\leq C_R^p \; h^{\min\{s,k\}+\frac{1-\alpha}{2}} \; \|\mathtt{C}_{22}^{\frac{1}{2}}[\![\vec{\mathbf{v}}]\!]\|_{0,\mathcal{E}_h} \, \|p\|_{1+s,\Omega}\,, \\ |\mathcal{R}^q(\vec{\mathbf{q}},r)| &\leq C_R^q \; h^{\min\{s,k\}+\frac{1-\alpha}{2}} \; \|\mathtt{C}_{11}^{\frac{1}{2}}[\![r]\!]\|_{0,\mathcal{E}_h} \, \|\vec{\mathbf{q}}\|_{1+s,\Omega}\,, \end{aligned}$$

with constants C_R^p and C_R^q independent of h, which depend only on $\alpha_{11}, \alpha_{12}, \alpha_{22}$, and the constant in Lemma 5.5.

(ii) The bilinear forms are estimated by following :

$$\tilde{a}_h(e^p, \Pi_h p - p) \leq C_a \; h^{\min\{s,k\}+\frac{1+\alpha}{2}} \left(h^{\frac{1-\alpha}{2}} \|e^p\|_{0,\Omega} + \|\mathtt{C}_{11}^{\frac{1}{2}}[\![e^p]\!]\|_{0,\mathcal{E}_h} \right) \|p\|_{1+s,\Omega},$$

$$\tilde{c}_h(e^{\vec{\mathbf{q}}}, \mathbf{\Pi}_h\vec{\mathbf{q}} - \vec{\mathbf{q}}) \leq C_c \; h^{\min\{s,k\}+\frac{1+\alpha}{2}} \left(h^{\frac{1-\alpha}{2}} \|e^{\vec{\mathbf{q}}}\|_{0,\Omega} + \|\mathtt{C}_{22}^{\frac{1}{2}}[\![e^{\vec{\mathbf{q}}}]\!]\|_{0,\mathcal{E}_h} \right) \|\vec{\mathbf{q}}\|_{1+s,\Omega},$$

with constants C_a and C_c independent of h, which depend only on α_{11}, α_{22}, and the constant in Lemma 5.5.

Proof. (i) To show the first estimate we begin with the definition of $\mathcal{R}^p$ in (5.36), and apply the Cauchy-Schwarz inequality and approximation properties in Lemma 5.5 to obtain that

$$\begin{aligned} |\mathcal{R}^p(p,\vec{\mathbf{v}})|^2 \leq & \sum_{e\in\mathcal{E}_h} \int_e |\mathtt{C}_{22}^{\frac{1}{2}}[\![\vec{\mathbf{v}}]\!]|^2 ds \cdot \sum_{e\in\mathcal{E}_h} \int_e \mathtt{C}_{22}^{-1} \left| (\vec{\mathtt{C}}_{12} \cdot [\![\Pi_h p - p]\!] + \{\!\{\Pi_h p - p\}\!\}) \right|^2 ds \\ \leq & \alpha_{22}^{-1} \|\mathtt{C}_{22}^{\frac{1}{2}}[\![\vec{\mathbf{v}}]\!]\|_{0,\mathcal{E}_h}^2 \sum_{K\in\mathcal{T}_h} h_K^{-\alpha}(1 + |\vec{\mathtt{C}}_{12}|) \, \|p - \Pi_h p\|_{0,\partial K}^2 \\ \leq & C_R^{p\,2} h^{2\min\{s,k\}+1-\alpha} \|\mathtt{C}_{22}^{\frac{1}{2}}[\![\vec{\mathbf{v}}]\!]\|_{0,\mathcal{E}_h}^2 \, \|p\|_{1+s,\Omega}^2 \quad \textit{by} \quad (5.44). \end{aligned}$$

This completes the first estimate. Similarly we have the second bound in (i).

(ii) From the definition of $\tilde{a}_h$ in (5.32) we apply Hölder's inequality, the definition of α_{11},

Cauchy-Schwarz inequality, and Lemma 5.5,

$$
\begin{aligned}
&\tilde{a}_h(e^p, \Pi_h p - p) \\
&= \sum_{K\in\mathcal{T}_h}\int_K \sigma_p e^p(\Pi_h p - p)dx + \sum_{e\in\mathcal{E}_h}\int_e \mathtt{C}_{11}[\![e^p]\!]\cdot[\![\Pi p - p]\!]ds \\
&\le \sigma^* \sum_{K\in\mathcal{T}_h}\|e^p\|_{0,K}\|\Pi_h p - p\|_{0,K} + \sum_{e\in\mathcal{E}_h}\|\mathtt{C}_{11}^{\frac{1}{2}}[\![e^p]\!]\|_{0,e}\|\mathtt{C}_{11}^{\frac{1}{2}}[\![\Pi_h p - p]\!]\|_{0,e} \\
&\le \sigma^* \sum_{K\in\mathcal{T}_h}\|e^p\|_{0,K}\|\Pi_h p - p\|_{0,K} + \alpha_{11}^{\frac{1}{2}} \sum_{e\in\mathcal{E}_h}\|\mathtt{C}_{11}^{\frac{1}{2}}[\![e^p]\!]\|_{0,e}\|\mathtt{h}^{\frac{\alpha}{2}}[\![\Pi_h p - p]\!]\|_{0,e} \\
&\le C\left(\sigma^* h^{\frac{1-\alpha}{2}} \sum_{K\in\mathcal{T}_h}\|e^p\|_{0,K} + \alpha_{11}^{\frac{1}{2}}\|\mathtt{C}_{11}^{\frac{1}{2}}[\![e^p]\!]\|_{0,\mathcal{E}_h}\right)\left(\sum_{K\in\mathcal{T}_h} h_K^{\min\{s,k\}+\frac{1+\alpha}{2}}\|p\|_{1+s,K}\right),
\end{aligned}
$$

since

$$
\sum_{e\in\mathcal{E}_h}\|\mathtt{C}_{11}^{\frac{1}{2}}[\![e^p]\!]\|_{0,e}\|\mathtt{h}^{\frac{\alpha}{2}}[\![\Pi_h p - p]\!]\|_{0,e} \le \left(\sum_{e\in\mathcal{E}_h}\|\mathtt{C}_{11}^{\frac{1}{2}}[\![e^p]\!]\|_{0,e}^2\right)^{\frac{1}{2}}\left(\kappa\sum_{K\in\mathcal{T}_h} h_K^{\alpha}\|\Pi_h p - p\|_{0,\partial K}^2\right)^{\frac{1}{2}},
$$

where κ is shape-regularity constant in (5.6). This completes the first estimate of (ii). Similarly we can bound of $\tilde{c}_h(e^{\vec{\mathbf{q}}}, \boldsymbol{\Pi}\vec{\mathbf{q}} - \vec{\mathbf{q}})$ with the same order of h.

□

5.3.5 Proof of Theorem 5.3

Proof. From Theorem 5.1, we have that

$$
e^p \in C^0(\bar{I};\mathcal{P}(h))\cap C^1(\bar{I};L^2(\Omega)) \quad \text{and} \quad e^{\vec{\mathbf{q}}} \in C^0(\bar{I};\mathcal{Q}(h))\cap C^1(\bar{I};\mathbb{L}^2(\Omega)).
$$

Since $e^p = p - \Pi p + \Pi p - p^h$, $e^{\vec{\mathbf{q}}} = \vec{\mathbf{q}} - \boldsymbol{\Pi}\vec{\mathbf{q}} + \boldsymbol{\Pi}\vec{\mathbf{q}} - \vec{\mathbf{q}}^h$, using the error equations (5.38) and (5.39), we have that

$$
\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|\frac{1}{c}e^p\|_{0,\Omega}^2 + \frac{1}{2}\frac{d}{dt}\|e^{\vec{\mathbf{q}}}\|_{0,\Omega}^2 \\
&= (\frac{1}{c^2}e_t^p, p - \Pi_h p) - \tilde{a}_h(e^p, \Pi_h p - p^h) + (e_t^{\vec{\mathbf{q}}}, \boldsymbol{\Pi}_h\vec{\mathbf{q}} - \vec{\mathbf{q}}^h) - \tilde{c}_h(e^{\vec{\mathbf{q}}}, \boldsymbol{\Pi}_h\vec{\mathbf{q}} - \vec{\mathbf{q}}^h) \\
&\quad + \mathcal{R}^q(\vec{\mathbf{q}}, \Pi_h p - p^h) + \mathcal{R}^p(p, \boldsymbol{\Pi}_h\vec{\mathbf{q}} - \vec{\mathbf{q}}^h),
\end{aligned}
$$

by the property in (5.4).

Now we fix $\tau \in I$ and integrate over the time interval $(0, \tau)$. This yields

$$\begin{aligned}
&\frac{1}{2}\|\frac{1}{c}e^p(\tau)\|_{0,\Omega}^2 + \frac{1}{2}\|e^{\vec{\mathbf{q}}}(\tau)\|_{0,\Omega}^2 + \int_0^\tau \left(\tilde{a}_h(e^p, e^p) + \tilde{c}_h(e^{\vec{\mathbf{q}}}, e^{\vec{\mathbf{q}}})\right) dt \\
&= \frac{1}{2}\|\frac{1}{c}e^p(0)\|_{0,\Omega}^2 + \frac{1}{2}\|e^{\vec{\mathbf{q}}}(0)\|_{0,\Omega}^2 + \int_0^\tau \left[(\frac{1}{c^2}e_t^p, p - \Pi_h p) + (e_t^{\vec{\mathbf{q}}}, \vec{\mathbf{q}} - \mathbf{\Pi}_h\vec{\mathbf{q}})\right] dt \qquad (5.45) \\
&+ \int_0^\tau \left[\tilde{a}_h(e^p, p - \Pi_h p) + \tilde{c}_h(e^{\vec{\mathbf{q}}}, \vec{\mathbf{q}} - \mathbf{\Pi}_h\vec{\mathbf{q}})\right] dt + \int_0^\tau \left[\mathcal{R}^p(p, \mathbf{\Pi}_h\vec{\mathbf{q}} - \vec{\mathbf{q}}^h) + \mathcal{R}^q(\vec{\mathbf{q}}, \Pi_h p - p^h)\right] dt.
\end{aligned}$$

Integration by parts in the first integral on the right hand side and standard Hölder's inequality yield that

$$\begin{aligned}
\int_0^\tau \left[(\frac{1}{c^2}e_t^p, p - \Pi_h p) + (e_t^{\vec{\mathbf{q}}}, \vec{\mathbf{q}} - \mathbf{\Pi}_h\vec{\mathbf{q}})\right] dt &= -\int_0^\tau \left[(\frac{1}{c^2}e^p, (p - \Pi_h p)_t) + (e^{\vec{\mathbf{q}}}, (\vec{\mathbf{q}} - \mathbf{\Pi}_h\vec{\mathbf{q}})_t)\right] dt \\
&+ \left[(\frac{1}{c^2}e^p, p - \Pi p)\right]_{t=0}^{t=\tau} + \left[(e^{\vec{\mathbf{q}}}, \vec{\mathbf{q}} - \mathbf{\Pi}_h\vec{\mathbf{q}})\right]_{t=0}^{t=\tau} \\
&\le \|\frac{1}{c}e^p\|_{L^\infty(I;L^2(\Omega))}\|\frac{1}{c}(p - \Pi_h p)_t\|_{L^1(I;L^2(\Omega))} + \|e^{\vec{\mathbf{q}}}\|_{L^\infty(I;L^2(\Omega))}\|(\vec{\mathbf{q}} - \mathbf{\Pi}_h\vec{\mathbf{q}})_t\|_{L^1(I;L^2(\Omega))} \\
&+2\|\frac{1}{c}e^p\|_{L^\infty(I;L^2(\Omega))}\|\frac{1}{c}(p - \Pi_h p)\|_{L^\infty(I;L^2(\Omega))} + 2\|e^{\vec{\mathbf{q}}}\|_{L^\infty(I;L^2(\Omega))}\|\vec{\mathbf{q}} - \mathbf{\Pi}_h\vec{\mathbf{q}}\|_{L^\infty(I;L^2(\Omega))} \\
&:= T_1.
\end{aligned}$$

From the definition of $\tilde{a}_h$ and $\tilde{c}_h$ and standard Hölder's inequality in the second integral on the right hand side in (5.45), we have that

$$\begin{aligned}
&\int_0^\tau \left[\tilde{a}_h(e^p, p - \Pi_h p) + \tilde{c}_h(e^{\vec{\mathbf{q}}}, \vec{\mathbf{q}} - \mathbf{\Pi}\vec{\mathbf{q}})\right] dt \\
&\le \sigma^* \int_0^\tau \left(\|\frac{1}{c}e^p\|_{0,\Omega}\|c(p - \Pi_h p)\|_{0,\Omega} + \|e^{\vec{\mathbf{q}}}\|_{0,\Omega}\|\vec{\mathbf{q}} - \mathbf{\Pi}\vec{\mathbf{q}}\|_{0,\Omega}\right) dt \\
&+ \int_0^\tau \left(\|\mathtt{C}_{11}^{\frac{1}{2}}[\![e^p]\!]\|_{0,\mathcal{E}_h}\|\mathtt{C}_{11}^{\frac{1}{2}}[\![p - \Pi_h p]\!]\|_{0,\mathcal{E}_h} + \|\mathtt{C}_{22}^{\frac{1}{2}}[\![e^{\vec{\mathbf{q}}}]\!]\|_{0,\mathcal{E}_h}\|\mathtt{C}_{22}^{\frac{1}{2}}[\![\vec{\mathbf{q}} - \mathbf{\Pi}_h\vec{\mathbf{q}}]\!]\|_{0,\mathcal{E}_h}\right) dt \\
&\le \sigma^* T \left(\|\frac{1}{c}e^p\|_{L^\infty(I;L^2(\Omega))}\|c(p - \Pi_h p)\|_{L^\infty(I;L^2(\Omega))} + \|e^{\vec{\mathbf{q}}}\|_{L^\infty(I;L^2(\Omega))}\|\vec{\mathbf{q}} - \mathbf{\Pi}_h\vec{\mathbf{q}}\|_{L^\infty(I;L^2(\Omega))}\right) \\
&+\|\mathtt{C}_{11}^{\frac{1}{2}}[\![e^p]\!]\|_{L^1(I;L^2(\mathcal{E}_h))}\|\mathtt{C}_{11}^{\frac{1}{2}}[\![p - \Pi_h p]\!]\|_{L^\infty(I;L^2(\mathcal{E}_h))} + \|\mathtt{C}_{22}^{\frac{1}{2}}[\![e^{\vec{\mathbf{q}}}]\!]\|_{L^1(I;L^2(\mathcal{E}_h))}\|\mathtt{C}_{22}^{\frac{1}{2}}[\![\vec{\mathbf{q}} - \mathbf{\Pi}_h\vec{\mathbf{q}}]\!]\|_{L^\infty(I;L^2(\mathcal{E}_h))} \\
&:= T_2.
\end{aligned}$$

Now we combine T_1 and T_2 together and rewrite the left hand side (5.45) with the new bounds,

$$\frac{1}{2}\|\frac{1}{c}e^p(\tau)\|_0^2+\frac{1}{2}\|e^{\vec{\mathbf{q}}}(\tau)\|_0^2+\int_0^\tau\left(\|\frac{\sigma_p}{c^2}e^p\|_{0,\Omega}^2+\|\mathtt{C}_{11}^{\frac{1}{2}}[\![e^p]\!]\|_{0,\mathcal{E}_h}^2+\|\sigma_q e^{\vec{\mathbf{q}}}\|_{0,\Omega}^2+\|\mathtt{C}_{22}^{\frac{1}{2}}[\![e^{\vec{\mathbf{q}}}]\!]\|_{0,\mathcal{E}_h}^2\right)dt$$
$$\leq\frac{1}{2}\|\frac{1}{c}e^p(0)\|_{0,\Omega}^2+\frac{1}{2}\|e^{\vec{\mathbf{q}}}(0)\|_{0,\Omega}^2+T_1+T_2+\int_0^\tau\left|\mathcal{R}^p(p,\mathbf{\Pi}_h\vec{\mathbf{q}}-\vec{\mathbf{q}}^h)\right|+\left|\mathcal{R}^q(\vec{\mathbf{q}},\Pi_h p-p^h)\right|dt.$$

Since this inequality holds for any $\tau\in I$, it also holds for the supremum over I, that is

$$\frac{1}{2}\sup_{t\in I}\left(\|\frac{1}{c}e^p(t)\|_{L^2(\Omega)}^2+\|e^{\vec{\mathbf{q}}}(t)\|_{\mathbb{L}^2(\Omega)}^2\right)+\|\frac{\sigma_p}{c^2}e^p\|_{L^1(I;L^2(\Omega))}^2+\|\mathtt{C}_{11}^{\frac{1}{2}}[\![e^p]\!]\|_{L^1(I;L^2(\mathcal{E}_h))}^2$$
$$+\|\sigma_q e^{\vec{\mathbf{q}}}\|_{L^1(I;\mathbb{L}^2(\Omega))}^2+\|\mathtt{C}_{22}^{\frac{1}{2}}[\![e^{\vec{\mathbf{q}}}]\!]\|_{L^1(I;\mathbb{L}^2(\mathcal{E}_h))}^2$$
$$\leq\frac{1}{2}\|\frac{1}{c}e^p(0)\|_{0,\Omega}^2+\frac{1}{2}\|e^{\vec{\mathbf{q}}}(0)\|_{0,\Omega}^2+T_1+T_2+\int_I\left|\mathcal{R}^p(p,\mathbf{\Pi}_h\vec{\mathbf{q}}-\vec{\mathbf{q}}^h)\right|dt+\int_I\left|\mathcal{R}^q(\vec{\mathbf{q}},\Pi_h p-p^h)\right|dt.$$

Using the geometric-arithmetic mean inequality $|ab|\leq\frac{1}{2\varepsilon}a^2+\frac{\varepsilon}{2}b^2$, valid for $\varepsilon>0$, $(a+b)^2\leq 2(a^2+b^2)$, and the approximation results in Lemma (5.5), we obtain that

$$\begin{aligned}
T_1 &= \|\frac{1}{c}e^p\|_{L^\infty(I;L^2(\Omega))}\left(\|\frac{1}{c}(p-\Pi_h p)_t\|_{L^1(I;L^2(\Omega))}+2\|\frac{1}{c}(p-\Pi_h p)\|_{L^\infty(I;L^2(\Omega))}\right)\\
&\quad+\|e^{\vec{\mathbf{q}}}\|_{L^\infty(I;L^2(\Omega))}\left(\|(\vec{\mathbf{q}}-\mathbf{\Pi}_h\vec{\mathbf{q}})_t\|_{L^1(I;L^2(\Omega))}+2\|\vec{\mathbf{q}}-\mathbf{\Pi}_h\vec{\mathbf{q}}\|_{L^\infty(I;L^2(\Omega))}\right)\\
&\leq\frac{1}{2\varepsilon}\|\frac{1}{c}e^p\|_{L^\infty(I;L^2(\Omega))}^2+\varepsilon\left(\|\frac{1}{c}(p-\Pi_h p)_t\|_{L^1(I;L^2(\Omega))}^2+4\|\frac{1}{c}(p-\Pi_h p)\|_{L^\infty(I;L^2(\Omega))}^2\right)\\
&\quad+\frac{1}{2\varepsilon}\|e^{\vec{\mathbf{q}}}\|_{L^\infty(I;L^2(\Omega))}^2+\varepsilon\left(\|(\vec{\mathbf{q}}-\mathbf{\Pi}_h\vec{\mathbf{q}})_t\|_{L^1(I;L^2(\Omega))}^2+4\|\vec{\mathbf{q}}-\mathbf{\Pi}_h\vec{\mathbf{q}}\|_{L^\infty(I;L^2(\Omega))}^2\right)\\
&\leq\frac{1}{2\varepsilon}\|\frac{1}{c}e^p\|_{L^\infty(I;L^2(\Omega))}^2\quad+\varepsilon Ch^{2\min\{s,k+1\}}\left(\frac{1}{c_*^2}\|p_t\|_{L^1(I;H^s(\Omega))}^2+\|\vec{\mathbf{q}}_t\|_{L^1(I;\mathbb{H}^s(\Omega))}^2\right)\\
&\quad+\frac{1}{2\varepsilon}\|e^{\vec{\mathbf{q}}}\|_{L^\infty(I;L^2(\Omega))}^2+4\varepsilon Ch^{2\min\{s,k\}+2}\left(\frac{1}{c_*^2}\|p\|_{L^\infty(I;H^{1+s}(\Omega))}^2+\|\vec{\mathbf{q}}\|_{L^\infty(I;\mathbb{H}^{1+s}(\Omega))}^2\right),
\end{aligned}$$

and

$$
\begin{aligned}
&T_2\\
&= \frac{1}{2\varepsilon}\left(\|\frac{1}{c}e^p\|^2_{L^\infty(I;L^2(\Omega))} + \|e^{\vec{\mathbf{q}}}\|^2_{L^\infty(I;L^2(\Omega))}\right) + \frac{1}{2\mu}\left(\|\mathsf{C}_{11}^{\frac{1}{2}}[\![e^p]\!]\|^2_{L^1(I;L^2(\mathcal{E}_h))} + \|\mathsf{C}_{22}^{\frac{1}{2}}[\![e^{\vec{\mathbf{q}}}]\!]\|^2_{L^1(I;L^2(\mathcal{E}_h))}\right)\\
&\qquad + \frac{\varepsilon}{2}\sigma^{*2}T^2\left(c^{*2}\|p-\Pi_h p\|^2_{L^\infty(I;L^2(\Omega))} + \|\vec{\mathbf{q}}-\mathbf{\Pi}_h\vec{\mathbf{q}}\|^2_{L^\infty(I;L^2(\Omega))}\right)\\
&\qquad + \frac{\mu}{2}\left(\|\mathsf{C}_{11}^{\frac{1}{2}}[\![p-\Pi_h p]\!]\|^2_{L^\infty(I;L^2(\mathcal{E}_h))} + \|\mathsf{C}_{22}^{\frac{1}{2}}[\![\vec{\mathbf{q}}-\mathbf{\Pi}_h\vec{\mathbf{q}}]\!]\|^2_{L^\infty(I;L^2(\mathcal{E}_h))}\right)\\
&\leq \frac{1}{2\varepsilon}\left(\|\frac{1}{c}e^p\|^2_{L^\infty(I;L^2(\Omega))} + \|e^{\vec{\mathbf{q}}}\|^2_{L^\infty(I;L^2(\Omega))}\right) + \frac{1}{2\mu}\left(\|\mathsf{C}_{11}^{\frac{1}{2}}[\![e^p]\!]\|^2_{L^1(I;L^2(\mathcal{E}_h)} + \|\mathsf{C}_{22}^{\frac{1}{2}}[\![e^{\vec{\mathbf{q}}}]\!]\|^2_{L^1(I;L^2(\mathcal{E}_h))}\right)\\
&\qquad + \frac{\varepsilon}{2}C\sigma^{*2}T^2h^{2\min\{s,k\}+2}\left(c^{*2}\|p\|^2_{L^\infty(I;H^{1+s}(\Omega))} + \|\vec{\mathbf{q}}\|^2_{L^\infty(I;\mathbb{H}^{1+s}(\Omega))}\right)\\
&\qquad + \frac{\mu}{2}Ch^{2\min\{s,k\}+1+\alpha}\left(\|p\|^2_{L^\infty(I;H^{1+s}(\Omega))} + \|\vec{\mathbf{q}}\|^2_{L^\infty(I;\mathbb{H}^{1+s}(\Omega))}\right).
\end{aligned}
$$

Using Lemma 5.5 and Lemma 5.7 we can also bound the error equations

$$
\begin{aligned}
&\int_I \left|\mathcal{R}^p(p,\mathbf{\Pi}\vec{\mathbf{q}}-\vec{\mathbf{q}}^h)\right| dt \leq \int_I \left|\mathcal{R}^p(p,e^{\vec{\mathbf{q}}})\right| dt + \int_I \left|\mathcal{R}^p(p,\mathbf{\Pi}\vec{\mathbf{q}}-\vec{\mathbf{q}})\right| dt\\
&\leq C_R^p h^{\min\{s,k\}+\frac{1-\alpha}{2}}\int_I\left[\left(\|\mathsf{C}_{22}^{\frac{1}{2}}[\![e^{\vec{\mathbf{q}}}]\!]\|_{0,\mathcal{E}_h} + \|\mathsf{C}_{22}^{\frac{1}{2}}[\![\mathbf{\Pi}\vec{\mathbf{q}}-\vec{\mathbf{q}}]\!]\|_{0,\mathcal{E}_h}\right)\|p\|_{1+s,\Omega}\right] dt\\
&\leq \frac{\mu}{2}{C_R^p}^2h^{2\min\{s,k\}+1-\alpha}\|p\|^2_{L^\infty(I;H^{1+s}(\Omega))} + \frac{1}{2\mu}\|\mathsf{C}_{22}^{\frac{1}{2}}[\![e^{\vec{\mathbf{q}}}]\!]\|^2_{L^1(I;H^{1+s}(\mathcal{E}_h))}\\
&\qquad + \frac{\mu}{2}{C_R^p}^2T^2h^{2\min\{s,k\}+1-\alpha}\|p\|^2_{L^\infty(I;H^{1+s}(\Omega))} + \frac{1}{2\mu}\|\mathsf{C}_{22}^{\frac{1}{2}}[\![\mathbf{\Pi}\vec{\mathbf{q}}-\vec{\mathbf{q}}]\!]\|^2_{L^\infty(I;\mathbb{L}^2(\mathcal{E}_h))}\\
&\leq \frac{\mu}{2}{C_R^p}^2(1+T^2)h^{2\min\{s,k\}+1-\alpha}\|p\|^2_{L^\infty(I;H^{1+s}(\Omega))} + \frac{1}{2\mu}\|\mathsf{C}_{22}^{\frac{1}{2}}[\![e^{\vec{\mathbf{q}}}]\!]\|^2_{L^1(I;H^{1+s}(\mathcal{E}_h))}\\
&\qquad + \frac{1}{2\mu}h^{2\min\{s,k\}+1+\alpha}\|\vec{\mathbf{q}}\|^2_{L^\infty(I;\mathbb{H}^{1+s}(\Omega))},
\end{aligned}
$$

and

$$
\begin{aligned}
\int_I \left|\mathcal{R}^q(\vec{\mathbf{q}},\Pi_h p-p^h)\right| dt \;&\leq \frac{\mu}{2}{C_R^q}^2(1+T^2)h^{2\min\{s,k\}+2}\|\vec{\mathbf{q}}\|^2_{L^\infty(I;\mathbb{H}^{1+s}(\Omega))}\\
&\quad + \frac{1}{2\mu}\|\mathsf{C}_{11}^{\frac{1}{2}}[\![e^p]\!]\|^2_{L^1(I;H^{1+s}(\mathcal{E}_h))}\\
&\quad + \frac{1}{2\mu}h^{2\min\{s,k\}+1+\alpha}\|p\|^2_{L^\infty(I;H^{1+s}(\Omega))}.
\end{aligned}
$$

Combining the above estimates and T_1, T_2, with $\varepsilon = 4$, and $\mu = 2$, then we have that

$$\frac{1}{4}\sup_{t\in I}\left(\|\frac{1}{c}e^p\|^2_{L^2(\Omega)} + \|e^{\vec{\mathbf{q}}}\|^2_{\mathbb{L}^2(\Omega)}\right) + \frac{1}{2}\|\mathtt{C}_{11}^{\frac{1}{2}}[\![e^p]\!]\|^2_{L^1(I;L^2(\mathcal{E}_h))} + \frac{1}{2}\|\mathtt{C}_{22}^{\frac{1}{2}}[\![e^{\vec{\mathbf{q}}}]\!]\|^2_{L^1(I;\mathbb{L}^2(\mathcal{E}_h))}$$

$$\leq \frac{1}{2}\|\frac{1}{c}e^p(0)\|^2_{0,\Omega} + \frac{1}{2}\|e^{\vec{\mathbf{q}}}(0)\|^2_{0,\Omega}$$

$$+Ch^{2\min\{s,k+1/2\}}\left(\|p\|^2_{L^\infty(I;H^{1+s}(\Omega))} + \|\vec{\mathbf{q}}\|^2_{L^\infty(I;\mathbb{H}^{1+s}(\Omega))} + \|p_t\|^2_{L^1(I;H^s(\Omega))} + \|\vec{\mathbf{q}}_t\|^2_{L^1(I;\mathbb{H}^s(\Omega))}\right),$$

with a constant that is independent of the mesh size h taking $\alpha = 0$. Using the bound $\frac{1}{c^{*2}}\|e^p\|^2_{L^2(\Omega)} \leq \|\frac{1}{c}e^p\|^2_{L^2(\Omega)}$, we conclude the proof of Theorem 5.3. □

Remark 5.5 *In our DG method the parameters are independent of mesh size h which gives higher accuracy of the L^2-norms of errors in p and $\vec{\mathbf{q}}$ with $k+\frac{1}{2}$, $k+\frac{1}{2}$, respectively for smooth solutions $(p, \vec{\mathbf{q}})$. But, in IP methods, i.e., the stabilization parameters $\mathtt{C}_{11}, \mathtt{C}_{22}$ of order $\mathcal{O}(h^{-1})$, we lose accuracy $\frac{1}{2}$ for p and $\vec{\mathbf{q}}$, respectively, from the interior penalty flux. (see proof of Theorem* 5.3 *with $\alpha = -1$).*

6. FULLY DISCRETIZED SCHEME ERROR ESTIMATION

6.1. Fully Discretized Discontinuous Galerkin Method for the system

In this section, we present the fully discrete discontinuous Galerkin method for the system (5.22)-(5.23), which extends the spatial discretization in Chapter 5.2. to fully discrete scheme. The discretization in space is based on the discontinuous Galerkin method while the time discretization is based on the *θ-scheme* finite difference approximation. Especially we apply the Trapezoidal method in time discretization, i.e., when $\theta = \frac{1}{2}$. We show an *a priori* L^2-norm error estimate for the scheme, following [31].

6.1.1 Time discretization

We now use the θ-scheme to discretize in time the system of equations (5.22)-(5.23). To that end we introduce a time step $\triangle t = T/N$ and define the discrete times $t_n = n\triangle t$ for $n = 0, \cdots, N$. For a (sufficiently smooth) function $r(x,t)$, $\vec{\mathbf{v}}(x,t)$, we set

$$\partial_t r^n = \partial_t r(\cdot, t_n), \qquad \partial_t \vec{\mathbf{v}}^n = \partial_t \vec{\mathbf{v}}(\cdot, t_n).$$

Let $(p, \vec{\mathbf{q}})$ be the solution to the system (5.2). We wish to find DG approximations $\{(P^n, \vec{\mathbf{Q}}^n)\}$ such that $(P^n, \vec{\mathbf{Q}}^n) \approx (p^n, \vec{\mathbf{q}}^n)$ at the discrete times t_n. To do so, we introduce the finite difference operator

$$\Delta P^n = \frac{P^{n+1} - P^n}{\triangle t}, \qquad \boldsymbol{\Delta} \vec{\mathbf{Q}}^n = \frac{\vec{\mathbf{Q}}^{n+1} - \vec{\mathbf{Q}}^n}{\triangle t}, \ n = 0, \cdots, N-1. \tag{6.1}$$

The fully discrete numerical solution to the system (5.22)-(5.23) is then defined by finding $\{(P^n, \vec{\mathbf{Q}}^n)\}$ such that

$$(\frac{1}{c^2}\Delta P^n, r^h) + a_h((1-\theta)P^n + \theta P^{n+1}, r^h) - b_h'((1-\theta)\vec{\mathbf{Q}}^n + \theta \vec{\mathbf{Q}}^{n+1}, r^h) = 0, \tag{6.2}$$

$$(\boldsymbol{\Delta} \vec{\mathbf{Q}}^n, \mathbf{v}^h) + b_h((1-\theta)P^n + \theta P^{n+1}, \vec{\mathbf{v}}^h) + c_h((1-\theta)\vec{\mathbf{Q}}^n + \theta \vec{\mathbf{Q}}^{n+1}, \vec{\mathbf{v}}^h) = 0, \tag{6.3}$$

for all $n = 1, \cdots, N-1$, and for all $(r^h, \vec{\mathbf{v}}^h) \in \mathcal{P}^h(\Omega) \times \mathcal{Q}^h(\Omega)$. Especially when $\theta = \frac{1}{2}$ it gives the Trapezoidal method

$$(\frac{1}{c^2}\Delta P^n, r^h) + \frac{1}{2}a_h(P^n + P^{n+1}, r^h) - \frac{1}{2}b_h'(\vec{\mathbf{Q}}^n + \vec{\mathbf{Q}}^{n+1}, r^h) = 0, \tag{6.4}$$

$$(\boldsymbol{\Delta}\vec{\mathbf{Q}}^n, \vec{\mathbf{v}}^h) + \frac{1}{2}b_h(P^n + P^{n+1}, \vec{\mathbf{v}}^h) + \frac{1}{2}c_h(\vec{\mathbf{Q}}^n + \vec{\mathbf{Q}}^{n+1}, \vec{\mathbf{v}}^h) = 0, \tag{6.5}$$

for all $n = 0, \cdots, N-1$, and for all $(r^h, \vec{\mathbf{v}}^h) \in \mathcal{P}^h(\Omega) \times \mathcal{Q}^h(\Omega)$.

The initial conditions $P^0 \in \mathcal{P}^h(\Omega)$ and $\mathbf{Q}^0 \in \mathcal{Q}^h(\Omega)$ are given by

$$P^0 = \Pi_h f, \quad \mathbf{Q}^0 = \vec{\mathbf{0}}.$$

In the above equations in (6.4), (6.5), every time step involves the inversion of the DG mass matrix. Since it is an invertible block matrix $\begin{bmatrix} A & -B \\ B^T & C \end{bmatrix}$, where A and C are symmetric positive definite, the new approximations P^{n+1} and $\vec{\mathbf{Q}}^{n+1}$ are well-defined for $n \geq 1$. Therefore the fully discrete DG approximations $\{(P^n, \vec{\mathbf{Q}}^{n+1})\}$ are uniquely defined, which completes the definition of fully discrete DG methods for the system.

Remark 6.1 *The fully discretized scheme* (6.4), (6.5) *is unconditionally stable. Choosing* $r^h = P^n + P^{n+1}, \vec{\mathbf{v}}^h = \vec{\mathbf{Q}}^n + \vec{\mathbf{Q}}^{n+1}$ *with the property in* (5.43)*, we have that*

$$\begin{aligned}
\|\frac{1}{c}P^{n+1}\|_{0,\Omega}^2 + \|\vec{\mathbf{Q}}^{n+1}\|_{0,\Omega}^2 =& \|\frac{1}{c}P^n\|_{0,\Omega}^2 + \|\vec{\mathbf{Q}}^n\|_{0,\Omega}^2 \\
& -\frac{\triangle t}{2}\left(a_h(P^n + P^{n+1}, P^n + P^{n+1}) + c_h(\vec{\mathbf{Q}}^n + \vec{\mathbf{Q}}^{n+1}, \vec{\mathbf{Q}}^n + \vec{\mathbf{Q}}^{n+1})\right) \\
\leq& \|\frac{1}{c}P^n\|_{0,\Omega}^2 + \|\vec{\mathbf{Q}}^n\|_{0,\Omega}^2,
\end{aligned}$$

for all $n = 0, \cdots, N-1$. *Therefore we obtain that* $\|\frac{1}{c}P^N\|_{0,\Omega}^2 + \|\vec{\mathbf{Q}}^N\|_{0,\Omega}^2 \leq \|\frac{1}{c}P^0\|_{0,\Omega}^2 + \|\vec{\mathbf{Q}}^0\|_{0,\Omega}^2$, *which is independent of size* $\triangle t$ *and* h.

6.1.2 An *A priori* Estimate

In this section, we state *a priori* error estimate for the fully discrete DG method introduced above. We decompose the error e^n at time t_n into

$$\begin{aligned} e^n &= e^{p^n} + e^{\vec{\mathbf{q}}^n} \\ &= (p^n - P^n) + (\vec{\mathbf{q}}^n - \vec{\mathbf{Q}}^n) \\ &= (p^n - \Pi_h p^n + \Pi_h p^n - P^n) + (\vec{\mathbf{q}}^n - \mathbf{\Pi}_h \vec{\mathbf{q}}^n + \mathbf{\Pi}_h \vec{\mathbf{q}}^n - \vec{\mathbf{Q}}^n)\ , n = 0, \cdots, N, \end{aligned}$$

where $p^n = p(\cdot, t_n)$ and $\vec{\mathbf{q}}^n = \vec{\mathbf{q}}(\cdot, t_n)$.

Our main result establishes an error estimate of the L^2-norm of the error. The following result holds.

Theorem 6.1 *Let the solution $(p, \vec{\mathbf{q}})$ of the system satisfy the following properties for a regularity constant $s > \frac{1}{2}$.*

$$p \in C(\bar{I}; H^{1+s}(\Omega)), \quad p_{tt} \in C(\bar{I}; H^s(\Omega)), \quad \partial_t^3 p \in L^1(I; L^2(\Omega)),$$

$$\vec{\mathbf{q}} \in C(\bar{I}; \mathbb{H}^{1+s}(\Omega)), \quad \vec{\mathbf{q}}_{tt} \in C(\bar{I}; \mathbb{H}^s(\Omega)), \quad \partial_t^3 \vec{\mathbf{q}} \in L^1(I; \mathbb{L}^2(\Omega)).$$

Then there holds the error estimate

$$\max_{n=0}^{N}(\|p^n - P^n\| + \|\mathbf{q}^n - \vec{\mathbf{Q}}^n\|) \ \leq \ C(h^{\min\{s, k+\frac{1}{2}\}} + \triangle t^2),$$

with a constant $C > 0$ that is independent of the mesh size h and time step size $\triangle t$. Then the numerical solution holds the accuracy up to $h^{k+\frac{1}{2}} + \triangle t^2$ for a smooth solution $(p, \vec{\mathbf{q}})$.

We denotes the differences between numerical solutions and the projection of analytical solutions for $n = 0, \cdots, N-1$, by

$$\Phi_p^n \ = \ P^n - \Pi_h p^n, \quad \Phi_{\mathbf{q}}^n \ = \ \vec{\mathbf{Q}}^n - \mathbf{\Pi}_h \vec{\mathbf{q}}^n,$$

$$R_p^n \ = \ \Pi_h p^n - p^n, \quad R_{\mathbf{q}}^n \ = \ \mathbf{\Pi}_h \vec{\mathbf{q}}^n - \vec{\mathbf{q}}^n.$$

Note that the initial condition $\Phi_p^0 \ = \ \Pi_h p^0 - P^0 = 0$ and $\Phi_{\mathbf{q}}^0 \ = \ \mathbf{\Pi}_h \vec{\mathbf{q}}^0 - \vec{\mathbf{Q}}^0 = \vec{\mathbf{0}}$. The following approximation properties hold by Lemma 5.6.

Lemma 6.1 *For $0 \le n \le N-1$, the following holds:*

$$\begin{aligned}
\|[\![R_p^n + R_p^{n+1}]\!]\|_{0,\Omega} &\le Ch^{\min\{s,k\}+\frac{1}{2}} \left(\|p^n\|_{1+s,\Omega} + \|p^{n+1}\|_{1+s,\Omega}\right), \\
\|\{\!\{R_p^n + R_p^{n+1}\}\!\}\|_{0,\Omega} &\le Ch^{\min\{s,k\}+\frac{1}{2}} \left(\|p^n\|_{1+s,\Omega} + \|p^{n+1}\|_{1+s,\Omega}\right),
\end{aligned}$$

and

$$\|R_p^n + R_p^{n+1}\|_{0,\Omega} \le Ch^{\min\{s,k\}+1} \left(\|p^n\|_{1+s,\Omega} + \|p^{n+1}\|_{1+s,\Omega}\right),$$

with a constant $C > 0$ that is independent of $h, \triangle t$, and T.

Proof. Using the definition of R_p^n and the property in Lemma 5.6 and Lemma 5.5 we have that

$$\begin{aligned}
\|[\![R_p^n + R_p^{n+1}]\!]\|_{0,\Omega}^2 &\le 2\left(\|[\![\Pi_h p^n - p^n]\!]\|_{0,\Omega}^2 + \|[\![\Pi_h p^{n+1} - p^{n+1}]\!]\|_{0,\Omega}^2\right) \\
&\le Ch^{2\min\{s,k\}+1} \left(\|p^n\|_{1+s}^2 + \|p^{n+1}\|_{1+s,\Omega}^2\right),
\end{aligned}$$

$$\begin{aligned}
\|\{\!\{R_p^n + R_p^{n+1}\}\!\}\|_{0,\Omega}^2 &\le 2\left(\|\{\!\{\Pi_h p^n - p^n\}\!\}\|_{0,\Omega}^2 + \|\{\!\{\Pi_h p^{n+1} - p^{n+1}\}\!\}\|_{0,\Omega}^2\right) \\
&\le Ch^{2\min\{s,k\}+1} \left(\|p^n\|_{1+s,\Omega}^2 + \|p^{n+1}\|_{1+s,\Omega}^2\right),
\end{aligned}$$

and

$$\begin{aligned}
\|R_p^n + R_p^{n+1}\|_{0,\Omega}^2 &\le 2\left(\|R_p^n\|_{0,\Omega}^2 + \|R_p^{n+1}\|_{0,\Omega}^2\right) \\
&\le 2\left(\|(\Pi_h p^n - p^n\|_{0,\Omega}^2 + \|\Pi_h p^{n+1} - p^{n+1}\|_{0,\Omega}^2\right) \\
&\le Ch^{2\min\{s,k\}+2} \left(\|p^n\|_{1+s,\Omega}^2 + \|p^{n+1}\|_{1+s,\Omega}^2\right)
\end{aligned}$$

for $n = 0, \cdots, N-1$. □

The previous approximation results are satisfied by $R_{\mathbf{q}}^n$ similarily.

Now we consider that

$$\begin{aligned}
b_h(p, \vec{\mathbf{v}}) &= -\sum_{K\in\mathcal{T}_h} \int_K p\nabla\cdot\vec{\mathbf{v}}dx - \sum_{e\in\mathcal{E}_h} \int_e [\![\vec{\mathbf{v}}]\!] \left(\vec{\mathtt{C}}_{12}\cdot[\![p]\!] - \{\!\{p\}\!\}\right) ds \\
&= \sum_{K\in\mathcal{T}_h} \int_K \nabla p\cdot\vec{\mathbf{v}}dx - \sum_{e\in\mathcal{E}_h} \int_e [\![p]\!]\cdot\left(\vec{\mathtt{C}}_{12}[\![\vec{\mathbf{v}}]\!] + \{\!\{\vec{\mathbf{v}}\}\!\}\right) ds
\end{aligned}$$

by Green's identity and the trace identity (5.8). Then we have the following bounds by the property of L^2-projection $\int_K R_p \nabla \cdot \Phi_{\mathbf{q}} dx = 0$ and $\int_K R_{\mathbf{q}} \nabla \cdot \Phi_p dx = 0$.

Remark 6.2 *It holds that for* $n = 0, \cdots, N-1$, *for any* $\varepsilon_0 > 0,\ \varepsilon_1 > 0$,

$$b_h(\Phi_p^n, \Phi_{\mathbf{q}}^n) \ - \ b_h'(\Phi_{\mathbf{q}}^n, \Phi_p^n) \ = \ 0,$$

and

$$\begin{aligned} & b_h(R_p^n, \Phi_{\mathbf{q}}^n) \ - \ b_h'(R_{\mathbf{q}}^n, \Phi_p^n) \\ = & -\sum_{e \in \mathcal{E}_h} \int_e [\![\Phi_{\mathbf{q}}^n]\!] \left(\vec{\mathtt{C}}_{12} \cdot [\![R_p^n]\!] - \{\!\{R_p^n\}\!\} \right) ds + \sum_{e \in \mathcal{E}_h} \int_e \left(\vec{\mathtt{C}}_{12} [\![R_{\mathbf{q}}^n]\!] + \{\!\{R_{\mathbf{q}}^n\}\!\} \right) \cdot [\![\Phi_p^n]\!] ds. \end{aligned}$$

Next the finite difference operators $\Delta_p^n, \boldsymbol{\Delta}_{\mathbf{q}}^n$ are bounded as follows.

Lemma 6.2 *We set*

$$\frac{1}{c} \Delta_p^n \ = \ \frac{1}{c} \Delta \Pi_h p^n \ - \ \frac{p_t^n + p_t^{n+1}}{2c}, \quad n = 0, \cdots, N-1,$$

and

$$\boldsymbol{\Delta}_{\mathbf{q}}^n \ = \ \boldsymbol{\Delta} \boldsymbol{\Pi}_h \vec{\mathbf{q}}^n \ - \ \frac{\vec{\mathbf{q}}_t^n + \vec{\mathbf{q}}_t^{n+1}}{2}, \quad n = 0, \cdots, N-1,$$

For $0 \le n \le N-1$, *there holds*

$$\left\| \frac{1}{c} \Delta_p^n \right\|_{0,\Omega} \ \le \ C \left(\frac{1}{c_* \triangle t} \int_{t_n}^{t_{n+1}} h^{\min\{s,k+1\}} \|p_t\|_{s,\Omega} d\tau + \frac{\triangle t}{4 c_*} \int_{t_n}^{t_{n+1}} \|\partial_t^3 p\|_{0,\Omega} d\tau \right),$$

and

$$\left\| \boldsymbol{\Delta}_{\mathbf{q}}^n \right\|_{0,\Omega} \ \le \ C \left(\frac{1}{\triangle t} \int_{t_n}^{t_{n+1}} h^{\min\{s,k+1\}} \|\vec{\mathbf{q}}_t\|_{s,\Omega} d\tau + \frac{\triangle t}{4} \int_{t_n}^{t_{n+1}} \|\partial_t^3 \vec{\mathbf{q}}\|_{0,\Omega} d\tau \right).$$

Proof. We can split the expression to estimate it

$$\left\| \frac{1}{c} \Delta \Pi_h p^n - \frac{p_t^n + p_t^{n+1}}{2c} \right\|_{0,\Omega} \ \le \ \left\| \frac{1}{c} \Delta (\Pi_h p^n - p^n) \right\|_{0,\Omega} + \left\| \frac{1}{c} \Delta p^n - \frac{p_t^n + p_t^{n+1}}{2c^2} \right\|_{0,\Omega}. \quad (6.6)$$

To bound the first term on the right-hand side of (6.6), we use the identity

$$p(\cdot, t_{n+1}) \ - \ p(\cdot, t_n) \ = \ \int_{t_n}^{t_{n+1}} p_t(\cdot, \tau) d\tau,$$

which is Fundamental Theorem of calculus. By the property $\partial_t(\Pi_h p - p) = \Pi_h p_t - p_t$ and Lemma 5.5, we obtain that

$$\left\|\frac{1}{c}\Delta(\Pi_h p^n - p^n)\right\|_{0,\Omega} \leq \frac{1}{c_* \triangle t}\int_{t_n}^{t_{n+1}} \|(\Pi_h p - p)_t\|_{0,\Omega} ds \leq C\frac{1}{c_* \triangle t}\int_{t_n}^{t_{n+1}} h^{\min\{s,k+1\}}\|p_t\|_{s,\Omega} d\tau.$$

To estimate the second term on the right-hand side of the equation in (6.6), we also use the following identity,

$$-\frac{p_t^n + p_t^{n+1}}{2} + \Delta p^n = \frac{1}{2\triangle t}\int_{t_n}^{t_{n+1}} \left((t_{n+\frac{1}{2}} - \tau)^2 - \left(\frac{\triangle t}{2}\right)^2\right)\partial_t^3 p(\cdot,\tau) d\tau,$$

which is obtained from Taylor's formula with integral remainder,

$$\begin{aligned}
-\frac{p_t^n + p_t^{n+1}}{2} + \Delta p^n &= \frac{1}{\triangle t}\left[(t_{n+\frac{1}{2}} - \tau)p_t(\cdot,\tau)\right]_{t_n}^{t_{n+1}} + \frac{1}{\triangle t}\int_{t_n}^{t_{n+1}} p_t(\cdot,\tau) d\tau \\
&= \frac{1}{2\triangle t}\left[\left\{(t_{n+\frac{1}{2}} - \tau)^2 - (\frac{1}{2}\triangle t)^2\right\} p_{tt}(\cdot,\tau)\right]_{t_n}^{t_{n+1}} + \frac{1}{\triangle t}\int_{t_n}^{t_{n+1}} (t_{n+\frac{1}{2}} - \tau)p_{tt}(\cdot,\tau) d\tau \\
&= \frac{1}{2\triangle t}\int_{t_n}^{t_{n+1}} \left((t_{n+\frac{1}{2}} - \tau)^2 - \left(\frac{\triangle t}{2}\right)^2\right)\partial_t^3 p(\cdot,\tau) d\tau.
\end{aligned}$$

Since $|t_{n+\frac{1}{2}} - \tau| \leq \frac{\triangle t}{2}$, we can deduce that

$$\begin{aligned}
\left\|\frac{1}{c}\left(\Delta p^n - \frac{p_t^n + p_t^{n+1}}{2}\right)\right\|_{0,\Omega} &\leq \frac{1}{c_*}\left\|\frac{1}{2\triangle t}\int_{t_n}^{t_{n+1}} \left(\left(t_{n+\frac{1}{2}} - \tau\right)^2 - \left(\frac{\triangle t}{2}\right)^2\right)\partial_t^3 p(\cdot,\tau)\ d\tau\right\|_{0,\Omega} \\
&\leq \frac{\triangle t}{4c_*}\int_{t_n}^{t_{n+1}} \|\partial_t^3 p\|_{0,\Omega} d\tau.
\end{aligned}$$

Similarly, we have the same estimation for $\vec{\mathbf{q}}$. □

6.1.3 Proof of the main Theorem 6.1

We are now ready to complete the proof of Theorem 6.1 . By the triangle inequality, we have that

$$\begin{aligned}
&\max_{n=0}^{N} \|e^n\|_{0,\Omega} \\
\leq \max_{n=0}^{N}&\left\{\|\Pi_h p^n - P^n\|_0 + \|\Pi_h \vec{\mathbf{q}}^n - \vec{\mathbf{Q}}^n\|_{0,\Omega}\right\} + \max_{n=0}^{N}\left\{\|p^n - \Pi_h p^n\|_{0,\Omega} + \|\vec{\mathbf{q}}^n - \mathbf{\Pi}_h \vec{\mathbf{q}}^n\|_{0,\Omega}\right\} \\
&\leq \max_{n=0}^{N}\{\|\Phi_p^n\|_{0,\Omega} + \|\Phi_{\mathbf{q}}^n\|_{0,\Omega}\} + \max_{n=0}^{N}\{\|R_p^n\|_{0,\Omega} + \|R_{\mathbf{q}}^n\|_{0,\Omega}\}.
\end{aligned}$$

In this equation, the approximation properties of the L^2-projection show that

$$\|R_p^n\|_{0,\Omega} \;=\; \|p^n - \Pi_h p^n\|_{0,\Omega} \;\leq\; Ch^{\min\{s,k\}+1}\|p^n\|_{1+s,\Omega}, \tag{6.7}$$

and

$$\|R_{\mathbf{q}}^n\|_{0,\Omega} \;=\; \|\vec{\mathbf{q}}^n - \mathbf{\Pi}_h\vec{\mathbf{q}}^n\|_{0,\Omega} \;\leq\; Ch^{\min\{s,k\}+1}\|\vec{\mathbf{q}}^n\|_{1+s,\Omega}. \tag{6.8}$$

for all $n = 0, \cdots, N$. Therefore we only need to estimate Φ_p^n and $\Phi_{\mathbf{q}}^n$ in order to estimate the error in L^2-norm for all $n = 1, \cdots, N$.

Recall that for $(p, \vec{\mathbf{q}})$ the exact solution of (5.2) it holds that

$$\begin{aligned}
(\frac{1}{c^2}p_t, r^h) \;+\; a_h(p, r^h) \;-\; b_h'(\vec{\mathbf{q}}, r^h) \;&=\; 0, \qquad && \forall r^h \in \mathcal{P}^h(\Omega), t \in I,\\
(\vec{\mathbf{q}}_t, \vec{\mathbf{v}}^h) \;+\; b_h(p, \vec{\mathbf{v}}^h) \;+\; c_h(\vec{\mathbf{q}}, \vec{\mathbf{v}}^h) \;&=\; 0, \qquad && \forall \vec{\mathbf{v}}^h \in \mathcal{Q}^h(\Omega), t \in I.
\end{aligned}$$

Trivially it holds for all $n = 0, \cdots, N$,

$$(\frac{1}{c^2}p_t^n, r^h) \;+\; a_h(p^n, r^h) \;-\; b_h'(\vec{\mathbf{q}}^n, r^h) \;=\; 0, \qquad \forall r^h \in \mathcal{P}^h(\Omega), t \in I, \tag{6.9}$$

$$(\vec{\mathbf{q}}_t^n, \vec{\mathbf{v}}^h) \;+\; b_h(p^n, \vec{\mathbf{v}}^h) \;+\; c_h(\vec{\mathbf{q}}^n, \vec{\mathbf{v}}^h) \;=\; 0, \qquad \forall \vec{\mathbf{v}}^h \in \mathcal{Q}^h(\Omega), t \in I. \tag{6.10}$$

We subtract (6.9), (6.10) from (6.2), (6.3), add $\Pi_h p^n, \Pi_h p^{n+1}$, and subtract them again, respectively, in order to use the notations $\Phi_p^n, \Phi_{\mathbf{q}}^n, R_p^n$, and $R_{\mathbf{q}}^n$ for each time step n, then we can obtain the following:

$$\left(\frac{1}{c^2}\left(\Delta\Phi_p^n + \Delta\Pi p^n - \frac{p_t^n + p_t^{n+1}}{2}\right), r^h\right)$$

$$+\frac{1}{2}a_h(\Phi_p^n + \Phi_p^{n+1} + R_p^n + R_p^{n+1}, r^h) - \frac{1}{2}b_h'(\Phi_{\mathbf{q}}^n + \Phi_{\mathbf{q}}^{n+1} + R_{\mathbf{q}}^n + R_{\mathbf{q}}^{n+1}, r^h) \;=\; 0,$$

and

$$\left(\Delta\Phi_{\mathbf{q}}^n + \Delta\mathbf{\Pi}_h\vec{\mathbf{q}}^n - \frac{\vec{\mathbf{q}}_t^n + \vec{\mathbf{q}}_t^{n+1}}{2}, \vec{\mathbf{v}}^h\right)$$

$$+\frac{1}{2}b_h(\Phi_p^n + \Phi_p^{n+1} + R_p^n + R_p^{n+1}, \vec{\mathbf{v}}^h) + \frac{1}{2}c_h(\Phi_{\mathbf{q}}^n + \Phi_{\mathbf{q}}^{n+1} + R_{\mathbf{q}}^n + R_{\mathbf{q}}^{n+1}, \vec{\mathbf{v}}^h) \;=\; 0.$$

Next, we choose $r^h = \Phi_p^n + \Phi_p^{n+1} \in \mathcal{P}^h(\Omega)$ and $\vec{\mathbf{v}}^h = \Phi_{\mathbf{q}}^n + \Phi_{\mathbf{q}}^{n+1} \in \mathcal{Q}^h(\Omega)$, multiply the resulting expression by $\triangle t$, and add two equations to have that

$$\|\frac{1}{c}\Phi_p^{n+1}\|_{0,\Omega}^2 + \|\Phi_{\mathbf{q}}^{n+1}\|_{0,\Omega}^2 - \|\frac{1}{c}\Phi_p^n\|_{0,\Omega}^2 - \|\Phi_{\mathbf{q}}^n\|_{0,\Omega}^2 + \frac{\triangle t}{2}\left(\mathtt{A}_{\Phi\Phi}^n + \mathtt{A}_{R\Phi}^n - \mathtt{B}'^n_{R\Phi} + \mathtt{B}_{R\Phi}^n + \mathtt{C}_{\Phi\Phi}^n + \mathtt{C}_{R\Phi}^n\right)$$
$$+\frac{\triangle t}{c^2}\left(\Delta\Pi_h p^n - \frac{p_t^n + p_t^{n+1}}{2}, \Phi_p^n + \Phi_p^{n+1}\right) + \triangle t\left(\Delta\Pi_h \vec{\mathbf{q}}^n - \frac{\vec{\mathbf{q}}_t^n + \vec{\mathbf{q}}_t^{n+1}}{2}, \Phi_{\mathbf{q}}^n + \Phi_{\mathbf{q}}^{n+1}\right) = 0,$$

where

$$\begin{aligned}
\mathtt{A}_{\Phi\Phi}^n &= a_h\left(\Phi_p^n + \Phi_p^{n+1},\ \Phi_p^n + \Phi_p^{n+1}\right), \qquad \mathtt{A}_{R\Phi}^n = a_h\left(R_p^n + R_p^{n+1},\ \Phi_p^n + \Phi_p^{n+1}\right),\\
\mathtt{B}_{R\Phi}^n &= b_h\left(R_p^n + R_p^{n+1},\ \Phi_{\mathbf{q}}^n + \Phi_{\mathbf{q}}^{n+1}\right), \qquad \mathtt{B}'^n_{R\Phi} = b_h'\left(R_{\mathbf{q}}^n + R_{\mathbf{q}}^{n+1},\ \Phi_p^n + \Phi_p^{n+1}\right), \quad (6.11)\\
\mathtt{C}_{\Phi\Phi}^n &= c_h\left(\Phi_{\mathbf{q}}^n + \Phi_{\mathbf{q}}^{n+1},\ \Phi_{\mathbf{q}}^n + \Phi_{\mathbf{q}}^{n+1}\right), \qquad \mathtt{C}_{R\Phi}^n = c_h\left(R_{\mathbf{q}}^n + R_{\mathbf{q}}^{n+1},\ \Phi_{\mathbf{q}}^n + \Phi_{\mathbf{q}}^{n+1}\right).
\end{aligned}$$

Here we have also used that $b_h\left(\Phi_p^n + \Phi_p^{n+1}, \Phi_{\mathbf{q}}^n + \Phi_{\mathbf{q}}^{n+1}\right) - b_h'\left(\Phi_{\mathbf{q}}^n + \Phi_{\mathbf{q}}^{n+1}, \Phi_p^n + \Phi_p^{n+1}\right) = 0$. Summation from $n = 0$ to $n = m$, for $0 \le m \le N-1$, shows that

$$\|\frac{1}{c}\Phi_p^{m+1}\|_{0,\Omega}^2 + \|\Phi_{\mathbf{q}}^{m+1}\|_{0,\Omega}^2 + \frac{\triangle t}{2}\sum_{n=0}^{m}(\mathtt{A}_{\Phi\Phi}^n + \mathtt{C}_{\Phi\Phi}^n) \le \|\frac{1}{c}\Phi_p^0\|_{0,\Omega}^2 + \|\Phi_q^0\|_{0,\Omega}^2$$
$$+\frac{\triangle t}{2}\sum_{n=0}^{m}\left(|\mathtt{A}_{R\Phi}^n| + |\mathtt{C}_{R\Phi}^n| + |\mathtt{B}_{R\Phi}^n| + |\mathtt{B}'^n_{R\Phi}|\right)$$
$$+\triangle t\sum_{n=0}^{m}\left|\left(\frac{1}{c}\Delta_p^n, \frac{1}{c}(\Phi_p^n + \Phi_p^{n+1})\right)\right| + \triangle t\sum_{n=0}^{m}\left|\left(\Delta_{\mathbf{q}}^n, \Phi_{\mathbf{q}}^n + \Phi_{\mathbf{q}}^{n+1}\right)\right|.$$

Now we use the definitions in (6.11), the Cauchy-Schwarz inequality, the geometric-arithmetic inequality $ab \le \frac{a^2}{2\varepsilon} + \frac{\varepsilon b^2}{2}$ for any $\varepsilon > 0$, and also Remark 6.2 on right hand side in the previous equation,

$$\begin{aligned}
\mathtt{A}_{\Phi\Phi}^n &= \|\frac{\sigma_p^{\frac{1}{2}}}{c}(\Phi_p^n + \Phi_p^{n+1})\|_{0,\Omega}^2 + \|\mathtt{C}_{11}^{\frac{1}{2}}[\![\Phi_p^n + \Phi_p^{n+1}]\!]\|_{0,\mathcal{E}_h}^2,\\
\mathtt{C}_{\Phi\Phi}^n &= \|\sigma_q^{\frac{1}{2}}(\Phi_{\mathbf{q}}^n + \Phi_{\mathbf{q}}^{n+1})\|_{0,\Omega}^2 + \|\mathtt{C}_{22}^{\frac{1}{2}}[\![\Phi_{\mathbf{q}}^n + \Phi_{\mathbf{q}}^{n+1}]\!]\|_{0,\mathcal{E}_h}^2,
\end{aligned}$$

$$\mathtt{A}_{R\Phi}^n$$
$$\le \|\frac{\sigma_p}{c}(R_p^n + R_p^{n+1})\|_{0,\Omega}\|\frac{1}{c}(\Phi_p^n + \Phi_p^{n+1})\|_{0,\Omega} + \|\mathtt{C}_{11}^{\frac{1}{2}}[\![R_p^n + R_p^{n+1}]\!]\|_{0,\mathcal{E}_h}\|\mathtt{C}_{11}^{\frac{1}{2}}[\![\Phi_p^n + \Phi_p^{n+1}]\!]\|_{0,\mathcal{E}_h}$$
$$\le \|\frac{\sigma_p}{c}(R_p^n + R_p^{n+1})\|_{0,\Omega}\|\frac{1}{c}(\Phi_p^n + \Phi_p^{n+1})\|_{0,\Omega} + \frac{1}{2\varepsilon}\|\mathtt{C}_{11}^{\frac{1}{2}}[\![R_p^n + R_p^{n+1}]\!]\|_{0,\mathcal{E}_h}^2 + \frac{\varepsilon}{2}\|\mathtt{C}_{11}^{\frac{1}{2}}[\![\Phi_p^n + \Phi_p^{n+1}]\!]\|_{0,\mathcal{E}_h}^2,$$

$$\mathtt{C}^n_{R\Phi}$$

$$\leq \|\sigma_q(R^n_{\mathbf{q}}+R^{n+1}_{\mathbf{q}})\|_{0,\Omega}\|\Phi^n_{\mathbf{q}}+\Phi^{n+1}_{\mathbf{q}}\|_{0,\Omega}+\|\mathtt{C}_{22}^{\frac{1}{2}}[\![R^n_{\mathbf{q}}+R^{n+1}_{\mathbf{q}}]\!]\|_0\|\mathtt{C}_{22}^{\frac{1}{2}}[\![\Phi^n_{\mathbf{q}}+\Phi^{n+1}_{\mathbf{q}}]\!]\|_{0,\mathcal{E}_h}$$

$$\leq \|\sigma_q(R^n_{\mathbf{q}}+R^{n+1}_{\mathbf{q}})\|_{0,\Omega}\|\Phi^n_{\mathbf{q}}+\Phi^{n+1}_{\mathbf{q}}\|_{0,\Omega}+\frac{1}{2\varepsilon}\|\mathtt{C}_{22}^{\frac{1}{2}}[\![R^n_{\mathbf{q}}+R^{n+1}_{\mathbf{q}}]\!]\|^2_{0,\mathcal{E}_h}+\frac{\varepsilon}{2}\|\mathtt{C}_{22}^{\frac{1}{2}}[\![\Phi^n_{\mathbf{q}}+\Phi^{n+1}_{\mathbf{q}}]\!]\|^2_{0,\mathcal{E}_h},$$

and

$$\begin{aligned}\mathtt{B}^n_{R\Phi} \;-\; \mathtt{B}'^n_{R\Phi} \;\leq\; & \frac{\varepsilon}{2}\left(\|\mathtt{C}_{11}^{\frac{1}{2}}[\![\Phi^n_p+\Phi^{n+1}_p]\!]\|^2_{0,\mathcal{E}_h}+\|\mathtt{C}_{22}^{\frac{1}{2}}[\![\Phi^n_{\mathbf{q}}+\Phi^{n+1}_{\mathbf{q}}]\!]\|^2_{0,\mathcal{E}_h}\right)\\ & +\frac{1}{2\varepsilon}\left(2|\vec{\mathtt{C}}_{12}|^2\|\mathtt{C}_{11}^{-\frac{1}{2}}[\![R^n_p+R^{n+1}_p]\!]\|^2_{0,\mathcal{E}_h}+2\|\mathtt{C}_{11}^{-\frac{1}{2}}\{\!\{R^n_p+R^{n+1}_p\}\!\}\|^2_{0,\mathcal{E}_h}\right.\\ & \left.+\,2|\vec{\mathtt{C}}_{12}|^2\|\mathtt{C}_{22}^{-\frac{1}{2}}[\![R^n_{\mathbf{q}}+R^{n+1}_{\mathbf{q}}]\!]\|^2_{0,\mathcal{E}_h}+2\|\mathtt{C}_{22}^{-\frac{1}{2}}\{\!\{R^n_{\mathbf{q}}+R^{n+1}_{\mathbf{q}}\}\!\}\|^2_{0,\mathcal{E}_h}\right).\end{aligned}$$

With the following note that

$$\begin{aligned}\sum_{n=0}^{m}\|\frac{\sigma_p}{c}(R^n_p+R^{n+1}_p)\|_{0,\Omega}\|\frac{1}{c}(\Phi^n_p+\Phi^{n+1}_p)\|_{0,\Omega} \;&\leq\; 2\max_{1\leq n\leq m+1}\|\Phi^n_p\|_{0,\Omega}\sum_{n=0}^{m}\|\frac{\sigma_p}{c}(R^n_p+R^{n+1}_p)\|_{0,\Omega}\\ &\leq\; 4\max_{1\leq n\leq m+1}\|\Phi^n_p\|_{0,\Omega}\sum_{n=0}^{m+1}\|\frac{\sigma_p}{c}R^n_p\|_{0,\Omega},\end{aligned}$$

and the previous estimations we can have that

$$\begin{aligned}&\|\frac{1}{c}\Phi^{m+1}_p\|^2_{0,\Omega}\;+\;\|\Phi^{m+1}_{\mathbf{q}}\|^2_{0,\Omega}\;+\;\frac{\triangle t}{2}\sum_{n=0}^{m}\left(\|\mathtt{C}_{11}^{\frac{1}{2}}[\![\Phi^n_p\;+\;\Phi^{n+1}_p]\!]\|^2_{0,\mathcal{E}_h}+\|\mathtt{C}_{22}^{\frac{1}{2}}[\![\Phi^m_{\mathbf{q}}+\Phi^{n+1}_{\mathbf{q}}]\!]\|^2_{0,\mathcal{E}_h}\right)\\ &\leq\;\|\frac{1}{c}\Phi^0_p\|^2_{0,\Omega}+\|\Phi^0_{\mathbf{q}}\|^2_{0,\Omega}+\frac{\varepsilon\triangle t}{4}\sum_{n=0}^{m}\left(\|\mathtt{C}_{11}^{\frac{1}{2}}[\![\Phi^n_p+\Phi^{n+1}_p]\!]\|^2_{0,\mathcal{E}_h}+\|\mathtt{C}_{22}^{\frac{1}{2}}[\![\Phi^n_{\mathbf{q}}+\Phi^{n+1}_{\mathbf{q}}]\!]\|^2_{0,\mathcal{E}_h}\right)\\ &\quad+2\left(\max_{0\leq n\leq m+1}\|\frac{1}{c}\;\Phi^n_p\|_{0,\Omega}\right)\triangle t\sum_{n=0}^{m+1}\left(\|\frac{1}{c}\Delta^n_p\|_{0,\Omega}+\|\frac{\sigma_p}{c}R^n_p\|_{0,\Omega}\right)\\ &\quad+2\left(\max_{0\leq n\leq m+1}\|\Phi^n_{\mathbf{q}}\|_{0,\Omega}\right)\triangle t\sum_{n=0}^{m+1}\left(\|\Delta^n_{\mathbf{q}}\|_{0,\Omega}+\|\sigma_q R^n_{\mathbf{q}}\|_{0,\Omega}\right)+\frac{\triangle t}{2\varepsilon}\sum_{n=0}^{m}T^n_B,\end{aligned}$$

where T^n_B is defined by

$$\begin{aligned}T^n_B:=\Big(&2|\vec{\mathtt{C}}_{12}|^2\|\mathtt{C}_{11}^{-\frac{1}{2}}[\![R^n_p+R^{n+1}_p]\!]\|^2_{0,\mathcal{E}_h}+\|\mathtt{C}_{11}^{\frac{1}{2}}[\![R^n_p+R^{n+1}_p]\!]\|^2_{0,\mathcal{E}_h}+2\|\mathtt{C}_{11}^{-\frac{1}{2}}\{\!\{R^n_p+R^{n+1}_p\}\!\}\|^2_{0,\mathcal{E}_h}\\ &+2|\vec{\mathtt{C}}_{12}|^2\|\mathtt{C}_{22}^{-\frac{1}{2}}[\![R^n_{\mathbf{q}}+R^{n+1}_{\mathbf{q}}]\!]\|^2_{0,\mathcal{E}_h}+\|\mathtt{C}_{22}^{\frac{1}{2}}[\![R^n_{\mathbf{q}}+R^{n+1}_{\mathbf{q}}]\!]\|^2_{0,\mathcal{E}_h}+2\|\mathtt{C}_{22}^{-\frac{1}{2}}\{\!\{R^n_{\mathbf{q}}+R^{n+1}_{\mathbf{q}}\}\!\}\|^2_{0,\mathcal{E}_h}\Big).\end{aligned}$$

for $0 \leq m \leq N-1$.

We subtract the term $\frac{\varepsilon \triangle t}{2} \sum_{n=0}^{m} \left(\|\mathtt{C}_{11}^{\frac{1}{2}} [\![\Phi_p^n + \Phi_p^{n+1}]\!] \|_{0,\mathcal{E}_h}^2 + \|\mathtt{C}_{22}^{\frac{1}{2}} [\![\Phi_{\mathbf{q}}^n + \Phi_{\mathbf{q}}^{n+1}]\!] \|_{0,\mathcal{E}_h}^2 \right)$ from both sides taking $\varepsilon = 1$, and by using the geometric-arithmetic inequality $ab \leq \frac{\mu}{2} a^2 + \frac{b^2}{2\mu}, \mu = 2$ we obtain that

$$\begin{aligned} \|\frac{1}{c}\Phi_p^{m+1}\|_{0,\Omega}^2 + \|\Phi_{\mathbf{q}}^{m+1}\|_{0,\Omega}^2 \leq \|\frac{1}{c}\Phi_p^0\|_{0,\Omega}^2 + \|\Phi_{\mathbf{q}}^0\|_{0,\Omega}^2 \\ + \frac{1}{4}\left(\max_{0\leq n\leq N} \|\frac{1}{c}\ \Phi_p^n\|_{0,\Omega} \right)^2 + \left(\triangle t \sum_{n=0}^{N} \left(\|\frac{1}{c}\Delta_p^n\|_{0,\Omega} + \|\frac{\sigma_p}{c} R_p^n\|_{0,\Omega} \right) \right)^2 \\ + \frac{1}{4}\left(\max_{0\leq n\leq N} \|\Phi_{\mathbf{q}}^n\|_{0,\Omega} \right)^2 + \left(\triangle t \sum_{n=0}^{N} \left(\|\Delta_{\mathbf{q}}^n\|_{0,\Omega} + \|\sigma_q R_{\mathbf{q}}^n\|_{0,\Omega} \right) \right)^2 + \frac{\triangle t}{2} \sum_{n=0}^{N} T_B^n. \end{aligned}$$

Since the right-hand side is independent of m, we take maximum on the left hand side and subtract $\frac{1}{4}\left(\max_{m=0}^{N-1} \|\frac{1}{c}\Phi_p^m\|_{0,\Omega} \right)^2$ and $\frac{1}{4}\left(\max_{m=0}^{N-1} \|\Phi_{\mathbf{q}}^m\|_{0,\Omega} \right)^2$, and multiply by 2 on both sides we readily obtain that

$$\begin{aligned} \max_{0\leq n\leq N} \left(\|\frac{1}{c}\Phi_p^n\|_{0,\Omega}^2 + \|\Phi_{\mathbf{q}}^n\|_{0,\Omega}^2 \right) \leq 2\|\frac{1}{c}\Phi_p^0\|_{0,\Omega}^2 + 2\|\Phi_{\mathbf{q}}^0\|_{0,\Omega}^2 \\ +2\left(\triangle t \sum_{n=0}^{N} \left(\|\Delta_{\mathbf{q}}^n\|_{0,\Omega} + \|\sigma_q R_{\mathbf{q}}^n\|_{0,\Omega} \right) \right)^2 + 2\left(\triangle t \sum_{n=0}^{N} \left(\|\frac{1}{c}\Delta_p^n\|_{0,\Omega} + \|\frac{\sigma_p}{c} R_p^n\|_{0,\Omega} \right) \right)^2 + \triangle t \sum_{n=0}^{N} T_B^n. \end{aligned}$$

Taking square roots on both sides we obtain that

$$\begin{aligned} \max_{0\leq n\leq N} \left(\|\Phi_p^n\|_{0,\Omega} + \|\Phi_{\mathbf{q}}^n\|_{0,\Omega} \right) \leq \sqrt{2}\|\frac{1}{c}\Phi_p^0\|_{0,\Omega} + \sqrt{2}\|\Phi_{\mathbf{q}}^0\|_{0,\Omega} \\ +\sqrt{2}\triangle t \sum_{n=0}^{N} \left(\|\Delta_{\mathbf{q}}^n\|_{0,\Omega} + \|\sigma_q R_{\mathbf{q}}^n\|_{0,\Omega} \right) + \sqrt{2}\triangle t \sum_{n=0}^{N} \left(\|\frac{1}{c}\Delta_p^n\|_{0,\Omega} + \|\frac{\sigma_p}{c} R_p^n\|_{0,\Omega} \right) + \left(\triangle t \sum_{n=0}^{N} T_B^n \right)^{\frac{1}{2}}. \end{aligned}$$

We can bound the right-hand side using (6.7) and (6.8) that

$$\begin{aligned} \triangle t \sum_{n=1}^{N} (\|\frac{\sigma_p}{c} R_p^n\|_{0,\Omega} + \|\sigma_q R_{\mathbf{q}}^n\|_{0,\Omega}) \leq T \left(\max_{1\leq n\leq N} \|\frac{\sigma_p}{c} R_p^n\|_{0,\Omega} + \max_{0}^{N} \|\sigma_q R_{\mathbf{q}}^n\|_{0,\Omega} \right) \\ \leq C h^{\min\{s,k\}+1} T(\|p\|_{C(\overline{I};H^{1+s}(\Omega))} + \|\vec{\mathbf{q}}\|_{C(\overline{I};\mathbb{H}^{1+s}(\Omega))}), \quad (6.12) \end{aligned}$$

and $\Delta_p^n, \Delta_{\mathbf{q}}^n$ can be bounded as well using Lemma 6.2 ,

$$\triangle t \sum_{n=0}^{N} \|\Delta_p^n\|_{0,\Omega} \leq C \left(h^{\min\{s,k+1\}} \|p_t\|_{L^1(I;H^s(\Omega))} + \triangle t^2 \|\partial_t^3 p\|_{L^1(I;L^2(\Omega))} \right), \tag{6.13}$$

$$\triangle t \sum_{n=0}^{N} \|\Delta_{\mathbf{q}}^n\|_{0,\Omega} \leq C \left(h^{\min\{s,k+1\}} \|\vec{\mathbf{q}}_t\|_{L^1(I;H^s(\Omega))} + \triangle t^2 \|\partial_t^3 \vec{\mathbf{q}}\|_{L^1(I;L^2(\Omega))} \right). \tag{6.14}$$

By Lemma 6.1 we estimate the term T_B^n,

$$T_B^n \leq C h^{2\min\{s,k\}+1} \left(\|p^n + p^{n+1}\|_{1+s,\Omega}^2 + \|\vec{\mathbf{q}}^n + \vec{\mathbf{q}}^{n+1}\|_{1+s,\Omega}^2 \right),$$

and

$$\triangle t \sum_{n=0}^{N} \|T_B^n\|_{0,\Omega} \leq T \max_{0\leq n\leq N} \|T_B^n\|_{0,\Omega} \leq C h^{2\min\{s,k\}+1} T \left(\|p\|_{C(\bar{I};H^{1+s}(\Omega))}^2 + \|\vec{\mathbf{q}}\|_{C(\bar{I};\mathbb{H}^{1+s}(\Omega))}^2 \right),$$

so that

$$\left(\triangle t \sum_{n=0}^{N} T_B^n \right)^{\frac{1}{2}} \leq C h^{\min\{s,k\}+\frac{1}{2}} T^{\frac{1}{2}} \left(\|p\|_{C(\bar{I};H^{1+s}(\Omega))} + \|\vec{\mathbf{q}}\|_{C(\bar{I};\mathbb{H}^{1+s}(\Omega))} \right).$$

We combine (6.12), (6.13), (6.14), and (6.1.3) to get

$$\begin{aligned} &\max_{0\leq n\leq N} \left(\|\Phi_p^n\|_{0,\Omega}^2 + \|\Phi_{\mathbf{q}}^n\|_{0,\Omega}^2 \right) \\ \leq\ & 2\|\frac{1}{c}\Phi_p^0\|_{0,\Omega}^2 + 2\|\Phi_{\mathbf{q}}^0\|_{0,\Omega}^2 \ +\ C\triangle t^2 \left(\|\partial_t^3 p\|_{L^1(I;L^2(\Omega))} + \|\partial_t^3 \vec{\mathbf{q}}\|_{L^1(I;L^2(\Omega))} \right) \\ +C h^{\min\{s,k+\frac{1}{2}\}} & \left(\|p\|_{C(\bar{I};H^{1+s}(\Omega))} + \|\vec{\mathbf{q}}\|_{C(\bar{I};\mathbb{H}^{1+s}(\Omega))} + \|p_t\|_{L^1(I;H^s(\Omega))} + \|\vec{\mathbf{q}}_t\|_{L^1(I;\mathbb{H}^s(\Omega))} \right), \end{aligned}$$

in which the constant C grows linearly with T, which completes the proof.

Remark 6.3 *We also take staggered scheme with centered difference in time of the systems* (5.22)-(5.23) *for all* $n = 1, \cdots, N$,

$$\begin{cases} (\frac{1}{c^2}\Delta^n P^{n-\frac{1}{2}}, r^h) + \frac{1}{2}a_h(P^n + P^{n-1}, r^h) - b_h'(\vec{\mathbf{Q}}^{n-\frac{1}{2}}, r^h) &= 0, \quad \forall r^h \in \mathcal{P}^h(\Omega) \\ (\boldsymbol{\Delta}^n \vec{\mathbf{Q}}^n, \vec{\mathbf{v}}^h) + b_h(P^n, \vec{\mathbf{v}}^h) \ +\ \frac{1}{2}c_h(\vec{\mathbf{Q}}^{n+\frac{1}{2}} + \vec{\mathbf{Q}}^{n-\frac{1}{2}}, \vec{\mathbf{v}}^h) &= 0, \quad \forall \vec{\mathbf{v}}^h \in \mathcal{Q}^h(\Omega), \end{cases} \tag{6.15}$$

where

$$\Delta^n P^{n-\frac{1}{2}} = \frac{P^n - P^{n-1}}{\triangle t}, \quad \boldsymbol{\Delta}^n \vec{\mathbf{Q}}^n = \frac{\vec{\mathbf{Q}}^{n-\frac{1}{2}} - \vec{\mathbf{Q}}^{n+\frac{1}{2}}}{\triangle t}.$$

Choosing $r^h = P^n + P^{n-1}, \vec{\mathbf{v}}^h = \vec{\mathbf{Q}}^{n-\frac{1}{2}} + \vec{\mathbf{Q}}^{n+\frac{1}{2}}$ *with the property in (5.43), we can have that*

$$\begin{aligned}\|\frac{1}{c}P^n\|_{0,\Omega}^2 + \|\vec{\mathbf{Q}}^{n+\frac{1}{2}}\|_{0,\Omega}^2 =& \|\frac{1}{c}P^{n-1}\|_{0,\Omega}^2 + \|\vec{\mathbf{Q}}^{n-\frac{1}{2}}\|_{0,\Omega}^2 + \frac{\triangle t}{2}\left(-b_h'(\vec{\mathbf{Q}}^{n-\frac{1}{2}}, P^{n-1}) + b_h(P^n, \vec{\mathbf{Q}}^{n+\frac{1}{2}})\right) \\ & -\frac{\triangle t}{2}\left(a_h(P^n + P^{n-1}, P^n + P^{n-1}) + c_h(\vec{\mathbf{Q}}^{n-\frac{1}{2}} + \vec{\mathbf{Q}}^{n+\frac{1}{2}}, \vec{\mathbf{Q}}^{n-\frac{1}{2}} + \vec{\mathbf{Q}}^{n+\frac{1}{2}})\right) \\ \leq & \|\frac{1}{c}P^{n-1}\|_{0,\Omega}^2 + \|\vec{\mathbf{Q}}^{n-\frac{1}{2}}\|_{0,\Omega}^2 + \frac{\triangle t}{2}\left(-b_h'(\vec{\mathbf{Q}}^{n-\frac{1}{2}}, P^{n-1}) + b_h(P^n, \vec{\mathbf{Q}}^{n+\frac{1}{2}})\right),\end{aligned}$$

by the property of $b_h'(\vec{\mathbf{Q}}^{n-\frac{1}{2}}, P^n) = b_h(P^n, \vec{\mathbf{Q}}^{n-\frac{1}{2}})$.

Summation of the equations above from $n = 1$ *to* $n = N$ *gives that*

$$\|\frac{1}{c}P^{N+1}\|_{0,\Omega}^2 + \|\vec{\mathbf{Q}}^{N+\frac{1}{2}}\|_{0,\Omega}^2 \leq \|\frac{1}{c}P^0\|_{0,\Omega}^2 + \|\vec{\mathbf{Q}}^{\frac{1}{2}}\|_{0,\Omega}^2 + \frac{\triangle t}{2}\left(-b_h'(\vec{\mathbf{Q}}^{\frac{1}{2}}, P^0) + b_h(P^N, \vec{\mathbf{Q}}^{N+\frac{1}{2}})\right).$$

In the case of CG(Continuous Galerkin) method, we have

$$|b_h(P^N, \vec{\mathbf{Q}}^{N+\frac{1}{2}})| \leq \frac{1}{2}\|\nabla P^N\|_{0,\Omega}^2 + \frac{1}{2}\|\vec{\mathbf{Q}}^{N+\frac{1}{2}}\|_{0,\Omega}^2.$$

The inverse inequality (Appendix 0.1) *allows to have constant* $C_{inv}(h) > 0$ *satisfying* $\|\nabla P^N\|_{0,\Omega}^2 \leq C_{inv}(h)\|P^N\|_{0,\Omega}^2$, *where* $C_{inv}(h)$ *depends on shape-regularity. Therefore we can obtain CFL condition choosing* $\triangle t \leq 4{c^*}^2 C_{inv}(h)$ *for the scheme* (6.15).

7. DISCUSSION AND CONCLUSIONS

PMLs for the acoustic wave equation with variable sound speed appear in several different forms. But stability and well-posedness are not clearly answered mathematically in higher dimensions. Even though the energy decay rate of 1-d acoustic PML wave was solved for constant speed [46], it remains unknown for the variable sound speed case. We showed the exponential energy decay of 1-d PML wave with variable sound speed in Chapter 4, but the energy decay rate is still unknown.

The second order regularized 2-d PML wave equation was introduced in Chapter 2, and we showed the well-posedness of the system and efficiency by numerical experiments. But the stability of both classical and regularized system is still not clear, and so remains a question for further research.

In chapter 3 the multi directional PMLs are introduced with additional damping terms. We showed well-posedness of the regularized system, and numerical experiments indicate that the multi-directional PMLs are more effective than the classical PML. There are many remaining questions such as well-posedness, stability, general efficiency, etc.

We introduce a general stable formulation of a first order hyperbolic system with lower order damping in chapter 5. Construction of effective damping terms, possibly time dependent, in order to have numerically desirable absorption in the layers is still largely open. In chapter 6 we constructed a locally discontinuous Galerkin method (LDG) for the system. An *a priori* L^2-error estimate under additional regularity assumptions was shown for the semi-discrete DG method as well as for the fully discretized scheme in Chapter 7.

BIBLIOGRAPHY

1. A. Ern, J.-L. Guermond *Theory and practice of finite elements* Springer Verlag, New York, (2004).

2. A. F. Oskooi, L. Zhang, Y. Avniel, and S. G. Johnson *The failure of perfectly matched layers, and towards their redemption by adiabatic absorbers* Optics Express, Vol. 16, 11376-11392, July, (2008).

3. A. S. Omar, K. F. Schunemann *Complex and backward-wave modes in inhomogeneously and anisotropically filled waveguides* IEEE Trans. Microwave Theory Tech., Vol. MTT-35, no. 3, 268-275, (1987).

4. B. Cockburn *Discontinuous Galerkin Methods* ZAMM·Z. Angew. Math. Mech. 83, No.11, 731-754, (2003).

5. B. E. Petersen *Introduction to the Fourier Transform and Pseudo-Differential Operators* Pitman Advanced Pub., (1983).

6. B. Cockburn, G. Kanschat, I. Perugia, D. Schötzau *Superconvergence of the Local Discontinuous Galerkin Method for Elliptic Problems on Cartesian Grids*, SIAM J. Numer. Anal. Vol. 39, No. 1, 264-285, (2001).

7. B. Kaltenbacher, M. Kaltenbacher, I. Sim *A modified and stable version of a perfectly matched layer technique for the 3-d second order wave equation in time domain with an application to aeroacoutics* J. Comput. Phys. 235 407-422, (2013).

8. B. Bidégaray-Fesquet *Stability of FD-TD schemes for Maxwell-Debye and Maxwell-Lorentz equations* SIAM Journal on Numerical Analysis 46.5 (2008): 2551-2566, (2008).

9. D. Boffi, F. Brezzi, M. Fortin *Mixed Finite Element Methods and Applications*, Springer Series in Computational Mathematics, Vol. 44, (2013).

10. D. N. Arnold, F. Brezzi, B. Cockburn, L. D. Marini *Unified Analysis of Discontinuous Galerkin Methods for Elliptic Problems*, SIAM J. Numer. Anal. Vol. 39, No. 5, 1749-1779, (2002).

11. D. Appelö, T. Hagstrom, G. Kress *Perfectly matched layer for hyperbolic systems: general formulation, well-posedness and stability*, J. Appl. Math., 67:1-23, (2006).

12. E. Magenes, G. Stampacchia *I problemi al contorno per le equazioni differenziali di tipo ellitico*, Annali della Scuola Normale Superiore di Pisa 12, 247-358, (1958).

13. E. Becache, S. Fauquex, P. Joly *Stability of perfectly matched layers, group velocities and anisotropic waves* J. Comput. Phys. 188, 399-433, (2003).

14. E. Becache, P. Joly *On the analysis of Berebers perfectly matched layers for Maxwell equations* M2AN 36 (1), 87-120, (2002).

15. E. Bécache, A. Prieto *Remarks on the stability of Cartesian PMLs in corners* [Research Report] RR-7620, pp.18.<inria-00593182>, (2011).

16. E. Turkei, A. Yefet *Absorbing PML boundary layers for wave-like equations* App. Num. Math. Vol 27 (4), 533-557, (1998).

17. F. Q. Hu *On Absorbing Boundary Conditions for Linearized Euler Equations by a Perfectly Matched Layer* J. Comput. Phys. 129, 201-219, (1996).

18. H. Assi, R. S. C. Coobold *Perfectly matched layer for second-order time-domain elastic wave equation: formulation and stability* arXiv:1312.3722v1[physics.comp-ph], (2013).

19. I. Perugia, D. Schötzau *An hp-analysis of the local discontinuous Galerkin method for diffusion problems*, J. Sci. Comput., 17, 561-571, (2002).

20. J.-L. Lions, J. Métral, O. Vacus *Well-posed absorbing layer for hyperbolic problems* Numer. Math. 92:535-562, (2002).

21. J. P. Berenger *A perfectly matched layer for the absorption of electromagnetic waves* J. Comput. Phys. 114: 185-200, (1994).

22. J. C. Strikwerda *Finite Difference Schemes and Partial Differential Equations, Second Edition*, SIAM, Other Titles in Applied Mathematics, (2004).

23. J. J. H. Miller *On the Location of Zeros of Certain Classes of Polynomials with Applications to Numerical Analysis*, J. Inst. Maths Applics, Vol. 8, 397-406, (1971).

24. J. Rauch, M. Taylor *Exponential Decay of Solutions to Hyperbolic Equations in Bounded Domain* Indiana Unv. Math. J. Vol. 24 (1), No.1 79-86, (1974).

25. L. C. Evans *Partial Differential Equations: Second Edition* Graduate Series in Mathematics, Vol. 19. R (2010).

26. L. Halpern, S. Petit-Bergez, J. Rauch *The analysis of Matched Layers* Confluentes Mathematici, Vol. 3, No. 2, 159236, (2011).

27. L. Zhao, A.C. Cangellaris *A general approach for the development of unsplit-field time-domain implementations of perfectly matched layers for FDTD grid truncation* IEEE Microwave and Guided Lett. Vol. 6 (5), 209-211, (1996).

28. M. G. Larson, A. Målqvist, *A posteriori error estimates for mixed finite element approximations of elliptic equations*, Numer. Math., 108 , 487-500, (2008).

29. M. J. Grote, A. Schneebeli, D. Schötzau *Discotinuous Galerkin Finite Element Method for the Wave Equation*, SIAM J. Numer. Anal. Vol. 44, No. 6, 2408-2431, (2006).

30. M. J. Grote, I. Sim *Efficient PML for the wave equation* arXiv:1001.0319v1 [math.NA] 2 Jan (2010).

31. M. J. Grote, D. Schötzau *Optimal Error Estimates for the Fully Discrete Interior Penalty DG Mehtod for the Wave Equation*, J. Sci. Comput. 40, 257-272, (2009).

32. J. Necas *Les méthodes directes en théorie des équations elliptiques*, Academia, Prague, (1976).

33. P. Stefanov, G. Uhlmann *Thermoacoustic Tomagraphy with Variable Sound Speed* Inverse Problems, Vol. 25, 075011, Number 7, (2009).

34. P. G. Ciarlet *The Finite Element Methods for Elliptic Problems*, North-Holland, Amsterdam, (1978).

35. P. J. B. Clarricoats, R. A. Waldron *Non-periodic slow-wave and backward-wave structures* J. Electron. Contr., Vol. 8, 455-458, (1960).

36. P. Loh, A. F. Oskooi, M. Ibanescu, M. Skorobogatiy, S. G. Johnson *Fundamental relation between phase and group velocity, and application to the failure of perfectly matched layers in backward-wave structures* Phys. Rev. E 79 065601(R), (2009).

37. P. G. Petropoulos, L. Zhao, A. C. Cangellaris *A reflecless sponge layer absorbing boundary condition for the solution of Maxwell's equations with high order staggered finite difference schemes* J. Comput. Phys. 139 (1) 184-208, (1998).

38. Q. Liu, J. Tao *The perfectly matched layer for acoustic waves in absorptive media* J. Acoust. Soc. Am. 102 (4) October, 2072-208, (1997).

39. R.E. Showalter *Hilbert Space Methods for Partial Differential Equations* Dover Publications, Inc. Mineola, New York (1979).

40. R. J. LeVeque *Finite Difference Methods for Ordinary and Partial Differential Equations* SIAM. (2007).

41. R. A. Waldron *Theory and potential applications of backward waves in non-periodic inhomogeneous waveguides* Proc. IEE, Vol. 111, 1659 1667, (1964).

42. R. W. Carroll *Abstract Methods in Partial Differential Equations* Harper and Row, Publishers, New York, Evanston, and London (1969).

43. S. Prudhomme, F. Pascal, J. T. Oden, A. Romkes *Review of* A Priori *Error Estimation for Discontinuous* Galerkin *Methods* Tech. report 00-27, TICAM, Austin, TX, (2000).

44. S. D. Gedney *An anisotropic perfectly matched layer absorbing media for the truncation of FDTD latices* Antennas and Propagation, IEEE Transactions on 44 (12): 16301639, (1996).

45. S. Abarbanel, D. Gottlieb, J. S. Hesthaven *Well-posed perfectly matched layers for advective acoustics* J. Comput. Phys. 154, 266283, (1999).

46. S. Ervedoza, E. Zuazua *Perfectly matched layers 1-d : energy decay for continuous and semi-discrete waves* Numer. Math. 109:597-634, (2008).

47. S. Johnson *Notes on Perfectly Matched Layers* Technical report, Massachusetts Institute of Technology, Cambridge, MA, (2010).

48. S. A. Cummer *Perfectly Matched Layer Behavior in Negative Refractive Index Materials* IEEE Ant. Wireless Prop. Lett. 3, 172-175, (2004).

49. S. V. Tsynkov, E. Turkel *A Cartesian Perfectly Matched Layer for the Helmholtz Equation, in: Absorbing Boundaries and Layers, Domain Decomposition Methods* Applications to Large Scale Computations, Loïc Tourrette and Lorance Halpern, eds., Nova Science Publishers, Inc., New York, 279-309, (2001).

50. W. C. Chew, W. H. Weedon *A 3D Perfectly Matched Medium from Modified Maxwell's Equations with Stretched Coordinates* Microwave and Optical Technology Letters, Vol. 7 (13), 599604, September (1994).

51. W. C. Chew, Q. H. Liu *Perfectly Matched Layers for Elastodynamics: A New Absorbing Boundary Condition* J. Comp. Asoust., Vol. 4, 341-359 (1996).

APPENDICES

A APPENDIX Inverse Inequality

One obtains as penalty a factor with negative powers of the diameter of the mesh size to estimate a norm of a higher order derivative of a finite element function by a norm of a lower order. These are so-called inverse estimates [1].

Consider an affine family of finite elements $\{K\}_{K\in\mathcal{T}_h}$ whose mesh cells are generated by affine mappings $F_K : \hat{K} \to K$ by

$$F_K\hat{\mathbf{x}} = B\hat{\mathbf{x}} + \mathbf{b},$$

where $\hat{K}$ is a reference cell and B is a non-singular $d \times d$ matrix and $\mathbf{b}$ is a d vector.

Lemma 0.1 *For each matrix norm $\|\cdot\|$ we have the estimates*

$$\|B\| \le ch_K, \quad \|B^{-1}\| \le ch_K^{-1},$$

where the constants depend on the matrix norm and on K.

Proof. Since $\hat{K}$ is a Lipschitz domain, it contains a ball $B(\hat{\mathbf{x}}_0, r)$ with $\hat{\mathbf{x}}_0 \in \hat{K}$ and some $r > 0$. Then $\hat{\mathbf{x}}_0 + \hat{\mathbf{y}} \in \hat{K}$ for all $\|\hat{\mathbf{y}}\|_2 = r$. It follows that

$$\mathbf{x}_0 = B\hat{\mathbf{x}}_0 + \mathbf{b} \in K, \quad \mathbf{x} = B(\hat{\mathbf{x}}_0 + \hat{\mathbf{y}}) + \mathbf{b} = \mathbf{x}_0 + B\hat{\mathbf{y}} \in K.$$

Then we obtain that for all $\hat{\mathbf{y}}$

$$\|B\hat{\mathbf{y}}\|_2 = \|\mathbf{x} - \mathbf{x}_0\|_2 \le C_R h_K.$$

Now, it holds for the spectral norm that

$$\|B\|_2 = \sup_{\hat{\mathbf{z}}\neq 0} \frac{\|B\hat{\mathbf{z}}\|_2}{\|\hat{\mathbf{z}}\|_2} = \frac{1}{r} \sup_{\|\hat{\mathbf{z}}\|_2 = r} \|B\hat{\mathbf{z}}\|_2 \le \frac{C_R}{r} h_K,$$

where C_R depends on K, but it can be independent on K with the assumption of quasi-uniform mesh. An estimate of this form, with a possible different constant, holds also for

all other matrix norms since all matrix norms are equivalent. The estimate for $\|B^{-1}\|$ proceeds in the same way with interchanging the roles of K and $\hat{K}$. □

Remark 0.1 *We can get the estimate for the determinants of B and B^{-1} from the previous Lemma 0.1 and Leibniz formula for determinants*

$$|\det B| \leq Ch_K^d, \quad |\det B^{-1}| \leq Ch_K^{-d}.$$

Using this bounds we can get the following for all $\hat{v} \in P(\hat{K}), v \in P(K)$

$$\int_K \|D_{\mathbf{x}}^k v(\mathbf{x})\|_2^p d\mathbf{x} \leq Ch_K^{-kp} |\det B| \int_{\hat{K}} \|D_{\hat{\mathbf{x}}}^k \hat{v}(\hat{\mathbf{x}})\|_2^p d\hat{\mathbf{x}} \leq Ch_K^{-kp+d} \int_{\hat{K}} \|D_{\hat{\mathbf{x}}}^k \hat{v}(\hat{\mathbf{x}})\|_2^p d\hat{\mathbf{x}},$$

and

$$\int_{\hat{K}} \|D_{\hat{\mathbf{x}}}^k \hat{v}(\hat{\mathbf{x}})\|_2^p d\hat{\mathbf{x}} \leq Ch_K^{kp} |\det B^{-1}| \int_K \|D_{\mathbf{x}}^k v(\mathbf{x})\|_2^p d\mathbf{x} \leq Ch_K^{kp-d} \int_K \|D_{\mathbf{x}}^k v(\mathbf{x})\|_2^p d\mathbf{x},$$

where $P(\hat{K})$ the space of polynomials of order N over $\hat{K}$, and $P(K) = \{p \in K \to \mathbb{R} : p = \hat{p} \circ F_K^{-1}, \hat{p} \in P(\hat{K})\}$.

Theorem 0.1 *(Inverse Estimate). Let $0 \leq k \leq l$ be natural numbers and let $p, q \in [1, \infty]$. Then there is a constant C_{inv}, which depends only on $k, l, p, q, \hat{K}, P(\hat{K})$ such that*

$$\|D^l v^h\|_{L^q(K)} \leq C_{inv} h_K^{(k-l)-d(p^{-1}-q^{-1})} \|D^k v^h\|_{L^p(K)} \quad \forall v^h \in P(K).$$

Proof. Assume $h_{\hat{K}} = 1$ on the reference mesh cell. For $k = 0$, we obtain that

$$\|D^l v^h\|_{L^q(\hat{K})} \leq \|\hat{v}^h\|_{W^{l,q}(\hat{K})} \leq C\|\hat{v}^h\|_{L^p(\hat{K})} \quad \forall v^h \in P(\hat{K}),$$

since all norms are equivalent in finite dimensional space. For $k > 0$, consider the space of polynomials such as

$$\tilde{P}(\hat{K}) = \{\partial_\alpha \hat{v}^h : \hat{v}^h \in P(\hat{K}), |\alpha| = k\}.$$

Then we apply $\tilde{P}(\hat{K})$ to obtain that

$$\begin{aligned}\|D^l \hat{v}^h\|_{L^q(\hat{K})} \leq & \sum_{|\alpha|=k} \|D^{l-k}(\partial_\alpha \hat{v}^h)\|_{L^q(\hat{K})} \leq C \sum_{|\alpha|=k} \|\partial_\alpha \hat{v}^h\|_{L^q(\hat{K})} \\ = & C\|D^k \hat{v}^h\|_{L^q(\hat{K})}.\end{aligned}$$

From the estimates for the transformations, we obtain that

$$\begin{aligned}\|D^l v^h\|_{L^q(K)} \leq & C h_K^{-l+d/q} \|D^l(\partial_\alpha v^h)\|_{L^q(\hat{K})} \leq C h_K^{-l+d/q} \|D^k v^h\|_{L^p(\hat{K})} \\ = & C_{inv} h_K^{k-l+d/q-d/p} \|D^k v^h\|_{L^p(K)}.\end{aligned}$$

□

For example, the inequality (5.18) holds when $p = q = d = 2, k = 0, l = 1$.

Remark 0.2 *One obtains the global inverse inequality with the assumption of quasi-uniform mesh (Definition 5.3),*

$$\|D^l v^h\|_{L^q(\mathcal{T}_h)} \leq C h^{(k-l)-d(p^{-1}-q^{-1})} \|D^k v^h\|_{L^p(\mathcal{T}_h)} \quad \forall v^h \in P(\Omega).$$

B APPENDIX Figures

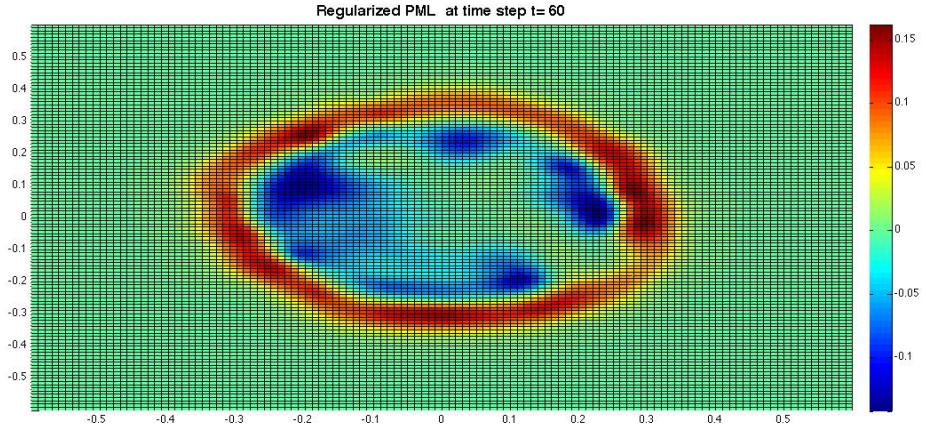

Figure 0.1: Regularized Acoustic PML wave with variable sound speed at time steps 60

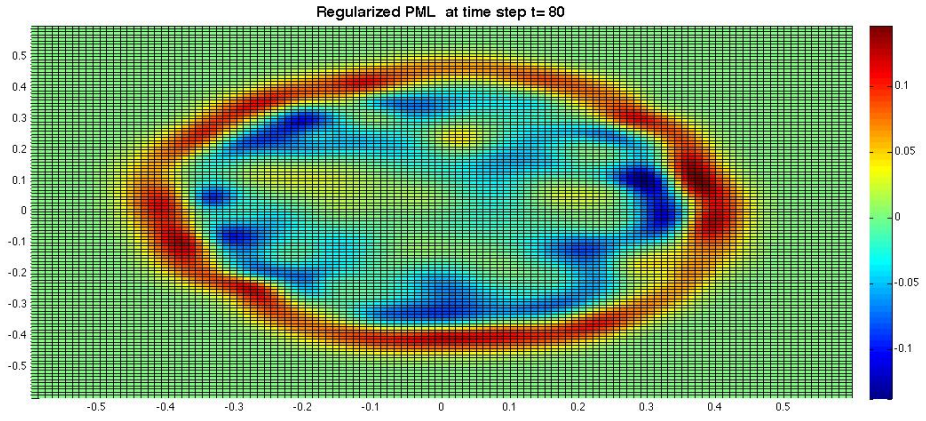

Figure 0.2: Regularized Acoustic PML wave with variable sound speed at time steps 80

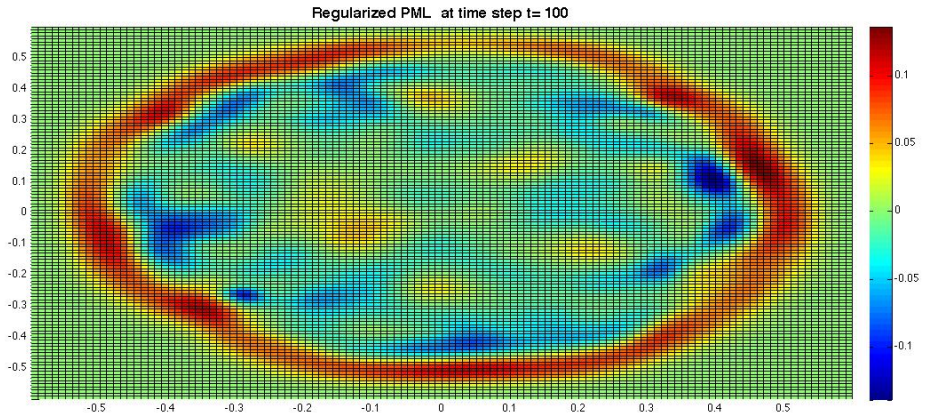

Figure 0.3: Regularized Acoustic PML wave with variable sound speed at time steps 100

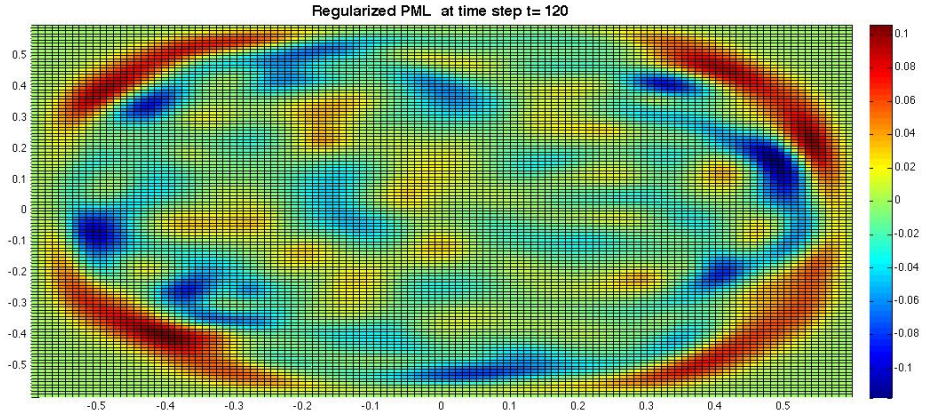

Figure 0.4: Regularized Acoustic PML wave with variable sound speed at time steps 120

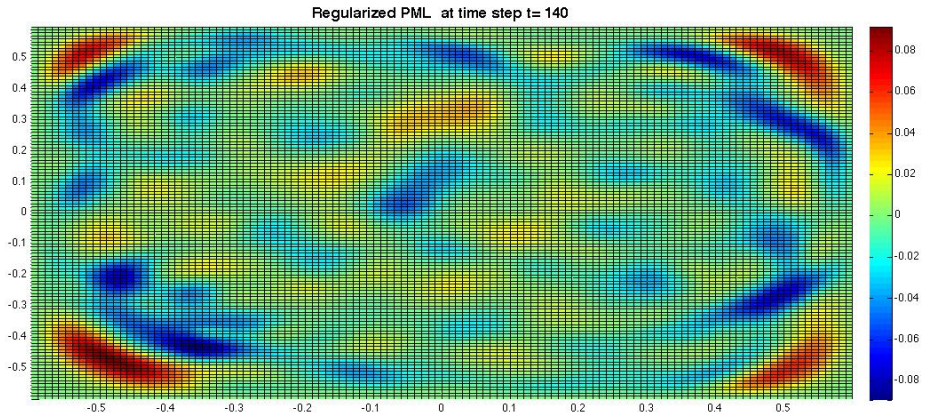

Figure 0.5: Regularized Acoustic PML wave with variable sound speed at time steps 140

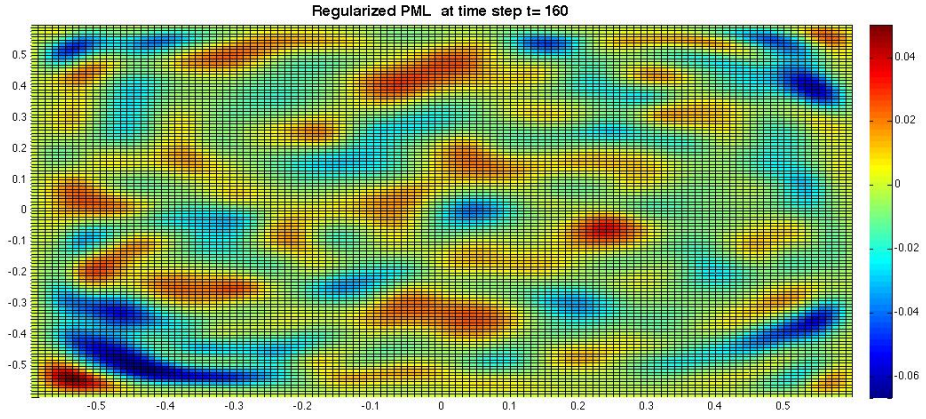

Figure 0.6: Regularized Acoustic PML wave with variable sound speed at time steps 160

C APPENDIX Codes

```
function System 2nd order PML regularization

h = 0.01;

L = 0.1; a = .5;

dx = h; dy = dx;

dt = dx/2;

tf = 180;

sigma = @getsigma;

x = −L − a : h : a + L;

npic = length(a : h : a + L);

ncom = length(−a : h : a);

y = x;

nx = length(x); ny = length(y);

f = zeros(nx);

ka = −.0;

kb = .0;

for xi = 1 : nx

    for yj = 1 : nx

      f(xi, yj) = 1 ∗ exp(−(20. ∗ (x(xi) − ka))^2 − (20. ∗ (y(yj) − kb))^2);

    end

end u = f; uold = u;

unew = zeros(nx);

qx = zeros(nx − 1);

qy = zeros(nx − 1);
```

```
qxnew = zeros(nx − 1);
qynew = zeros(nx − 1); convM = convolution(nx − 1, ny − 1);
uconv = zeros(tf, (ncom − 2)^2);
load velocitycomp
crand = ones(nx);
crand(npic : npic + ncom − 1, npic : npic + ncom − 1) = velocitycomp;
fort = 1 : tf,
    forxi = 1 : nx − 1
        foryj = 1 : nx − 1
            sigmax = sigma(x(xi) + dx/2, 0);
            sigmay = sigma(0, y(yj) + dy/2);
            lhsqx = 1 + .5 ∗ sigmax ∗ dt;
            lhsqy = 1 + .5 ∗ sigmay ∗ dt;
            dxu = u(xi + 1, yj + 1) + u(xi + 1, yj) − u(xi, yj + 1) − u(xi, yj);
            dxu = .5 ∗ (dxu)/dx;
            dyu = u(xi + 1, yj + 1) + u(xi, yj + 1) − u(xi + 1, yj) − u(xi, yj);
            dyu = .5 ∗ (dyu)/dy;
            rhsqx = (1 − .5 ∗ sigmax ∗ dt) ∗ qx(xi, yj) − dt ∗ (sigmax − sigmay) ∗ dxu;
            rhsqy = (1 − .5 ∗ sigmay ∗ dt) ∗ qy(xi, yj) − dt ∗ (sigmay − sigmax) ∗ dyu;
            qxnew(xi, yj) = rhsqx/lhsqx;
            qynew(xi, yj) = rhsqy/lhsqy;
        end
    end
    Convolq2xn = convM ∗ reshape(qxnew, (nx − 1) ∗ (ny − 1), 1);
    Convolq2yn = convM ∗ reshape(qynew, (nx − 1) ∗ (ny − 1), 1);
```

```
Convqxnew = reshape(Convolq2xn, nx − 1, ny − 1);

Convqynew = reshape(Convolq2yn, nx − 1, ny − 1);

Convqx = qx;

Convqy = qy;

for xi = 2 : nx − 1

    for yj = 2 : nx − 1

        cij = crand(xi, yj)^2;

        ddxu = (u(xi + 1, yj) − 2 ∗ u(xi, yj) + u(xi − 1, yj))/(dx ∗ dx);

        ddyu = (u(xi, yj + 1) − 2 ∗ u(xi, yj) + u(xi, yj − 1))/(dy ∗ dy);

        dxqx = ((Convqx(xi, yj − 1) + Conv_q x(xi, yj)) − (Convqx(xi − 1, yj)

              + Convqx(xi − 1, yj − 1)))/(2 ∗ dx);

        dxqxnew = ((Convqxnew(xi, yj − 1) + Convqxnew(xi, yj))

              − (Convqxnew(xi − 1, yj) + Convqxnew(xi − 1, yj − 1)))/(2 ∗ dx);

        dyqy = ((Convqy(xi, yj) + Convqy(xi − 1, yj)) − (Convqy(xi, yj − 1)

              + Convqy(xi − 1, yj − 1)))/(2 ∗ dy);

        dyqynew = ((Convqynew(xi, yj) + Convqynew(xi − 1, yj))

          − (Convqynew(xi, yj − 1) + Convqynew(xi − 1, yj − 1)))/(2 ∗ dy);

        divqConv = .5 ∗ (dxqx + dxqxnew + dyqy + dyqynew);

        delu = ddxu + ddyu;

        sigmax = sigma(x(xi), 0);

        sigmay = sigma(0, y(yj));

        lhs = 1 + (sigmax + sigmay) ∗ dt/2;

        rhs = (2 − dt ∗ dt ∗ sigmax ∗ sigmay) ∗ u(xi, yj) + (−1 + 0.5 ∗ (sigmax...

         + sigmay) ∗ dt) ∗ uold(xi, yj) + dt ∗ dt ∗ cij ∗ (divqConv + delu);

        unew(xi, yj) = rhs/lhs;

    end

end
```

```
    ucomp = unew(npic + 1 : npic + ncom − 2, npic + 1 : npic + ncom − 2);
    uconv(t, :) = reshape(ucomp, (ncom − 2)^2, 1);
    uold = u; u = unew;
    qx = qxnew; qy = qynew;
end
```