Mixed hybrid numerical methods for convected acoustics. Applications in helioseismology.

Nathan Rouxelin

THÈSE

UNIVERSITE DE PAU ET DES PAYS DE L'ADOUR

École doctorale des Sciences Exactes et de leurs Applications

ED 211 SEA

Présentée et soutenue le 10 décembre 2021

par **Nathan ROUXELIN**

pour obtenir le grade de docteur
de l'Université de Pau et des Pays de l'Adour
Spécialité : Mathématiques

Méthodes numériques mixtes condensées pour l'étude de la propagation des ondes acoustiques en écoulement. Applications en héliosismologie.

MEMBRES DU JURY

RAPPORTEURS

• Xavier ANTOINE	Professeur, Université de Lorraine
• Xavier FERRIERES	Maître de Recherche, ONERA

EXAMINATEURS

• Juliette CHABASSIER	Chargée de Recherche, Inria Bordeaux – Sud-Ouest
• Robert EYMARD	Professeur, Université Gustave Eiffel (Marne-la-Vallée) – Président du jury
• Florian FAUCHER	Chargé de Recherche, Inria Bordeaux – Sud-Ouest
• Laurent GIZON	Professeur, Georg-August Universität, Göttingen Directeur, Max-Planck-Institut für Sonnensystemforschung

DIRECTEURS

• Hélène BARUCQ	Directrice de Recherche, Inria Bordeaux – Sud-Ouest
• Sébastien TORDEUX	Maître de Conférences HDR, Université de Pau et des Pays de l'Adour

Inria

UNIVERSITÉ DE PAU ET DES PAYS DE L'ADOUR

e2s UPPA energy environment solutions

Résumé en français

Dans cette thèse, nous nous intéressons à des problèmes de propagation d'ondes acoustiques dans un écoulement en vue d'applications en physique solaire. En effet, la structure interne du Soleil peut être étudiée à partir de l'observation de ces ondes sur la surface solaire.

Dans un premier temps, nous rappelons le procédé de linéarisation des équations de la mécanique des fluides permettant de construire des modèles vectoriels de propagation d'ondes en écoulement. Nous étudions alors l'équivalence entre les modèles obtenus à partir des linéarisations eulériennes et lagrangiennes et nous montrons que cette équivalence ne peut pas toujours être garantie en régime harmonique. D'un point de vue pratique, l'exploitation numérique de cette équivalence semble inefficace, notamment pour la reconstruction des perturbations lagrangiennes à partir des perturbations eulériennes. Nous nous concentrons ensuite sur l'étude d'un phénomène de résonance en régime harmonique lorsque l'écoulement porteur a des lignes de courant fermées. Une étude modale montre que les équations considérées dégénèrent sur certaines lignes de courant. Sur ces lignes, il n'est alors pas possible de résoudre les équations. Il semble toutefois possible de résoudre ce problème en étudiant ces équations dans un espace hilbertien dont les propriétés de régularités sont plus faibles que pour les espaces de Hilbert habituellement utilisés pour l'étude des équations d'ondes aéroacoustiques. D'un point de vue numérique, la recherche de la solution dans un tel espace nécessite la construction de nouvelles méthodes numériques.

Dans la deuxième partie de ce travail, nous nous concentrons sur la construction de méthodes numériques pour un modèle aéroacoustique simple : l'équation de Helmholtz convectée. Ce modèle scalaire peut être obtenu à partir des modèles vectoriels lorsque l'écoulement porteur est irrotationnel. Pour cette équation, nous construisons trois variantes de la méthode de Galerkine Discontinue Hybride (HDG). Les méthodes HDG sont des méthodes de Galerkine Discontinues mixtes dont le coût numérique reste raisonnable grâce à un procédé de condensation statique qui permet de réduire le problème à un problème posé uniquement sur le squelette du maillage. Nous avons effectué une analyse détaillée de ces méthodes, en particulier nous avons montré le caractère bien posé des méthodes, ainsi que des estimations de la vitesse de convergence pour des solutions régulières. Enfin, nous avons également discuté le choix du paramètre de pénalisation qui peut exercer une influence importante sur la qualité des résultats numériques. Ces méthodes ont été implémentées dans le code open-source Hawen et les résultats numériques ont permis d'illustrer les conclusions de notre étude théorique. Nous avons également construit des conditions aux limites absorbantes (CLA) d'ordre faible pour l'équation de Helmholtz convectée. Ces CLA sont obtenues par transformation de PrandtL-Glauert-Lorentz de CLA pour l'équation de Helmholtz standard lorsque l'écoulement porteur est uniforme à l'extérieur du domaine. Ces CLA sont performantes pour des écoulements dont le nombre de Mach est faible ou modéré et leur mise en œuvre dans un code éléments finis, notamment HDG, est simple.

Enfin, la troisième partie de ce travail est consacrée à l'extension des méthodes HDG construites pour l'équation de Helmholtz convectée à des modèles plus réalistes. Dans un premier temps, nous décrivons les changements à apporter pour traiter les cas vectoriels. La con-

struction d'une méthode HDG pour le cas vectoriel sans convection semble relativement aisée, mais la prise en compte des phénomènes convectifs pose de nombreuses questions tant théoriques que pratiques. Finalement, nous illustrons les possibilités des méthodes HDG sur un problème scalaire issu de l'astérosismologie qui prend en compte une partie des effets liés à la gravité. Le modèle obtenu est semblable à une équation d'Helmholtz convectée par la gravité.

Summary

In this thesis, we consider wave propagation problems in a flow arising in helioseismology. Indeed, the Sun's internal structure can be studied thanks to surface observations of those waves.

In the first part of this work, we recall how vectorial aeroacoustic models can be derived by linearizing the equations of fluid dynamics around an equilibrium state. We then study the equivalence between the resulting equations when both lagrangian and eulerian linearization are used. We show that this equivalence cannot always be guaranteed in the frequency domain. From a practical point of view, exploiting this equivalence seems inefficient, when trying to reconstruct the lagrangian perturbation from the eulerian ones. We then focus on the study of a resonant-like phenomenon that occurs in the frequency domain when the background flow has closed streamlines. The aeroacoustic equations are decomposed as a modal system, which degenerates on some of those streamlines. On those resonant streamlines it is therefore not possible to solve the equations. However, it seems possible to overcome this difficulty by studying these equations in a Hilbert space with lower regularity than the one commonly used from harmonic wave propagation problem. From a numerical point of view, new numerical methods are required to look for the solution in this low-regularity space.

In the second part of this work, we focus on the construction of numerical methods for a simple aeroacoustic model: the convected Helmholtz equation. This scalar model can be derived from the vectorial ones when the carrier flow is irrotational. For this simple equation, we construct three variants of the Hybridizable Discontinuous Galerkin (HDG) method. HDG methods are mixed Discontinuous Galerkin (DG) methods whose numerical cost is reduced thanks to a static condensation process leading to a problem only posed on the skeleton of the mesh. For those methods, we performed a detailed analysis which includes well-posedness results when the mesh is fine enough, as well as convergence estimates for regular solutions. We also discussed the choice of penalization parameter, as it can impact the quality of the numerical results. Those methods were implemented in the open-source software Hawen and the theoretical results were illustrated by the numerical simulations. Finally, we also constructed low-order Absorbing Boundary Conditions (ABCs) for the convected Helmholtz equation by using the Prandtl-Glauert-Lorentz transformation of ABCs for the standard Helmholtz equation when the background flow is locally uniform in the vicinity of the boundary of the domain. These ABCs are accurate for low intermediate Mach numbers, easy to implement in a finite-element solver and well-suited to the HDG formulations that we have constructed.

Finally, the last part of this work is devoted to the extension of those HDG methods to more realistic models. We first describe the changes required to devise HDG methods for vectorial aeroacoustic models. Even if it seems straightforward to construct HDG schemes for the vectorial equation without convection, many theoretical and practical questions arise when we try to take this phenomenon into account. We then illustrate the HDG methods that we have constructed on a scalar problem arising in asteroseismology. This scalar model is similar

to a convected Helmholtz equation where the velocity field is replaced by the gravitational one.

Remerciements

Je tiens tout d'abord à adresser mes remerciements Hélène Barucq et Sébastien Tordeux qui ont encadré ce travail. Je suis très reconnaissant d'avoir pu travailler à leurs côtés pendant cette thèse et de bénéficier de leur expertise et de leurs conseils. Je les remercie pour leur disponibilité et leur implication, même à distance.

Je souhaite remercier Xavier Antoine, Juliette Chabassier, Robert Eymard, Florian Faucher, Xavier Ferrieres et Laurent Gizon qui ont accepté de participer à mon jury de thèse. Je remercie particulièrement les rapporteurs Xavier Antoine et Xavier Ferrieres pour leurs commentaires sur mon travail. Je souhaite également remercier Robert Eymard qui a accepté de présider le jury. Un merci particulier à Juliette et Florian à qui ce travail doit beaucoup. Enfin, je remercie également Laurent Gizon qui m'a accueilli dans son équipe à Göttingen.

Je souhaite remercier tous les membres des équipes MAGIQUE-3D : Elvira, Justine, Alain, Yder, Victor puis MAKUTU : Julien, Ha, Augustin, Chengyi, Vinduja, Rose, Pierre, Aurélien, Margot, Stefano, Julien, Ibrahima, Arjeta, Johan, Nicolas, Sylvie, Algiane, Marc pour tous les moments que nous avons partagé pendant ces trois années.

Enfin, je souhaite remercier ma famille et mes ami(e)s pour leur présence et leur soutien. Merci à tous de vous être connecté pour assister à ma soutenance, et un grand merci à mes parents et à Mathilde et Noël qui ont fait le déplacement jusqu'à Pau.

Contents

Introduction

Contents

In this chapter, we give a general introduction to this work. Our goal is to give some information about the context of this thesis and more precisely on *helioseismology* as most readers may not be familiar with this field. As the subsequent chapters all have their introduction, we do not go into details here and we keep this general introduction relatively short.

1 Mechanical waves characterize the propagation medium

Helioseismology focuses on solar oscillations to probe the solar interior. These oscillations can be understood as aeroacoustic waves propagating inside the Sun. We would like to recall that there is a strong link between mechanical waves and their propagation medium: indeed, measurements of those waves on the surface of an object can be used to probe the interior of this object. This property has been used in many applications, such as

- *seismology:* the interior of the Earth has been imaged using measurements of seismic waves,
- *medical imaging:* doctors can "see" inside a patient using electromagnetic (MRI) or acoustic (ultrasound imaging) waves,
- *nondestructive testing:* it is used to assess the quality of some industrial products.

The idea behind wave-based imaging is that waves propagating inside an object will be reflected or scattered by the internal structure of the object, and this will reflect on the surface measurements.

Illustration of the interaction between waves and the propagation medium. Here, we will illustrate the influence of the nature of the medium on wave propagation with some numerical simulations with the Helmholtz equation

$$-\omega^2 p - c_0^2 \Delta p = \delta,$$

where ω is the frequency, c_0 is the sound-speed and δ is a point-source, solved in the following cases:

- Case #1: uniform sound-speed,

- Case #2: one obstacle and uniform sound-speed,
- Case #3: no obstacle and an inclusion with a change in sound speed.

For each of those cases, a point-source is located at the center of the domain, which is a disk of radius $R = 1.5$, and an *Absorbing Boundary Condition* is used on the exterior boundary of the domain. In FIGURE 1, the results in the whole domain and measured on the circle of radius $r = 1.2$ are depicted. We can clearly see that changes in the propagation medium lead to changes in the resulting waves.

Among those three cases, the closest one to helioseismology is Case #3. Indeed, the goal of helioseismology is to use measurements of the solar oscillations to understand how the physical parameters (such as density, pressure, ...) change inside the Sun. As illustrated with Case #3, wave propagation is impacted by changes in the physical parameters, even if those changes are located in a small area in the domain.

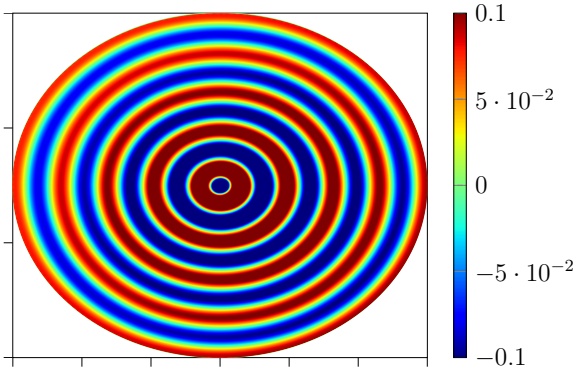
Inverse problems. Imaging the interior of an object from surface measurements of waves is called an *inverse problem*, while determining how waves propagate in a known object is the associated *direct problem* that was illustrated right before.

An inverse problem can be solved either in a *qualitative* or *quantitative* way. Qualitative inversion only focuses on the localization of structure elements. A usual example of qualitative inversion is found in medical imaging: when using X-rays, doctors want to determine where the fracture in a bone is, but they do not need to obtain information about the bone's density. In helioseismic inversion, solar physicists are actually interested in quantitative information on the solar interior, *i.e.* they want to obtain values for some parameters of interest such as physical properties (sound speed, temperature, pressure, ...) or on the chemical composition (abundance of certain elements).

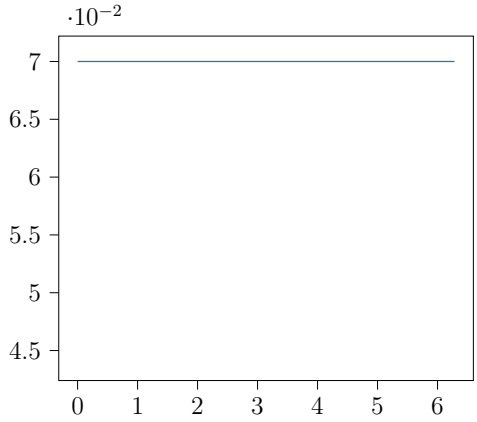
One of the most advanced quantitative inversion procedure is the so-called *full waveform inversion* (FWI) which can be summarized as follows:

1. An initial guess is made for the parameters of interest, this model is denoted by m_0
2. The direct problem is solved using the initial guess leading to the numerical solution $u_h(\mathsf{m}_0)$,
3. The results from this simulation is compared with the experimental (or observational) data u_{exp}
$$\mathcal{J}(\mathsf{m}_0) := \frac{1}{2} \|u_{\text{exp}} - u_h(\mathsf{m}_0)\|^2 + \text{regularisation},$$
where $\mathcal{J}$ is called the *misfit functional*,
4. The parameters of interest are updated
$$\mathsf{m}_{k+1} = \mathsf{m}_k + \alpha_k \mathsf{s}_k,$$
where s_k is the direction and α_k is the step-size, both are determined using a gradient descent procedure on $\mathcal{J}$,
5. Steps 2-3-4 are repeated until $\mathcal{J}(\mathsf{m}_k) <$ tolerance or $|\mathcal{J}(\mathsf{m}_k) - \mathcal{J}(\mathsf{m}_{k-1})| <$ tolerance.

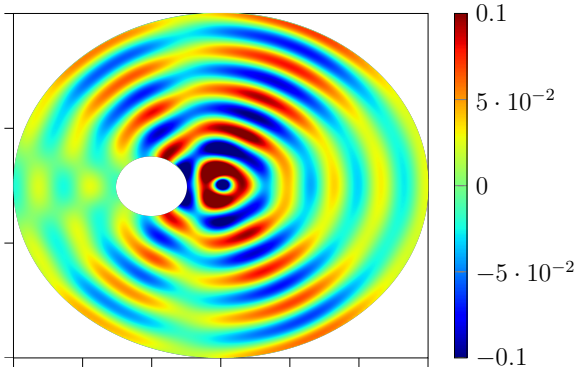
The choice of the initial guess m_0 is quite important: indeed, as $\mathcal{J}$ may not have a unique minimum, the algorithm could converge to a *local minimum* instead of the global one. If prior information is available, it should therefore be incorporated into m_0 to ensure a fast convergence to the real model. If no such information is available, a uniform initial modal can be used, but this may lead to a local minimum and take much more time to obtain convergence. For more details on inverse problems, we refer the reader to [Fau17].



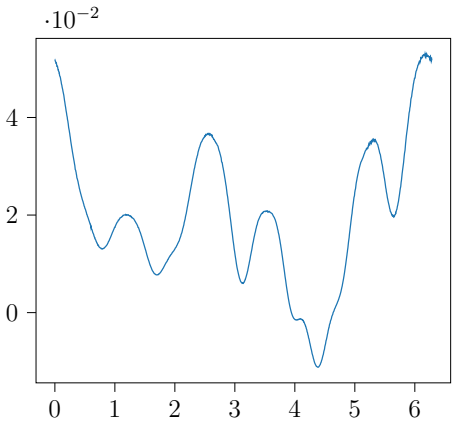
(a) Case #1: $\mathfrak{Re}(p)$



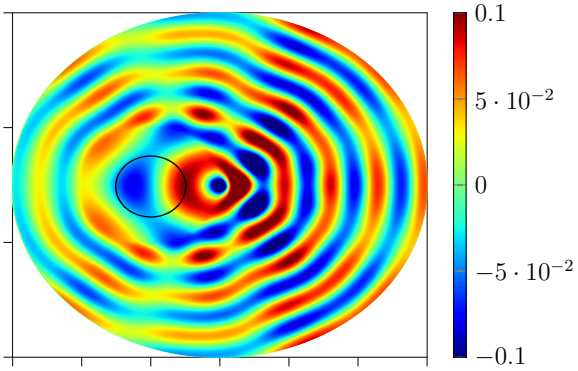
(b) Case #1: $\mathfrak{Re}(p)$ measured on the circle $\{r = 1.2\}$



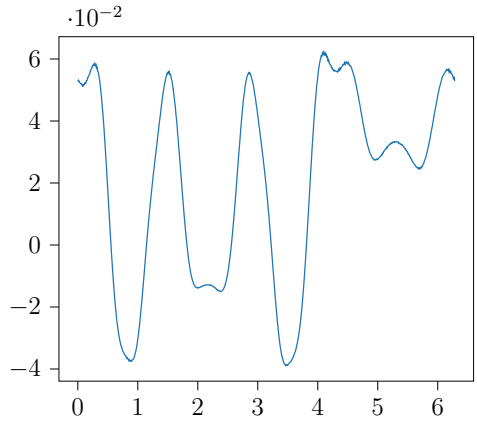
(c) Case #2: $\mathfrak{Re}(p)$



(d) Case #2: $\mathfrak{Re}(p)$ measured on the circle $\{r = 1.2\}$



(e) Case #3: $\mathfrak{Re}(p)$ – change of sound-speed inside the black circle



(f) Case #3: $\mathfrak{Re}(p)$ measured on the circle $\{r = 1.2\}$

Figure 1: Snapshots and measurements of the wave propagation the three cases described above. The computational domain is a disk of radius $R = 1.5$.

2 Solar models

Before moving to the description of the solar oscillations, we give a quick description of the internal structure of the Sun, the interested reader can found more details in [Chr04, Sec. 3.2] or in [Chr02]. As the goal of helioseismology is to infer details about the solar structure, we summarize how *standard solar models*, that describe the global structure of the Sun, are constructed. Those models are usually constructed using Assumption 1. Even if they are quite simple, those models are very useful in the pipeline of helioseismic inversion as they can be used as an initial guess for the inversion procedure.

Assumption 1 (Standard solar models):
To construct a "standard" solar model, the following assumptions are made
- the internal structure is spherically symmetric,
- magnetic effects are neglected,
- rotation is neglected.

As the rotation is not taken into account by the standard model, it is usually one of the parameters that are reconstructed by the inversion procedure.
To describe the solar structure, we introduce the *solar radius* $R_\odot$, which is approximately 700 000km. The structure described by standard solar models is stratified under gravity with three main layers:

- *The nuclear core:* Deepest part of the Sun, from 0 to $0.3R_\odot$, where the nuclear fusion happens. This layer cannot be imaged using helioseismology.
- *The radiative zone:* Intermediate layer, from $0.3R_\odot$ to $0.7R_\odot$, where the temperature is so high that the energy transfer happens through radiative transfer.
- *The convective zone:* Outside layer, from $0.7R_\odot$ to $R_\odot$, where most of the energy transfer happens through convection.

A sketch of this structure is given in Figure 2. It is interesting to notice that the equations governing this structure are quite simple

$$\begin{aligned}
&\frac{\mathrm{d}p}{\mathrm{d}r} = -\frac{Gm\rho}{r^2}, \\
&\frac{\mathrm{d}m}{\mathrm{d}r} = 4\pi r^2 \rho, \\
&\frac{\mathrm{d}T}{\mathrm{d}r} = \frac{\mathrm{d}\ln T}{\mathrm{d}\ln p}\frac{\mathrm{d}p}{\mathrm{d}r}, && \text{in the convective zone,} \\
&\frac{\mathrm{d}T}{\mathrm{d}r} = \frac{3}{16\pi a\tilde{c}G}\frac{\kappa p}{GT^4}\frac{L(r)}{m(r)}, && \text{in the radiative zone,} \\
&\frac{\mathrm{d}L}{\mathrm{d}r} = 4\pi r^2 \rho\varepsilon,
\end{aligned}$$

where p is the pressure, m is the mass contained in the ball of radius r, ρ is the density, T is the temperature, L is the flow energy per unit time through the sphere of radius r, ε is the rate of nuclear energy generation per unit mass and time, u is the internal energy per unit volume, G is the gravitational constant, $\tilde{c}$ is the speed of light, a is the radiation density constant and κ is the opacity.

Even if those equations are quite simple, some complexity is actually hidden. Indeed, to use those equations, source terms must be added. Those source terms are usually expressed in terms of the basic unknowns p, m, T, L and of the abundance of chemical species. Those source terms are designed to fit some properties of the Sun (*e.g.* chemical composition, opacity, ...) and may lead to different results even if the same "basic" equations are used.

As an example, the density of one of the most common solar model, called *Model S*, described in [CDA+96] is depicted on FIGURE 3. In this plot, we can clearly see that it is exponentially decreasing with r and that there is a huge drop close to the solar surface. As several orders of magnitudes are lost between the solar interior and the surface, it may be challenging to perform numerical simulation in those settings. To illustrate this further, we also give the values of some of the physical properties of the Sun in TABLE 1. Once again we can see a huge drop between the core and the surface as the temperature goes from 10^7K to 10^3K.

Property	Value
Age	$4.6 \cdot 10^9$ years
Mass ($M_\odot$)	$2 \cdot 10^{30}$ kg
Radius ($R_\odot$)	700 000 km
Temperature (surf)	5 770 K
Temperature (core)	$1.5 \cdot 10^7$ K

Table 1: Some physical properties of the Sun.

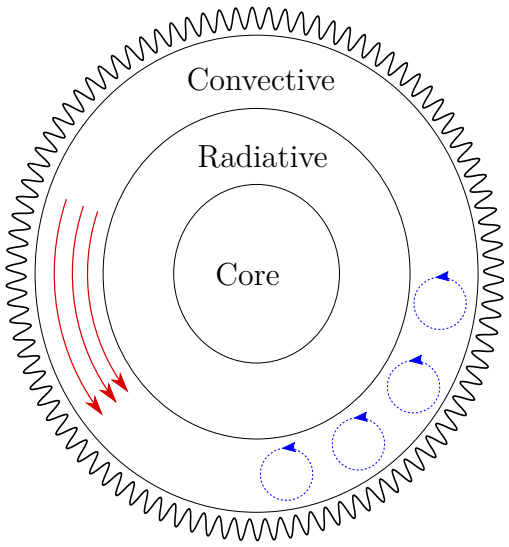


Figure 2: Sketch of the internal structure of the Sun. Red arrows represent the rotation and blue arrows represent convection bubbles that are not taken into account into the standard models.

3 Describing solar oscillations

As the Sun is constituted of plasma, it can be described by *Magneto-Hydro Dynamics* (MHD) equations, which are a coupling between Navier-Stokes' or Euler's equations describing the fluid and Maxwell's equations describing the magnetic field. However, as the MHD equations contain both the solar oscillations and the other MHD effects occurring in the Sun, they are not suitable for helioseismology. Indeed, the oscillations are small when compared to other phenomena occurring in the Sun, and using MHD equations to study them would require a very fine scale resolution of a non-linear system of PDEs. As this would be very expensive from a computational point of view, we need to obtain a set of equations describing only the solar oscillations.

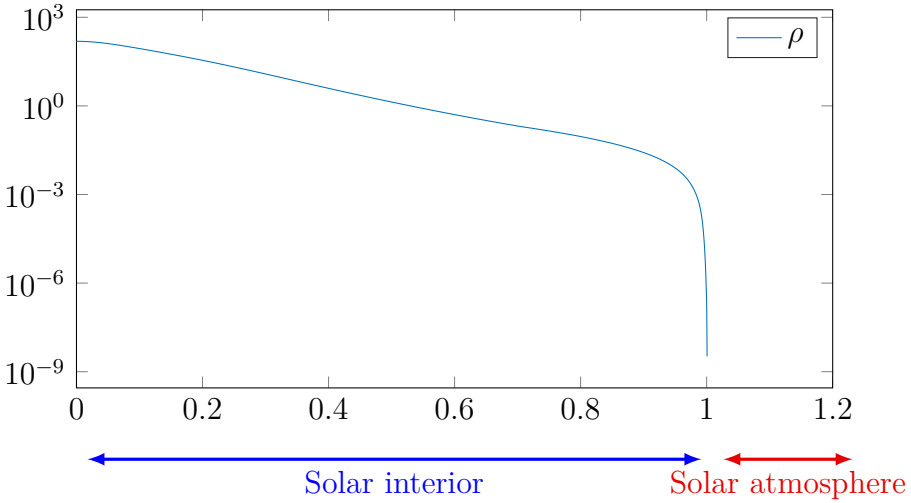


Figure 3: Density ρ (in g·cm^{-3}) inside the Sun as predicted by Model S plotted against the scaled radius $r/R_\odot$ in semi-log scale. Notice that Model S expands into the solar atmosphere, where the density is very low.

To obtain this equation, we need to make some more simplifications. Depending on the phenomena that we wish to study, we need to make one of those assumptions:

- *Aeroacoustic approximation:* the magnetic effects are neglected,
- *Magneto-acoustic approximation:* the hydrodynamic effects are neglected.

The first approximation can be used to probe deep into the solar interior and will be used throughout this thesis. The second one is useful to study surface phenomena such as sun spots. Once again, we only briefly recall the main ideas here, and we refer the interested readers to [Chr04, Chr02] for more details.

In the *aeroacoustic approximation*, the Sun can be described as a perfect fluid satisfying Euler's equations. To obtain a set of equations for the solar oscillations, the Euler's equations are linearized around an equilibrium state. Each physical quantity is decomposed using the following asymptotic development

$$q = q_0 + \varepsilon q' + \mathcal{O}(\varepsilon^2), \quad \varepsilon \ll 1,$$

where q is the *total quantity*, q_0 is the quantity in the equilibrium state or *background quantity* and q' is the *perturbation* produced by solar oscillations. When this asymptotic development is introduced into the Euler's equations and after identifying the powers of ε, we obtain to set of equations:

- the Euler's equations for the equilibrium state at the zeroth order in ε,
- an equation for solar oscillations, also called *aeroacoustic equation*, at the first order in ε.

If the *displacement vector* is used as the main unknown, the obtained *aeroacoustic* equation is the so-called *Galbrun's equation*

$$\rho_0 \left(\frac{\partial}{\partial t} + \nabla_{\boldsymbol{v_0}} \right)^2 \boldsymbol{\xi} - \nabla \left[\rho_0 c_0^2 \mathrm{div}\,(\boldsymbol{\xi}) \right] + \mathrm{div}\,(\boldsymbol{\xi})\, \nabla p_0 - \nabla \left[\nabla_{\boldsymbol{\xi}} p_0 \right] + \nabla_{\boldsymbol{\xi}} \nabla p_0 - \rho_0 \nabla_{\boldsymbol{\xi}} \boldsymbol{g_0} = \boldsymbol{s}, \quad (1)$$

where

- $\boldsymbol{\xi}$ is the displacement vector, which describes the solar oscillations,
- ρ_0, p_0, c_0 and $\boldsymbol{g_0}$ are the background density, pressure, adiabatic sound-speed and gravity respectively,

- $\boldsymbol{s}$ is the source of the solar oscillations,
- $\nabla_{\boldsymbol{v_0}}\boldsymbol{\xi}$ is the directional derivative of $\boldsymbol{\xi}$ along $\boldsymbol{v_0}$.

This equation will be derived in CHAPTER 1. From a historical point of view, it was derived for the first time in [Gal31] and it was introduced in astrophysics in [LO67]. It can be decomposed into two parts. The first one,

$$\rho_0 \left(\frac{\partial}{\partial t} + \nabla_{\boldsymbol{v_0}} \right)^2 \boldsymbol{\xi} - \nabla \left[\rho_0 c_0^2 \mathrm{div}\,(\boldsymbol{\xi}) \right],$$

is a vectorial convected wave operator and the remaining terms,

$$\mathrm{div}\,(\boldsymbol{\xi})\,\nabla p_0 - \nabla\left[\nabla_{\boldsymbol{\xi}} p_0\right] + \nabla_{\boldsymbol{\xi}} \nabla p_0 - \rho_0 \nabla_{\boldsymbol{\xi}} \boldsymbol{g_0},$$

describe the interaction between the oscillations and the background flow.

Different types of waves are supported by this vectorial wave equation. As the Sun is (almost) spherical, it is convenient to decompose the waves onto the basis of *spherical harmonics*. We can then obtain a *power spectrum*, which is a graphical representation of dispersion relations linking the frequency of the wave and the harmonic degree l. An example of such a power spectrum is given in FIGURE 4. Following [Chr04, Sec. 3], we give a quick description of the three types of waves shown on FIGURE 4 below for the simple case where $\boldsymbol{v_0} = \boldsymbol{0}$.

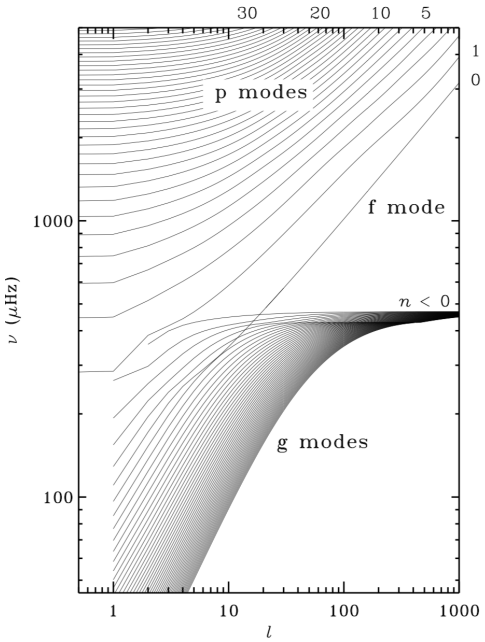


Figure 4: Power spectrum computed for Model S – Frequency is plotted vs the harmonic degree l – Extracted from [Chr02]

Acoustic waves (p-modes). The so-called *p-modes* are standard acoustic waves whose restoring force is pressure. For a wave vector $\boldsymbol{\kappa}$, the acoustic waves satisfy the following *dispersion relation*

$$\omega^2 = c_0^2 |\boldsymbol{\kappa}|^2.$$

Notice that the *adiabatic sound-speed* can be expressed as

$$c_0^2 = \frac{\gamma k_B T_0}{m_u \mu},$$

where $\gamma = 5/3$ is the heat capacity ratio, k_B is the Boltzmann constant, T_0 the background temperature, μ is the mean molecular weight (which describes the presence of heavy elements inside the Sun), m_u is the atomic mass unit. The sound-speed therefore depends on T_0/μ as the other quantities are constant.
Those acoustic waves can propagate deep inside the Sun through the convective and radiative zones. To describe their radial variation, we write the wave vector $\boldsymbol{\kappa}$ as

$$\boldsymbol{\kappa} = \kappa_r \boldsymbol{e}_r + \boldsymbol{\kappa}_h,$$

and we can show that

$$\kappa_r^2 = \frac{\omega^2}{c_0^2}\left(1 - \frac{S_l^2}{\omega^2}\right),$$

where $S_l = l(l+1)c_0^2/r^2 \simeq |\boldsymbol{\kappa}_h|^2 c_0^2$ is the so-called *Lamb frequency*. This equation has a geometric interpretation. As the temperature increases with depth, so does the sound-speed, we can therefore see that κ_r decreases with depth. On the other hand, as depicted on FIGURE 6, S_l increases with depth. As a result, the waves are refracted, and their propagation rays are bent, this is illustrated on FIGURE 5. The turning point is located at the depth r_t where $S_l(r_t) = \omega$.

Finally, we would like to point out that most of the measured solar oscillations consist in p-modes.

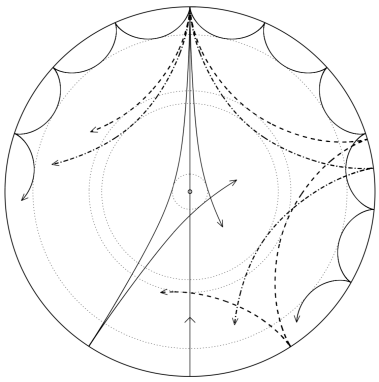


Figure 5: Propagation rays of p-modes inside the Sun – Extracted from [Chr04]

Gravity waves (f- and g-modes). The second important type of waves propagating inside the Sun are *gravity waves*[1]. The restoring force of those waves is buoyancy. They can be separated into two types: *f-modes* that are surfacic gravity waves and *g-modes* that are internal gravity waves. All of those modes are approximately divergence-free.

The *f-modes* satisfy the following dispersion relation

$$\omega^2 = g_s |\boldsymbol{\kappa}_h|,$$

where g_s is the surface gravity and $\boldsymbol{\kappa} = \kappa_r \boldsymbol{e}_r + \boldsymbol{\kappa}_h$ is the wave vector. As this dispersion relation only depends on the surface gravity, the f-modes are independent of the internal structure of the Sun. In the observations, they can therefore be identified even if there are some model uncertainties. On the power spectrum of FIGURE 4, a f-mode is present. Its frequency is similar to the frequencies of p-modes, but its harmonic degree is higher.

[1] Gravity waves of fluid dynamics should not be mistaken with the *gravitational waves* of general relativity.

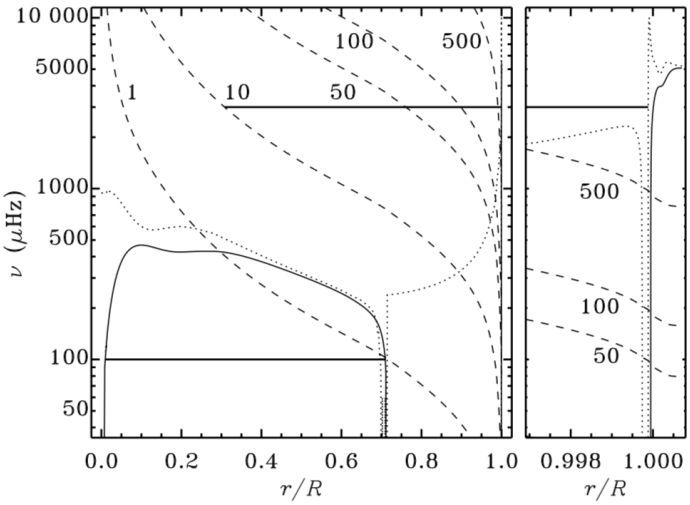


Figure 6: Characteristic frequencies inside the Sun computed with Model S: $N/2\pi$ (solid line), $S_l/2\pi$ (dashed lines, labelled for l) – Heavy horizontal lines describe the trapping region of a g-mode of frequency 100 μHz and a p-mode of frequency 3000 μHz – Extracted from [Chr04]

To describe the *g-modes*, we need to introduce the so-called *Brunt-Vaisala frequency*, or *buoyancy frequency*

$$N^2 = |\boldsymbol{g_0}| \left(\frac{1}{\gamma p_0} \frac{\mathrm{d}p_0}{\mathrm{d}r} - \frac{1}{\rho_0} \frac{\mathrm{d}\rho_0}{\mathrm{d}r} \right).$$

Notice that, despite being a squared quantity, N^2 can be negative. The g-modes can only propagate when $N^2 > 0$. This can be understood by a simple argument given in [Chr04]:

> When a fluid element is displaced upwards in an adiabatic motion, its behavior depends on the density of its new surroundings. If $N^2 > 0$, the element is heavier than the fluid and buoyancy forces it back into its original position leading to an oscillatory motion. On the other hand if $N^2 < 0$, the element is lighter than the fluid and buoyancy acts to enhance the motion.

As it can be seen on FIGURE 6, where N is depicted by a solid line, the condition $N^2 > 0$ is only achieved deep inside the Sun (in the core and in the radiative zone). In particular, g-modes are evanescent the convective zone which is the outer layer of the Sun. As a result, solar g-mode are difficult to measure as we only have access to surface observations.
The g-modes are described by the following dispersion relation

$$\omega^2 = \frac{N^2}{1 + \frac{\kappa_r^2}{|\boldsymbol{\kappa}_h|^2}},$$

where $\boldsymbol{\kappa} = \kappa_r \boldsymbol{e}_r + \boldsymbol{\kappa}_h$ is the wave vector. Notice that we have $\omega^2 < N^2$, so the g-modes tend to have a low frequency as it can be seen on the power spectrum of FIGURE 4.

Damping. In the previous paragraphs, we have written dispersion relations without damping. However, it was found out in [Bal92] that all solar modes are damped. It seems that an important contribution to the damping effect comes from the turbulence. As the modelling of the damping effect is still an open question, we only consider a very simple damping function in this thesis. In the frequency domain, the frequency ω is replaced by

$$\omega \longleftarrow \omega + i\sigma,$$

where σ is the damping parameter. This formally mimics the *limiting absorption principle* coming from the analysis of the Helmholtz equation, we refer to [Wil75] for more details on the analysis of the Helmholtz equation.

Effects of rotation. In this section, we have written down the equations without taking the solar rotation into accounts. Considering this phenomenon can be done by simply adding terms corresponding to the centrifugal and Coriolis forces to Galbrun's equation (1), see [LO67].

Taking the rotation into account has two main effects. Firstly, it breaks the spherical symmetry of the model that was obtained when the background flow only depends on the radial coordinate. Secondly, new modes are supported by the equations such as *Rossby waves*, or *r-modes*, whose restoring force is the Coriolis one. Those Rossby waves were recently observed in the Sun, see [LGB+18].

Reduction to a scalar equation in simple cases. We would like to point out that it is sometimes possible to obtain a scalar equation instead of the vectorial Galbrun's equation (1). Indeed, as described in [GBD+17], by taking the divergence of (1) and neglecting the terms describing the interactions between the flow and the waves and some quadratic terms, the following equation can be obtained

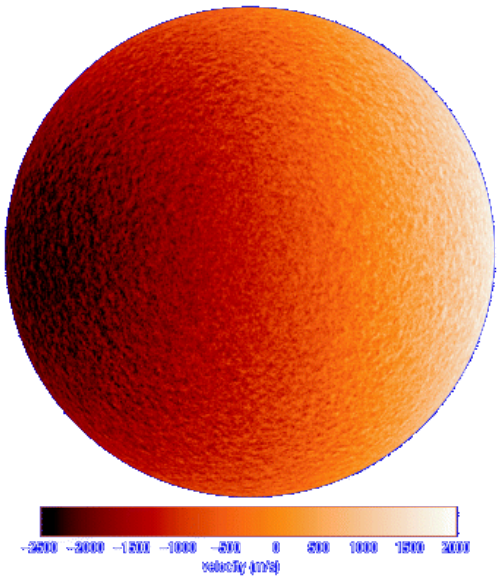
$$-(\omega^2 + 2i\omega\sigma)\psi - 2i\omega\boldsymbol{v_0}\cdot\nabla\psi - c_0\mathrm{div}\left(\frac{1}{c_0}\nabla[\rho_0 c_0\psi]\right) = s, \tag{2}$$

where $s := \mathrm{div}\,(\boldsymbol{s})$ and $\psi := c_0\mathrm{div}\,(\boldsymbol{\xi})$. The quantity $\rho_0 c_0\psi$ is homogenous to a pressure perturbation. When ρ_0 and c_0 are solar-like, the scalar equation (2) captures most of the physics linked with p-modes.

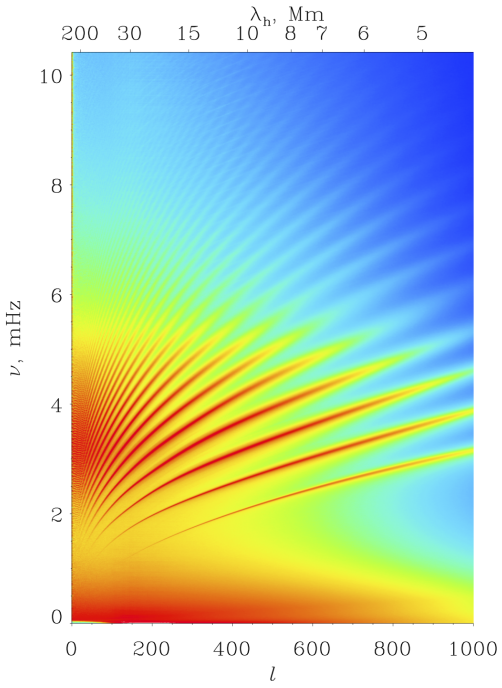
Measuring the solar oscillations. At this point, we have described both a model for solar oscillations and the kind of modes that we expect to find inside the Sun. However all of this cannot be used if there is no observational data to compare with the results of numerical simulation. As we already mentioned, the solar oscillations were observed for the first time in the 1950s. Most of those observations rely on the *relativistic Doppler effect*: the Doppler shift in the solar spectrum can be linked to the projection of the displacement vector $\boldsymbol{\xi}$ on the line-of-sight. Some observational data for the Sun are depicted on FIGURE 7. Notice that the bottom-right ridge on the power spectrum corresponds to the f-mode of the theoretical power spectrum of FIGURE 4.

Solar data can come from:

- ground-based observations, in this case a network of observatory is needed to continuously observe the Sun, the most important ground-based solar observation program is the *Global Oscillations Network Group* (GONG) which relies on six identical telescopes located in Hawaii, California, Chile, Tenerife, India and Australia,
- space-based observations, using satellites located at the *Lagrange points* of the Sun-Earth system, among others some of the famous space mission dedicated to solar observations are
  - the *Solar and Heliospheric Observatory* (SOHO) which has been operating for 25 years (the planned mission duration was two years),
  - the *Solar Dynamics Observatory* (SDO), launched in 2010, which is the follow-up mission to SOHO,
  - the *Solar Orbiter* (SoIO), launched in 2020, which aims at studying the link between magnetic effects and solar eruptions.



(a) Doppler image of the Sun taken by the MDI instrument aboard the SOHO satellite. (Credit: NASA)



(b) Power spectrum obtained from MDI data (Credit: Stanford University)

Figure 7: Observational data for the Sun.

Timeline of the discovery of solar oscillations. We end this section by recalling the timeline[2] of the discovery of solar oscillations. In particular, we would like to point out that helioseismology is quite a new branch of solar physics as the first certain observation of solar oscillations dates back to the 1950s and the first helioseismic inference of solar properties to the middle of the 1980s.

1962: first definitive observations by Leighton:
- roughly periodic oscillations in local Doppler velocity with period about 300s and a lifetime of a few periods,
- convective motion at the supergranular scale,

confirmed the same year by Evans & Michard,

1970s: acoustic nature of the oscillations was understood in the works of Ulrich (1970), Leibacher & Stein (1971), Wolff (1972), Ando & Osaki (1975),

1975: Deubner identified ridges in the wavenumber - frequency diagram describing the modal structure of the oscillations confirmed by Rhodes *et al.* (1977) and constraints on the properties of the convection zone were obtained,

1980: identification of global solar modes from observations at the South Pole by Grec, Fossat & Pomerantz,

1983: observation of intermediate degree observations by Duvall & Harvey,

1984: first helioseismic inference of the internal solar rotation by Duvall *et al.*,

1985: first helioseismic inference of the internal solar sound speed by Christensen-Dalsgaard *et al.*.

[2]This timeline has been extracted from [Chr02] where more precise references can be found.

4 Numerical methods

For a long time, solar physicits relied on *semi-analytical methods* to solve Galbrun's equation (1) by reducing the model to a scalar ODE. Even if those kinds of methods are very efficient, they can only be used under restrictive assumptions.

As those methods can only be used for a spherically symmetric Sun, it is convenient to introduce the *spherical coordinates* (r, θ, φ) and the associated basis vector-fields $(\boldsymbol{e}_r, \boldsymbol{e}_\theta, \boldsymbol{e}_\varphi)$. Vector-fields can then be decomposed into their *radial* and *horizontal* parts as follows

$$\boldsymbol{\xi} = \xi_r \boldsymbol{e}_r + \boldsymbol{\xi}_h, \tag{3}$$

where $\boldsymbol{\xi}_h \in \operatorname{span}(\boldsymbol{e}_\theta, \boldsymbol{e}_\varphi)$. A similar decomposition also holds for the gradient operator, indeed let u be a regular function, we have

$$\nabla u = \frac{\partial u}{\partial r}\boldsymbol{e}_r + \nabla_h u, \qquad \text{where} \quad \nabla_h u := \frac{1}{r}\frac{\partial u}{\partial \theta}\boldsymbol{e}_\theta + \frac{1}{r\sin\theta}\frac{\partial u}{\partial \varphi}\boldsymbol{e}_\varphi.$$

Assumption 2 (Deriving a semi-analytical method):
To derive a semi-analytical method to solve (1), we assume

- a null velocity field: $\boldsymbol{v_0} = \boldsymbol{0}$,
- a radial background flow: $\nabla_h q_0 = \boldsymbol{0}$ for $q_0 \in \{\rho_0, p_0, c_0\}$.

Spherical harmonics. Let l be a non-negative integer and m an integer such that $-l \leqslant m \leqslant l$. We denote by S^2 the unit sphere. The function $Y_l^m : S^2 \to \mathbb{C}$ satisfying

$$-\Delta_S Y_l^m(\theta, \varphi) = l(l+1) Y_l^m(\theta, \varphi),$$

where Δ_S is the Laplace-Beltrami operator over S^2, is called the *spherical harmonic function of degree l and order m*. The number l can be interpreted as the number of parallels and $|m|$ as the number of meridians in the graphical representation of Y_l^m. The spherical harmonics are the eigenfunctions of the spherical Laplace-Beltrami operator. The first few spherical harmonics are depicted on Figure 8.

The spherical harmonics can be expressed in terms of the *associated Legendre polynomials* P_l^m defined by

$$P_l^m(x) = \frac{(-1)^m}{2^l l!}(1-x^2)^{\frac{m}{2}}\frac{\mathrm{d}^{l+m}}{\mathrm{d}x^{l+m}}(x^2-1)^l,$$

the following formula links Y_l^m and P_l^m

$$Y_l^m(\theta, \varphi) = c_{l,m} P_l^m(\cos\theta) e^{im\varphi},$$

where $c_{l,m}$ is a normalization constant.

As spherical harmonics constitute an orthogonal Hilbert basis of $\mathcal{C}(S^2)$, any continuous function over S^2 can be decomposed as a series of spherical harmonics: let $f : S^2 \to \mathbb{C}$ be a continuous function, we have

$$f(\theta, \varphi) = \sum_{l=0}^{+\infty} \sum_{m=-l}^{l} f_l^m Y_l^m(\theta, \varphi).$$

This property is a key ingredient to devise semi-analytical methods for Galbrun's equation (1).

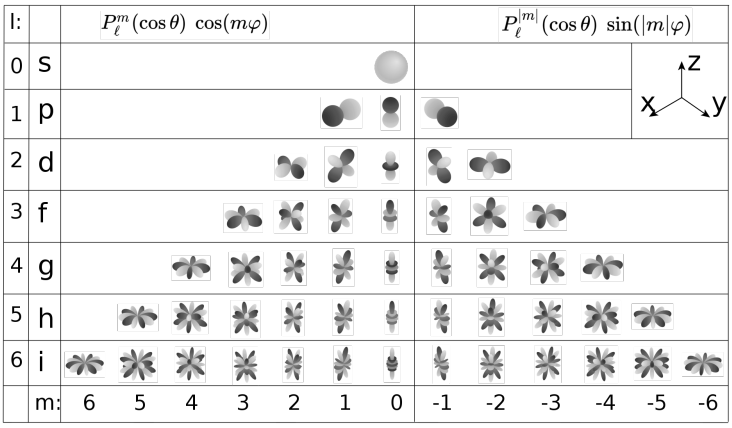


Figure 8: Representation of the first spherical harmonics. The letters s,p,d,f... corresponds to the valence shells in chemistry. Extracted from Wikipedia.

Reduction to a system of ODEs. The construction presented in this paragraph is a summary of [Chr04, Chap. 4] to which we refer the interested reader for more details. To devise a semi-analytical method, it is not really convenient to directly work with Galbrun's equation (1). Instead we consider the following system

$$\rho_0 \frac{\partial^2 \boldsymbol{\xi}}{\partial t^2} + \frac{\rho'}{\rho_0} \nabla p_0 + \nabla p' = \boldsymbol{s}, \tag{4}$$

$$\rho' + \operatorname{div}(\rho_0 \boldsymbol{\xi}) = 0, \tag{5}$$

$$p' + \rho_0 c_0^2 \operatorname{div}(\boldsymbol{\xi}) + \boldsymbol{\xi} \cdot \nabla p_0 = 0, \tag{6}$$

where the *eulerian perturbations* of density ρ' and of pressure p' explicitly appear. As it will be detailed in Chapter 1, Galbrun's equation (1) is actually derived from the system (4)–(5)–(6).
Using the decomposition (3) of $\boldsymbol{\xi}$ into its radial and horizontal parts into (4), we can obtain an equation for the horizontal component

$$\rho_0 \frac{\partial^2 \boldsymbol{\xi}_h}{\partial t^2} + \nabla_h p' = \boldsymbol{s}_h,$$

as $\nabla_h p_0 = \mathbf{0}$. By taking the horizontal divergence, we obtain

$$\frac{\partial^2 \operatorname{div}_h(\boldsymbol{\xi}_h)}{\partial t^2} + \Delta_h p' = \operatorname{div}_h(\boldsymbol{s}_h). \tag{7}$$

Using (5), we also have

$$\rho' = -\frac{1}{r^2} \frac{\partial r^2 \rho_0 \xi_r}{\partial r} - \rho_0 \operatorname{div}_h(\boldsymbol{\xi}_h),$$

so (7) becomes

$$-\frac{\partial^2}{\partial t^2} \left[\frac{\rho'}{\rho_0} - \frac{1}{r^2 \rho_0} \frac{\partial r^2 \rho_0 \xi_r}{\partial r} \right] + \Delta_h p' = \operatorname{div}_h(\boldsymbol{s}_h). \tag{8}$$

The radial part of (4) is

$$\rho_0 \frac{\partial^2 \xi_r}{\partial t^2} + \frac{\rho'}{\rho_0} \frac{\mathrm{d} p_0}{\mathrm{d} r} + \frac{\partial p'}{\partial r} = s_r, \tag{9}$$

and (6) can be written as

$$p' + \rho_0 c_0^2 \left(\frac{1}{r^2} \frac{\partial r^2 \xi_r}{\partial r} + \mathrm{div}_h \left(\boldsymbol{\xi}_h \right) \right) + \xi_r \frac{\mathrm{d}p_0}{\mathrm{d}r} = 0,$$

which leads to

$$p' + \rho_0 c_0^2 \left(\frac{1}{r^2} \frac{\partial r^2 \xi_r}{\partial r} - \frac{1}{r^2 \rho_0} \frac{\partial r^2 \rho_0 \xi_r}{\partial r} + \frac{\rho'}{\rho_0} \right) + \xi_r \frac{\mathrm{d}p_0}{\mathrm{d}r} = 0. \tag{10}$$

At this point, we notice that the partial derivatives with respect to θ and φ in (8)–(9)–(10) only appear in the Δ_h term in (8). We therefore introduce the following decompositions onto the spherical harmonics

$$\begin{aligned}
\xi_r(t, r, \theta, \varphi) &= \mathfrak{Re} \sum_{l=0}^{+\infty} \sum_{m=-l}^{l} \xi_{r,lm}(r) Y_l^m(\theta, \varphi) e^{-i\omega t}, \\
p'(t, r, \theta, \varphi) &= \mathfrak{Re} \sum_{l=0}^{+\infty} \sum_{m=-l}^{l} p'_{lm}(r) Y_l^m(\theta, \varphi) e^{-i\omega t}, \\
\rho'(t, r, \theta, \varphi) &= \mathfrak{Re} \sum_{l=0}^{+\infty} \sum_{m=-l}^{l} \rho'_{lm}(r) Y_l^m(\theta, \varphi) e^{-i\omega t},
\end{aligned}$$

where we further assumed that the time dependance is harmonic. Plugging this decomposition into the system (8)–(9)–(10) and identifying each mode, we obtain the following set of *modal equations*

$$\omega^2 \left[\frac{\rho'_{lm}}{\rho_0} - \frac{1}{r^2 \rho_0} \frac{\mathrm{d}r^2 \rho_0 \xi_{r,lm}}{\mathrm{d}r} \right] - l(l+1) p'_{lm} = \mathrm{div}_h \left(\boldsymbol{s}_h \right), \tag{11}$$

$$-\rho_0 \omega^2 \xi_{r,lm} + \frac{\rho'_{lm}}{\rho_0} \frac{\mathrm{d}p_0}{\mathrm{d}r} + \frac{\mathrm{d}p'_{lm}}{\mathrm{d}r} = s_{r,lm}, \tag{12}$$

$$p'_{lm} + \rho_0 c_0^2 \left(\frac{1}{r^2} \frac{\mathrm{d}r^2 \xi_{r,lm}}{\mathrm{d}r} - \frac{1}{r^2 \rho_0} \frac{\mathrm{d}r^2 \rho_0 \xi_{r,lm}}{\mathrm{d}r} + \frac{\rho'_{lm}}{\rho_0} \right) + \xi_{r,lm} \frac{\mathrm{d}p_0}{\mathrm{d}r} = 0. \tag{13}$$

The modal system (11)–(12)–(13) is a linear system of ODEs that can be solved independently for each mode. A numerical solution can be obtained by truncating the series in l. We can notice that because of the spherical symmetry, it does not depend on the order m.

Our proposed solution to go further. In the previous paragraph, we have shown how the time-harmonic Galbrun's equation (1) can be reduced to a linear system of ODEs (11)–(12)–(13) for cases where Assumption 2 holds. For more realistic configurations, this cannot be done anymore. For example, when $\boldsymbol{v_0} \neq 0$, it may not be possible to obtain a simple modal system as the convection terms can couple different modes. To address those difficulties, we propose to use different numerical methods that belong to the family of *Galerkin methods*, or *finite element methods*.

Galerkin methods are more flexible and will allow to take more realistic configurations into account. They are based on the discretization of a *weak*, or *variational*, formulation on an unstructured mesh. In [GBD+17], Galerkin methods were introduced for the scalar model (2) and interesting results were obtained. In particular, we will consider *(Hybridizable) Discontinuous Galerkin methods* for which the continuity of the solution between two elements is only weakly enforced. Both theoretical and implementation details on the methods are given in Chapter 3.

5 Organization of this work

We end this introduction by saying a few words about the outline of this work.

- In Chapter 1: We describe how the aeroacoustic equations describing the solar oscillations can be derived from the equations describing an ideal gas.
- in Chapter 2: We study a resonant-like phenomenon that can occur when aeroacoustic waves interact with a recirculating flow.
- in Chapter 3: We introduce a framework for solving the convected Helmholtz equation using HDG methods, both numerical and theoretical results are obtained.
- in Chapter 4: We construct Absorbing Boundary Conditions (ABCs) for the convected Helmholtz equation, they are based on the Lorentz transformation of the Engquist-Madja ABCs for the standard Helmholtz equation.
- in Chapter 5: We give some details on how the HDG framework of Chapter 3 could be extended to the vectorial Galbrun's equation, and we also point out the main theoretical difficulties arising in this process.
- in Chapter 6: We use a scalar model coming from *asteroseismology* that takes some effects of the gravity into account to illustrate that the HDG framework of Chapter 3 can be used for models that are more complex than the convected Helmholtz equation.

Bibliography

[Bal92] Neil J. Balmforth. Solar pulsational stability - I. Pulsation-mode thermodynamics. *Monthly Notices of the Royal Astronomical Society*, 255:603–649, April 1992.

[CDA+96] Jorgen Christensen-Dalsgaard, W. Däppen, S. V. Ajukov, E. R. Anderson, H. M. Antia, S. Basu, V. A. Baturin, G. Berthomieu, B. Chaboyer, S. M. Chitre, A. N. Cox, P. Demarque, J. Donatowicz, W. A. Dziembowski, M. Gabriel, D. O. Gough, D. B. Guenther, J. A. Guzik, J. W. Harvey, F. Hill, G. Houdek, C. A. Iglesias, A. G. Kosovichev, J. W. Leibacher, P. Morel, C. R. Proffitt, J. Provost, J. Reiter, E. J. Rhodes, F. J. Rogers, I. W. Roxburgh, M. J. Thompson, and R. K. Ulrich. The Current State of Solar Modeling. *Science*, 272(5266):1286–1292, May 1996.

[Chr02] Jorgen Christensen-Dalsgaard. Helioseismology. *Reviews of Modern Physics*, 74(4):1073–1129, November 2002.

[Chr04] Jorgen Christensen-Dalsgaard. *Lecture Notes on Stellar Oscillations*. 2004.

[Fau17] Florian Faucher. *Contributions to Seismic Full Waveform Inversion for Time Harmonic Wave Equations: Stability Estimates, Convergence Analysis, Numerical Experiments Involving Large Scale Optimization Algorithms*. Theses, Université de Pau et des Pays de l'Adour, November 2017.

[Gal31] Henri Galbrun. *Propagation d'une Onde Sonore Dans l'atmosphère et Théorie Des Zones de Silence*. Gauthier-Villars, 1931.

[GBD+17] Laurent Gizon, Hélène Barucq, Marc Durufle, Chris Hanson, Michael Leguèbe, Aaron Birch, Juliette Chabassier, Damien Fournier, Thorsten Hohage, and Emanuele Papini. Computational helioseismology in the frequency domain: Acoustic waves in axisymmetric solar models with flows. *Astronomy and Astrophysics - A&A*, 600:A35, 2017.

[LGB+18] Björn Löptien, Laurent Gizon, Aaron C. Birch, Jesper Schou, Bastian Proxauf, Thomas L. Duvall, Richard S. Bogart, and Ulrich R. Christensen. Global-scale equatorial Rossby waves as an essential component of solar internal dynamics. *Nature Astronomy*, 2(7):568–573, July 2018.

[LO67] Donald Lynden-Bell and Jeremiah Paul Ostriker. On the Stability of Differentially Rotating Bodies. *Monthly Notices of the Royal Astronomical Society*, 136(3):293–310, July 1967.

[Wil75] Calvin H. Wilcox. *Scattering Theory for the d'Alembert Equation in Exterior Domains*. Lecture Notes in Mathematics. Springer-Verlag, Berlin Heidelberg, 1975.

Part I

Aeroacoustic modelling & resonance

Chapter 1

Derivation of aeroacoustic models

Contents

Introduction

In this chapter, we recall how *aeroacoustic wave equations* are derived from the Euler's equations describing an ideal gas. Those equations are used to model the propagation of waves in complex flows. We also discuss the equivalence between two aeroacoustic models in time-harmonic domain: Galbrun's equation and LEE. Finally, we give a brief reviews of the different well-posedness results that are available for those equations.

1.1 Description of an ideal fluid

Before being able to describe the propagation of acoustic waves in a fluid, we need to have a model for the fluid of interest.

Physical quantities : In all the document, we will use the following notations for the physical quantities :

- ρ : mass density
- p : pressure

- $\boldsymbol{v}$: velocity field
- $\boldsymbol{g}$: gravity field
- e : internal energy
- c : sound speed

Directional derivative for vector-fields : We will use the notation $\nabla_{\boldsymbol{u}}\boldsymbol{v}$ to describe the convective part of the Euler's equations. In cartesian coordinates, it should be understood as the product between $\boldsymbol{u}$ and the jacobian matrix of $\boldsymbol{v}$:

$$\nabla \boldsymbol{v} = \begin{bmatrix} \frac{\partial v_1}{\partial x_1} & \cdots & \frac{\partial v_1}{\partial x_d} \\ \vdots & & \vdots \\ \frac{\partial v_d}{\partial x_1} & \cdots & \frac{\partial v_d}{\partial x_d} \end{bmatrix}, \quad \text{and} \quad [\nabla_{\boldsymbol{u}}\boldsymbol{v}]_\ell = \sum_{i=1}^{d} u_i \frac{\partial v_\ell}{\partial x_i}.$$

Those formulas cannot be used for other systems of coordinates, as we need to take into account the derivatives of the basis vector fields:

$$[\nabla_{\boldsymbol{u}}\boldsymbol{v}]_\ell = \sum_{i=1}^{d} u_i \frac{\partial v_\ell}{\partial x_i} + \sum_{i,j=1}^{d} \Gamma^\ell_{ij} v_i u_j,$$

where $(\Gamma^\ell)^\ell$ should be interpreted as the jacobian matrices of the basis vector fields. For example, in polar coordinates, the Γs[1] are given by

$$\Gamma^r = \begin{bmatrix} 0 & 0 \\ 0 & -\frac{1}{r} \end{bmatrix}, \quad \text{and} \quad \Gamma^\theta = \begin{bmatrix} 0 & \frac{1}{r} \\ 0 & 0 \end{bmatrix}.$$

We have chosen to use the notation $\nabla_{\boldsymbol{u}}\boldsymbol{v}$, which is the notation for the *covariant derivative*, instead of the notation $(\boldsymbol{u}\cdot\nabla)\boldsymbol{v}$, which is widely used in fluid dynamics, as we find the second notation to be quite misleading. Indeed it should be interpreted as a matrix-vector product, but it is written as a vector-matrix product, and one should be very careful when using this notation in non-cartesian coordinate systems. However, for scalar quantities we have

$$\nabla_{\boldsymbol{u}} p := \boldsymbol{u} \cdot \nabla p,$$

by definition of the covariant derivative.

Euler's equation : The classical Euler's equations read

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \boldsymbol{v}) = 0, \tag{Eul-ρ}$$

$$\frac{\partial \rho \boldsymbol{v}}{\partial t} + \operatorname{div}(\rho \boldsymbol{v} \otimes \boldsymbol{v}) + \nabla p = \rho \boldsymbol{g}, \tag{Eul-$\rho\boldsymbol{v}$}$$

$$\frac{\partial}{\partial t}\left[\rho\left(\frac{|\boldsymbol{v}|^2}{2} + e\right)\right] + \operatorname{div}\left(\left(\rho\frac{|\boldsymbol{v}|^2}{2} + \rho e + p\right)\boldsymbol{v}\right) = 0. \tag{Eul-e}$$

Here, $\otimes$ denotes the tensor product of two vector fields, and the derivative operator acts on second-order tensor fields. In the followings, we will only use n alternative form of (Eul-$\rho\boldsymbol{v}$) where this term does not appear. We there only give the expression of (Eul-$\rho\boldsymbol{v}$) in an orthonormal basis

$$\frac{\partial \rho v_i}{\partial t} + \sum_{j=1}^{d} \frac{\partial}{\partial x_j}\left[\rho v_i v_j\right] + \frac{\partial p}{\partial x_i} = \rho g_i, \quad \forall i \in [\![1, d]\!].$$

[1] More precisely, they are the *Christoffel symbols of the second kind* computed the covention $\Gamma^\ell_{ij} = \langle \boldsymbol{\omega}^\ell, \nabla_{\boldsymbol{e}_j} \boldsymbol{e}_i \rangle$, where $(\boldsymbol{\omega}^\ell)^\ell$ is the dual basis of $(\boldsymbol{e}_j)_j$.

A derivation of those equations can be found for instance in [Lio96, Ch 1].
In order to form a closed system of equations, we need to add an equation of state

$$p = \mathcal{P}(\rho, e)$$

The balance of momentum (Eul-$\rho\boldsymbol{v}$) can be rewritten as

$$\rho\left(\frac{\partial \boldsymbol{v}}{\partial t} + \nabla_{\boldsymbol{v}}\boldsymbol{v}\right) + \nabla p = \rho \boldsymbol{g} \tag{Eul-$\boldsymbol{v}$}$$

Indeed one has

$$\begin{aligned}\frac{\partial \rho \boldsymbol{v}}{\partial t} + \operatorname{div}(\rho \boldsymbol{v} \otimes \boldsymbol{v}) &= \frac{\partial \rho}{\partial t}\boldsymbol{v} + \rho\frac{\partial \boldsymbol{v}}{\partial t} + (\rho \operatorname{div}(\boldsymbol{v}))\boldsymbol{v} + \rho\nabla_{\boldsymbol{v}}\boldsymbol{v} + (\boldsymbol{v}\cdot\nabla\rho)\boldsymbol{v} \\ &= \rho\left(\frac{\partial \boldsymbol{v}}{\partial t} + \nabla_{\boldsymbol{v}}\boldsymbol{v}\right) + \underbrace{\left(\frac{\partial \rho}{\partial t} + \rho\operatorname{div}(\boldsymbol{v}) + \boldsymbol{v}\cdot\nabla\rho\right)}_{=0 \text{ by (Eul-}\rho\text{)}}\boldsymbol{v}\end{aligned}$$

The balance of total energy (Eul-e) leads to the following balance of internal energy e

$$\frac{\partial \rho e}{\partial t} + \operatorname{div}(\rho e \boldsymbol{v}) + p\operatorname{div}(\boldsymbol{v}) = 0 \tag{1.1}$$

Using (1.1) together with adiabatic equation of state for a polytropic ideal gas

$$e = \frac{p}{\rho(\gamma - 1)},$$

we obtain the following equation for pressure

$$\frac{\partial p}{\partial t} + \boldsymbol{v}\cdot\nabla p + \gamma p \operatorname{div}(\boldsymbol{v}) = 0. \tag{Eul-p}$$

In the following we will consider the system (Eul-ρ)–(Eul-$\boldsymbol{v}$)–(Eul-p). This system is actually used in helioseismology, see [Chr04, Ch. 3].

1.2 Linearized Euler's Equations

The system (Eul-ρ)–(Eul-$\boldsymbol{v}$)–(Eul-p) can be used to describe the evolution of a fluid. However the variations due to wave propagation are small when compared to other fluid phenomena, and the system (Eul-ρ)–(Eul-$\boldsymbol{v}$)–(Eul-p) is therefore not suited to study wave propagation in a fluid. From a numerical point of view, using a *Computational Fluid Dynamics* (CFD) solver for this purpose would require a very fine mesh or a very high order and would lead to a prohibitive numerical cost. To address this difficulty, we consider a perturbative approach to derive an equation for wave propagation by linearizing (Eul-ρ)–(Eul-$\boldsymbol{v}$)–(Eul-p) around an equilibrium state.

Eulerian perturbations Let $q \in \{\rho, \boldsymbol{v}, p\}$ be a physical quantity, we assume that the following *perturbation expansion* holds

$$q(\boldsymbol{x}, t) = q_0(\boldsymbol{x}, t) + \varepsilon q'(\boldsymbol{x}, t) + \mathcal{O}(\varepsilon^2), \tag{1.2}$$

where q_0 is the equilibrium (or background) value of q, q' is its perturbation due to wave propagation, and $\varepsilon \ll 1$ is a small parameter.
As the space variable, $\boldsymbol{x}$, is the same for all the quantities in (1.2), we say that q' is an *eulerian perturbation*. Introducing the perturbation expansion (1.2) in (Eul-ρ)–(Eul-$\boldsymbol{v}$)–(Eul-p) leads to two sets of equations: one for the equilibrium state and one for wave propagation.

Basic assumptions on wave propagation We can now state the assumptions under which we can derive a model for aeroacoustic wave propagation.

Assumption 3 (Basic assumptions):
To derive an equation for wave propagation, we make the following assumptions on the physical processes

- *Linear waves*: (1.2) holds,
- *Steady-state background flow*: $\frac{\partial q_0}{\partial t} = 0$ for $q_0 \in \{\rho_0, \boldsymbol{v_0}, p_0\}$,
- *Adiabatic waves*: (Eul-p) holds for the eulerian settings and $\delta p = c_0^2 \delta\rho$ for the lagrangian settings (detailed later),
- *Cowling approximation*: there is no perturbation of gravitation, *i.e.* $\boldsymbol{g'} = \boldsymbol{0}$.

Sound waves are usually assumed to be adiabatic. To justify this assumption, the following reason is often quoted: the propagation of sound waves is so fast that heat exchanges do not have enough time to take place. As discussed in [Pie19, Ch. 1], the wave propagation always produces a thermal gradient. This heat gradient can only be neglected for low and medium frequencies. At those frequencies sound propagation is almost an adiabatic process. As discussed in [MMMP17], the rigorous assumption to derive an aeroacoustic wave equations is *homentropicity*: the entropy of the background flow is uniform. With this assumption, the entropy perturbation satisfies a decoupled equation and only acts as a source term in the other equations. Notice that the assuming that the background flow is homentropic is more restrictive than assuming it is isentropic: indeed, for in an isentropic flow the entropy density can change between streamlines, which is not the case for a homentropic flow.

Background equations Introducing the perturbation expansion (1.2) in (Eul-ρ)–(Eul-$\boldsymbol{v}$)–(Eul-p) and collecting the zeroth-order terms in ε leads to the following set of equations

$$\begin{aligned}
\frac{\partial \rho_0}{\partial t} + \operatorname{div}(\rho_0 \boldsymbol{v_0}) &= 0, \\
\rho_0 \left(\frac{\partial \boldsymbol{v_0}}{\partial t} + \nabla_{\boldsymbol{v_0}} \boldsymbol{v_0} \right) + \nabla p_0 &= \rho_0 \boldsymbol{g_0}, \\
\frac{\partial p_0}{\partial t} + \boldsymbol{v_0} \cdot \nabla p_0 + \gamma p_0 \operatorname{div}(\boldsymbol{v_0}) &= 0.
\end{aligned}$$

Furthermore, following Assumption 3, we assume that the background flow is *steady-state*, *i.e.*

$$\frac{\partial q_0}{\partial t} = 0, \qquad \forall q_0 \in \{\rho_0, \boldsymbol{v_0}, p_0\},$$

we obtain the following set of *background equations*

$$\begin{aligned}
\operatorname{div}(\rho_0 \boldsymbol{v_0}) &= 0, && \text{(BG-}\rho_0\text{)} \\
\rho_0 \nabla_{\boldsymbol{v_0}} \boldsymbol{v_0} + \nabla p_0 &= \rho_0 \boldsymbol{g_0}, && \text{(BG-}\boldsymbol{v_0}\text{)} \\
\boldsymbol{v_0} \cdot \nabla p_0 + \gamma p_0 \operatorname{div}(\boldsymbol{v_0}) &= 0. && \text{(BG-}p_0\text{)}
\end{aligned}$$

Linearized Euler's Equations We are now ready to obtain a system of equations for wave propagation. Once again, this is done by introducing the perturbation expansion (1.2) in (Eul-ρ)–(Eul-$\boldsymbol{v}$)–(Eul-p). We then collect the first-order terms in ε to obtain

$$\begin{aligned}
\mathrm{D}_t \rho' + \rho' \operatorname{div}(\boldsymbol{v_0}) + \operatorname{div}(\rho_0 \boldsymbol{v'}) &= 0 && \text{(LEE-}\rho'\text{)} \\
\rho_0 \mathrm{D}_t \boldsymbol{v'} + \rho_0 \nabla_{\boldsymbol{v'}} \boldsymbol{v_0} + \rho' (\nabla_{\boldsymbol{v_0}} \boldsymbol{v_0} - \boldsymbol{g_0}) + \nabla p' &= \boldsymbol{s} && \text{(LEE-}\boldsymbol{v'}\text{)} \\
\mathrm{D}_t p' + \boldsymbol{v'} \cdot \nabla p_0 + \gamma p_0 \operatorname{div}(\boldsymbol{v'}) + \gamma p' \operatorname{div}(\boldsymbol{v_0}) &= 0 && \text{(LEE-}p'\text{)}
\end{aligned}$$

where the *material derivative* is defined as

$$\mathrm{D}_t := \frac{\partial}{\partial t} + \nabla_{\boldsymbol{v_0}},$$

and $\boldsymbol{s}$ is the source of sound that encompasses:

- sound generated by the moving flow, these kind of sources were introduced by Lighthill in [LN52] and specialized to LEE in [BBJ02], roughly speaking this models the "sound of turbulence",
- sound generated by other sources, such as artificial devices.

Using (BG-$\boldsymbol{v_0}$), we can rewrite (LEE-$\boldsymbol{v}'$) as

$$\rho_0 \mathrm{D}_t \boldsymbol{v}' + \rho_0 \nabla_{\boldsymbol{v}'} \boldsymbol{v_0} - \left(\frac{1}{\rho_0} \nabla p_0\right) \rho' + \nabla p' = \boldsymbol{s},$$

so the gravity term $\boldsymbol{g_0}$ does not explicitly appear in the equation. This is interesting as the gravity term is useful for helioseismology, but usually neglected in other applications.
We also notice that, following the *Cowling approximation* of Assumption 3, there is no perturbation of gravity.

1.3 Galbrun's Equation

In the previous section, we obtained a first set of equations (LEE-ρ')–(LEE-$\boldsymbol{v}'$)–(LEE-p') for wave propagation in a flow. Those equations are based on an eulerian linearization around an equilibrium state. It is possible to obtain an alternative model by considering *lagrangian perturbations* instead of eulerian ones.

Lagrangian perturbations and displacement To define the eulerian perturbation in (1.2), we considered the difference between the total and background quantities at a fixed position. In the lagrangian framework, we fix a *fluid parcel*[2] and follow its motion through the flow. At time t, the parcel is at position $\boldsymbol{x}_0$ in the background flow, and due to the wave propagation, this position changes to $\boldsymbol{x} = \boldsymbol{x}_0 + \varepsilon \boldsymbol{\xi}$ in the total flow. The quantity $\boldsymbol{\xi}$ is called the *lagrangian displacement* and is sketched in Figure 1.1. The *lagrangian perturbation* is then defined by subtracting the total and background quantities following this given parcel

$$\varepsilon \delta q(\boldsymbol{x}_0, t) = q(\boldsymbol{x}_0 + \varepsilon \boldsymbol{\xi}, t) - q_0(\boldsymbol{x}_0, t) + \mathcal{O}(\varepsilon^2), \tag{1.3}$$

for any physical quantity $q \in \{\rho, \boldsymbol{v}, p\}$. We can rewrite (1.3) as

$$q(\boldsymbol{x}_0 + \varepsilon \boldsymbol{\xi}, t) = q_0(\boldsymbol{x}_0, t) + \varepsilon \delta q(\boldsymbol{x}, t) + \mathcal{O}(\varepsilon^2), \tag{1.4}$$

to mimic the structure of the perturbation expansion (1.2).

Remark 1.3.1: We have chosen to follow the convention of [Chr04] where q' denotes the eulerian perturbation of q and δq denotes its lagrangian perturbation. However there is no standard notations for those perturbations and other conventions are sometimes used in the literature.

[2] A *fluid parcel* is a very small amount of fluid, from a physical point of view, it is an averaged quantity at the mesoscopic scale over the actual fluid particles.

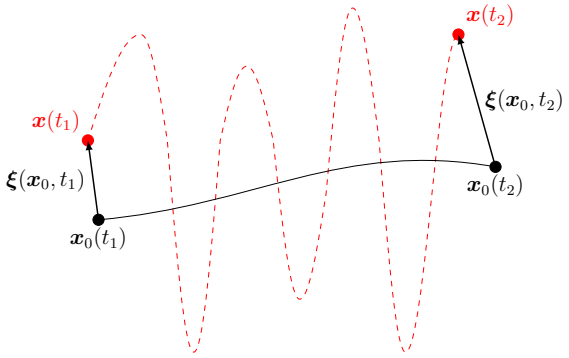


Figure 1.1: Sketch of the lagrangian displacement $\boldsymbol{\xi}$. The trajectory of the parcel in the background flow is depicted in solid black and its trajectory in the total flow is drawn in dashed red.

Mixed Eulerian-Lagrangian representation We assume that there is no perturbation at the initial time $t = 0$, leading to

$$\boldsymbol{x}(0) = \boldsymbol{x}_0(0).$$

With this assumption, we can identify a given parcel by its position in the background flow $\boldsymbol{x}_0$. Owing to its lagrangian nature, the displacement $\boldsymbol{\xi}$ is therefore a function of the time t and of the position of the parcel in background flow $\boldsymbol{x}_0$.
It is also possible to express the displacement as a function of the time t and the position $\boldsymbol{x}$ leading to the so-called *mixed eulerian-lagrangian representation*. In this framework, lagrangian perturbations are expressed in terms of eulerian coordinates. Using the chain rule, we have

$$\frac{\partial \boldsymbol{\xi}(\boldsymbol{x}_0, t)}{\partial t} = \frac{\partial \boldsymbol{\xi}(\boldsymbol{x}, t)}{\partial t} + \nabla_{\boldsymbol{v_0}} \boldsymbol{\xi}(\boldsymbol{x}, t) =: \mathrm{D}_t \boldsymbol{\xi}, \tag{1.5}$$

where $\boldsymbol{\xi}(\boldsymbol{x}_0, t)$ is the displacement expressed in lagrangian coordinates and $\boldsymbol{\xi}(\boldsymbol{x}, t)$ is the displacement expressed in eulerian coordinates. Notice that the notation $\boldsymbol{\xi}$ is slightly overloaded, but unless explicitly stated otherwise, we will always use the mixed eulerian-lagrangian representation in the remaining of this work.

Relating the displacement and the velocity For a given parcel, we denote by $\boldsymbol{x}_0$ its position in the background flow and by $\boldsymbol{x} = \boldsymbol{x}_0 + \varepsilon \boldsymbol{\xi}(\boldsymbol{x}_0, t)$ its position in the total flow. Notice that the displacement $\boldsymbol{\xi}(\boldsymbol{x}_0, t)$ is expressed in lagrangian coordinates for now. The velocity $\boldsymbol{v}$ of a parcel is given by

$$\boldsymbol{v} = \frac{\partial \boldsymbol{x}}{\partial t} = \frac{\partial \boldsymbol{x}_0}{\partial t} + \varepsilon \frac{\partial \boldsymbol{\xi}(\boldsymbol{x}_0, t)}{\partial t},$$

and using (1.4) we also have

$$\boldsymbol{v} = \boldsymbol{v_0} + \varepsilon \boldsymbol{\delta v} + \mathcal{O}(\varepsilon^2).$$

Identifying the powers of ε leads to

$$\boldsymbol{\delta v} = \frac{\partial \boldsymbol{\xi}(\boldsymbol{x}_0, t)}{\partial t} + \mathcal{O}(\varepsilon) = \mathrm{D}_t \boldsymbol{\xi}(\boldsymbol{x}, t) + \mathcal{O}(\varepsilon), \tag{1.6}$$

where we used (1.5) to obtain the last equality linking the lagrangian and mixed representations.

Assuming that $\boldsymbol{v}$ is regular enough, we may write the following Taylor expansion around $\boldsymbol{x}_0$

$$\boldsymbol{v}(\boldsymbol{x}) = \boldsymbol{v}(\boldsymbol{x}_0 + \varepsilon\boldsymbol{\xi}) = \boldsymbol{v}(\boldsymbol{x}_0) + \varepsilon\nabla_{\boldsymbol{\xi}}\boldsymbol{v}(\boldsymbol{x}_0) + \mathcal{O}(\varepsilon^2),$$

and subtracting $\boldsymbol{v_0}(\boldsymbol{x}_0)$ leads to

$$\boldsymbol{v}(\boldsymbol{x}_0 + \varepsilon\boldsymbol{\xi}) - \boldsymbol{v_0}(\boldsymbol{x}_0) = \boldsymbol{v}(\boldsymbol{x}_0) - \boldsymbol{v_0}(\boldsymbol{x}_0) + \varepsilon\nabla_{\boldsymbol{\xi}}\boldsymbol{v}(\boldsymbol{x}_0) + \mathcal{O}(\varepsilon^2).$$

Recalling the definitions of the lagrangian (1.3) and eulerian (1.2) perturbations, we have

$$\varepsilon\boldsymbol{\delta v} = \varepsilon\boldsymbol{v}' + \varepsilon\nabla_{\boldsymbol{\xi}}\boldsymbol{v}(\boldsymbol{x}_0) + \mathcal{O}(\varepsilon^2),$$

and using the definition of the eulerian perturbation (1.2) inside the directional derivative yields

$$\boldsymbol{\delta v} = \boldsymbol{v}' + \nabla_{\boldsymbol{\xi}}\left[\boldsymbol{v_0} + \varepsilon\boldsymbol{v}'\right] + \mathcal{O}(\varepsilon).$$

Then, collecting the zeroth-order terms leads to

$$\boldsymbol{v}' = \boldsymbol{\delta v} - \nabla_{\boldsymbol{\xi}}\boldsymbol{v_0} + \mathcal{O}(\varepsilon),$$

and using (1.6), we finally have

$$\boldsymbol{v}' = \mathrm{D}_t\boldsymbol{\xi} - \nabla_{\boldsymbol{\xi}}\boldsymbol{v_0} + \mathcal{O}(\varepsilon). \tag{1.7}$$

Derivation of Galbrun's equation We will now derive Galbrun's equation from LEE, by adapting the construction from [God97]. To make the notations lighter, the $\mathcal{O}(\varepsilon)$ terms will be omitted from now on, however one must keep in mind that the following equations are only exact to first-order. Using (1.7), a direct computation shows that

$$\mathrm{div}\left(\boldsymbol{v}'\right) = \mathrm{D}_t\left[\mathrm{div}\left(\boldsymbol{\xi}\right)\right] - \nabla_{\boldsymbol{\xi}}\left[\mathrm{div}\left(\boldsymbol{v_0}\right)\right], \tag{1.8}$$

$$\nabla_{\boldsymbol{v}'}T = \mathrm{D}_t\left[\nabla_{\boldsymbol{\xi}}T\right] - \nabla_{\boldsymbol{\xi}}[\mathrm{D}_tT], \tag{1.9}$$

where T is an arbitrary function of space and time with enough regularity for (1.9) to make sense.
Using (1.7), (1.8) and (1.9) into (LEE-$\boldsymbol{v}'$) leads to

$$\rho_0\mathrm{D}_t^2\boldsymbol{\xi} - \rho_0\nabla_{\boldsymbol{\xi}}[\mathrm{D}_t\boldsymbol{v_0}] + \frac{\rho'}{\rho_0}\nabla p_0 + \nabla p' = \boldsymbol{s}. \tag{1.10}$$

In the same time, using (1.7), (1.8), and (1.9) into (LEE-ρ') leads to

$$\mathrm{D}_t\left[\rho' + \mathrm{div}\left(\rho_0\boldsymbol{\xi}\right)\right] + \mathrm{div}\left(\boldsymbol{v_0}\right)\left[\rho' + \mathrm{div}\left(\rho_0\boldsymbol{\xi}\right)\right] = 0.$$

Together with the identity

$$\mathrm{D}_t\left[\frac{\rho_0\left(\rho' + \mathrm{div}\left(\rho_0\boldsymbol{\xi}\right)\right)}{\rho_0}\right] = \rho_0\mathrm{D}_t\left[\frac{\rho' + \mathrm{div}\left(\rho_0\boldsymbol{\xi}\right)}{\rho_0}\right] + \frac{\rho' + \mathrm{div}\left(\rho_0\boldsymbol{\xi}\right)}{\rho_0}\mathrm{D}_t\rho_0,$$

we obtain

$$\rho_0\mathrm{D}_t\left[\frac{\rho' + \mathrm{div}\left(\rho_0\boldsymbol{\xi}\right)}{\rho_0}\right] + \left(\frac{\mathrm{D}_t\rho_0}{\rho_0} + \mathrm{div}\left(\boldsymbol{v_0}\right)\right)\left(\rho' + \mathrm{div}\left(\rho_0\boldsymbol{\xi}\right)\right) = 0.$$

Then, the background continuity equation (BG-ρ_0) finally yields

$$\rho_0\mathrm{D}_t\left[\frac{\rho' + \mathrm{div}\left(\rho_0\boldsymbol{\xi}\right)}{\rho_0}\right] = 0.$$

At this point, we need to make the following assumption:

Assumption 4 (No-resonance):
The following implication is true

$$\rho_0 \mathrm{D}_t \left[\frac{\rho' + \operatorname{div}(\rho_0 \boldsymbol{\xi})}{\rho_0} \right] = 0 \Longrightarrow \rho' + \operatorname{div}(\rho_0 \boldsymbol{\xi}) = 0.$$

Assumption 4 actually translates to constraints on the geometry of the background flow that will be discussed in Section 1.8 and in Chapter 2.
In Assumption 3, we assumed that the wave propagation process is adiabatic, we therefore have

$$p' + \nabla_{\boldsymbol{\xi}} p_0 = c_0^2 \left(\rho' + \nabla_{\boldsymbol{\xi}} \rho_0 \right). \tag{1.11}$$

Together with Assumption 4, we have

$$p' = -\rho_0 c_0^2 \operatorname{div}(\boldsymbol{\xi}) - \nabla_{\boldsymbol{\xi}} p_0. \tag{1.12}$$

Using Assumption 4 and (1.12) into (1.10) leads to

$$\rho_0 \mathrm{D}_t^2 \boldsymbol{\xi} - \rho_0 \nabla_{\boldsymbol{\xi}} \left[\mathrm{D}_t \boldsymbol{v_0} \right] - \frac{\operatorname{div}(\rho_0 \boldsymbol{\xi})}{\rho_0} \nabla p_0 - \nabla \left[\rho_0 c_0^2 \operatorname{div}(\boldsymbol{\xi}) + \nabla_{\boldsymbol{\xi}} p_0 \right] = \boldsymbol{s},$$

and using the background continuity equation (BG-ρ_0) we have

$$\rho_0 \mathrm{D}_t^2 \boldsymbol{\xi} - \nabla \left[\rho_0 c_0^2 \operatorname{div}(\boldsymbol{\xi}) \right] + \operatorname{div}(\boldsymbol{\xi}) \nabla p_0 - \nabla \left[\nabla_{\boldsymbol{\xi}} p_0 \right] + \frac{\nabla_{\boldsymbol{\xi}} \rho_0}{\rho_0} \nabla p_0 - \rho_0 \nabla_{\boldsymbol{\xi}} \left[\mathrm{D}_t \boldsymbol{v_0} \right] = \boldsymbol{s}. \tag{1.13}$$

We notice that

$$-\rho_0 \nabla_{\boldsymbol{\xi}} \mathrm{D}_t \boldsymbol{v_0} = -\rho_0 \nabla_{\boldsymbol{\xi}} [\nabla_{\boldsymbol{v_0}} \boldsymbol{v_0}] = -\nabla_{\boldsymbol{\xi}} [\rho_0 \nabla_{\boldsymbol{v_0}} \boldsymbol{v_0}] + (\nabla_{\boldsymbol{\xi}} \rho_0) \nabla_{\boldsymbol{v_0}} \boldsymbol{v_0}, \tag{1.14}$$

and if we combine the terms in $\nabla_{\boldsymbol{\xi}} \rho_0$, we obtain

$$\frac{\nabla_{\boldsymbol{\xi}} \rho_0}{\rho_0} \nabla p_0 + (\nabla_{\boldsymbol{\xi}} \rho_0) \nabla_{\boldsymbol{v_0}} \boldsymbol{v_0} = \nabla_{\boldsymbol{\xi}} \rho_0 \left(\frac{1}{\rho_0} \nabla p_0 + \nabla_{\boldsymbol{v_0}} \boldsymbol{v_0} \right) = (\nabla_{\boldsymbol{\xi}} \rho_0) \boldsymbol{g_0}. \tag{1.15}$$

We also notice that the second term of (1.14) gives

$$-\nabla_{\boldsymbol{\xi}} [\nabla_{\boldsymbol{v_0}} \boldsymbol{v_0}] = -\nabla_{\boldsymbol{\xi}} [\rho_0 \boldsymbol{g_0} - \nabla p_0], \tag{1.16}$$

and combining the first term of (1.16) with (1.15), we have

$$(\nabla_{\boldsymbol{\xi}} \rho_0) \boldsymbol{g_0} - \nabla_{\boldsymbol{\xi}} [\rho_0 \boldsymbol{g_0}] = -\rho_0 \nabla_{\boldsymbol{\xi}} \boldsymbol{g_0}. \tag{1.17}$$

Introducing (1.14), (1.15), (1.16) and (1.17) into (1.13) we obtain a first form of Galbrun's equation

$$\rho_0 \mathrm{D}_t^2 \boldsymbol{\xi} - \nabla \left[\rho_0 c_0^2 \operatorname{div}(\boldsymbol{\xi}) \right] + \operatorname{div}(\boldsymbol{\xi}) \nabla p_0 - \nabla \left[\nabla_{\boldsymbol{\xi}} p_0 \right] + \nabla_{\boldsymbol{\xi}} \nabla p_0 - \rho_0 \nabla_{\boldsymbol{\xi}} \boldsymbol{g_0} = \boldsymbol{s}. \tag{1.18}$$

This form of Galbrun's equation is obtained in [LO67, Eq. (22), (25) & (28)] and in [God97, Eq. (29)]. If we assume that the coordinates are cartesian, we can simplify (1.18). Indeed we have

$$-\nabla \left[\nabla_{\boldsymbol{\xi}} p_0 \right] + \nabla_{\boldsymbol{\xi}} \nabla p_0 = -(\nabla \boldsymbol{\xi})^T \nabla p_0 - \nabla_{\boldsymbol{\xi}} \nabla p_0 + \nabla_{\boldsymbol{\xi}} \nabla p_0 = -(\nabla \boldsymbol{\xi})^T \nabla p_0,$$

leading to the usual form of Galbrun's equation.

Proposition 1.3.1:
The *lagrangian displacement* $\boldsymbol{\xi}$ satisfies the following equation

$$\rho_0 \mathrm{D}_t^2 \boldsymbol{\xi} - \nabla \left[\rho_0 c_0^2 \operatorname{div}(\boldsymbol{\xi}) \right] + \operatorname{div}(\boldsymbol{\xi}) \nabla p_0 - (\nabla \boldsymbol{\xi})^T \nabla p_0 - \rho_0 \nabla_{\boldsymbol{\xi}} \boldsymbol{g_0} = \boldsymbol{s}, \tag{Gal}$$

called *Galbrun's equation.*

Remark 1.3.2: If we assume that the gravity is irrotational *i.e.* $\boldsymbol{g_0} = \nabla\varphi_0$, we have

$$\rho_0\nabla_{\boldsymbol{\xi}}\boldsymbol{g_0} = \rho_0\nabla_{\boldsymbol{\xi}}\nabla\varphi_0.$$

It is interesting to notice that this term accounts for the lagrangian perturbation of the gravity, even if there is no eulerian perturbation of gravity under *Cowling's approximation*. Indeed, if we fix a given parcel, as its position has changed because of the wave propagation, it does not "see" the same gravitational field, as depicted on FIGURE 1.2.

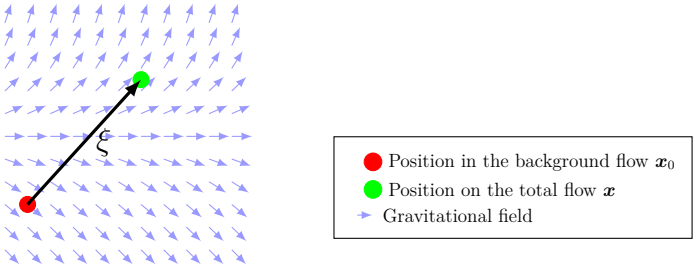


Figure 1.2: Lagrangian perturbation of the gravity

Other derivations of Galbrun's equation In the previous paragraph, we derived Galbrun's equation from LEE, as it was formerly done in [God97, HB21]. However it is also possible to derive Galbrun's equation using a *continuum mechanics* point of view where the Euler's equations (Eul-ρ)–(Eul-$\boldsymbol{v}$)–(Eul-p) are written in a lagrangian frame. It is then possible to linearize those equation using lagrangian perturbations and to obtain Galbrun's equation by going back to eulerian coordinates. This construction can be found in [Leg03, GM14] and in [Poi85] where second-order perturbations are also considered. Finally we would like to mention [LO67], where Galbrun's equation is derived by directly computing lagrangian perturbations in an eulerian frame, leading to a construction which is similar to the two previous ones. In [**?**] a construction similar to the one in [LO67] is used for solar cases when $\boldsymbol{v_0} = \boldsymbol{0}$.

On the necessity of the no-resonance assumption To derive Galbrun's equation we had to use ASSUMPTION 4 to relate ρ' and $\boldsymbol{\xi}$. This is one of the weaknesses of this construction: indeed when ASSUMPTION 4 does not hold, we cannot obtain Galbrun's equation from LEE. As there are other constructions of Galbrun's equation, we may wonder if ASSUMPTION 4 is really required. In [LO67, Eq. (18) & (19)] the relationship between ρ' and $\boldsymbol{\xi}$

$$\rho' = -\mathrm{div}\,(\rho_0\boldsymbol{\xi}) \tag{1.19}$$

is given but not proven. For the constructions based on lagrangian linearization, it seems that ASSUMPTION 4 is not required. In [GM14, Sec. 2.3.5], a weaker version of ASSUMPTION 4 is used:

$$\mathrm{D}_t\left[\frac{\rho'}{\rho_0}\right] = 0 \implies \frac{\rho'}{\rho_0} = \text{constant}.$$

In [Poi85, Leg03] however no such assumptions are made. It is actually possible to obtain an algebraic relationship that translates mass conservation in lagrangian coordinates which leads to (1.19) after linearization, for more details we refer to [Sal05, Sec. II.4 & II.5]. Finally, we would like to point out that even if ASSUMPTION 4 is not required to derive Galbrun's equation, similar problems will occur as $\boldsymbol{\xi}$ must satisfy the *vectorial transport equation* (1.7).

1.4 Equivalence between those models

We will now investigate the relationship between Galbrun's equation and LEE. The following theorem holds.

Theorem 1 : *Relationship between Galbrun's equation and LEE*

The solution $\boldsymbol{\xi}$ of Galbrun's equation (Gal) and $(\rho', \boldsymbol{v}', p')$ of (LEE-ρ')–(LEE-$\boldsymbol{v}'$)–(LEE-p') are linked by the following system

$$\begin{aligned} \rho' &= -\mathrm{div}\,(\rho_0 \boldsymbol{\xi})\,, && (\rho' \leftrightarrow \boldsymbol{\xi}) \\ \boldsymbol{v}' &= \mathrm{D}_t \boldsymbol{\xi} - \boldsymbol{\xi} \cdot \nabla \boldsymbol{v_0}, && (\boldsymbol{v}' \leftrightarrow \boldsymbol{\xi}) \\ p' &= -\rho_0 c_0^2 \mathrm{div}\,(\boldsymbol{\xi}) - \boldsymbol{\xi} \cdot \nabla p_0. && (p' \leftrightarrow \boldsymbol{\xi}) \end{aligned}$$

Proof : The system $(\rho' \leftrightarrow \boldsymbol{\xi})$–$(\boldsymbol{v}' \leftrightarrow \boldsymbol{\xi})$–$(p' \leftrightarrow \boldsymbol{\xi})$ has already be obtained, see (1.7) for $(\boldsymbol{v}' \leftrightarrow \boldsymbol{\xi})$, Assumption 4 or (1.19) for $(\rho' \leftrightarrow \boldsymbol{\xi})$ and (1.12) for $(p' \leftrightarrow \boldsymbol{\xi})$.

To show that Galbrun's equation and LEE are actually equivalent, we need to prove that

- When the solution $(\rho', \boldsymbol{v}', p')$ of (LEE-ρ')–(LEE-$\boldsymbol{v}'$)–(LEE-p') is known, then the solution $\boldsymbol{\xi}$ of the system $(\rho' \leftrightarrow \boldsymbol{\xi})$–$(\boldsymbol{v}' \leftrightarrow \boldsymbol{\xi})$–$(p' \leftrightarrow \boldsymbol{\xi})$ is also a solution of Galbrun's equation (Gal). This is what we have done when we derived Galbrun's equation in the previous section. Notice that we have to solve an overdetermined system of PDEs to construct $\boldsymbol{\xi}$, and that there may be no solution to $(\rho' \leftrightarrow \boldsymbol{\xi})$–$(\boldsymbol{v}' \leftrightarrow \boldsymbol{\xi})$–$(p' \leftrightarrow \boldsymbol{\xi})$.
- When the solution $\boldsymbol{\xi}$ of Galbrun's equation (Gal) is known, then $(\rho', \boldsymbol{v}', p')$ defined by $(\rho' \leftrightarrow \boldsymbol{\xi})$–$(\boldsymbol{v}' \leftrightarrow \boldsymbol{\xi})$–$(p' \leftrightarrow \boldsymbol{\xi})$is a solution of (LEE-$\rho'$)–(LEE-$\boldsymbol{v}'$)–(LEE-$p'$). This can be done by a lengthy direct computation detailed in Appendix 1.A.

∎

Remark 1.4.1: Notice that because of the adiabaticity condition (1.11), $(p' \leftrightarrow \boldsymbol{\xi})$ and $(\rho' \leftrightarrow \boldsymbol{\xi})$ are equivalent.

The issues with Theorem 1 At first glance, it may seem that Theorem 1 gives an equivalence result between Galbrun's equation and LEE, however some issues remain to be dealt with:

1. As the system $(\rho' \leftrightarrow \boldsymbol{\xi})$–$(\boldsymbol{v}' \leftrightarrow \boldsymbol{\xi})$–$(p' \leftrightarrow \boldsymbol{\xi})$ is overdetermined, it is not clear wether or not it can be used to define $\boldsymbol{\xi}$ when $(\rho', \boldsymbol{v}', p')$ are known, as there may be no solution. For $\boldsymbol{\xi}$ to be a solution of Galbrun's equation (Gal), it must satisfy the three equations of $(\rho' \leftrightarrow \boldsymbol{\xi})$–$(\boldsymbol{v}' \leftrightarrow \boldsymbol{\xi})$–$(p' \leftrightarrow \boldsymbol{\xi})$.
2. To ensure the solvability of $(\boldsymbol{v}' \leftrightarrow \boldsymbol{\xi})$ an assumption similar to Assumption 4 will be required, this can be critical in the time-harmonic domain.
3. Compatible boundary conditions must be derived to use this result in a practical settings.

On the relationship between Galbrun's equation and LEE There is an interesting analogy that illustrates the relationship between Galbrun's equation and LEE: Galbrun's equation can be seen as the "antiderivative" of LEE. Indeed, when the solution of Galbrun's equation is known, getting a solution of LEE only amounts to differentiating, which is easy and always possible. On the other hand, when the solution of LEE is known, an overdetermined system must be solved to obtain a solution of Galbrun's equation, which is harder and not always possible.

We can therefore conclude that when Galbrun's equation is well-posed, then so does LEE, but the reciprocal may not be true.

1.5 Boundary conditions

As discussed in the previous section, the choice of boundary conditions is a difficult problem.

Usual boundary conditions in aeroacoustics In TABLE 1.1, we recall the usual boundary conditions for the LEE and for Galbrun's equation on the *sliding boundary*

$$\Gamma_0 := \left\{ \boldsymbol{x} \in \partial\mathcal{O} \;\middle|\; \boldsymbol{v_0} \cdot \boldsymbol{n} = 0 \right\}$$

of the domain. From a physical point of view, the sliding boundary cannot be crossed by the background flow and is therefore used to models structures immersed in the flow. One of the advantages of Galbrun's equation over LEE is that the boundary conditions for wave-structure interactions are naturally expressed in terms of the displacement $\boldsymbol{\xi}$. For the derivation of such boundary conditions, we refer *e.g.* to [Mye80] and to [God97, Sec. 3]. For complex structure such as walls with acoustic treatment, the relevant boundary conditions are obtained by combining the terms in TABLE 1.1. Once again, we refer to [Mye80, God97] for more details.
In this section $\boldsymbol{n}$ always denotes the outward-oriented unit normal vector to the boundary of the domain.

Boundary condition	LEE	Galbrun
Dirichlet	$\boldsymbol{v}' \cdot \boldsymbol{n} = 0$	$\boldsymbol{\xi} \cdot \boldsymbol{n} = 0$
Neumann	$p' = 0$	$\operatorname{div}(\boldsymbol{\xi}) = 0$
BG condition	$\boldsymbol{v_0} \cdot \boldsymbol{n} = 0$	$\boldsymbol{v_0} \cdot \boldsymbol{n} = 0$ $\nabla p_0 \times \boldsymbol{n} = 0$

Table 1.1: Usual boundary conditions for aeroacoustics on the *sliding boundary* Γ_0

Other parts of the boundary where $\boldsymbol{v_0} \cdot \boldsymbol{n} \neq 0$ usually are *artificial boundaries* on which domain truncation techniques, such as *Perfectly Matched Layers* (PMLs) or *Absorbing Boundary Conditions* (ABCs). We would like to point out that the sign of $\boldsymbol{v_0} \cdot \boldsymbol{n}$ may play an important part in the treatment of those boundary conditions, indeed

- if $\boldsymbol{v_0} \cdot \boldsymbol{n} < 0$: the flow goes into the domain,
- if $\boldsymbol{v_0} \cdot \boldsymbol{n} > 0$: the flows goes outside the domain.

In particular, it was noted in [BBL03] that the presence of convection may render the PMLs unstable as it may create *backward propagating modes* that go back into the computational domain. Finally, we would like to point out that other boundary conditions can be used on those parts of the boundary. For example, there are some cases in [Gab03] where the full displacement $\boldsymbol{\xi}$ is prescribed on some parts of the boundary where $\boldsymbol{v_0} \cdot \boldsymbol{n} < 0$.

Compatibility of the boundary conditions If we are interested in using THEOREM 1 to reconstruct the solution $\boldsymbol{\xi}$ of Galbrun's equation from the solution $(\rho', \boldsymbol{v}', p')$ of LEE, we need to ensure the compatibility of the boundary conditions. Indeed, as noted in [HB21], Galbrun's equation and LEE are equivalent in time-domain if the system $(\rho' \leftrightarrow \boldsymbol{\xi})$–$(\boldsymbol{v}' \leftrightarrow \boldsymbol{\xi})$–$(p' \leftrightarrow \boldsymbol{\xi})$ is satisfied at the initial time $t = 0$ and on the boundaries of the domain $\mathcal{O}$.

Assuming that we want to reconstruct the solution of Galbrun's equation (Gal) with the Neumann boundary condition $\operatorname{div}(\boldsymbol{\xi}) = 0$ on Γ_0 from the solution of LEE (LEE-ρ')–(LEE-$\boldsymbol{v}'$)–(LEE-p'). Plugging the Neumann boundary condition into ($p' \leftrightarrow \boldsymbol{\xi}$) leads to

$$\operatorname{div}(\boldsymbol{\xi}) = 0 \implies p' = -\boldsymbol{\xi} \cdot \nabla p_0,$$

the Neumann boundary condition of Galbrun's equation only leads to a Neumann boundary condition for LEE when $\nabla p_0|_{\Gamma_0} = \mathbf{0}$, *i.e.* when the background pressure is uniform in a neighborhood of Γ_0. Notice that once again the other way around is easier: if we want to reconstruct a solution of LEE (LEE-ρ')–(LEE-$\boldsymbol{v}'$)–(LEE-p') with Neumann boundary conditions from a solution of Galbrun's equation (Gal), we obtain the following Robin-like boundary condition

$$p' = 0 \implies \rho_0 c_0^2 \operatorname{div}(\boldsymbol{\xi}) + \boldsymbol{\xi} \cdot \nabla p_0 = 0,$$

by plugging the Neumann boundary condition for LEE into ($p' \leftrightarrow \boldsymbol{\xi}$).

1.6 Time-harmonic solutions

In this section, we describe the time-harmonic conventions that we use to study Galbrun's equation (Gal) and LEE (LEE-ρ')–(LEE-$\boldsymbol{v}'$)–(LEE-p') in the frequency domain.

The time-domain formulation is natural as it is easy to understand the time dependance of the underlying physical phenomena. As the time-harmonic formulation leads to studying the wave propagation problem at fixed frequencies, things become harder to visualise. However this formulation can be very useful despite being less natural.

As this work takes place in a project whose long-term goal is to solve realistic *inverse problems* to image the solar interior, it is natural to consider time-harmonic wave propagation. Indeed, inversion problems are easier to solve in the frequency domain as different frequencies provide different information on the propagation medium: low frequencies are useful to reconstruct large structure, whereas high frequencies can be used to reconstruct finer details. For a comprehensive introduction to inverse problems, we refer to [Fau17]. We would also like to mention that a recent work has been done to adapt the methodology of large-scale inverse problems to time-domain, see [Jac21].

Convention for time-harmonic solutions To study the wave propagation in the frequency domain, we use the following *ansatz* for the solution of (Gal)

$$\boldsymbol{\xi}(\boldsymbol{x}, t) \longleftarrow \mathfrak{Re}\left[\boldsymbol{\xi}(\boldsymbol{x}) e^{-i\omega t}\right],$$

where i is the imaginary unit $i^2 = -1$ and $\omega > 0$ is the *angular frequency*.
Similarly, for the solutions of (LEE-ρ')–(LEE-$\boldsymbol{v}'$)–(LEE-p') we have

$$\begin{aligned}
\rho'(\boldsymbol{x}, t) &\longleftarrow \mathfrak{Re}\left[\rho'(\boldsymbol{x}) e^{-i\omega t}\right], \\
\boldsymbol{v}'(\boldsymbol{x}, t) &\longleftarrow \mathfrak{Re}\left[\boldsymbol{v}'(\boldsymbol{x}) e^{-i\omega t}\right], \\
p'(\boldsymbol{x}, t) &\longleftarrow \mathfrak{Re}\left[p'(\boldsymbol{x}) e^{-i\omega t}\right].
\end{aligned}$$

In the frequency domain, partial derivatives with respect to time are transformed to multiplications, indeed with $e^{-i\omega t}$ convention, we have

$$\frac{\partial \boldsymbol{\xi}}{\partial t} \longleftarrow -i\omega \boldsymbol{\xi}(\boldsymbol{x}) e^{-i\omega t},$$

and the material derivative becomes

$$\mathrm{D}_t\boldsymbol{\xi} := \frac{\partial \boldsymbol{\xi}}{\partial t} + \nabla_{\boldsymbol{v_0}}\boldsymbol{\xi} \longleftarrow \left(-i\omega\boldsymbol{\xi}(\boldsymbol{x}) + \nabla_{\boldsymbol{v_0}}\boldsymbol{\xi}(\boldsymbol{x})\right) e^{-i\omega t}.$$

If we further assume that the source term is also harmonic, *i.e.*

$$\boldsymbol{s}(\boldsymbol{x},t) \longleftarrow \mathfrak{Re}\left[\boldsymbol{s}(\boldsymbol{x})e^{-i\omega t}\right],$$

then the exponential term $e^{-i\omega t}$ can be factored out of the equations, leading to a family of equations parametrized by the frequency that do not depend on time anymore. As there is one equation for each frequency, the problem can be solved independently for each frequency of interest.

Damping term Aeroacoustic waves propagating inside the Sun are actually damped. Modelling the damping inside the Sun is still an open problem, and in this work we will only consider a very simple damping term formally coming from the *limiting absorption principle*. We introduce a parameter $\sigma > 0$, and we replace the frequency by the following complex number

$$\omega \longleftarrow \omega + i\sigma. \tag{1.20}$$

For the standard Helmholtz equation, (1.20) models a damping term that allows to recover the causality that was lost by going from time to frequency domain. However its physical relevance for aeroacoustic wave propagation inside the Sun is not clear.

1.7 Numerical investigation of the equivalence in simple cases

In this section, we will investigate the equivalence result of Theorem 1. To do so, we solve both Galbrun's equation and LEE, and we then reconstruct the solutions using ($\boldsymbol{v}' \leftrightarrow \boldsymbol{\xi}$). We are therefore able to compare:

- $\boldsymbol{\xi}$ obtained from solving Galbrun's equation, and $\boldsymbol{\xi}$ reconstructed from the solution of LEE,
- $\boldsymbol{v}'$ computed from solving LEE, and $\boldsymbol{v}'$ reconstructed from the solution of Galbrun's equation.

We use the numerical methods from [CD18] in the `montjoie` solver[3], in particular we use the PQ-formulation for Galbrun's equation, and we will therefore only consider the simple case of a uniform background flow for which those numerical methods have been validated. For the transport equation ($\boldsymbol{v}' \leftrightarrow \boldsymbol{\xi}$), we use an upwind DG method similar to the one of [Pey13, Sec. 2 & 3].

We would like to point out that the process of reconstructing the solution of Galbrun's equation from the solution of LEE has a practical interest. Indeed, as noted in [CD18], using LEE seems to give better numerical results in solar-like cases than solving Galbrun's equation. As $\boldsymbol{\xi}$ is the quantity of interest in helioseismic inversion, this could provide a more reliable direct solver.

[3] `montjoie` is a versatile high-order finite-element solver, see http://montjoie.gforge.inria.fr/.

Settings We consider the following background flow :

$$\rho_0 = 2.5, \quad c_0 = 0.8, \quad p_0 = 1, \quad \boldsymbol{v_0} = \begin{bmatrix} v_0 \\ 0 \end{bmatrix}, \quad \omega = 2\pi \cdot 0.78, \quad \sigma = 0.1 \tag{1.21}$$

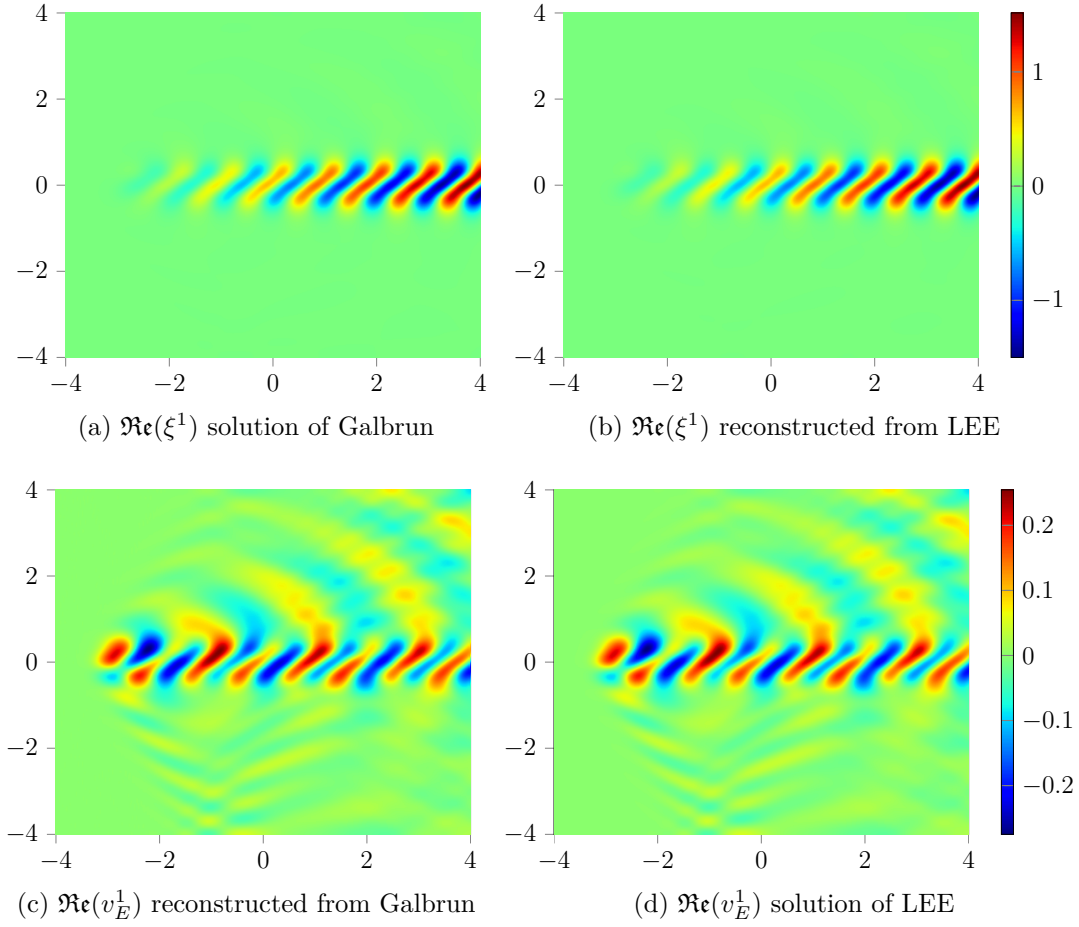
We choose a gaussian source centered at (x_c, y_c) of the form

$$\boldsymbol{s} = g(x,y)\boldsymbol{e}_x, \qquad \text{where} \qquad g(x,y) = \sqrt{\frac{\alpha}{\pi}} \exp\left[-\alpha\left((x-x_c)^2 + (y-y_c)^2\right)\right]$$

and the constant α is chosen so that g is equal to 10^{-6} when

$$(x-x_c)^2 + (y-y_c)^2 > 1$$

As it is shown in [CD18, Section 3] that the so-called PQ formulation of Galbrun's equation exhibits good convergence for the flow (1.21), we will use it compute a reference solution $\boldsymbol{\xi}_{\text{ref}}$ with $\mathbb{Q}_{12}$ elements on a regular mesh with $N = 51$ points in each direction.



(a) $\mathfrak{Re}(\xi^1)$ solution of Galbrun (b) $\mathfrak{Re}(\xi^1)$ reconstructed from LEE

(c) $\mathfrak{Re}(v_E^1)$ reconstructed from Galbrun (d) $\mathfrak{Re}(v_E^1)$ solution of LEE

Figure 1.3: Results for a uniform background flow – PMLs case – PMLs not displayed

Case 1: Domain surrounded by PMLs To avoid dealing with the compatibility of boundary conditions, we consider wave propagation in an infinite medium. From a practical point of view, we need to truncate the domain to a finite one to perform numerical simulations and we therefore surround the domain with *Perfectly Matched Layers* (PMLs). In general, one should be careful when using PMLs for convected wave propagation, but as we only consider simple background flows they provide reliable results. Indeed, as noted in [BBL03], PMLs can become unstable in the presence of a background flow.

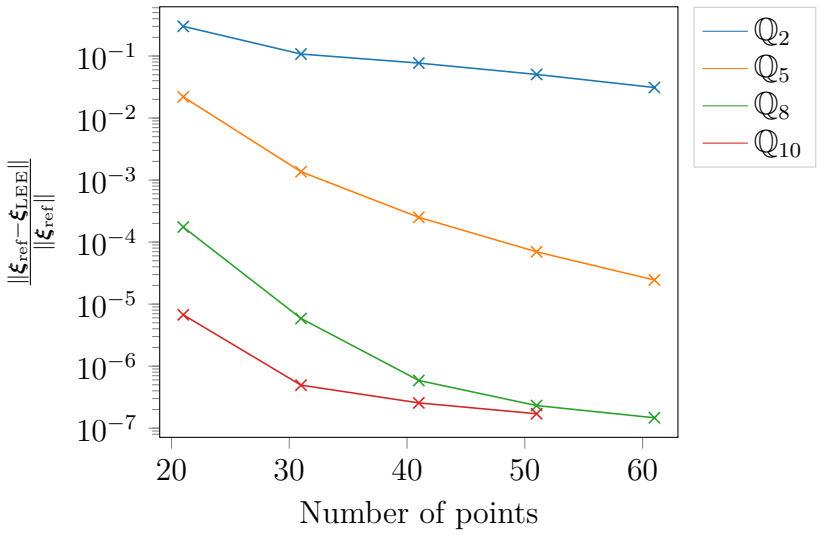
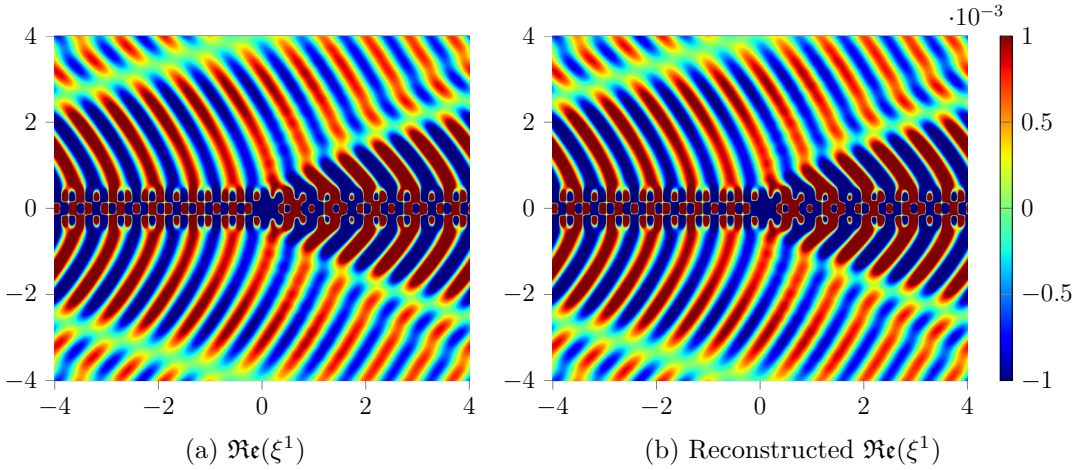


Figure 1.4: Convergence of the reconstructed displacement – PMLs case

On FIGURE 1.4, the relative error between the reconstructed field and the reference solution for different orders of approximation is depicted. The numerical results are depicted on FIGURE 1.3

We can see that solving LEE and reconstructing the solution of Galbrun's equation is inefficient, even if the results look promising from a visual inspection: indeed by using a numerical method based on $\mathbb{Q}_8$ elements for the reconstruction process we obtain the same level of accuracy as directly solving Galbrun's equation with $\mathbb{Q}_5$-based method, see [CD18, Sec. 3]. Furthermore the convergence rate of the method seems to be lost in the process.

Case 2: Periodic boundary conditions For this case, we still use PMLs on the horizontal boundaries, but we enforce a periodicity condition on the vertical boundaries. Once again, this allows us to avoid the problem of dealing with boundary conditions.
The numerical results are depicted in FIGURE 1.5 and the convergence curve of the reconstructed displacement is depicted in FIGURE 1.6. Once again, we can see that the reconstruction process of $\boldsymbol{\xi}$ from $\boldsymbol{v}'$ is inefficient as the error level seem to be locked at 10^{-5}, whereas 10^{-6} could be achieved in the previous case and there was no locking effect in the convergence of the numerical method in [CD18].



(a) $\mathfrak{Re}(\xi^1)$ (b) Reconstructed $\mathfrak{Re}(\xi^1)$

Figure 1.5: Results for a uniform flow – Periodic case

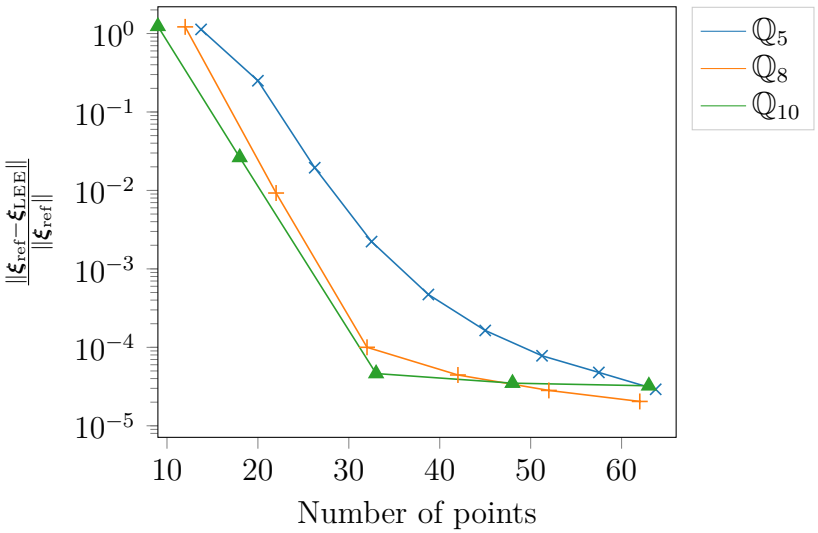


Figure 1.6: Convergence of the reconstructed displacement – Periodic case

Conclusion Even in the very simple case of a uniform background flow, the reconstruction of $\boldsymbol{\xi}$ from $\boldsymbol{v}'$ is an inefficient process, as it does not exhibit a good convergence behavior. For more complicated background flows, the well-posedness of the transport equation ($\boldsymbol{v}' \leftrightarrow \boldsymbol{\xi}$) will become an issue. So this process should probably not be used in practice, even if it may look appealing to reconstruct $\boldsymbol{\xi}$ from $\boldsymbol{v}'$ when the numerical solution of LEE seems nicer than the numerical solution for Galbrun's equation: indeed, we cannot guarantee the quality of $\boldsymbol{\xi}$ obtained by this process.

1.8 Review of well-posedness results

Now that we have derived Galbrun's equation, a natural question is whether or not this equation leads to well-posed problem. In this section, we therefore review different well-posedness results. We recall that problem is said to be *well-posed* in the sense of Hadamard if it has the following properties:

1. a solution exists,
2. the solution is unique,
3. the solution's behaviour changes continuously with the initial or boundary data.

Time-harmonic propagation for industrial applications The time-harmonic Galbrun's equation has been studied for applications in aeronautics. Although this led to many theoretical [BMM+12, Pey13, Leg03] and numerical [Jou10, Duc07, Gab03, Bér08] results, they usually come with restrictive assumptions on the geometry of background flow. Figure 1.7 describes a typical configuration : the background is assumed to be uniform outside of a bounded domain of interest $\mathcal{O}$. Inflow and outflow boundaries Γ_- and Γ_+ are therefore defined and the well-posedness is achieved when the flow «fills» the domain $\mathcal{O}$ in the following sense: each point of the domain $\mathcal{O}$ is reached

- following a streamline,
- starting from the inflow boundary Γ_-,
- in a bounded time.

Such flows are said to be $\mathcal{O}$-filling, and in particular configurations with recirculations (closed streamlines) or stopping points are forbidden. For a $\mathcal{O}$-filling background flow, well-posedness is achieved in $\boldsymbol{H}^1(\mathcal{O})$ using a regularization technique. Standard finite-element methods can therefore be used to solve Galbrun's equation when the background flow is $\mathcal{O}$-filling.

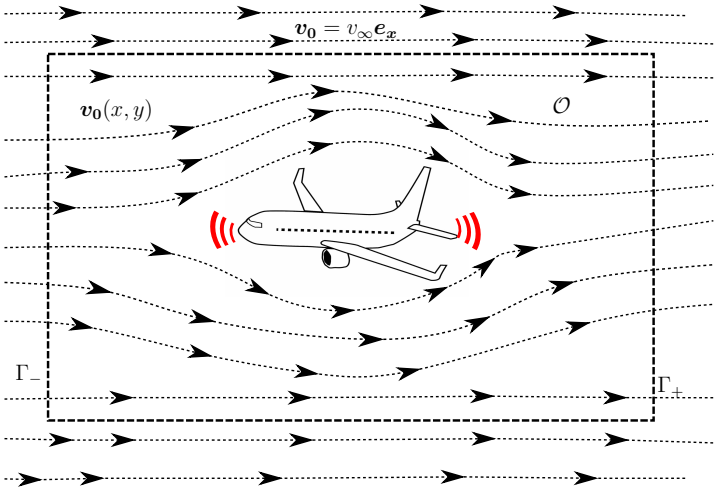


Figure 1.7: A typical configuration for applications in aeronautics

Time-domain propagation To the best of our knowledge, the first results for time-domain propagations were obtained in [Ber06] for uniform flows. In particular existence of strong solutions is obtained through standard Hille-Yosida theory. The case of general background flows was considered in [HB21] where the authors showed the well-posedness of LEE in the framework of Friedrichs' systems and then exploited the equivalence between Galbrun's equation and LEE to obtain a well-posedness result for Galbrun's equation. Those results hold without the geometric assumptions on the background flow, but exponential growth of the solution in time cannot be excluded.

Time-harmonic propagation for helioseismology As realistic solar background flows do not belong to the class of $\mathcal{O}$-filling flows, well-posedness in Hilbert space with low regularity has been considered in [HH21]. In this work, the authors have shown that Galbrun's equation with modified damping is weakly T-coercive, which is more or less equivalent to an inf-sup condition, in the Hilbert space

$$\boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O}) := \left\{ \boldsymbol{\xi} \in \boldsymbol{L}^2(\mathcal{O}) \;\middle|\; \operatorname{div}(\boldsymbol{\xi}) \in L^2(\mathcal{O}),\ \nabla_{\boldsymbol{v_0}} \boldsymbol{\xi} \in \boldsymbol{L}^2(\mathcal{O}) \right\}.$$

Those findings are consistent with those of [Cha19] for the case where $\boldsymbol{v_0} = \boldsymbol{0}$. We would like to point out that it is not possible to use standard finite-element methods to exploit those theoretical results. Indeed, for a function $\boldsymbol{\xi} \in \boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O})$, the continuity of the solution on the interface between two elements depends on the interface:

- if $\boldsymbol{v_0} \cdot \boldsymbol{n} = 0$: only the normal part of $\boldsymbol{\xi}$ is continuous (coming from the divergence),
- if $\boldsymbol{v_0} \cdot \boldsymbol{n} \neq 0$: $\boldsymbol{\xi}$ is fully continuous (coming from the directional derivative),

where $\boldsymbol{n}$ is a unitary normal vector to the interface. This leads us to consider (Hybridizable) Discontinuous Galerkin methods to weakly enforce those continuity requirements on the different interfaces of the mesh.

Long-term behaviour We would also like to mention that a lot of theoretical works have been done to study the long-term behaviour of the solution of Galbrun's equation to adress the question of the stability of stars. This is usually done by studying the spectrum of Galbrun's equation and details can be found *e.g.* in [LO67, DSC79, Bey02].

Conclusion

In this chapter, we have discussed the derivation of two aeroacoustic models, Galbrun's equation and LEE, by linearizing fluid dynamic equations around an equilibrium state.

Even if those two models are said to be equivalent, their unknowns are linked by an overdetermined system of equations. It is therefore not clear wether or not the solution of Galbrun's equation can be reconstructed from the solution of LEE. It is also important to point out that the equivalence can only be achieved for compatible boundary conditions that may not be well-defined. From a numerical point of view, we have shown that trying to use the "equivalence" leads to a larger error than directly solving Galbrun's equation, and to a loss of the convergence properties of the numerical methods.

We also review the existing theoretical work on Galbrun's equation. We have noticed that the use of a standard hilbertian settings, $\boldsymbol{H}^1(\mathcal{O})$, in time-harmonic domain requires strong assumptions on the background flow: closed streamlines are forbidden. As closed streamlines arise in helioseismology, we will study this case more precisely in the next chapter. We will then determine if we can use the $\boldsymbol{H}^1(\mathcal{O})$-settings despite the closed streamlines, or if we should resort to the low-regularity $\boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O})$-settings.

Appendix

1.A Proof of Theorem 1

In this appendix, we detailed the direct computations that show that a solution of LEE can be reconstructed from a solution of Galbrun's equation.
For ρ' : Plugging ($\rho' \leftrightarrow \boldsymbol{\xi}$) and ($\boldsymbol{v}' \leftrightarrow \boldsymbol{\xi}$) into (LEE-$\rho'$) leads to

$$
\begin{aligned}
&- \mathrm{D}_t \left[\mathrm{div}\left(\rho_0 \boldsymbol{\xi}\right)\right] - \mathrm{div}\left(\rho_0 \boldsymbol{\xi}\right) \mathrm{div}\left(\boldsymbol{v_0}\right) + \mathrm{div}\left(\rho_0 \mathrm{D}_t \boldsymbol{\xi} - \rho_0 \nabla_{\boldsymbol{\xi}} \boldsymbol{v_0}\right) \\
=&- \frac{\partial}{\partial t}\left[\mathrm{div}\left(\rho_0 \boldsymbol{\xi}\right)\right] - \nabla_{\boldsymbol{v_0}}\left[\mathrm{div}\left(\rho_0 \boldsymbol{\xi}\right)\right] - \mathrm{div}\left(\rho_0 \boldsymbol{\xi}\right) \mathrm{div}\left(\boldsymbol{v_0}\right) + \mathrm{div}\left(\rho_0 \frac{\partial \boldsymbol{\xi}}{\partial t} + \rho_0 \nabla_{\boldsymbol{v_0}} \boldsymbol{\xi} - \rho_0 \nabla_{\boldsymbol{\xi}} \boldsymbol{v_0}\right) \\
=&-\mathrm{div}\left(\cancel{\rho_0 \frac{\partial \boldsymbol{\xi}}{\partial t}}\right) - \nabla_{\boldsymbol{v_0}}\left[\mathrm{div}\left(\rho_0 \boldsymbol{\xi}\right)\right] - \mathrm{div}\left(\rho_0 \boldsymbol{\xi}\right) \mathrm{div}\left(\boldsymbol{v_0}\right) + \mathrm{div}\left(\cancel{\rho_0 \frac{\partial \boldsymbol{\xi}}{\partial t}} + \rho_0 \nabla_{\boldsymbol{v_0}} \boldsymbol{\xi} - \rho_0 \nabla_{\boldsymbol{\xi}} \boldsymbol{v_0}\right) \\
=&- \nabla_{\boldsymbol{v_0}}\left[\mathrm{div}\left(\rho_0 \boldsymbol{\xi}\right)\right] - \mathrm{div}\left(\rho_0 \boldsymbol{\xi}\right) \mathrm{div}\left(\boldsymbol{v_0}\right) + \mathrm{div}\left(\rho_0 \nabla_{\boldsymbol{v_0}} \boldsymbol{\xi} - \rho_0 \nabla_{\boldsymbol{\xi}} \boldsymbol{v_0}\right) \\
=&- \nabla_{\boldsymbol{v_0}}\left[\mathrm{div}\left(\rho_0 \boldsymbol{\xi}\right)\right] - \mathrm{div}\left(\rho_0 \boldsymbol{\xi}\right) \mathrm{div}\left(\boldsymbol{v_0}\right) + \rho_0 \left(\nabla_{\boldsymbol{v_0}}\left[\mathrm{div}\left(\boldsymbol{\xi}\right)\right] - \nabla_{\boldsymbol{\xi}}\left[\mathrm{div}\left(\boldsymbol{v_0}\right)\right]\right) \\
&+ \left(\nabla_{\boldsymbol{v_0}} \boldsymbol{\xi} - \nabla_{\boldsymbol{\xi}} \boldsymbol{v_0}\right) \cdot \nabla \rho_0 \\
=&- \nabla_{\boldsymbol{v_0}}\left[\mathrm{div}\left(\rho_0 \boldsymbol{\xi}\right)\right] - \mathrm{div}\left(\mathrm{div}\left(\boldsymbol{v_0}\right) \rho_0 \boldsymbol{\xi}\right) + \rho_0 \nabla_{\boldsymbol{v_0}}\left[\mathrm{div}\left(\boldsymbol{\xi}\right)\right] + \left(\nabla_{\boldsymbol{v_0}} \boldsymbol{\xi} - \nabla_{\boldsymbol{\xi}} \boldsymbol{v_0}\right) \cdot \nabla \rho_0 \\
=&- \mathrm{div}\left(\rho_0 \mathrm{div}\left(\boldsymbol{v_0}\right) \boldsymbol{\xi}\right) - \mathrm{div}\left(\boldsymbol{\xi}\right) \nabla_{\boldsymbol{v_0}} \rho_0 - \boldsymbol{v_0} \cdot \left(\nabla \boldsymbol{\xi}^T \nabla \rho_0 + \nabla_{\boldsymbol{\xi}} \nabla \rho_0\right) + \left(\nabla_{\boldsymbol{v_0}} \boldsymbol{\xi} - \nabla_{\boldsymbol{\xi}} \boldsymbol{v_0}\right) \cdot \nabla \rho_0 \\
=&- \left(\rho_0 \mathrm{div}\left(\boldsymbol{v_0}\right) + \nabla_{\boldsymbol{v_0}} \rho_0\right) \mathrm{div}\left(\boldsymbol{\xi}\right) - \nabla_{\boldsymbol{\xi}}\left[\rho_0 \mathrm{div}\left(\boldsymbol{v_0}\right)\right] - \boldsymbol{v_0} \cdot \left(\nabla \boldsymbol{\xi}^T \nabla \rho_0 + \nabla_{\boldsymbol{\xi}} \nabla \rho_0\right) \\
&+ \left(\nabla_{\boldsymbol{v_0}} \boldsymbol{\xi} - \nabla_{\boldsymbol{\xi}} \boldsymbol{v_0}\right) \cdot \nabla \rho_0 \\
=&- \nabla_{\boldsymbol{\xi}}\left[\rho_0 \mathrm{div}\left(\boldsymbol{v_0}\right)\right] - \boldsymbol{v_0} \cdot \left(\nabla \boldsymbol{\xi}^T \nabla \rho_0 + \nabla_{\boldsymbol{\xi}} \nabla \rho_0\right) + \left(\nabla_{\boldsymbol{v_0}} \boldsymbol{\xi} - \nabla_{\boldsymbol{\xi}} \boldsymbol{v_0}\right) \cdot \nabla \rho_0 \\
=&\nabla_{\boldsymbol{\xi}}\left[\nabla_{\boldsymbol{v_0}} \rho_0\right] - \boldsymbol{v_0} \cdot \left(\nabla \boldsymbol{\xi}^T \nabla \rho_0 + \nabla_{\boldsymbol{\xi}} \nabla \rho_0\right) + \left(\nabla_{\boldsymbol{v_0}} \boldsymbol{\xi} - \nabla_{\boldsymbol{\xi}} \boldsymbol{v_0}\right) \cdot \nabla \rho_0 \\
=&\nabla_{\boldsymbol{\xi}} \boldsymbol{v_0}^T \nabla \rho_0 - \boldsymbol{v_0} \cdot \left(\nabla \boldsymbol{\xi}\right)^T \nabla \rho_0 + \left(\nabla_{\boldsymbol{v_0}} \boldsymbol{\xi} - \nabla_{\boldsymbol{\xi}} \boldsymbol{v_0}\right) \cdot \nabla \rho_0 \\
=&\left(\boldsymbol{\xi} \cdot \left(\nabla \boldsymbol{v_0}\right)^T - \boldsymbol{v_0} \cdot \left(\nabla \boldsymbol{\xi}\right)^T\right) \nabla \rho_0 + \left(\nabla_{\boldsymbol{v_0}} \boldsymbol{\xi} - \nabla_{\boldsymbol{\xi}} \boldsymbol{v_0}\right) \cdot \nabla \rho_0 \\
=&0
\end{aligned}
$$

As

$$
\boldsymbol{v_0} \cdot \left(\nabla \boldsymbol{\xi}\right)^T \nabla \rho_0 = \sum_i v_0^i \left(\sum_j \frac{\partial \xi^j}{\partial x^i} \frac{\partial \rho_0}{\partial x^j}\right) = \sum_{i,j} v_0^i \frac{\partial \xi^j}{\partial x^i} \frac{\partial \rho_0}{\partial x^j} = \sum_j \left(\sum_i v_0^i \frac{\partial \xi^j}{\partial x^i}\right) \frac{\partial \rho_0}{\partial x^j} = \nabla_{\boldsymbol{v_0}} \boldsymbol{\xi} \cdot \nabla \rho_0
$$

For p' :

$$
\begin{aligned}
&\mathrm{D}_t\left[-\gamma p_0 \mathrm{div}\left(\boldsymbol{\xi}\right) - \nabla_{\boldsymbol{\xi}} p_0\right] + \left(\mathrm{D}_t\boldsymbol{\xi} - \nabla_{\boldsymbol{\xi}} v_0\right)\cdot\nabla p_0 + \gamma p_0 \mathrm{div}\left(\mathrm{D}_t\boldsymbol{\xi} - \nabla_{\boldsymbol{\xi}}\boldsymbol{v_0}\right)\\
&-\gamma\left(\gamma p_0 \mathrm{div}\left(\boldsymbol{\xi}\right) + \nabla_{\boldsymbol{\xi}} p_0\right)\mathrm{div}\left(\boldsymbol{v_0}\right)\\
&=\nabla_{\boldsymbol{v_0}}\left[-\gamma p_0 \mathrm{div}\left(\boldsymbol{\xi}\right)\right] - \boldsymbol{\xi}\cdot\nabla_{\boldsymbol{v_0}}\nabla p_0 - \nabla_{\boldsymbol{\xi}}\nabla_{\boldsymbol{v_0}} p_0 + \gamma p_0 \mathrm{div}\left(\nabla_{\boldsymbol{v_0}}\boldsymbol{\xi} - \nabla_{\boldsymbol{\xi}}\boldsymbol{v_0}\right)\\
&-\gamma\left(\gamma p_0 \mathrm{div}\left(\boldsymbol{\xi}\right) + \nabla_{\boldsymbol{\xi}} p_0\right)\mathrm{div}\left(\boldsymbol{v_0}\right)\\
&=\nabla_{\boldsymbol{v_0}}\left[-\gamma p_0 \mathrm{div}\left(\boldsymbol{\xi}\right)\right] - \boldsymbol{\xi}\cdot\nabla_{\boldsymbol{v_0}}\nabla p_0 - \nabla_{\boldsymbol{\xi}}\nabla_{\boldsymbol{v_0}} p_0 + \gamma p_0\left(\nabla_{\boldsymbol{v_0}}\left[\mathrm{div}\left(\boldsymbol{\xi}\right)\right] - \nabla_{\boldsymbol{\xi}}\left[\mathrm{div}\left(\boldsymbol{v_0}\right)\right]\right)\\
&-\gamma\left(\gamma p_0 \mathrm{div}\left(\boldsymbol{\xi}\right) + \nabla_{\boldsymbol{\xi}} p_0\right)\mathrm{div}\left(\boldsymbol{v_0}\right)\\
&=-\gamma \mathrm{div}\left(\boldsymbol{\xi}\right)\nabla_{\boldsymbol{v_0}} p_0 - \boldsymbol{\xi}\cdot\nabla_{\boldsymbol{v_0}}\nabla p_0 - \nabla_{\boldsymbol{\xi}}\nabla_{\boldsymbol{v_0}} p_0 - \gamma p_0\nabla_{\boldsymbol{\xi}}\left[\mathrm{div}\left(\boldsymbol{v_0}\right)\right] - \gamma\left(\gamma p_0 \mathrm{div}\left(\boldsymbol{\xi}\right) + \nabla_{\boldsymbol{\xi}} p_0\right)\mathrm{div}\left(\boldsymbol{v_0}\right)\\
&=-\gamma \mathrm{div}\left(\boldsymbol{\xi}\right)\left(\nabla_{\boldsymbol{v_0}} p_0 + \gamma p_0 \mathrm{div}\left(\boldsymbol{v_0}\right)\right) - \boldsymbol{\xi}\cdot\nabla_{\boldsymbol{v_0}}\nabla p_0 - \nabla_{\boldsymbol{\xi}}\nabla_{\boldsymbol{v_0}} p_0 - \gamma p_0\nabla_{\boldsymbol{\xi}}\left[\mathrm{div}\left(\boldsymbol{v_0}\right)\right] - \gamma\nabla_{\boldsymbol{\xi}} p_0 \mathrm{div}\left(\boldsymbol{v_0}\right)\\
&=-\boldsymbol{\xi}\cdot\nabla_{\boldsymbol{v_0}}\nabla p_0 - \nabla_{\boldsymbol{\xi}}\nabla_{\boldsymbol{v_0}} p_0 - \gamma p_0\nabla_{\boldsymbol{\xi}}\left[\mathrm{div}\left(\boldsymbol{v_0}\right)\right] - \gamma\nabla_{\boldsymbol{\xi}} p_0 \mathrm{div}\left(\boldsymbol{v_0}\right)\\
&=-\boldsymbol{\xi}\cdot\nabla_{\boldsymbol{v_0}}\nabla p_0 - \nabla_{\boldsymbol{\xi}}\nabla_{\boldsymbol{v_0}} p_0 - \nabla_{\boldsymbol{\xi}}\left[\gamma p_0 \mathrm{div}\left(\boldsymbol{v_0}\right)\right]\\
&=-\nabla_{\boldsymbol{\xi}}\left[\nabla_{\boldsymbol{v_0}} p_0 + \gamma p_0 \mathrm{div}\left(\boldsymbol{v_0}\right)\right]\\
&=0
\end{aligned}
$$

For $\boldsymbol{v}'$:

$$
\begin{aligned}
\boldsymbol{s} =&\rho_0 \mathrm{D}_t^2\boldsymbol{\xi} - \rho_0\mathrm{D}_t\nabla_{\boldsymbol{\xi}}\boldsymbol{v_0} - \rho_0\boldsymbol{\xi}\cdot\nabla_{\boldsymbol{v_0}}\nabla\boldsymbol{v_0} - \rho_0\nabla_{\boldsymbol{\xi}}\nabla_{\boldsymbol{v_0}}\boldsymbol{v_0} + \rho_0\mathrm{D}_t\nabla_{\boldsymbol{\xi}}\boldsymbol{v_0} + \mathrm{div}\left(\rho_0\boldsymbol{\xi}\right)\frac{\nabla p_0}{\rho_0}\\
&-\nabla\left[\rho_0 c_0^2 \mathrm{div}\left(\boldsymbol{\xi}\right) + \nabla_{\boldsymbol{\xi}} p_0\right]\\
=&\rho_0 \mathrm{D}_t^2\boldsymbol{\xi} - \rho_0\boldsymbol{\xi}\cdot\nabla_{\boldsymbol{v_0}}\nabla\boldsymbol{v_0} - \rho_0\nabla_{\boldsymbol{\xi}}\nabla_{\boldsymbol{v_0}}\boldsymbol{v_0} + \mathrm{div}\left(\boldsymbol{\xi}\right)\nabla p_0 + \nabla_{\boldsymbol{\xi}}\rho_0\frac{\nabla p_0}{\rho_0}\\
&-\nabla\left[\rho_0 c_0^2 \mathrm{div}\left(\boldsymbol{v_0}\right)\right] - \nabla\boldsymbol{\xi}^T\nabla p_0 - \nabla_{\boldsymbol{\xi}}\nabla p_0\\
=&\rho_0 \mathrm{D}_t^2\boldsymbol{\xi} - \nabla\left[\rho_0 c_0^2 \mathrm{div}\left(\boldsymbol{\xi}\right)\right] + \mathrm{div}\left(\boldsymbol{\xi}\right)\nabla p_0 - \nabla\boldsymbol{\xi}^T\nabla p_0 - \boldsymbol{\xi}\cdot\left(\rho_0\nabla_{\boldsymbol{v_0}}\nabla\boldsymbol{v_0} + \rho_0\nabla\nabla_{\boldsymbol{v_0}}\boldsymbol{v_0} + \nabla\nabla p_0\right)\\
&-\nabla_{\boldsymbol{\xi}}\rho_0\left(\nabla_{\boldsymbol{v_0}}\boldsymbol{v_0} - \boldsymbol{g_0}\right)\\
=&\rho_0 \mathrm{D}_t^2\boldsymbol{\xi} - \nabla\left[\rho_0 c_0^2 \mathrm{div}\left(\boldsymbol{\xi}\right)\right] + \mathrm{div}\left(\boldsymbol{\xi}\right)\nabla p_0 - \nabla\boldsymbol{\xi}^T\nabla p_0 - \nabla_{\boldsymbol{\xi}}\left[\rho_0\nabla_{\boldsymbol{v_0}}\boldsymbol{v_0} + \nabla p_0\right] + \nabla_{\boldsymbol{\xi}}\rho_0\boldsymbol{g_0}\\
=&\rho_0 \mathrm{D}_t^2\boldsymbol{\xi} - \nabla\left[\rho_0 c_0^2 \mathrm{div}\left(\boldsymbol{\xi}\right)\right] + \mathrm{div}\left(\boldsymbol{\xi}\right)\nabla p_0 - \nabla\boldsymbol{\xi}^T\nabla p_0 - \nabla_{\boldsymbol{\xi}}\left[\rho_0\boldsymbol{g_0}\right] + \nabla_{\boldsymbol{\xi}}\rho_0\boldsymbol{g_0}\\
=&\rho_0 \mathrm{D}_t^2\boldsymbol{\xi} - \nabla\left[\rho_0 c_0^2 \mathrm{div}\left(\boldsymbol{\xi}\right)\right] + \mathrm{div}\left(\boldsymbol{\xi}\right)\nabla p_0 - \nabla\boldsymbol{\xi}^T\nabla p_0 - \rho_0\nabla_{\boldsymbol{\xi}}\boldsymbol{g_0}\\
=&\boldsymbol{s}
\end{aligned}
$$

Bibliography

[BBJ02] Christophe Bogey, Christophe Bailly, and Daniel Juve. Computation of Flow Noise Using Source Terms in Linearized Euler's Equations. page 9, 2002.

[BBL03] Eliane Bécache, Anne-Sophie Bonnet-Ben Dhia, and Guillaume Legendre. Perfectly matched layers for the convected Helmholtz equation. Technical report, 2003.

[Ber06] Kamel Berriri. *Approche Analytique et Numérique Pour l'aéroacoustique En Régime Transitoire Par Le Modèle de Galbrun.* PhD thesis, 2006.

[Bér08] Hadrien Bériot. *Eléments Finis d'ordre Élévé Pour l'opérateur de Galbrun En Régime Harmonique*. Thesis, Compiègne, January 2008.

[Bey02] Horst R. Beyer. A Framework for Perturbations and Stability of Differentially Rotating Stars. *Proceedings: Mathematical, Physical and Engineering Sciences*, 458(2018):359–380, 2002.

[BMM+12] Anne-Sophie Bonnet-Ben Dhia, Jean-François Mercier, Florence Millot, Sébastien Pernet, and Emilie Peynaud. Time-Harmonic Acoustic Scattering in a Complex Flow: A Full Coupling Between Acoustics and Hydrodynamics. *Communications in Computational Physics*, 11(2):555–572, February 2012.

[CD18] Juliette Chabassier and Marc Duruflé. Solving time-harmonic Galbrun's equation with an arbitrary flow. Application to Helioseismology. *Rapport de recherche Inria*, 2018.

[Cha19] Théophile Chaumont-Frelet. Mixed finite element discretizations of acoustic Helmholtz problems with high wavenumbers. *Calcolo*, 2019.

[Chr04] Jorgen Christensen-Dalsgaard. *Lecture Notes on Stellar Oscillations*. 2004.

[DSC79] James Dyson, Bernard Schutz, and Sivaramakrishna Chandrasekhar. Perturbations and stability of rotating stars. I. Completeness of normal modes. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 368(1734):389–410, November 1979.

[Duc07] Eve-Marie Duclairoir. *Rayonnement Acoustique Dans Un Écoulement Cisaille : Une Méthode d'éléments Finis Pour La Simulation Du Régime Harmonique*. Thesis, Palaiseau, Ecole polytechnique, January 2007.

[Fau17] Florian Faucher. *Contributions to Seismic Full Waveform Inversion for Time Harmonic Wave Equations: Stability Estimates, Convergence Analysis, Numerical Experiments Involving Large Scale Optimization Algorithms*. Theses, Université de Pau et des Pays de l'Adour, November 2017.

[Gab03] Gwenael Gabard. *Méthodes Numériques et Modèles de Sources Aéroacoustiques Fondées Sur l'équation de Galbrun*. PhD thesis, 2003.

[GM14] Marcus Guettler and Steffen Marburg. The nonlinear inhomogeneous Galbrun-Equation: Derivation and possible Ways to solve numerically. 2014.

[God97] Oleg A. Godin. Reciprocity and energy theorems for waves in a compressible inhomogeneous moving fluid. *Wave Motion*, 25(2):143–167, March 1997.

[HB21] Hägg, Linus and Berggren, Martin. On the well-posedness of Galbrun's equation. *Journal des Mathématiques Pures et Appliquées*, 2021.

[HH21] Martin Halla and Thorsten Hohage. On the Well-posedness of the Damped Time-harmonic Galbrun Equation and the Equations of Stellar Oscillations. *SIAM Journal on Mathematical Analysis*, pages 4068–4095, January 2021.

[Jac21] Pierre Jacquet. *Inversion Par Forme d'ondes Complète En Domaine Temporel Utilisant Des Méthodes de Galerkine Discontinues Avancées.* These de doctorat, Pau, February 2021.

[Jou10] Lauris Joubert. *Approche Asymptotique Pour l'étude Mathématique et La Simulation Numérique de La Propagation Du Son En Présence d'un Écoulement Fortement Cisaillé*. Thesis, Palaiseau, Ecole polytechnique, January 2010.

[Leg03] Guillaume Legendre. *Rayonnement acoustique dans un fluide en écoulement : analyse mathématique et numérique de l'équation de Galbrun*. PhD thesis, ENSTA ParisTech, September 2003.

[Lio96] Pierre-Louis Lions. *Mathematical Topics in Fluid Mechanics: Volume 1: Incompressible Models*. Oxford Lecture Series in Mathematics and Its Applications. Oxford University Press, Oxford, New York, June 1996.

[LN52] Michael James Lighthill and Maxwell Herman Alexander Newman. On sound generated aerodynamically I. General theory. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 211(1107):564–587, March 1952.

[LO67] Donald Lynden-Bell and Jeremiah Paul Ostriker. On the Stability of Differentially Rotating Bodies. *Monthly Notices of the Royal Astronomical Society*, 136(3):293–310, July 1967.

[MMMP17] Jean-François Mercier, Colin Mietka, Florence Millot, and Vincent Pagneux. Acoustic propagation in a vortical homentropic flow. page 19, 2017.

[Mye80] Michael K. Myers. On the acoustic boundary condition in the presence of flow. *Journal of Sound and Vibration*, 71(3):429–434, August 1980.

[Pey13] Emilie Peynaud. *Rayonnement Sonore Dans Un Écoulement Subsonique Complexe En Régime Harmonique : Analyse et Simulation Numérique Du Couplage Entre Les Phénomènes Acoustiques et Hydrodynamiques*. Thesis, Toulouse, INSA, June 2013.

[Pie19] Allan D. Pierce. *Acoustics: An Introduction to Its Physical Principles and Applications*. Springer International Publishing, third edition, 2019.

[Poi85] Poirée, Bernard. Les équations de l'acoustique linéaire et non-linéaire dans un écoulement de fluide parfait. *Acustica*, 1985.

[Sal05] Jean Salençon. *Mécanique Des Milieux Continus - Tome 1 - Concepts Généraux*. Éditions d l'École Polytechnique, 2005.

Chapter 2

A resonant-like phenomena

Contents

Introduction

In this chapter, we are interested in solving Galbrun's equation with general background flow. In particular, special care will be devoted to *recirculating flows* with closed streamlines, as they are very common in helioseismology. They are however excluded from the well-posedness results of [BMM+12, Pey13]. In [Ben18, BJM18] special care was devoted to the study of Goldstein's equation, which is an alternative aeroacoustic model, and it was proven that a resonant-like phenomenon can occur for recirculating flows. In this chapter, we extend this approach to LEE and Galbrun's equation. We will show that it is not possible to solve the aeroacoustic models (either Galbrun's equation or LEE) at a given frequency ω if a resonant line exists. Those frequencies are not the eigenvalues of some generalized Laplace operator, as it is usually the case for resonant frequencies of time-harmonic wave equations. In particular, they can constitute an uncountable set depending on the background flow.

This resonant-like phenomenon also gives a mathematical interpretation to the *spurious modes* that can occur when using Lagrange finite elements to solve Galbrun's equation. This pollution effect was noted in [CD18] and in [DHM+14].

2.1 Model problem for recirculating flows

To study aeroacoustic wave propagation in recirculating background flow, we follow [Ben18] and we introduce a simple model problem: the orthoradial flow in an annulus.
In this chapter, we will use cylindrical coordinates (r, θ) and the classical cylindrical basis $(\boldsymbol{e}_r, \boldsymbol{e}_\theta)$.

The domain $\mathcal{O}$ is an annulus

$$\mathcal{O} := \left\{ (r, \theta) \in \mathbb{R}_+ \times [0, 2\pi) \;\middle|\; r_i < r < r_e \right\},$$

where $0 < r_i < r_e$, and the flow $\boldsymbol{v_0}$ is *orthoradial*

$$\boldsymbol{v_0}(r, \theta) := v_0(r)\boldsymbol{e}_\theta.$$

We will also assume that

$$\forall r \in \mathbb{R}_+, \quad v_0(r) > 0$$

This situation is depicted in FIGURE 2.1. Notice that considering an annulus instead of a circular domain prevents from dealing with a geometric singularity at $r = 0$.

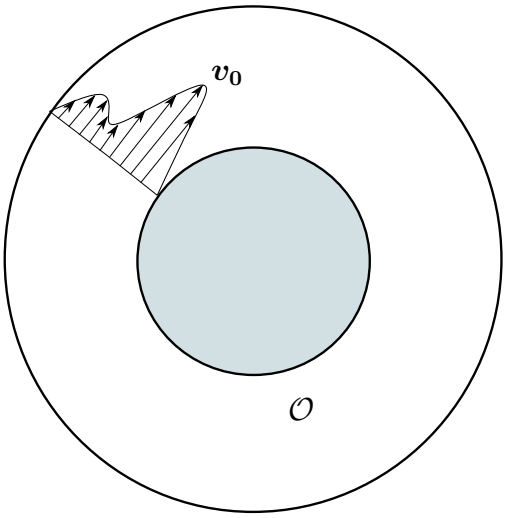


Figure 2.1: Orthoradial flow in an annulus

In this configuration, choosing $\boldsymbol{v_0}$ determines ρ_0 and p_0. It is therefore rather easy to produce solutions to the steady-state Euler's equations in this case. Indeed for a given $\boldsymbol{v_0}$ and assuming that the background flow is an isentropic ideal gas, the following equation of state holds

$$p_0 = K\rho_0^\gamma, \tag{2.1}$$

where γ is the *heat capacity ratio*, defined by

$$\gamma := \frac{c_P}{c_V},$$

with c_P and c_V the specific heat capacities at constant pressure and volume respectively. In this case, the density and pressure do not depend on θ and ρ_0 solves

$$\frac{\partial \rho_0}{\partial r} - \frac{r}{K\gamma}\Omega_0^2(r)\rho_0^{2-\gamma} = 0 \tag{2.2}$$

where we have introduced the angular velocity $\Omega_0 = \frac{v_0(r)}{r}$.
Equation (2.2) is an ordinary differential equation which belongs to the family of Bernoulli's differential equations. It can be solved by a change a function. By substituting

$$u = \rho_0^{\gamma-1}$$

in (2.2) we get

$$\frac{1}{\gamma - 1}\frac{\partial u}{\partial r} = \frac{r\Omega_0^2(r)}{K\gamma}$$

which is a linear differential equation, that can easily be solved compared to the steady-state Euler's equations (BG-ρ_0)–(BG-$\boldsymbol{v_0}$)–(BG-p_0), which is a set of non-linear PDEs, satisfied by the background flow.

2.2 Existence of resonance for the vectorial transport equation

Before moving to time-harmonic solution, let us have a look at what happens in time domain. From the previous chapter, we know that $\boldsymbol{\xi}$ and $\boldsymbol{v}'$ are linked by the following vectorial transport equation

$$\mathrm{D}_t\boldsymbol{\xi} - \nabla_{\boldsymbol{\xi}}\boldsymbol{v_0} = \boldsymbol{v}',$$

and for an orthoradial flow, this specializes to

$$\frac{\partial \xi_r}{\partial t} + \Omega_0 \frac{\partial \xi_r}{\partial \theta} = v_r', \tag{2.3}$$

$$\frac{\partial \xi_\theta}{\partial t} + \Omega_0 \frac{\partial \xi_\theta}{\partial \theta} + \left(\Omega_0 - \frac{\partial v_0}{\partial r}\right)\xi_r = v_\theta'. \tag{2.4}$$

As the system (2.3)–(2.4) is only weakly coupled, we can solve (2.3) and use the result to solve (2.4). We also notice that the unknowns ξ_r and ξ_θ are not differentiated with respect to r, so the system (2.3)–(2.4) is actually a family of equations parametrized by r, and we can independently solve the equations for fixed values of r.

Using the standard theory of transport equations, we know that

$$\xi_r(r,\theta,t) = \Phi(r,\theta-\Omega_0(r)t) + \int_0^t v_r'(r,\theta+(s-t)\Omega_0(r),s)\mathrm{d}s, \tag{2.5}$$

where Φ is an arbitrary function which is 2π-periodic in its second argument. A similar expression can be obtained for ξ_θ.

From time-domain to harmonic solutions To obtain harmonic solutions of (2.3), we use the *ansatz*

$$\xi_r(r,\theta,t) \longleftarrow \mathfrak{Re}\left[\xi_r(r,\theta)e^{-i\omega t}\right], \quad \text{and} \quad \boldsymbol{v}'(r,\theta,t) \longleftarrow \mathfrak{Re}\left[\boldsymbol{v}'(r,\theta)e^{-i\omega t}\right]. \tag{2.6}$$

To study the solvability of (2.3) in the time-harmonic domain, let us first focus on the homogenous problem whose solution is denoted by u

$$\frac{\partial u}{\partial t} + \Omega_0 \frac{\partial u}{\partial \theta} = 0,$$

according to (2.5), we have

$$u(r,\theta,t) = \Phi(r,\theta-\Omega_0(r)t),$$

where Φ is 2π-periodic in its second argument. We can therefore use a decomposition into Fourier series, leading to

$$u(r,\theta,t) = \sum_{n\in\mathbb{Z}} \Phi_n(r)e^{in(\theta-\Omega_0(r)t)} = \sum_{n\in\mathbb{Z}} \Phi_n(r)e^{in\theta}e^{-in\Omega_0(r)t}.$$

Let $\mathcal{F}_t$ denotes the Fourier transform with respect to time, we have

$$\mathcal{F}_t[u](r,\theta,\omega) = \sum_{n\in\mathbb{Z}} \Phi_n(r)e^{in\theta}\delta(\omega - n\Omega_0(r)),$$

where δ is the Dirac distribution. We can therefore conclude that (2.3) with the *ansatz* (2.6) is not uniquely solvable when $\omega = n\Omega_0$. Indeed if $\xi_r(r,\theta)e^{-i\omega t}$ is a solution of (2.6) then

$w := (\xi_r(r,\theta) + \Phi_n(r)e^{in\theta})e^{-i\omega t}$ is also one of them. To understand this, we plug (2.6) into (2.3) to obtain

$$-i\omega w + \Omega_0 \frac{\partial w}{\partial \theta} = v_r',$$

and using that $\omega = n\Omega_0$, we have

$$-in\Omega_0 w + \Omega_0 \frac{\partial w}{\partial \theta} = v_r'.$$

Finally using the definition of $w(r,\theta) := \xi_r(r,\theta) + \Phi_n(r)e^{in\theta}$ yields

$$-in\Omega_0 \xi_r + \frac{\partial \xi_r}{\partial \theta} + \underbrace{-in\Omega_0 \Phi_n e^{in\theta} + in\Omega_0 \Phi_n e^{in\theta}}_{=0} = v_r'.$$

Directly solving the equation in time-harmonic domain In this section, we will show that those resonant lines also appear when we try to solve the vectorial transport equation in time-harmonic domain. Using the same *ansatz* as before for time harmonic solutions

$$\boldsymbol{\xi}(\boldsymbol{x}, t) \longleftarrow \boldsymbol{\xi}(\boldsymbol{x})e^{-i\omega t},$$

the vectorial transport equation (2.3)–(2.4) becomes

$$-i\omega \xi_r + \Omega_0 \frac{\partial \xi_r}{\partial \theta} = v_r', \tag{2.7}$$

$$-i\omega \xi_\theta + \Omega_0 \frac{\partial \xi_\theta}{\partial \theta} + \left(\Omega_0 - \frac{\partial v_0}{\partial r}\right)\xi_r = v_\theta', \tag{2.8}$$

with the periodicity conditions

$$\xi_r(\cdot, 0) = \xi_r(\cdot, 2\pi), \quad \text{and} \quad \xi_\theta(\cdot, 0) = \xi_\theta(\cdot, 2\pi). \tag{2.9}$$

Notice that there is no derivative with respect to r, which therefore is a parameter. For a given r, (2.7)–(2.8) is a system of weakly coupled ODEs, and we can proceed as in the previous case by solving (2.7) and then (2.8).
Using Duhamel's principle, we can write the solution of (2.7) as

$$\xi_r(r,\theta) = K(r)\exp\left[i\frac{\omega}{\Omega_0}\theta\right] + \Omega_0^{-1}\int_0^\theta \exp\left[i\frac{\omega}{\Omega_0}(\theta - \vartheta)\right] v_r'(r,\vartheta)\mathrm{d}\vartheta,$$

where K is a function that must be determined using the periodicity condition (2.9). We can see that

$$\left(1 - \exp\left[2i\pi\frac{\omega}{\Omega_0}\right]\right) K(r) = \Omega_0^{-1}\int_0^{2\pi} \exp\left[i\frac{\omega}{\Omega_0}(2\pi - \vartheta)\right] v_r'(r,\vartheta)\mathrm{d}\vartheta.$$

It is clear that K can only be determined when

$$\begin{aligned} \exp\left[2i\pi\frac{\omega}{\Omega_0}\right] \neq 1 \quad &\Longleftrightarrow \quad \frac{\omega}{\Omega_0} \notin \mathbb{Z}, \\ &\Longleftrightarrow \quad \omega \neq n\Omega_0, \ \forall n \in \mathbb{Z}. \end{aligned}$$

As in the previous case, we can see that the time-harmonic vectorial transport equation is not solvable when $\omega = n\Omega_0$ for some $n \in \mathbb{Z}$.

Decomposition into Fourier series The two previous techniques relied on the simplicity of the transport equation allowing us to solve it analytically. When moving to more complicated systems such as Galbrun's equation and LEE, this is not possible anymore, and we therefore need another way to compute the resonant lines.
As the solution is 2π-periodic in its second argument, we can perform a decomposition into Fourier series with respect to θ, leading to

$$\xi_r(r,\theta) = \sum_{n\in\mathbb{Z}} \xi_{r,n}(r)e^{in\theta}, \quad \text{and} \quad \xi_\theta(r,\theta) = \sum_{n\in\mathbb{Z}} \xi_{\theta,n}(r)e^{in\theta}. \tag{2.10}$$

Introducing (2.10) into (2.7)–(2.8) leads to the following *modal system*

$$-i\omega\xi_{r,n} + in\Omega_0\xi_{r,n} = v'_{r,n}, \tag{2.11}$$

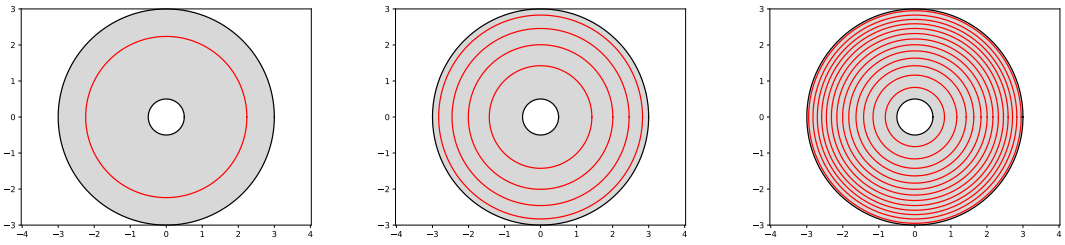
$$-i\omega\xi_{\theta,n} + in\Omega_0\xi_{\theta,n} + \left(\Omega_0 + \frac{\partial v_0}{\partial r}\right)\xi_{r,n} = v'_{\theta,n}, \tag{2.12}$$

for all $n \in \mathbb{Z}$. We can see that (2.11)–(2.12) is a purely algebraic system, whereas (2.7)–(2.8) was a system of ODEs.
Once again, it is clear that (2.11)–(2.12) can only be solved when $\omega \neq n\Omega_0$ for all $n \in \mathbb{Z}$.

Distribution of the resonant lines. Following [Ben18], we consider a simple example to illustrate the distribution of resonant lines. We consider a *potential flow* in the annulus with $r_i = 0.5$ and $r_e = 3$, this flow is given by

$$\boldsymbol{v_0} = \frac{1}{r}\boldsymbol{e}_\theta.$$

In FIGURE 2.2, we can clearly see that the number of resonant lines increases with the frequency. In particular, the set of resonant lines can become uncountable when the frequency is large enough. In [Ben18, Prop. 8.3.3], it is proven that if the background flow is not locally in solid rotation, then there exists a frequency ω_0 such that there is at least one resonant line for each $\omega > \omega_0$. As we will see in the following sections, it is not possible to solve the aeroacoustic models (either Galbrun's equation or LEE) at a given frequency ω if a resonant line on which $\omega = n\Omega_0(r^\star)$ exists.



(a) $\omega = 0.2$ (b) $\omega = 0.5$ (c) $\omega = 1.5$

Figure 2.2: Resonant lines (in red) in the annulus $\{0.5 < r < 3.0\}$ for three different frequencies.

2.3 Existence of resonance for Galbrun's equation

To study the existence of resonant lines for Galbrun's equation (Gal), we cannot solve the equation directly as we did for the transport equation. We therefore rely on the decomposi-

tion into Fourier series

$$\xi_r(r,\theta) = \sum_{n\in\mathbb{Z}} \xi_{r,n}(r)e^{in\theta}, \quad \text{and} \quad \xi_\theta(r,\theta) = \sum_{n\in\mathbb{Z}} \xi_{\theta,n}(r)e^{in\theta},$$

that led to the same resonant lines for the transport equation.
Writing down Galbrun's equation into polar coordinates in the annulus and performing the decomposition into Fourier series to get a modal system is a lengthy and cumberstone process that has been detailed in [Poi85].
The following *modal system* is a 2D reduction of the 3D system obtained by B. Poirée in [Poi85]

$$-\frac{\mathrm{d}^2\xi_{r,n}}{\mathrm{d}r^2} - \frac{1+X}{r}\frac{\mathrm{d}\xi_{r,n}}{\mathrm{d}r} + \left(\frac{1-X}{r^2} - \frac{\lambda^2}{c_0^2}\right)\xi_{r,n} + in\frac{\mathrm{d}}{\mathrm{d}r}\left[\frac{\xi_{\theta,n}}{r}\right] - \left(2iv_0\lambda + in\frac{M^2-X}{r}\right)\frac{\xi_{\theta,n}}{r} = s_{r,n}, \tag{2.13}$$

$$in\frac{\mathrm{d}\xi_{r,n}}{\mathrm{d}r} + \left(2i\frac{M\lambda}{c_0} + in\frac{1+M^2}{r}\right)\xi_{r,n} + \left(n^2 - \frac{\lambda^2 r^2}{c_0^2}\right)\frac{\xi_{\theta,n}}{r} = s_{\theta,n}, \tag{2.14}$$

where

$$\begin{aligned}\lambda(r) &:= \omega - n\Omega_0(r),\\ M(r) &:= \frac{v_0(r)}{c_0(r)},\\ X(r) &:= \frac{r}{\rho_0 c_0^2}\frac{\mathrm{d}\rho_0 c_0^2}{\mathrm{d}r}.\end{aligned}$$

Singular points of ODEs We recall the definition of a *singular point* in the context of ODEs. Let us consider the equation

$$A(r)\frac{\mathrm{d}^2u}{\mathrm{d}r^2} + B(r)\frac{\mathrm{d}u}{\mathrm{d}r} + C(r)u = S(r). \tag{2.15}$$

Définition 1 :

We say that $r^\star \in \mathbb{R}$ is a *singular point* of (2.15) if $A(r^\star) = 0$.

It is possible to study the behavior of the solution of (2.15) around a singular point $r^\star$ by extending the equation to the complex plane, *i.e.* solving for $u(z)$ with $z \in \mathbb{C}$ instead of $u(r)$ with $r \in \mathbb{R}$. If the singularity at $r^\star$ is regular enough, *i.e.* $\frac{B}{A}$ has a pole of order one and $\frac{C}{A}$ has a pole of order up to two, then Fuchs theorem states that (2.15) has a solution that can be written as a generalized power series and can be constructed using Frobenius method.

Even if it is possible to give meaning to the solution of an ODE in the vicinity of a singular point, let us point out some drawbacks:

- *Regularity of the coefficients:* as this result comes from complex analysis, the coefficients of (2.15) should be meromorphic. This will translate to regularity requirements on the physical parameters that are not realistic.
- *Regularity of the singularity:* in practice, it is not possible to guarantee that the singular points are regular.
- *Not variational:* the solutions defined by Fuchs theorem do not lie in the Hilbertian settings required to use finite-element methods.

A detailed study of an alternative aeroacoustic model called *Goldstein's equation* with recirculating flows has been performed in [Ben18]. Thanks to Fuchsian theory, the authors were able to devise semi-analytical methods. As the construction of those semi-analytical methods heavily relies on the particular form of the considered problem, it may be difficult to generalize them. Finally, we would like to recall that one of the goal of this work is to devise numerical methods that are flexible enough to handle cases where semi-analytical methods cannot be used.

Existence of singular points We will now show the existence of singular points for the modal equation associated with Galbrun's equation (Gal). Those singular points are closely related to the resonant lines of the previous sections. Assuming that

$$\frac{\lambda^2 r^2}{c_0^2} \neq n^2, \tag{2.16}$$

we can use (2.14) to eliminate $\xi_{\theta,n}$ from (2.13). In particular, we have

$$\frac{\mathrm{d}}{\mathrm{d}r}\left[\frac{\xi_{\theta,n}}{r}\right] = -in\left(n^2 - \frac{\lambda^2 r^2}{c_0^2}\right)^{-1} \frac{\mathrm{d}^2\xi_{r,n}}{\mathrm{d}r^2} + F\left(\frac{\mathrm{d}\xi_{r,n}}{\mathrm{d}r}, \xi_{r,n}\right),$$

leading to the following second-order term in (2.13)

$$\left(n^2\left(n^2 - \frac{\lambda^2 r^2}{c_0^2}\right)^{-1} - 1\right)\frac{\mathrm{d}^2\xi_{r,n}}{\mathrm{d}r^2} =: A(r)\frac{\mathrm{d}^2\xi_{r,n}}{\mathrm{d}r^2}.$$

We can clearly see that if there exists an $r^\star$ such that $\lambda(r^\star) = 0$, then $A(r^\star) = n^2 n^{-2} - 1 = 0$ and $r^\star$ is therefore a singular point of (2.13). Those singular points are closely related to the resonance of the previous section, indeed if $\lambda = 0$, then $\omega = n\Omega_0$ for some $n \in \mathbb{Z}$. We would like to point out that the resonant points of the modal system correspond to resonant streamlines for Galbrun's equation, and they are therefore sometimes referred to as *resonant lines*.

Notice that another set of points could lead to singularity, indeed to perform the previous computations we had to make the assumption (2.16) and the points where $\lambda^2 r^2 = c_0^2 n^2$ could also be singular. To see that they actually are not singular, it is easier to consider a first-order system. We therefore introduce a new unknown

$$\alpha_n := \frac{\mathrm{d}\xi_{r,n}}{\mathrm{d}r},$$

and we rewrite (2.13)–(2.14) as

$$\begin{aligned}
\frac{\mathrm{d}\xi_{r,n}}{\mathrm{d}r} &= \alpha_n, \\
\frac{\mathrm{d}\alpha_n}{\mathrm{d}r} - in\frac{\mathrm{d}}{\mathrm{d}r}\left[\frac{\xi_{\theta,n}}{r}\right] &= -\frac{1+X}{r}\alpha_n + \left(\frac{1-X}{r^2} - \frac{\lambda^2}{c_0^2}\right)\xi_{r,n} \\
&\quad - \left(2iv_0\lambda + in\frac{M^2 - X}{r}\right)\frac{\xi_{\theta,n}}{r} - s_{r,n}, \qquad (2.17) \\
in\frac{\mathrm{d}\xi_{r,n}}{\mathrm{d}r} &= -\left(2i\frac{M\lambda}{c_0} + in\frac{1+M^2}{r}\right)\xi_{r,n} - \left(n^2 - \frac{\lambda^2 r^2}{c_0^2}\right)\frac{\xi_{\theta,n}}{r} + s_{\theta,n}. \qquad (2.18)
\end{aligned}$$

If

$$\frac{\lambda^2 r^2}{c_0^2} = n^2,$$

then we obtain the following decoupled system for $\xi_{r,n}$ and α_n

$$\frac{\mathrm{d}\xi_{r,n}}{\mathrm{d}r} - \alpha_n = 0,$$
$$in\frac{\mathrm{d}\xi_{r,n}}{\mathrm{d}r} + \left(2i\frac{M\lambda}{c_0} + in\frac{1+M^2}{r}\right)\xi_{r,n} = s_{\theta,n},$$

and we can solve (2.17) for $\xi_{\theta,n}$.
With this formulation we can also retrieve the singular points from before. Indeed if we write (2.18) using α_n and if we assume that (2.16) holds, we have

$$in\alpha_n + \left(2i\frac{M\lambda}{c_0} + in\frac{1+M^2}{r}\right)\xi_{r,n} + \left(n^2 - \frac{\lambda^2 r^2}{c_0^2}\right)\frac{\xi_{\theta,n}}{r} = s_{\theta,n}.$$

Using this equation to eliminate $\xi_{\theta,n}$ from (2.17) leads to the following term

$$\left(n^2\left(n^2 - \frac{\lambda^2 r^2}{c_0^2}\right)^{-1} - 1\right)\frac{\mathrm{d}\alpha_n}{\mathrm{d}r},$$

and to the same singular points.

Designing numerical methods We have shown that it may not be possible to solve the time-harmonic Galbrun's equation because of resonant lines that appear when $\omega = n\Omega_0$. Indeed, for those frequencies there are resonant streamlines. Let $\mathcal{L}^\star = \{r = r^\star\}$, with $r^\star \in (r_i, r_e)$, be a resonant streamline and let $\boldsymbol{n}_{\mathcal{L}^\star}$ be an outward unit normal vector to $\mathcal{L}^\star$. As the background flow is orthoradial, *i.e.* $\boldsymbol{v_0} = v_0(r)\boldsymbol{e}_\theta$, we have $\boldsymbol{n}_{\mathcal{L}^\star} \propto \boldsymbol{e}_r$ and therefore $\boldsymbol{v_0} \cdot \boldsymbol{n}_{\mathcal{L}^\star} = 0$.

At this point, we recall that there are two different Hilbertian settings for Galbrun's equation: the usual $\boldsymbol{H}^1(\mathcal{O})$-settings of [BMM+12], and the low-regularity $\boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O})$-settings of [HH21], where

$$\boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O}) := \left\{\boldsymbol{\xi} \in \boldsymbol{L}^2(\mathcal{O}) \;\middle|\; \mathrm{div}\,(\boldsymbol{\xi}) \in L^2(\mathcal{O}),\ \nabla_{\boldsymbol{v_0}}\boldsymbol{\xi} \in \boldsymbol{L}^2(\mathcal{O})\right\}.$$

If we approximate $\boldsymbol{\xi}$ in $\boldsymbol{H}^1(\mathcal{O})$, then the numerical method will "try" to define the behavior of $\boldsymbol{\xi}$ on $\mathcal{L}^\star$, which will lead to bad-quality numerical results as the method will generate *spurious modes*. The bad quality of the numerical results has been pointed out in [CD18], and an example of resonant line is pictured in FIGURE 2.3 that has been obtained using the PQ-formulation of [CD18] in the `montjoie` solver. In [DHM+14], it was noted that a mixed finite-element discretization of Galbrun's equation can produce *spurious modes* that pollute the numerical solution. Those spurious modes, see *e.g.* [DHM+14, Fig. 17], are very similar to the resonant phenomenon depicted in FIGURE 2.3.

On the other hand, if we work with the low-regularity settings and approximate $\boldsymbol{\xi}$ in $\boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O})$, then the numerical method will not respect the behavior of $\boldsymbol{\xi}$ on $\mathcal{L}^\star$ as $\boldsymbol{v_0} \cdot \boldsymbol{n}_{\mathcal{L}^\star} = 0$. Indeed, as we discussed in the previous chapter, the continuity of $\boldsymbol{\xi} \in \boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O})$ is only enforced on curves where $\boldsymbol{v_0} \cdot \boldsymbol{n} \neq 0$.

The previous remarks seem to favor the second approach to design numerical methods for Galbrun's equation. Given the properties of the space $\boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O})$, this will naturally lead us to non-conforming numerical methods and to *(Hybridizable) Discontinuous Galerkin methods* in particular.

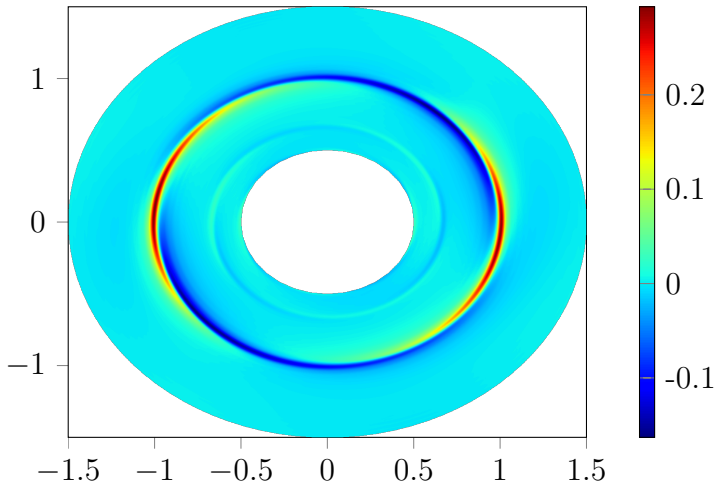


Figure 2.3: Resonant line for Galbrun's equation

2.4 Existence of resonance for LEE

In the previous sections, we have pointed out that a resonant-like phenomenon can happen for both the time-harmonic vectorial transport equation and Galbrun's equation. In this section, we would like to show that this phenomenon can also occur for the time-harmonic LEE, which is more widely used than Galbrun's equation.

We begin with writing the system (LEE-ρ')–(LEE-$\boldsymbol{v}'$)–(LEE-p') for an orthoradial background flow in an annulus

$$\begin{aligned}
-i\omega\rho' + \Omega_0\frac{\partial\rho'}{\partial\theta} + v'_r\frac{\mathrm{d}\rho_0}{\mathrm{d}r} + \rho_0\frac{\mathrm{d}v'_r}{\mathrm{d}r} + \frac{1}{r}\rho_0 v'_r + \frac{1}{r}\rho_0\frac{\partial v'_\theta}{\partial\theta} &= 0,\\
\rho_0\left(-i\omega v'_r + \Omega_0\frac{\partial v'_r}{\partial\theta} - 2\Omega_0 v'_\theta\right) + \frac{1}{\rho_0}\frac{\mathrm{d}p_0}{\mathrm{d}r}\rho' + \frac{\mathrm{d}p'}{\mathrm{d}r} &= s_r,\\
\rho_0\left(-i\omega v'_\theta + \Omega_0\frac{\partial v'_\theta}{\partial\theta} + \Omega_0 v'^r\right) + \rho_0\frac{\mathrm{d}v_0}{\mathrm{d}r}v'_\theta + \frac{\partial p'}{\partial\theta} &= s_\theta,\\
-i\omega p' + \Omega_0\frac{\partial p'}{\partial\theta} + \frac{\mathrm{d}p_0}{\mathrm{d}r}v'_r + \gamma p_0\left(\frac{\mathrm{d}v'_\theta}{\mathrm{d}r} + \frac{1}{r}v'_r + \frac{1}{r}\frac{\partial v'_\theta}{\partial\theta}\right) &= 0.
\end{aligned}$$

As before, the unknowns are written as Fourier series

$$q(r,\theta) = \sum_{n\in\mathbb{Z}} q_n(r)e^{in\theta}, \quad \forall q \in \{\rho', v'_r, v'_\theta, p'\},$$

leading to the following *modal system*

$$i(-\omega + n\Omega_0)\rho'_n + v'_{r,n}\frac{\mathrm{d}\rho_0}{\mathrm{d}r} + \rho_0\left(\frac{\mathrm{d}v'_{r,n}}{\mathrm{d}r} + \frac{1}{r}v'_{r,n} + \frac{in}{r}v'_{\theta,n}\right) = 0, \tag{2.19}$$

$$\rho_0\left(i(-\omega + n\Omega_0)v'_{r,n} - 2\Omega_0 v'_{\theta,n}\right) + \frac{1}{r}\frac{\mathrm{d}\rho_0}{\mathrm{d}r}\rho'_n + \frac{\mathrm{d}p'_n}{\mathrm{d}r} = s_{r,n}, \tag{2.20}$$

$$\rho_0\left(i(-\omega + n\Omega_0)v'_{\theta,n} + \Omega_0 v'_{r,n}\right) + \rho_0\frac{\mathrm{d}v_0}{\mathrm{d}r}v'_{r,n} + inp'_n = s_{\theta,n}, \tag{2.21}$$

$$i(-\omega + n\Omega_0)p'_n + \frac{\mathrm{d}p_0}{\mathrm{d}r}v'_{r,n} + \gamma p_0\left(\frac{\mathrm{d}v'_{r,n}}{\mathrm{d}r} + \frac{1}{r}v'_{r,n} + \frac{in}{r}v'_{\theta,n}\right) = 0. \tag{2.22}$$

The system (2.19)–(2.20)–(2.21)–(2.22) is a *Differential Algebraic System of Equations*[1] (DAE): indeed (2.21) is an algebraic equation, and ρ'_n and $v'_{\theta,n}$ are algebraic unknowns that are not differentiated. To study this system of equations, we need to eliminate ρ'_n and p'_n to obtain a usual differential system.

We begin with noticing that (2.19) and (2.22) lead to a second algebraic constraint. Indeed by using that $\rho_0 c_0^2 = \gamma p_0$ and by multiplying (2.19) by c_0^2, we have

$$i(-\omega + n\Omega_0)\rho'_n = i(-\omega + n\Omega_0)p'_n + \left(\frac{\mathrm{d}p_0}{\mathrm{d}r} - c_0^2\frac{\mathrm{d}\rho_0}{\mathrm{d}r}\right)v'_{r,n}. \tag{2.23}$$

Non-resonant lines If $\omega \neq n\Omega_0$, we can use (2.21) and (2.23) to obtain

$$v'_{\theta,n} = \frac{i}{\rho_0(-\omega + n\Omega_0)}\left[inp'_n + \left(\Omega_0 + \rho_0\frac{\mathrm{d}v_0}{\mathrm{d}r}\right)v'_{r,n} - s_{\theta,n}\right],$$
$$\rho'_n = \frac{1}{c_0^2}p'_n + \frac{i}{-\omega + n\Omega_0}\left(c_0^2\frac{\mathrm{d}\rho_0}{\mathrm{d}r} - \frac{\mathrm{d}p_0}{\mathrm{d}r}\right)v'_{r,n},$$

which allows us to eliminate $v'_{\theta,n}$ and ρ'_n. We can then obtain the two following differential equations for $v'_{r,n}$ and p'_n

$$\frac{i}{-\omega + n\Omega_0}\left[\rho_0(-\omega + n\Omega_0)^2 + \frac{1}{r}\frac{\mathrm{d}v_0}{\mathrm{d}r}\left(c_0^2\frac{\mathrm{d}\rho_0}{\mathrm{d}r} - \frac{\mathrm{d}p_0}{\mathrm{d}r}\right) - \Omega_0\left(\Omega_0 + \rho_0\frac{\mathrm{d}v_0}{\mathrm{d}r}\right)\right]v'_{r,n}$$
$$+\left[\frac{1}{rc_0^2}\frac{\mathrm{d}v_0}{\mathrm{d}r} - \frac{n\Omega_0}{-\omega + n\Omega_0}\right]p'_n + \frac{\mathrm{d}p'_n}{\mathrm{d}r} = s_n^r - \frac{i\Omega_0}{-\omega + n\Omega_0}s_{\theta,n}, \tag{2.24}$$

and

$$\gamma p_0\frac{\mathrm{d}v'_{r,n}}{\mathrm{d}r} + \left(\frac{\mathrm{d}v_0}{\mathrm{d}r} + \frac{\gamma p_0}{r} - \frac{\gamma p_0\rho_0 n}{r\rho_0(-\omega + n\Omega)}\left(\Omega_0 + \rho_0\frac{\mathrm{d}v_0}{\mathrm{d}r}\right)\right)v'_{r,n}$$
$$+ i\left(-\omega + n\Omega_0 - \frac{\gamma p_0 n^2}{r\rho_0(-\omega + n\Omega_0)}\right)p'_n = -\frac{\gamma p_0 n}{r\rho_0(-\omega + n\Omega)}s_{\theta,n}. \tag{2.25}$$

The system (2.24)–(2.25) is a differential system that can be solved if $\omega \neq n\Omega_0$. However, it is clear that points $r^\star$ such that $\omega = n\Omega_0(r^\star)$ are singular points of (2.24)–(2.25). It is natural to wonder if those singular points are singular points of (2.19)–(2.20)–(2.21)–(2.22), or if they are introduced by the reduction to the differential system (2.24)–(2.25). This will be investigated in the next paragraph.

Resonant lines We now assume that $\omega = n\Omega_0$, and the system (2.19)–(2.20)–(2.21)–(2.22) reduces to

$$v'_{r,n}\frac{\mathrm{d}\rho_0}{\mathrm{d}r} + \rho_0\left(\frac{\mathrm{d}v'_{r,n}}{\mathrm{d}r} + \frac{1}{r}v'_{r,n} + \frac{in}{r}v'_{\theta,n}\right) = 0, \tag{2.26}$$
$$-2\rho_0\Omega_0 v'_{\theta,n} + \frac{1}{r}\frac{\mathrm{d}\rho_0}{\mathrm{d}r}\rho'_n + \frac{\mathrm{d}p'_n}{\mathrm{d}r} = s_{r,n}, \tag{2.27}$$
$$\rho_0\left(\Omega_0 + \frac{\mathrm{d}v_0}{\mathrm{d}r}\right)v'_{r,n} + inp'_n = s_{\theta,n}, \tag{2.28}$$
$$\frac{\mathrm{d}p_0}{\mathrm{d}r}v'_{r,n} + \gamma p_0\left(\frac{\mathrm{d}v'_{r,n}}{\mathrm{d}r} + \frac{1}{r}v'_{r,n} + \frac{in}{r}v'_{\theta,n}\right) = 0. \tag{2.29}$$

We cannot use the algebraic constraint to eliminate ρ'_n and $v'_{\theta,n}$ as in the previous case, however (2.28) indicates that we may try to eliminate p'_n to obtain an equation for $v'_{r,n}$ only.

[1] Differential Algebraic systems of Equations should not be confused with Algebraic Differential Equations arising in differential algebra.

Indeed, using (2.26) we have

$$\rho_0 \frac{in}{r} v'_{\theta,n} = -\rho_0 \frac{\mathrm{d}v'_{r,n}}{\mathrm{d}r} - \left(\frac{\mathrm{d}\rho_0}{\mathrm{d}r} + \frac{\rho_0}{r} \right) v'_{r,n}, \tag{2.30}$$

then (2.28) says that

$$inp'_n = \rho_0 \left(\Omega_0 + \frac{\mathrm{d}v_0}{\mathrm{d}r} \right) v'_{r,n} - s_{\theta,n},$$

we can now rewrite (2.27) only in terms of $v'_{r,n}$ and ρ'_n

$$2\Omega_0 \left(\frac{\rho_0}{r} \frac{\mathrm{d}v'_{r,n}}{\mathrm{d}r} + \left(\frac{\mathrm{d}\rho_0}{\mathrm{d}r} + \frac{\rho_0}{r} \right) \frac{v'_{r,n}}{r} \right) + \frac{in}{r} \frac{\mathrm{d}\rho_0}{\mathrm{d}r} \rho'_n + \frac{\mathrm{d}}{\mathrm{d}r} \left[\rho_0 \left(\Omega_0 + \frac{\mathrm{d}v_0}{\mathrm{d}r} \right) v'_{r,n} \right] = ins_{r,n} + \frac{\mathrm{d}s_{\theta,n}}{\mathrm{d}r}. \tag{2.31}$$

Finally, multiplying (2.30) by c_0^2 and using the result in (2.29) leads to

$$\left(\frac{\mathrm{d}p_0}{\mathrm{d}r} - c_0^2 \frac{\mathrm{d}\rho_0}{\mathrm{d}r} \right) v'_{r,n} = 0. \tag{2.32}$$

So on the resonant lines, the modal system (2.19)–(2.20)–(2.21)–(2.22) reduces to (2.32). As the coefficient

$$\frac{\mathrm{d}p_0}{\mathrm{d}r} - c_0^2 \frac{\mathrm{d}\rho_0}{\mathrm{d}r}$$

in the left-hand side can vanish, it is not clear whether or not (2.32) is solvable. Moreover, we can see in (2.30) and in (2.31) that $v'_{\theta,n}$ and ρ'_n are expressed in terms of the derivative of $v'_{r,n}$, and this quantity is not defined by (2.32). So even if $v'_{r,n}$ is defined by (2.32), the solution of LEE may not be well-defined on resonant lines where $\omega = n\Omega_0$.

In particular, in this chapter we have considered the equation of state (2.1)

$$p_0 = K\rho_0^\gamma,$$

which describes an isentropic ideal gas. With this choice of equation of state, we have

$$\frac{\mathrm{d}p_0}{\mathrm{d}r} = K\gamma\rho_0^{\gamma-1} \frac{\mathrm{d}\rho_0}{\mathrm{d}r} = \frac{\gamma p_0}{\rho_0} \frac{\mathrm{d}\rho_0}{\mathrm{d}r} = c_0^2 \frac{\mathrm{d}\rho_0}{\mathrm{d}r},$$

as $\gamma p_0 = \rho_0 c_0^2$. We can conclude that (2.32) always degenerates, and that (2.26) and (2.29) are equivalent. The modal system (2.26)–(2.27)–(2.28)–(2.29) is therefore underdetermined. To compute the kernel associated with a resonant line, we consider the system (2.26)–(2.27)–(2.28) with $s_{r,n} = s_{\theta,n} = 0$. Using (2.28), we can eliminate p'_n as we have

$$\frac{\mathrm{d}p'_n}{\mathrm{d}r} = \frac{\rho_0 \left(\Omega_0 + \frac{\mathrm{d}v_0}{\mathrm{d}r} \right)}{in} \frac{\mathrm{d}v'_{r,n}}{\mathrm{d}r} + \frac{\mathrm{d}}{\mathrm{d}r} \left[\frac{\rho_0 \left(\Omega_0 + \frac{\mathrm{d}v_0}{\mathrm{d}r} \right)}{in} \right] v'_{r,n} = A \frac{\mathrm{d}v'_{r,n}}{\mathrm{d}r} + \frac{\mathrm{d}A}{\mathrm{d}r} v'_{r,n},$$

where we introduced

$$A := \frac{\rho_0 \left(\Omega_0 + \frac{\mathrm{d}v_0}{\mathrm{d}r} \right)}{in}$$

to make the notations lighter. Equations (2.26) and (2.27) can then be rewritten as

$$\frac{\mathrm{d}v'_{r,n}}{\mathrm{d}r} = \frac{2\rho_0\Omega_0}{A} v'_{\theta,n} - \frac{1}{Ar} \frac{\mathrm{d}\rho_0}{\mathrm{d}r} \rho'_n - \frac{\mathrm{d}A}{\mathrm{d}r} v'_{r,n}, \tag{2.33a}$$

$$\frac{\mathrm{d}v'_{r,n}}{\mathrm{d}r} = -\frac{in\rho_0}{r} v'_{\theta,n} - \left(\frac{1}{r} + \frac{1}{\rho_0} \frac{\mathrm{d}\rho_0}{\mathrm{d}r} \right) v'_{r,n}. \tag{2.33b}$$

This system can be solved only if

$$-\frac{1}{Ar}\frac{\mathrm{d}\rho_0}{\mathrm{d}r}\rho'_n = \left(\frac{\mathrm{d}A}{\mathrm{d}r} - \frac{1}{r} - \frac{1}{\rho_0}\frac{\mathrm{d}\rho_0}{\mathrm{d}r}\right) v'_{r,n} - \left(\frac{in\rho_0}{r} + \frac{2\rho_0\Omega_0}{A}\right) v'_{\theta,n},$$

which makes (2.33a) equivalent to (2.33b). So the modal system (2.26)–(2.27)–(2.28) reduces to the following system

$$\begin{aligned}
\frac{\mathrm{d}v'_{r,n}}{\mathrm{d}r} &= -\frac{in\rho_0}{r}v'_{\theta,n} - \left(\frac{1}{r} + \frac{1}{\rho_0}\frac{\mathrm{d}\rho_0}{\mathrm{d}r}\right) v'_{r,n},\\
p'_n &= Av'_{r,n},\\
-\frac{1}{Ar}\frac{\mathrm{d}\rho_0}{\mathrm{d}r}\rho'_n &= \left(\frac{\mathrm{d}A}{\mathrm{d}r} - \frac{1}{r} - \frac{1}{\rho_0}\frac{\mathrm{d}\rho_0}{\mathrm{d}r}\right) v'_{r,n} - \left(\frac{in\rho_0}{r} + \frac{2\rho_0\Omega_0}{A}\right) v'_{\theta,n},
\end{aligned}$$

which has a solution for any $v'_{\theta,n}$.

Finally we would like to point out that if we use the numerical method for LEE described in [CD18], we obtain spurious modes that are similar to those of Galbrun's equation. The resonant line for LEE is depicted on Figure 2.4 and can be compared to Figure 2.3.

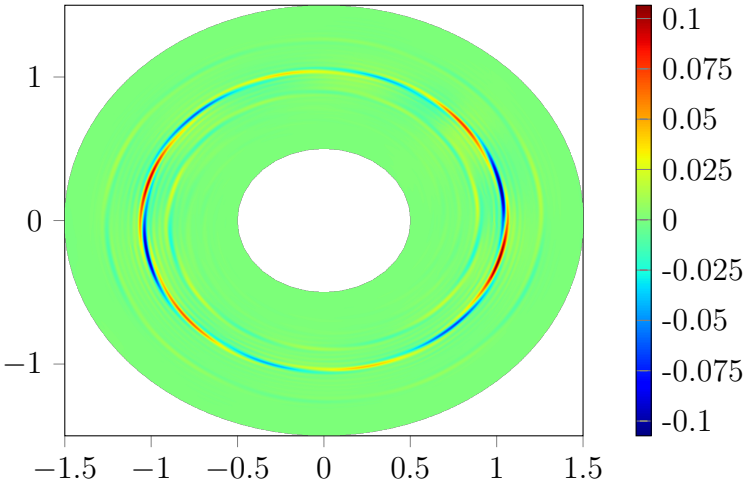


Figure 2.4: Resonant line for LEE

2.5 The case of the convected Helmholtz equation

In the two previous sections, we have seen that resonant lines exist for both Galbrun's equation and LEE. It is interesting to study this phenomenon for the *convected Helmholtz equation*, which is the simplest aeroacoustic model. LEE and Galbrun's equations are aeroacoustic models that describe acoustic and hydrodynamic phenomena, whereas the convected Helmholtz equation describes pure acoustic waves and is only valid for potential background flow.

We consider the following *convected Helmholtz equation*

$$\rho_0\left(-i\omega + \nabla_{\boldsymbol{v_0}}\right)^2 p - \operatorname{div}\left(\rho_0 c_0^2 \nabla p\right) = s, \tag{2.34}$$

where s is the acoustic source and p is an acoustic potential that is related to the eulerian perturbations by the following identities

$$\begin{aligned}
p' &= -\rho_0 c_0(-i\omega + \nabla_{\boldsymbol{v_0}})p,\\
\boldsymbol{v}' &= -c_0\nabla p.
\end{aligned} \tag{2.35}$$

For more details on the convected Helmholtz equation, we refer to [Pie90].
In polar coordinates, this equation reads

$$\rho_0\left(-\omega^2 p - 2i\omega\Omega_0\frac{\partial p}{\partial\theta} + \Omega_0^2\frac{\partial^2 p}{\partial\theta^2}\right) - \frac{1}{r}\left[\frac{\partial}{\partial r}\left(r\rho_0 c_0^2\frac{\partial p}{\partial r}\right) + \rho_0 c_0^2\frac{\partial^2 p}{\partial\theta^2}\right] = s$$

We proceed as before and we perform a decomposition of p into Fourier series with respect to θ

$$p(r,\theta) = \sum_{n\in\mathbb{Z}} p_n(r)e^{in\theta},$$

leading to the following modal equation

$$-\rho_0(\omega + n\Omega_0)^2 p_n + \frac{\rho_0 c_0^2 n^2}{r}p_n - \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\left(r\rho_0 c_0^2\frac{\mathrm{d}p_n}{\mathrm{d}r}\right) = s_n. \tag{2.36}$$

Contrary to the previous cases, it is clear that (2.36) cannot degenerate and that there is therefore no resonant line. As (2.34) models pure acoustics, we can conclude that the existence of resonant lines comes from the coupling between acoustic and hydrodynamic effects.

Finally, we would like to point that the modal counterpart of (2.35) is

$$p_n' = -i\rho_0 c_0(-\omega + n\Omega_0)p_n.$$

We can therefore see that even if the mode p_n associated with a resonant line where $\omega = n\Omega_0$ is well-defined, this mode does not contribute to the eulerian perturbation of pressure p' in the case of pure acoustics. This seems consistent with the hydrodynamic nature of the resonant-like phenomenon.

Conclusion & perspectives

From this section we can draw several conclusions. We studied a model problem for time-harmonic aeroacoustic wave propagation in recirculating background flows, *i.e.* flows with closed streamlines, and we have shown that if

$$\omega = n\Omega_0(r^\star),$$

for some $r^\star \in (r_i, r_e)$ and $n \in \mathbb{Z}$, then the vectorial transport equation, Galbrun's equation or LEE cannot be solved. It is important to point out that the frequencies ω for which this phenomenon happens are not eigenvalues of the Laplace operator, and that they can constitute an uncountable set. This resonant-like phenomenon gives a mathematical interpretation of spurious modes that have been observed in some numerical methods for time-harmonic aeroacoustic wave propagation.

As we discussed, this resonant-like phenomenon does not occur in the case of pure acoustic wave propagation and therefore comes from the coupling between acoustic and hydrodynamic effects in Galbrun's equation and LEE.

Fortunately, it seems possible to study Galbrun's equation with recirculating flows by using the well-posedness result of [HH21], where the following Hilbert space is used

$$\boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O}) := \left\{\boldsymbol{\xi} \in \boldsymbol{L}^2(\mathcal{O}) \;\middle|\; \mathrm{div}\,(\boldsymbol{\xi}) \in L^2(\mathcal{O}),\ \nabla_{\boldsymbol{v_0}}\boldsymbol{\xi} \in \boldsymbol{L}^2(\mathcal{O})\right\}.$$

This functional framework does not set the behavior of $\boldsymbol{\xi}$ on the resonant lines. Approximation of functions in this non-standard Hilbert space requires new numerical methods. To

the best of our knowledge, no such result exists for either LEE or the vectorial transport equation. Extending the work of [HH21] to those two equations seems very interesting. It could either lead to a similar result for LEE, and therefore prove the equivalence between LEE and Galbrun's equation in time-harmonic domain. Or, it could exhibits configurations where one of the model is well-posed and the other is not.

Bibliography

[Ben18] Antoine Bensalah. *Une Approche Nouvelle de La Modélisation Mathématique et Numérique En Aéroacoustique Par Les Équations de Goldstein et Applications En Aéronautique*. PhD thesis, July 2018.

[BJM18] Antoine Bensalah, Patrick Joly, and Jean-François Mercier. Well-posedness of a generalized time-harmonic transport equation for acoustics in flow. *Mathematical Methods in the Applied Sciences*, 41(8):3117–3137, 2018.

[BMM+12] Anne-Sophie Bonnet-Ben Dhia, Jean-François Mercier, Florence Millot, Sébastien Pernet, and Emilie Peynaud. Time-Harmonic Acoustic Scattering in a Complex Flow: A Full Coupling Between Acoustics and Hydrodynamics. *Communications in Computational Physics*, 11(2):555–572, February 2012.

[CD18] Juliette Chabassier and Marc Duruflé. Solving time-harmonic Galbrun's equation with an arbitrary flow. Application to Helioseismology. *Rapport de recherche Inria*, 2018.

[DHM+14] Felix Dietzsch, Luis Hervella-Nieto, Steff En Marburg, Rodolfo Rodríguez, and Hannah Weisbecker. Physical and Spurious Modes in Mixed Finite Element Formulation for the Galbrun Equation. 2014.

[HH21] Martin Halla and Thorsten Hohage. On the Well-posedness of the Damped Time-harmonic Galbrun Equation and the Equations of Stellar Oscillations. *SIAM Journal on Mathematical Analysis*, pages 4068–4095, January 2021.

[Pey13] Emilie Peynaud. *Rayonnement Sonore Dans Un Écoulement Subsonique Complexe En Régime Harmonique : Analyse et Simulation Numérique Du Couplage Entre Les Phénomènes Acoustiques et Hydrodynamiques*. Thesis, Toulouse, INSA, June 2013.

[Pie90] Allan Pierce. Wave equation for sound in fluids with unsteady inhomogeneous flow. *The Journal of the Acoustical Society of America*, 87(6):2292–2299, June 1990.

[Poi85] Bernard Poirée. Petites perturbations d'un écoulement tournant. *Acta acustica united with acustica*, Volume 59(Number 2):85–94, December 1985.

Part II

Numerical methods

Chapter 3

A HDG framework for the convected Helmholtz equation

Contents

Introduction

Nowadays the solar interior is studied by considering the propagation of aeroacoustic waves in time-harmonic domain. Realistic models of solar oscillations require to approximate non-standard Hilbert settings leading to non-conforming methods, such as *Discontinuous Galerkin Methods*. As those methods have a very important numerical cost, we consider the so-called *Hybridizable Discontinuous Galerkin Methods* (HDG), which relies on a static condensation process to reduce the number of degrees of freedom.

As a first step towards the use of HDG in helioseismology, we construct and study HDG for the simplest aeroacoustic model : the *convected Helmholtz equation*.

HDG have been used and validated by numerous authors for various problems such as elliptic equations in [CGL09, CDG+09, CC12, CC14], acoustic wave propagation in [GM11, GSV18, NPRC15], elastic wave propagation in [HPS17, BDMP21, CS13, FCS15, BCDL15], Maxwell equations in [CQSS17, CQS18, CLOS20]. These methods have also been used to implement the forward propagator in the context of quantitative inverse problems in [FS20] where a specific formulation of the adjoint method is developed. In this paper, we will consider the HDG+ variant of HDG, introduced in [Leh10], where different polynomial degrees are used for the different unknowns. This HDG+ has been considered for various applications in [CQSS17, Oik14, Oik16, Oik18, QSS16, QS16a, QS16b, Hun19] and to the best of our knowledge, the case of the convected Helmholtz equation has not been addressed yet.

Theory for HDGs is rather similar to the one for mixed finite elements and the actual connection was first established by Cockburn and his coworkers in [CGS10]. For a self-contained introduction to the theory of HDG, we refer to [DS19]. For a historical perspective on HDG, we refer to [Coc14].

For a comparison between HDG and Continuous Galerkin methods, we refer to [KSC12, YMKS16]. The relationship between HDG and HHO (Hybrid High-Order, another new generation of high-order face-based finite element method) has been studied in [CDPE16].

Main results: We construct three variants of the HDG method for convected acoustics in the frequency domain. Our main results include a detailed analysis of those methods where the most important properties of the method are proved including local and global solvability, convergence rate for regular solutions. The choice of the penalization parameter is also discussed. Finally, those three methods were implemented in `hawen` (see [Fau21]) and we also provide numerical experiments. Using those numerical experiments, we can conclude that the HDG+ and HDG-$\boldsymbol{\sigma}_h$ methods should be preferred to the HDG-$\boldsymbol{q}_h$ method as they seem more robust.

Organization of this chapter: This work is organized as follows:

- in SECTION 3.1: we present the convected Helmholtz equation and recall some results on this equation, we also present two ways to reach a first-order in space formulation;
- in SECTION 3.2: we introduce some notations and the approximation spaces needed to construct HDGs developed in this paper;
- in SECTION 3.3: we construct the HDG-$\boldsymbol{\sigma}_h$ method based on the *total flux formulation* of the convected Helmholtz equation, we also provide theoretical results and discuss the optimal choice of penalization parameter for this method;
- in SECTION 3.4: we construct the HDG-$\boldsymbol{q}_h$ and HDG+ methods based on the *diffusive flux formulation* of the convected Helmholtz equation, we also provide detailed analysis of those methods;
- in SECTION 3.5: we give details on how those methods can be implemented in a nodal settings;
- in SECTION 3.6: we present numerical experiments to illustrate our theoretical results, as well as some illustrative examples.

3.1 Model problem

As a model problem we consider the so-called *convected Helmholtz equation*

$$\rho_0 \left(-\omega^2 p - 2i\omega \boldsymbol{v_0} \cdot \nabla p + \boldsymbol{v_0} \cdot \nabla(\boldsymbol{v_0} \cdot \nabla p)\right) - \operatorname{div}\left(\rho_0 c_0^2 \nabla p\right) = s \tag{3.1}$$

where ω is the angular frequency, ρ_0 is the density of the fluid, $\boldsymbol{v_0}$ is the velocity of the fluid, c_0 is the adiabatic sound speed, and s is the acoustic source.

Validity of this equation: Equation (3.1) is the simplest aeroacoustic models and therefore has a limited validity. This equation can be used for

- a *uniform background flow*, in this case the unknown p can be interpreted as a pressure perturbation,
- a *potential background flow*, in this case the unknown p should be interpreted as an *acoustic potential* and the physical quantities can be retrieved using the following identities

$$\begin{aligned} &\text{Pressure perturbation:} && p' = -\rho_0 c_0 (-i\omega + \boldsymbol{v_0} \cdot \nabla) p, \\ &\text{Velocity perturbation:} && \boldsymbol{v}' = -c_0 \nabla p, \end{aligned}$$

see [Pie90, Sec. II.].

Combining the second-order differential operators: We will assume that the background flow is incompressible which leads to the following local mass conservation equation

$$\operatorname{div}(\rho_0 \boldsymbol{v_0}) = 0.$$

With this assumption, we have

$$\begin{aligned} \rho_0 \boldsymbol{v_0} \cdot \nabla(\boldsymbol{v_0} \cdot \nabla p) &= \operatorname{div}(\rho_0 (\boldsymbol{v_0} \cdot \nabla p) \boldsymbol{v_0}) - (\boldsymbol{v_0} \cdot \nabla p) \cancel{\operatorname{div}(\rho_0 \boldsymbol{v_0})} \\ &= \operatorname{div}(\rho_0 (\boldsymbol{v_0} \cdot \nabla p) \boldsymbol{v_0}) \\ &= \operatorname{div}\left(\rho_0 \boldsymbol{v_0} \boldsymbol{v_0}^T \nabla p\right) \end{aligned}$$

Leading to

$$\rho_0 \left(-\omega^2 p - 2i\omega \boldsymbol{v_0} \cdot \nabla p\right) - \operatorname{div}(\boldsymbol{K_0} \nabla p) = s \tag{3.2}$$

where $\boldsymbol{K_0} = \rho_0 \left(c_0^2 \mathbf{Id} - \boldsymbol{v_0} \boldsymbol{v_0}^T\right)$.
It is easy to prove that

Lemma 3.1.1:
$\boldsymbol{K_0}$ is symmetric positive-definite and

$$\mathrm{Sp}(\boldsymbol{K_0}) = \left\{\rho_0 c_0^2, \rho_0 (c_0^2 - |\boldsymbol{v_0}|^2)\right\}$$

Proof: $\boldsymbol{K_0 v_0} = \rho_0 \left(c_0^2 - |\boldsymbol{v_0}|^2\right) \boldsymbol{v_0}$ and $\boldsymbol{K_0 u} = \rho_0 c_0^2 \boldsymbol{u}$ for all $\boldsymbol{u} \in \boldsymbol{v_0}^\perp$.

Fredholm type : If the background flow is subsonic, ie.

$$\inf_{\mathcal{O}} \left(c_0^2 - |\boldsymbol{v_0}|^2\right) > 0, \tag{3.3}$$

then (3.2) leads to a problem of Fredholm type. Indeed, by using Lemma 3.1.1 we can conclude that $-\mathrm{div}\left(\boldsymbol{K_0} \nabla p\right)$ is a coercive operator, and that the convected Helmholtz equation therefore has a *coercive + compact* structure.

Boundary conditions: Let Γ be the boundary of the domain $\mathcal{O}$ and let $\boldsymbol{n}$ be the outward-facing normal vector.
We will use the following boundary conditions

$$\begin{aligned}
&\text{Neumann:} && (\boldsymbol{K_0} \nabla p) \cdot \boldsymbol{n} + 2i\omega(\rho_0 \boldsymbol{v_0} \cdot \boldsymbol{n}) p = g_N && \text{on } \Gamma_N && \text{(3.4a)}\\
&\text{Dirichlet:} && p = g_D && \text{on } \Gamma_D && \text{(3.4b)}\\
&\text{Impedance:} && (\boldsymbol{K_0} \nabla p) \cdot \boldsymbol{n} + \mathcal{Z} p = g_I && \text{on } \Gamma_I && \text{(3.4c)}
\end{aligned}$$

and

$$\Gamma = \Gamma_N \cup \Gamma_D \quad \Gamma_D \cap \Gamma_N = \varnothing.$$

Remark 3.1.1: In this report, we will only consider Dirichlet (3.4b) and Neumann (3.4a) boundary conditions. Impedance boundary condition (3.4c) is useful to consider local *absorbing boundary conditions* which will be considered in a future work.

3.1.1 First-order formulations

As it is usually done in the framework of HDG methods, we will rewrite (3.2) as a first-order in space system. Notice that we have chosen to keep a second-order dependance in frequency. Adaptation of our method to a first-order in frequency formulation is straightforward.
We will compare two different ways to reach a first-order in space formulation.

To lighten the notations in the remaining of this paper, we introduce the following vector field

$$\boldsymbol{b_0} := \rho_0 \boldsymbol{v_0},$$

that satisfies the following mass conservation equation

$$\mathrm{div}\left(\boldsymbol{b_0}\right) = 0.$$

Diffusive flux formulation:

We begin by introducing the *diffusive flux*

$$\boldsymbol{q} := -\boldsymbol{K_0}\nabla p$$

as a new unknown, leading to the following first-order in space system

$$\boldsymbol{W_0 q} + \nabla p = 0 \tag{3.5a}$$
$$-\rho_0\omega^2 p - 2i\omega \boldsymbol{b_0}\cdot\nabla p + \operatorname{div}(\boldsymbol{q}) = s \tag{3.5b}$$

where

$$\boldsymbol{W_0} := \boldsymbol{K_0}^{-1} = \frac{1}{\rho_0 c_0^2}\left[\mathbf{Id} + \frac{\boldsymbol{v_0}\boldsymbol{v_0}^T}{c_0^2 - |\boldsymbol{v_0}|^2}\right]. \tag{3.6}$$

Note that $\boldsymbol{K_0}$ is always invertible thanks to (3.3), indeed we have

$$\det \boldsymbol{K_0} = \rho_0 c_0^2\left(c_0^2 - |\boldsymbol{v_0}|^2\right) \neq 0.$$

The second equality in (3.6) comes from the *Sherman-Morrison formula*, see [SM50] :

Lemma 3.1.2:
If $\boldsymbol{A} \in GL_n(\mathbb{R})$ and $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$, then $\boldsymbol{A} + \boldsymbol{u}\boldsymbol{v}^T$ is invertible if and only if $1 + \boldsymbol{v}^T\boldsymbol{A}^{-1}\boldsymbol{u} \neq 0$ and

$$\left(\boldsymbol{A} + \boldsymbol{u}\boldsymbol{v}^T\right)^{-1} = \boldsymbol{A}^{-1} - \frac{\boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{v}^T\boldsymbol{A}^{-1}}{1 + \boldsymbol{v}^T\boldsymbol{A}^{-1}\boldsymbol{u}}.$$

With this formulation, the Neumann boundary condition (3.4a) becomes

$$\boldsymbol{q}\cdot\boldsymbol{n} - 2i\omega(\boldsymbol{b_0}\cdot\boldsymbol{n})p = -g_N.$$

Variational formulation: We can now write a variational formulation for (3.5a)–(3.5b) : Seek $(\boldsymbol{q}, p) \in \boldsymbol{H}_{\operatorname{div}}(\mathcal{O}) \times H^1(\mathcal{O})$ such that for all $(\boldsymbol{r}, w) \in \boldsymbol{H}_{\operatorname{div}}(\mathcal{O}) \times H^1(\mathcal{O})$

$$\int_{\mathcal{O}} \boldsymbol{W_0 q}\cdot\boldsymbol{r}^*\mathrm{d}\boldsymbol{x} - \int_{\mathcal{O}} p\operatorname{div}(\boldsymbol{r}^*)\,\mathrm{d}\boldsymbol{x} = -\int_{\partial\mathcal{O}} p\boldsymbol{r}^*\cdot\boldsymbol{n}\mathrm{d}\sigma \tag{3.7a}$$
$$\begin{aligned}-\omega^2\int_{\mathcal{O}}\rho_0 p w^*\mathrm{d}\boldsymbol{x} + 2i\omega\int_{\mathcal{O}} p\boldsymbol{b_0}\cdot\nabla w^*\mathrm{d}\boldsymbol{x} - \int_{\mathcal{O}}\boldsymbol{q}\cdot\nabla w^*\mathrm{d}\boldsymbol{x}\\ + \int_{\partial\mathcal{O}} w^*\boldsymbol{q}\cdot\boldsymbol{n} - 2i\omega p w^*\boldsymbol{b_0}\cdot\boldsymbol{n}\mathrm{d}\sigma = \int_{\mathcal{O}} s w^*\mathrm{d}\boldsymbol{x}\end{aligned} \tag{3.7b}$$

where the boundary integrals should formally be interpreted as the duality bracket $\langle\cdot,\cdot\rangle_{H^{-\frac{1}{2}}(\partial\mathcal{O}),H^{\frac{1}{2}}(\partial\mathcal{O})}$ between $H^{-\frac{1}{2}}(\partial\mathcal{O})$ and $H^{\frac{1}{2}}(\partial\mathcal{O})$. Notice that we have not enforced the boundary conditions in this weak formulation to keep it closer to he *local problems* of the HDG methods.

Total flux formulation:

As $\operatorname{div}(\boldsymbol{b_0}) = 0$, we notice that

$$2i\omega\boldsymbol{b_0}\cdot\nabla p = \operatorname{div}(2i\omega p\boldsymbol{b_0}),$$

and we can therefore rewrite (3.2) as

$$-\rho_0\omega^2 p - \operatorname{div}(\boldsymbol{K_0}\nabla p + 2i\omega p\boldsymbol{b_0}) = s.$$

This leads to another possible first-order in space formulation. We introduce the *total flux*

$$\boldsymbol{\sigma} := -\boldsymbol{K_0}\nabla p - 2i\omega p\boldsymbol{b_0},$$

leading to the following system

$$\boldsymbol{W_0}\boldsymbol{\sigma} + \nabla p + 2i\omega p\boldsymbol{W_0}\boldsymbol{b_0} = 0, \tag{3.8a}$$
$$-\rho_0\omega^2 p + \operatorname{div}(\boldsymbol{\sigma}) = s. \tag{3.8b}$$

With this formulation, the Neumann boundary condition (3.4a) becomes

$$\boldsymbol{\sigma}\cdot\boldsymbol{n} = -g_N.$$

Variational formulation: We can now write a variational formulation for (3.8a)–(3.8b) : Seek $(\boldsymbol{\sigma}, p) \in \boldsymbol{H}_{\mathrm{div}}(\mathcal{O}) \times H^1(\mathcal{O})$ such that for all $(\boldsymbol{r}, w) \in \boldsymbol{H}_{\mathrm{div}}(\mathcal{O}) \times H^1(\mathcal{O})$

$$\int_{\mathcal{O}} \boldsymbol{W_0}\boldsymbol{\sigma}\cdot\boldsymbol{r}^*\mathrm{d}\boldsymbol{x} - \int_{\mathcal{O}} p\operatorname{div}(\boldsymbol{r}^*)\,\mathrm{d}\boldsymbol{x} + 2i\omega\int_{\mathcal{O}} p\boldsymbol{W_0}\boldsymbol{b_0}\cdot\boldsymbol{r}^*\mathrm{d}\boldsymbol{x} = -\int_{\partial\mathcal{O}} p\boldsymbol{r}^*\cdot\boldsymbol{n}\mathrm{d}\sigma \tag{3.9a}$$
$$-\omega^2\int_{\mathcal{O}}\rho_0 p w^*\mathrm{d}\boldsymbol{x} - \int_{\mathcal{O}}\boldsymbol{\sigma}\cdot\nabla w^*\mathrm{d}\boldsymbol{x} + \int_{\partial\mathcal{O}} w^*\boldsymbol{\sigma}\cdot\boldsymbol{n}\mathrm{d}\sigma = \int_{\mathcal{O}} s w^*\mathrm{d}\boldsymbol{x} \tag{3.9b}$$

where the boundary integrals should formally be interpreted as the duality bracket $\langle\cdot,\cdot\rangle_{H^{-\frac{1}{2}}(\partial\mathcal{O}),H^{\frac{1}{2}}(\partial\mathcal{O})}$ between $H^{-\frac{1}{2}}(\partial\mathcal{O})$ and $H^{\frac{1}{2}}(\partial\mathcal{O})$. Notice that we have not enforced the boundary conditions in this weak formulation to keep it closer to he *local problems* of the HDG methods.

3.2 Notations

In this section, we introduce the notations and approximation spaces that will be used to construct the HDG methods considered in this paper.

3.2.1 Approximation spaces

We consider a mesh $\mathcal{T}_h$ of the domain $\mathcal{O}$ of dimension n. For an element $K \in \mathcal{T}_h$, we denote by $\mathcal{E}(K)$ the set of its edges. We also consider

$$\begin{aligned}
\text{The set of boundary edges:} &\quad \mathcal{E}_h^b := \{e = \partial K \cap \Gamma \mid K \in \mathcal{T}_h\},\\
\text{The set of interior edges:} &\quad \mathcal{E}_h^i := \{e = \partial K_+ \cap \partial K_- \mid K_+, K_- \in \mathcal{T}_h\},\\
\text{The set of all edges:} &\quad \mathcal{E}_h := \mathcal{E}_h^b \cup \mathcal{E}_h^i.
\end{aligned}$$

To study the convergence of the methods, we will assume that the mesh has the usual *shape-regularity* property, see [EG04, Def. 1.107].

For $K \in \mathcal{T}_h$, we denote by $\mathcal{P}_k(K)$ the space of polynomials of total degree at most k defined on K. We will also use the space of vectorial polynomials $\boldsymbol{\mathcal{P}}_k(K) = \mathcal{P}_k(K)^n$. Even if those spaces can be defined for $k > 0$, in this paper we will usually assume that $k \geqslant 2$ as HDG method of lower order have little interest from a computational point of view. Indeed, for $k \leqslant 2$, there is no internal degrees of freedom to eliminate. The static condensation process leads to a problem with only one scalar unknown and therefore eliminates the local vectorial unknown. The HDG methods with $k \leqslant 2$ are therefore mixed methods with a numerical cost similar to primal methods.

On each element $K \in \mathcal{T}_h$, we introduce the following approximation spaces for the pressure and the flux

$$\begin{aligned} \boldsymbol{V}_h(K) &:= \left\{ \boldsymbol{q} \in \boldsymbol{L}^2(K) \;\middle|\; \boldsymbol{q}|_K \in \boldsymbol{\mathcal{P}}_k(K) \right\} && \text{for the flux } \boldsymbol{q}_h \text{ or } \boldsymbol{\sigma}_h, \\ W_h(K) &:= \left\{ p \in L^2(K) \;\middle|\; p|_K \in \mathcal{P}_\ell(K) \right\} && \text{for the pressure } p_h. \end{aligned}$$

In this work we will consider both standard HDG formulation where $\ell = k$ and the so-called HDG+ formulation where $\ell = k + 1$.

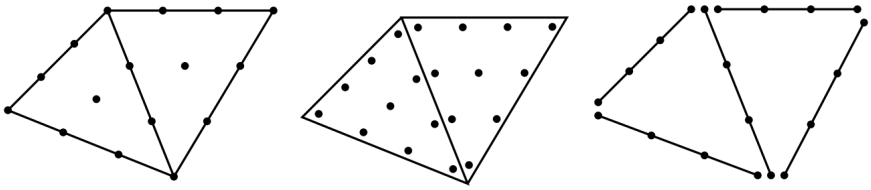
To construct HDG formulations, we will need to add a surfacic unknown, called the *numerical trace* and denoted by $\widehat{p_h}$, to the problem. This unknown will be the main unknown of the method as the *static condensation* process will allow to eliminate the local unknowns to obtain a so-called *global problem*. To approximate this new unknown we introduce the following space for $e \in \mathcal{E}(K)$

$$M_h(e) := \left\{ \mu \in L^2(e) \;\middle|\; \mu|_e \in \mathcal{P}_k(e) \right\}.$$

As those approximation spaces are discontinuous, we can construct the *global approximation spaces* as the cartesian product of the local ones

$$\begin{aligned} \boldsymbol{V}_h &:= \prod_{K \in \mathcal{T}_h} \boldsymbol{V}_h(K) && \text{for the flux } \boldsymbol{q}_h \text{ or } \boldsymbol{\sigma}_h, \\ W_h &:= \prod_{K \in \mathcal{T}_h} W_h(K) && \text{for the pressure } p_h, \\ M_h &:= \prod_{e \in \mathcal{E}_h} M_h(e) && \text{for the trace } \widehat{p_h}. \end{aligned}$$

In FIGURE 3.1, we have depicted the differences in the *degrees of freedom* for the continuous (CG), discontinuous (DG) and hybridizable discontinuous (HDG) Galerkin methods. The degrees of freedom of the HDG methods are the ones associated with the numerical trace $\widehat{p_h}$. As the numerical cost of the method is directly linked to the number of degrees of freedom, we can clearly see that the HDG method is less expensive than the DG method. We would like to recall one of the main advantages of mixed methods: as they are based on a first-order formulation, they can be used to obtain an approximation of the gradient of the unknown without the loss of order associated with the numerical evaluation of the derivatives. However mixed methods are usually more expensive than primal methods from a computational point of view and mixed DG methods usually have a prohibitive numerical cost. HDG methods have the advantages of both DG and mixed methods for a reasonable computational cost.



(a) CG (b) DG (c) HDG

Figure 3.1: Polynomial interpolation of degree 3

In TABLE 3.1, we give a summary of the choice of local spaces for the different variations of the HDG method considered in this report.

Variable	Space	HDG $(p_h, \boldsymbol{\sigma})$	HDG $(p_h, \boldsymbol{q}_h)$	HDG+ $(p_h, \boldsymbol{q}_h)$
Pressure p_h	$W_h(K)$	$\mathcal{P}_k(K)$		$\mathcal{P}_{k+1}(K)$
Flux $\boldsymbol{q}_h$ or $\boldsymbol{\sigma}_h$	$\boldsymbol{V}_h(K)$	$\boldsymbol{\mathcal{P}}_k(K)$		
Trace $\widehat{p_h}$	$M_h(e)$	$\mathcal{P}_k(e)$		

Table 3.1: Choice of local spaces and penalization parameter for the different methods

Let us now introduce the orthogonal projection on the space of piecewise polynomial functions on the edges

$$P_M : \prod_{K\in\mathcal{T}_h} L^2(\partial K) \longrightarrow \mathcal{R}_k(\partial\mathcal{T}_h) := \prod_{K\in\mathcal{T}_h}\prod_{e\in\mathcal{E}(K)} \mathcal{P}_k(e),$$

which will be required to construct the HDG+ formulation.
It is important to emphasize the difference between the space $\mathcal{R}_k(\partial\mathcal{T}_h)$ and M_h. Indeed as

$$M_h = \prod_{e\in\mathcal{E}_h} \mathcal{P}_k(e) \neq \prod_{K\in\mathcal{T}_h}\prod_{e\in\mathcal{E}(K)} \mathcal{P}_k(e),$$

the functions in M_h are single-valued on the skeleton of the mesh, whereas the functions in the other space are multi-valued on the interior edges. Functions in both spaces are discontinuous at the vertices.

Remark 3.2.1: It is also possible to choose a continuous space for $\widehat{p_h}$, this leads to the so-called *Locally Discontinuous but Globally Continuous method* (LDGC), see *eg.* [ALA13, FLd14]. However this choice does not seem to improve the convergence rate of the method.

3.2.2 Hermitian products and norms

For an element $K \in \mathcal{T}_h$, we denote the standard L^2-hermitian product[1] and its associated norm by

$$(u,v)_K := \int_K u \cdot v^* \mathrm{d}\boldsymbol{x} \quad \text{and} \quad \|u\|_K^2 := (u,u)_K\,,$$

we then introduce the broken hermitian product and norm

$$(u,v)_{\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h} (u,v)_K \quad \text{and} \quad \|u\|_{\mathcal{T}_h}^2 := \sum_{K\in\mathcal{T}_h} \|u\|_K^2\,.$$

On the boundary of an element K, we also denote the local hermitian product by

$$\langle u,v\rangle_{\partial K} := \sum_{e\in\mathcal{E}(K)} \int_e u \cdot v^* \mathrm{d}\sigma \quad \text{and} \quad \|u\|_{\partial K}^2 := \langle u,u\rangle_{\partial K}\,,$$

and the broken hermitian product is denoted by

$$\langle u,v\rangle_{\partial\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h} \langle u,v\rangle_{\partial K} \quad \text{and} \quad \|u\|_{\partial\mathcal{T}_h}^2 := \sum_{K\in\mathcal{T}_h} \|u\|_{\partial K}^2\,.$$

Here we would like to point out that, depending on the regularity of u and v, $\langle\cdot,\cdot\rangle_{\partial K}$ can denote either the hermitian product of $L^2(\partial K)$ or the duality bracket between $H^{-\frac{1}{2}}(\partial K)$ and $H^{\frac{1}{2}}(\partial K)$.

[1] For vector fields, the $\mathbb{R}^n$ dot-product is used inside the integral as the conjugate is already applied.

We also define the following weighted norms

$$\|u\|^2_{\rho_0,K} := (\rho_0 u, u)_K \qquad \text{which satisfies} \qquad \|u\|_{\rho_0,K} \leqslant \|\rho_0\|^{\frac{1}{2}}_{L^\infty(K)} \|u\|_K$$
$$\|\boldsymbol{q}\|^2_{\boldsymbol{W_0},K} := (\boldsymbol{W_0 q}, \boldsymbol{q})_K \qquad \text{which satisfies} \qquad \|\boldsymbol{q}\|_{\boldsymbol{W_0},K} \leqslant C_{\boldsymbol{W_0},K} \|\boldsymbol{q}\|_K$$

where

$$C_{\boldsymbol{W_0},K} = \left(\max_K \frac{1}{\rho_0 \left(c_0^2 - |\boldsymbol{v_0}|^2\right)} \right)^{\frac{1}{2}}$$

is the square root of the largest eigenvalue of $\boldsymbol{W_0}$ in K, see LEMMA 3.1.1.

3.2.3 Faces, jumps and averages

In this subsection, we will introduce notations for the face quantities. As usual with methods belonging to the DG family, we will need to define jumps and averages which link the unknowns between two elements.

Faces and normals: For an interior face $\mathcal{E}_h^i \ni e = \partial K_+ \cap \partial K_-$, we denote by $\boldsymbol{n}^+$ (resp. $\boldsymbol{n}^-$) a unitary outgoing normal vector of ∂K_+ (resp. ∂K_-). We will always assume that the flow $\boldsymbol{v_0}$ goes from K_- to K_+, as depicted on FIGURE 3.2.

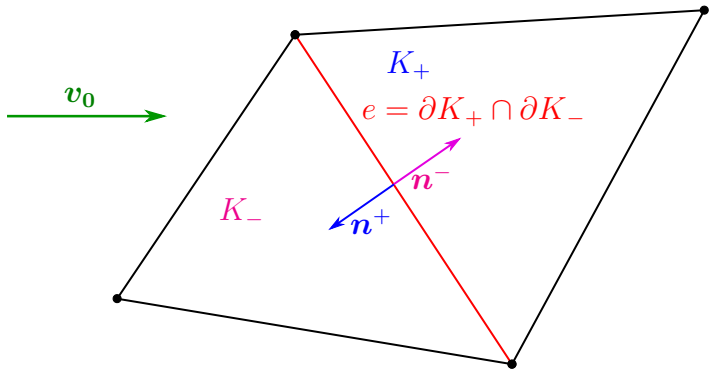


Figure 3.2: Outgoing normal vectors on an interior face

When the orientation of the face does not matter, we will denote by $\boldsymbol{n}$ any unitary normal vector to e.
If e is a boundary edge, then $\boldsymbol{n}$ denotes the outward-pointing unitary normal vector.

Jumps and averages: We will often use the *average operator* defined by

$$\text{On } \mathcal{E}_h^i \ni e = \partial K_+ \cap \partial K_-, \qquad \{\!\{\varphi\}\!\}_e := \frac{1}{2}\left(\varphi^+ + \varphi^-\right),$$
$$\text{On } \mathcal{E}_h^b \ni e = \partial K \cap \Gamma, \qquad \{\!\{\varphi\}\!\}_e := \frac{1}{2}\varphi,$$

where φ can either be a scalar or vectorial quantity.
We will also make frequent use of the *jump operator* defined by

$$\text{On } \mathcal{E}_h^i \ni e = \partial K_+ \cap \partial K_-, \qquad [\![\boldsymbol{q}]\!]_e := \boldsymbol{q}^+ \cdot \boldsymbol{n}^+ + \boldsymbol{q}^- \cdot \boldsymbol{n}^-,$$
$$\text{On } \mathcal{E}_h^b \ni e = \partial K \cap \Gamma, \qquad [\![\boldsymbol{q}]\!]_e := \boldsymbol{q} \cdot \boldsymbol{n},$$

for a vectorial quantity. Notice that with this definition, the jump operator only controls the normal part of the vector. For a scalar quantity, the *jump operator* is defined by

$$\begin{aligned} &\text{On } \mathcal{E}_h^i \ni e = \partial K_+ \cap \partial K_-, && [\![p]\!]_e := p^+ - p^-, \\ &\text{On } \mathcal{E}_h^b \ni e = \partial K \cap \Gamma, && [\![p]\!]_e := p, \end{aligned}$$

for a scalar quantity. A sketch of those quantities is given in FIGURE 3.3.

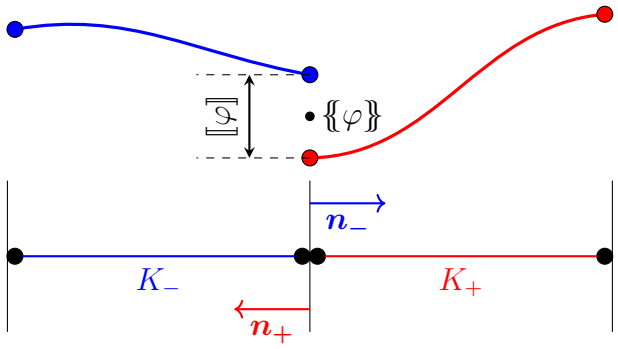


Figure 3.3: 1D-sketch of the jump and average on an interior node

3.3 HDG method for the total flux formulation

In this section, we will focus on the *total flux formulation*. We will first construct the HDG method and we will then discuss its most important properties.

3.3.1 Constructing the formulation

On an element $K \in \mathcal{T}_h$, the construction of the HDG method relies on the following integration by parts formula, which is related to the weak formulation (3.9a)–(3.9b)

$$\int_K \boldsymbol{W_0}\boldsymbol{\sigma} \cdot \boldsymbol{r}^* \mathrm{d}\boldsymbol{x} - \int_K p \mathrm{div}\,(\boldsymbol{r}^*)\,\mathrm{d}\boldsymbol{x} + 2i\omega \int_K p \boldsymbol{W_0}\boldsymbol{b_0} \cdot \boldsymbol{r}^* \mathrm{d}\boldsymbol{x} = -\langle p, \boldsymbol{r}\cdot\boldsymbol{n}\rangle_{\partial K}\,, \quad (3.11a)$$

$$-\omega^2 \int_K \rho_0 p w^* \mathrm{d}\boldsymbol{x} + \int_K \mathrm{div}\,(\boldsymbol{\sigma})\, w^* \mathrm{d}\boldsymbol{x} = \int_K s w^* \mathrm{d}\boldsymbol{x}, \quad (3.11b)$$

for all $(\boldsymbol{r}, w) \in \boldsymbol{H}_{\mathrm{div}}(K) \times H^1(K)$ and where $(\boldsymbol{\sigma}, p) \in \boldsymbol{H}_{\mathrm{div}}(\mathcal{O}) \times H^1(\mathcal{O})$ is the unknown.

Choice of approximation spaces: We denote by $\boldsymbol{\sigma}_h$ and p_h the approximations of $\boldsymbol{\sigma}$ and p on K.
For this method, we choose to use the following local approximation spaces

$$\begin{aligned} \boldsymbol{V}_h(K) &= \boldsymbol{\mathcal{P}}_k(K) && \text{for the flux } \boldsymbol{\sigma}_h, \\ W_h(K) &= \mathcal{P}_k(K) && \text{for the potential } p_h, \end{aligned}$$

where $k \geqslant 3$ is the degree of the method. We recall that if $k \leqslant 2$, then there are no interior degrees of freedom and the HDG method has little interest over the DG methods from a computational point of view. Notice that we use the same interpolation degree for both unknowns, which may lead to unstable continuous Galerkin methods, see *e.g.* [EG04, "Checkerboard-like instablitiy" p.188]. As we will discuss later, this is not a problem for HDG methods.

Introduction of the hybrid unknown: To reach a HDG formulation, we introduce a new unknown $\widehat{p_h}$ which is an approximation of p on $\mathcal{E}_h$, the skeleton of the mesh $\mathcal{T}_h$. We will usually refer to $\widehat{p_h}$ as the *numerical trace*. This unknown is the main unknown of the HDG method. Indeed, we will be able to use a *static condensation process* to eliminate the interior degrees of freedom and to obtain a so-called *global problem* for $\widehat{p_h}$ only. To introduce this unknown in the formulation, the boundary integral in (3.11a) is discretized as follows

$$\int_{\partial K} p\boldsymbol{r}^* \cdot \boldsymbol{n} \mathrm{d}\sigma \quad \text{becomes} \quad \int_{\partial K} \widehat{p_h}\boldsymbol{r}_h^* \cdot \boldsymbol{n} \mathrm{d}\sigma.$$

For this new unknown, we will use the following approximation space

$$M_h(e) = \mathcal{P}_k(e), \quad \forall e \in \mathcal{E}(K).$$

Penalization parameter: The unknown $\widehat{p_h}$ is often called a *Lagrange multiplier*. Indeed, when going from a continuous Galerkin method to a HDG one, the continuity of the numerical solution is not strongly enforced anymore and it is added in the method as a constraint. The quantity $\widehat{p_h}$ is therefore the Lagrange multiplier that enforces this weak continuity requirement.
To enforce this constraint, we introduce a *penalization parameter* denoted by τ, and the following boundary term

$$\langle \tau(p_h - \widehat{p_h}), w_h \rangle_{\partial K}$$

will be added to the local problem. This boundary term can be interpreted as a weak enforcement of the following Dirichlet boundary condition

$$p_h = \widehat{p_h}, \quad \text{on } \partial K.$$

Practical choice of τ will be discussed in Section 3.3.2.

Local problem: Using the integration by parts formula (3.11a)–(3.11b) on an element $K \in \mathcal{T}_h$, we define the *local problem* : seek $(\boldsymbol{\sigma}_h, p_h) \in \boldsymbol{V}_h(K) \times W_h(K)$ such that

$$(\boldsymbol{W_0}\boldsymbol{\sigma}_h, \boldsymbol{r}_h)_K - (p_h, \mathrm{div}\,(\boldsymbol{r}_h))_K + 2i\omega\,(p_h\boldsymbol{W_0}\boldsymbol{b_0}, \boldsymbol{r}_h)_K + \langle \widehat{p_h}, \boldsymbol{r}_h \cdot \boldsymbol{n} \rangle_{\partial K} = 0, \quad (3.12a)$$
$$-\omega^2\,(\rho_0 p_h, w_h)_K + (\mathrm{div}\,(\boldsymbol{\sigma}_h)\,, w_h)_K + i\omega\,\langle \tau(p_h - \widehat{p_h}), w_h \rangle_{\partial K} - (s, w_h)_K = 0, \quad (3.12b)$$

for all $(\boldsymbol{r}_h, w_h) \in \boldsymbol{V}_h(K) \times W_h(K)$.
Notice that (3.12a)–(3.12b) is the variational formulation of the convected Helmholtz equation on K with weak Dirichlet boundary conditions on ∂K, we can emphasize this fact by rewriting (3.12a) as

$$(\boldsymbol{W_0}\boldsymbol{\sigma}_h, \boldsymbol{r}_h)_K + (\nabla p_h, \boldsymbol{r}_h)_K + 2i\omega\,(p_h\boldsymbol{W_0}\boldsymbol{b_0}, \boldsymbol{r}_h)_K + \langle \widehat{p_h} - p_h, \boldsymbol{r}_h \cdot \boldsymbol{n} \rangle_{\partial K} = 0.$$

However we have to chosen keep (3.12a) as it is the usual way to implement this local equation.

Transmission condition: Due to the discontinuous nature of the approximation spaces, we need to link all the local problems together. To this end, we introduce the *numerical flux* for $\boldsymbol{\sigma}_h$

$$\widehat{\boldsymbol{\sigma}_h} \cdot \boldsymbol{n} := \boldsymbol{\sigma}_h \cdot \boldsymbol{n} + i\omega\tau(p_h - \widehat{p_h}), \quad (3.13)$$

which satisfies the following weak continuity equation

$$\langle \widehat{\boldsymbol{\sigma}_h}, \mu_h \rangle_{\partial \mathcal{T}_h \backslash \Gamma_D} = \sum_{e \in \mathcal{E}_h^i} \langle [\![\widehat{\boldsymbol{\sigma}_h}]\!], \mu_h \rangle_e = 0, \tag{3.14}$$

as $\boldsymbol{\sigma}_h \in \boldsymbol{V}_h$, $p_h \in W_h$ and $\widehat{p_h} \in M_h$, we have $[\![\widehat{\boldsymbol{\sigma}_h}]\!] \in M_h$ and we therefore conclude that $[\![\widehat{\boldsymbol{\sigma}_h}]\!] = 0$, as $[\![\widehat{\boldsymbol{\sigma}_h}]\!]$ is a polynomial of degree up to k orthogonal to all polynomials of degree up to k. We recall that on an interior edge $\mathcal{E}_h^i \ni e = \partial K_+ \cap \partial K_-$, the jump operator is defined as

$$[\![\boldsymbol{\sigma}_h]\!] := \boldsymbol{\sigma}_h^+ \cdot \boldsymbol{n}^+ + \boldsymbol{\sigma}_h^- \cdot \boldsymbol{n}^-.$$

This weak continuity requirement can be expressed as the following *transmission condition*

$$\langle \widehat{\boldsymbol{\sigma}_h}, \mu_h \rangle_{\partial \mathcal{T}_h \backslash \Gamma_D} + \langle \widehat{p_h} - g_D, \mu_h \rangle_{\Gamma_D} = \langle g_N, \mu_h \rangle_{\Gamma_N} \tag{3.15}$$

for all $\mu_h \in M_h$.
Notice that (3.15) enforces the normal continuity of $\widehat{\boldsymbol{\sigma}_h}$ on the interior faces as well as the Neumann and Dirichlet boundary conditions on Γ_N and Γ_D.

Remark 3.3.1: The transmission condition (3.15) can be understood as a weak requirement of $\boldsymbol{H}_{\mathrm{div}}(\mathcal{O})$-conformity. Indeed it is shown in [PE12, Lemma 1.2.4] that $\boldsymbol{\sigma}_h \in \boldsymbol{H}_{\mathrm{div}}(\mathcal{O})$ means

$$\forall K \in \mathcal{T}_h,\ \boldsymbol{\sigma}_h^K \in \boldsymbol{H}_{\mathrm{div}}(K) \quad \text{and} \quad \forall e \in \mathcal{E}_h^i,\ [\![\boldsymbol{\sigma}_h]\!]_e \equiv 0.$$

The former is a consequence of the polynomial nature of the approximation spaces, and we will now focus and the latter. Owing to the transmission condition, we have

$$\forall e \in \mathcal{E}_h^i,\ 0 = [\![\widehat{\boldsymbol{\sigma}_h}]\!] = [\![\boldsymbol{\sigma}_h]\!] + i\omega [\![\tau (p_h - \widehat{p_h})]\!].$$

As p_h and $\widehat{p_h}$ are two approximations of the same unknown p, the quantity $p_h - \widehat{p_h}$ is expected to be small. We can therefore conclude that $[\![\boldsymbol{\sigma}_h]\!]$ is small and that

$$[\![\boldsymbol{\sigma}_h]\!] \underset{h_K \to 0}{\longrightarrow} 0.$$

For applications where a precise approximation of the flux is required, a post-processing scheme producing a new approximate $\widetilde{\boldsymbol{\sigma}_h}$ with strong $\boldsymbol{H}_{\mathrm{div}}$-conformity was introduced in [CGS10, Sec. 5.1].

Compact formulation:

HDG methods are usually stated in a compact form that can be obtained by summing the local problems (3.12a)–(3.12b) over the mesh elements and by adding the transmission condition (3.15). This formulation reads : seek $(\boldsymbol{\sigma}_h, p_h, \widehat{p_h}) \in \boldsymbol{V}_h \times W_h \times M_h$ such that

$$(\boldsymbol{W_0}\boldsymbol{\sigma}_h, \boldsymbol{r}_h)_{\mathcal{T}_h} - (p_h, \mathrm{div}\,(\boldsymbol{r}_h))_{\mathcal{T}_h} + 2i\omega\,(p_h \boldsymbol{W_0}\boldsymbol{b_0}, \boldsymbol{r}_h)_{\mathcal{T}_h} + \langle \widehat{p_h}, \boldsymbol{r}_h \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} = 0, \tag{3.16a}$$

$$-\omega^2\,(\rho_0 p_h, w_h)_{\mathcal{T}_h} + (\mathrm{div}\,(\boldsymbol{\sigma}_h)\,, w_h)_{\mathcal{T}_h} + i\omega\,\langle \tau(p_h - \widehat{p_h}), w_h \rangle_{\partial \mathcal{T}_h} = (s, w_h)_{\mathcal{T}_h} \tag{3.16b}$$

$$\langle \boldsymbol{\sigma}_h \cdot \boldsymbol{n} + i\omega\tau(p_h - \widehat{p_h}), \mu_h \rangle_{\partial \mathcal{T}_h \backslash \Gamma_D} + \langle \widehat{p_h} - g_D, \mu_h \rangle_{\Gamma_D} = \langle g_N, \mu_h \rangle_{\Gamma_N} \tag{3.16c}$$

for all $(\boldsymbol{r}_h, w_h, \mu_h) \in \boldsymbol{V}_h \times W_h \times M_h$. This formulation will be useful to perform the numerical analysis of the method. We recall that (3.16c) should be understood as a rewriting of the weak continuity requirement (3.14).

Remark 3.3.2: At this point, to completely define the HDG method, it only remains to choose the penalization parameter τ, this will be done in the next section.

Condensed variational formulation

The compact formulation (3.16a)–(3.16b)–(3.16c) cannot directly be used to efficiently implement the HDG method. Indeed it is not clear how a formulation involving only $\widehat{p_h}$ can be reached. To emphasize how it can be done, we will now write a *condensed* variational formulation for $\widehat{p_h}$ only.
We introduce the so-called *local solvers*

$$\begin{aligned}
&\mathsf{P}^K : (\widehat{p_h}, s) \longmapsto p_h^K, \\
&\Sigma^K : (\widehat{p_h}, s) \longmapsto \boldsymbol{\sigma}_h^K, \\
&\widehat{\Sigma}^K : (\widehat{p_h}, s) \longmapsto \widehat{\boldsymbol{\sigma}_h}^K \cdot \boldsymbol{n}^K := \Sigma^K(\widehat{p_h}, s) \cdot \boldsymbol{n} + i\omega\tau \left(\mathsf{P}^K(\widehat{p_h}, s)\widehat{p_h}\right),
\end{aligned}$$

where $(\boldsymbol{\sigma}_h^K, p_h^K)$ is the solution of (3.12a)–(3.12b) and $\widehat{\boldsymbol{\sigma}_h}^K$ is defined by (3.13).
We can therefore rewrite the transmission condition (3.16c) as

$$a_h(\widehat{p_h}, \mu) = \ell_h(\mu), \tag{3.17}$$

where

$$\begin{aligned}
a_h(\widehat{p_h}, \mu_h) &:= \left\langle \Sigma^K(\widehat{p_h}, s) \cdot \boldsymbol{n} + i\omega\tau(\mathsf{P}^K(\widehat{p_h}, s) - \widehat{p_h}), \mu_h \right\rangle_{\partial\mathcal{T}_h \setminus \Gamma_D} + \langle \widehat{p_h}, \mu_h \rangle_{\Gamma_D}, \\
\ell_h(\mu_h) &:= \langle g_N, \mu_h \rangle_{\Gamma_N} + \langle g_D, \mu_h \rangle_{\Gamma_D}.
\end{aligned}$$

Equation (3.17) is the so-called *global problem* and is the main equation of the HDG method. From a computational point of view, we proceed as described in Algorithm 1.

Algorithm 1: Solving HDG-$\boldsymbol{\sigma}_h$

```
for K ∈ T_h do
    Construct the local solvers P^K, Σ^K, Σ̂^K
    Add local contribution to the global problem (3.17)
Solve the global problem (3.17) for p̂_h        // This is the most expensive step
for K ∈ T_h do
    Reconstruct the local unknowns p_h^K = P^k(p̂_h, s) and σ_h^K = Σ(p̂_h, s)
```

This algorithm is the blueprint of the practical implementation of the HDG method which will be discussed in Section 3.5.

3.3.2 Choice of penalization parameter

In this section, we will show how the penalization parameter can be chosen to obtain an upwinding mechanism with physical meaning. To do that, we will first rewrite the HDG method as a DG one, we will then solve a Riemann problem to obtain the value of τ.

DG formulation

In this section, we will rewrite the HDG method (3.16a)–(3.16b)–(3.16c) as a standard discontinuous Galerkin method.
We introduce the following bilinear form

$$\begin{aligned}
\mathcal{B}_h([\boldsymbol{\sigma}_h, p_h]; [\boldsymbol{r}_h, w_h]) := &(\boldsymbol{W_0}\boldsymbol{\sigma}_h, \boldsymbol{r}_h)_{\mathcal{T}_h} - (p_h, \operatorname{div}(\boldsymbol{r}_h))_{\mathcal{T}_h} + 2i\omega\,(p_h \boldsymbol{W_0}\boldsymbol{b_0}, \boldsymbol{r}_h)_{\mathcal{T}_h} \\
&- \omega^2\,(\rho_0 p_h, w_h)_{\mathcal{T}_h} + (\operatorname{div}(\boldsymbol{\sigma}_h), w_h)_{\mathcal{T}_h} \\
&+ \sum_{e \in \mathcal{E}_h} \left(\langle \widehat{p_h}, [\![\boldsymbol{r}_h]\!] \rangle_e + \langle \widehat{\boldsymbol{\sigma}_h} \cdot \boldsymbol{n}, [\![w_h]\!] \rangle_e \right),
\end{aligned} \tag{3.18}$$

from which all the mixed DG methods can be generated by choosing $\widehat{p_h}$ and $\widehat{\boldsymbol{\sigma}_h}$. Notice that $\widehat{p_h}$ was an unknown of the HDG method whereas it now should be chosen by the user of the method.
For example, the LDG method is obtained by choosing

$$\widehat{p_h} = \{\!\{p_h\}\!\} + \alpha[\![p_h]\!] \quad \text{and} \quad \widehat{\boldsymbol{\sigma}_h} = \{\!\{\boldsymbol{\sigma}_h\}\!\} + \beta[\![\boldsymbol{\sigma}_h]\!]\boldsymbol{n} + \gamma[\![p_h]\!]\boldsymbol{n},$$

and the DG method with central flux is obtained by choosing

$$\widehat{p_h} = \{\!\{p_h\}\!\} \quad \text{and} \quad \widehat{\boldsymbol{\sigma}_h} = \{\!\{\boldsymbol{\sigma}\}\!\} - \alpha[\![p_h]\!].$$

Proposition 3.3.1:
The HDG method (3.16a)–(3.16b)–(3.16c), with the following of numerical flux for $\boldsymbol{\sigma}_h$

$$\widehat{\boldsymbol{\sigma}_h} \cdot \boldsymbol{n}^{\pm} = \boldsymbol{\sigma}_h^{\pm} \cdot \boldsymbol{n}^{\pm} + i\omega\tau^{\pm}\left(p_h^{\pm} - \widehat{p_h}\right),$$

and the DG method associated to the bilinear form (3.18) are equivalent if and only if

$$\widehat{p_h} = \{\!\{p_h\}\!\} + \frac{\tau^+ - \tau^-}{2(\tau^+ + \tau^-)}[\![p_h]\!] + \frac{1}{i\omega(\tau^+ + \tau^-)}[\![\boldsymbol{\sigma}_h]\!], \tag{3.19a}$$

$$\widehat{\boldsymbol{\sigma}_h} \cdot \boldsymbol{n} = \{\!\{\boldsymbol{\sigma}_h\}\!\} \cdot \boldsymbol{n} + i\omega\frac{\tau^+\tau^-}{\tau^+ + \tau^-}[\![p_h]\!] - \frac{\tau^+ - \tau^-}{2(\tau^+ + \tau^-)}[\![\boldsymbol{\sigma}_h]\!], \tag{3.19b}$$

for all interior edges $\mathcal{E}_h^b \ni e = \partial K_+ \cap \partial K_-$ and where $\tau^{\pm} = \tau|_{\partial K_\pm}$.

This proposition is an intermediate result required to prove Proposition 3.3.2, where values of $\tau^{\pm}$ with physical meaning will be computed.

We recall the convention for the labelling $\partial K_\pm$: $\boldsymbol{v_0}$ is directed from ∂K_- toward ∂K_+ and we denote by $\boldsymbol{n}$ any normal vector to the face when the orientation does not matter.
Proof: Writing down the transmission condition (3.16c) on an interior face e, we have

$$[\![\boldsymbol{\sigma}_h]\!] + 2i\omega\left(\{\!\{\tau\}\!\}\{\!\{p_h\}\!\} + \frac{1}{4}[\![\tau]\!][\![p_h]\!]\right) - 2i\omega\{\!\{\tau\}\!\}\widehat{p_h} = 0,$$

which leads to

$$\widehat{p_h} = \{\!\{p_h\}\!\} + \frac{[\![\tau]\!]}{4\{\!\{\tau\}\!\}}[\![p_h]\!] + \frac{1}{2i\omega\{\!\{\tau\}\!\}}[\![\boldsymbol{\sigma}_h]\!],$$

and we obtain (3.19a) by developing the jumps and average terms.
As the numerical flux $\widehat{\boldsymbol{\sigma}_h}$ is continuous across the interface, we have

$$\begin{aligned}
\widehat{\boldsymbol{\sigma}_h} \cdot \boldsymbol{n} &= \{\!\{\widehat{\boldsymbol{\sigma}_h}\}\!\} \cdot \boldsymbol{n} \\
&= \{\!\{\boldsymbol{\sigma}_h\}\!\} \cdot \boldsymbol{n} + \frac{i\omega}{2}[\![\tau]\!]\left(\{\!\{p_h\}\!\} - \widehat{p_h}\right) + \frac{i\omega}{2}\{\!\{\tau\}\!\}[\![p_h]\!] \\
&= \{\!\{\boldsymbol{\sigma}_h\}\!\} \cdot \boldsymbol{n} - \frac{i\omega}{2}[\![\tau]\!]\left(\frac{[\![\tau]\!]}{4\{\!\{\tau\}\!\}}[\![p_h]\!] + \frac{1}{2i\omega\{\!\{\tau\}\!\}}[\![\boldsymbol{\sigma}_h]\!]\right) + \frac{i\omega}{2}\{\!\{\tau\}\!\}[\![p_h]\!]
\end{aligned}$$

and we obtain (3.19b) as

$$-\frac{[\![\tau]\!]^2}{4\{\!\{\tau\}\!\}} + \{\!\{\tau\}\!\} = -\frac{(\tau^+)^2 - 2\tau^+\tau^- + (\tau^-)^2}{2(\tau^+ + \tau^-)} + \frac{(\tau^+)^2 + 2\tau^+\tau^- + (\tau^-)^2}{2(\tau^+ + \tau^-)} = 2\frac{\tau^+\tau^-}{\tau^+ + \tau^-}.$$

We now have to show that this choice of numerical flux $\widehat{\boldsymbol{\sigma}_h}$ is compatible with the HDG method. Starting from (3.19b) we have

$$\begin{aligned}\widehat{\boldsymbol{\sigma}_h}\cdot\boldsymbol{n}^+ &= \boldsymbol{\sigma}_h^+\cdot\boldsymbol{n}^+ - \frac{1}{2}\left(1+\frac{\tau^+-\tau^-}{\tau^++\tau^-}\right)[\![\boldsymbol{\sigma}_h]\!] + i\omega\frac{\tau^+\tau^-}{\tau^++\tau^-}[\![p_h]\!]\\ &= \boldsymbol{\sigma}_h^+\cdot\boldsymbol{n}^+ - \frac{\tau^+}{\tau^++\tau^-}[\![\boldsymbol{\sigma}_h]\!] + i\omega\frac{\tau^+\tau^-}{\tau^++\tau^-}[\![p_h]\!],\end{aligned}$$

on the other hand, rewriting (3.19a) gives

$$\begin{aligned}-\frac{\tau^+}{\tau^++\tau^-}[\![\boldsymbol{\sigma}_h]\!] &= i\omega\tau^+\left[\{\!\{p_h\}\!\} - \widehat{p_h} + \frac{\tau^+-\tau^-}{2(\tau^++\tau^-)}[\![p_h]\!]\right]\\ &= i\omega\tau^+(p_h^+-\widehat{p_h}) + i\omega\frac{\tau^+}{2}\left(\frac{\tau^+-\tau^-}{\tau^++\tau^-}-1\right)[\![p_h]\!]\\ &= i\omega\tau^+(p_h^+-\widehat{p_h}) - i\omega\frac{\tau^+\tau^-}{\tau^++\tau^-}[\![p_h]\!],\end{aligned}$$

so we finally have

$$\widehat{\boldsymbol{\sigma}_h}\cdot\boldsymbol{n}^+ = \boldsymbol{\sigma}_h^+\cdot\boldsymbol{n}^+ + i\omega\tau^+(p_h^+-\widehat{p_h}).$$

Similar computations can be carried out on ∂K_-. ∎

Particular form of the HDG fluxes: We would like to point out that $\widehat{p_h}$ depends on $[\![\boldsymbol{\sigma}_h]\!]$, this is a distinctive feature of HDG methods among the family of DG methods. To understand this, let us consider DG method with the following fluxes

$$\widehat{p_h} = \{\!\{p_h\}\!\} + \alpha[\![p_h]\!] \quad \text{and} \quad \widehat{\boldsymbol{\sigma}_h} = \{\!\{\boldsymbol{\sigma}_h\}\!\} + \beta[\![\boldsymbol{\sigma}_h]\!]\boldsymbol{n} + \gamma[\![p_h]\!]\boldsymbol{n}, \tag{3.20}$$

where α,β,γ are arbitrary constants. This construction is adapted from [HW08, Sec 7.2.2]. Testing (3.18) with $[\boldsymbol{r}_h, 0]$ leads to

$$(\boldsymbol{W_0}\boldsymbol{\sigma}_h,\boldsymbol{r}_h)_{\mathcal{T}_h} - (p_h,\operatorname{div}(\boldsymbol{r}_h))_{\mathcal{T}_h} + 2i\omega\,(p_h\boldsymbol{W_0}\boldsymbol{b_0},\boldsymbol{r}_h)_{\mathcal{T}_h} + \sum_{e\in\mathcal{E}_h}\langle\widehat{p_h},[\![\boldsymbol{r}_h]\!]\rangle_e = 0.$$

Integrating by parts leads to

$$\begin{aligned}(\boldsymbol{W_0}\boldsymbol{\sigma}_h,\boldsymbol{r}_h)_{\mathcal{T}_h} = &-(\nabla p_h,\boldsymbol{r}_h)_{\mathcal{T}_h} - 2i\omega\,(p_h\boldsymbol{W_0}\boldsymbol{b_0},\boldsymbol{r}_h)_{\mathcal{T}_h} + \sum_{e\in\mathcal{E}_h}\left[\langle[\![p_h]\!]\boldsymbol{n},\{\!\{\boldsymbol{r}_h\}\!\}\rangle_e - \langle\widehat{p_h},[\![\boldsymbol{r}_h]\!]\rangle_e\right]\\ &+\sum_{e\in\mathcal{E}_h^i}\langle\{\!\{p_h\}\!\},[\![\boldsymbol{r}_h]\!]\rangle_e,\end{aligned} \tag{3.21}$$

where we used the identity

$$\langle p_h,\boldsymbol{r}_h\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = \sum_{e\in\mathcal{E}_h}\langle[\![p_h]\!]\boldsymbol{n},\{\!\{\boldsymbol{r}_h\}\!\}\rangle_e + \sum_{e\in\mathcal{E}_h^i}\langle\{\!\{p_h\}\!\},[\![\boldsymbol{r}_h]\!]\rangle_e,$$

coming from [HW08, Lemma 7.9]. Using the definition of $\widehat{p_h}$ given in (3.20), the surfacic terms in (3.21) become

$$-\sum_{e\in\mathcal{E}_h^i}\langle[\![p_h]\!],\{\!\{\boldsymbol{r}_h\}\!\}\cdot\boldsymbol{n}-\alpha[\![\boldsymbol{r}_h]\!]\rangle_e - \sum_{e\in\mathcal{E}_h^b}\langle p_h,\boldsymbol{r}_h\cdot\boldsymbol{n}\rangle_e.$$

We now introduce the *lifting operator* $\mathcal{L}$ defined by

$$(\mathcal{L}(p_h),\boldsymbol{r}_h)_{\mathcal{T}_h} = \sum_{e\in\mathcal{E}_h^i}\langle[\![p_h]\!],\{\!\{\boldsymbol{r}_h\}\!\}\cdot\boldsymbol{n}-\alpha[\![\boldsymbol{r}_h]\!]\rangle_e + \sum_{e\in\mathcal{E}_h^b}\langle p_h,\boldsymbol{r}_h\cdot\boldsymbol{n}\rangle_e,\quad\forall\boldsymbol{r}_h\in\boldsymbol{V}_h,$$

and (3.21) becomes

$$(\boldsymbol{W_0}\boldsymbol{\sigma}_h, \boldsymbol{r}_h)_{\mathcal{T}_h} = (-\nabla p_h - 2i\omega p_h \boldsymbol{W_0}\boldsymbol{b_0} - \mathcal{L}(p_h), \boldsymbol{r}_h)_{\mathcal{T}_h}\,.$$

We can see that $\boldsymbol{\sigma}_h$ is completely defined in terms of p_h and it is therefore not possible to add a transmission condition, which is required to allow the static condensation process. On the other hand, if the HDG flux

$$\widehat{p_h} = \{\!\{p_h\}\!\} + \alpha[\![p_h]\!] + \delta[\![\boldsymbol{\sigma}_h]\!]$$

is used, we will obtain an expression of $\boldsymbol{\sigma}_h$ in terms of p_h and $[\![\boldsymbol{\sigma}_h]\!]$. We will therefore need to add the transmission condition to close the discrete system and it will be possible to perform the static condensation.

Computing the penalization parameter

Proposition 3.3.2:
On an interior face $\mathcal{E}_h^i \ni e = \partial K_+ \cap \partial K_-$ the following choice of penalization parameter

$$\tau^{\pm} = \rho_0(c_0 + \boldsymbol{v_0}\cdot\boldsymbol{n}^{\pm}), \tag{3.22}$$

where $\tau^{\pm} = \tau|_{\partial K_{\pm}}$, leads to an upwinding mechanism.

To prove this proposition, we will need to solve a Riemann problem and compare its solution with Proposition 3.3.1 to obtain a value for $\tau^{\pm}$ with physical meaning. The first step to be able to solve the Riemann problem is to rewrite the original equation as a time-domain hyperbolic system.

Hyperbolic system: We start from the convected acoustic wave equation

$$\rho_0\left(\frac{\partial}{\partial t} + \boldsymbol{v_0}\cdot\nabla\right)^2 p - \operatorname{div}\left(\rho_0 c_0^2 \nabla p\right) = 0$$

and we write it as a hyperbolic system. First we have

$$\rho_0\frac{\partial^2 p}{\partial t^2} - \operatorname{div}\left(\boldsymbol{K_0}\nabla p - 2\rho_0\frac{\partial p}{\partial t}\boldsymbol{v_0}\right) = 0,$$

we therefore introduce the *total flux*

$$\frac{\partial\widetilde{\boldsymbol{\sigma}}}{\partial t} = -\boldsymbol{K_0}\nabla p + 2\rho_0\frac{\partial p}{\partial t}\boldsymbol{v_0},$$

leading to the following first-order formulation

$$\frac{\partial p}{\partial t} = -\frac{1}{\rho_0}\operatorname{div}\left(\widetilde{\boldsymbol{\sigma}}\right), \tag{3.23a}$$

$$\frac{\partial\widetilde{\boldsymbol{\sigma}}}{\partial t} = -\boldsymbol{K_0}\nabla p + 2\rho_0\frac{\partial p}{\partial t}\boldsymbol{v_0}. \tag{3.23b}$$

However this formulation does not have the form of a hyperbolic system.

Using (3.23a) in (3.23b), we have

$$\frac{\partial p}{\partial t} = -\frac{1}{\rho_0}\mathrm{div}\,(\tilde{\boldsymbol{\sigma}})\,, \tag{3.24a}$$

$$\frac{\partial \tilde{\boldsymbol{\sigma}}}{\partial t} = -\boldsymbol{K_0}\nabla p - 2\mathrm{div}\,(\tilde{\boldsymbol{\sigma}})\,\boldsymbol{v_0}. \tag{3.24b}$$

Notice that we need to work with a first-order in time formulation whereas our methods are written for second-order in time (or equivalently in frequency) formulations. However we have the following relationship between $\boldsymbol{\sigma}$ and $\tilde{\boldsymbol{\sigma}}$

$$\boldsymbol{\sigma} = i\omega\tilde{\boldsymbol{\sigma}},$$

making it possible to go back to a second-order formulation.
The system (3.24a)–(3.24b) can be written as

$$\frac{\partial \mathbf{U}}{\partial t} = \mathbb{A}_x\frac{\partial \mathbf{U}}{\partial x} + \mathbb{A}_y\frac{\partial \mathbf{U}}{\partial y}, \tag{3.25}$$

where

$$\mathbf{U} := \begin{bmatrix} p \\ \tilde{\boldsymbol{\sigma}} \end{bmatrix} \quad ; \quad \mathbb{A}_x := \begin{bmatrix} 0 & -\frac{1}{\rho_0} & 0 \\ -M_{0,xx} & -2v_{0,x} & 0 \\ -M_{0,yx} & -2v_{0,y} & 0 \end{bmatrix} \quad ; \quad \mathbb{A}_y := \begin{bmatrix} 0 & 0 & -\frac{1}{\rho_0} \\ -M_{0,xy} & 0 & -2v_{0,x} \\ -M_{0,yy} & 0 & -2v_{0,y} \end{bmatrix}.$$

To check that (3.25) is a hyperbolic system, one needs to show that for all $\alpha, \beta \in \mathbb{R}$ the matrix

$$\mathbb{A}_{\alpha,\beta} := \alpha\mathbb{A}_x + \beta\mathbb{A}_y = -\begin{bmatrix} 0 & \frac{\alpha}{\rho_0} & \frac{\beta}{\rho_0} \\ \alpha M_{0,xx} + \beta M_{0,xy} & 2\alpha v_{0,x} & 2\beta v_{0,x} \\ \alpha M_{0,yx} + \beta M_{0,yy} & 2\alpha v_{0,y} & 2\beta v_{0,y} \end{bmatrix}$$

is diagonalizable with real eigenvalues, see FIGURE 3.4.

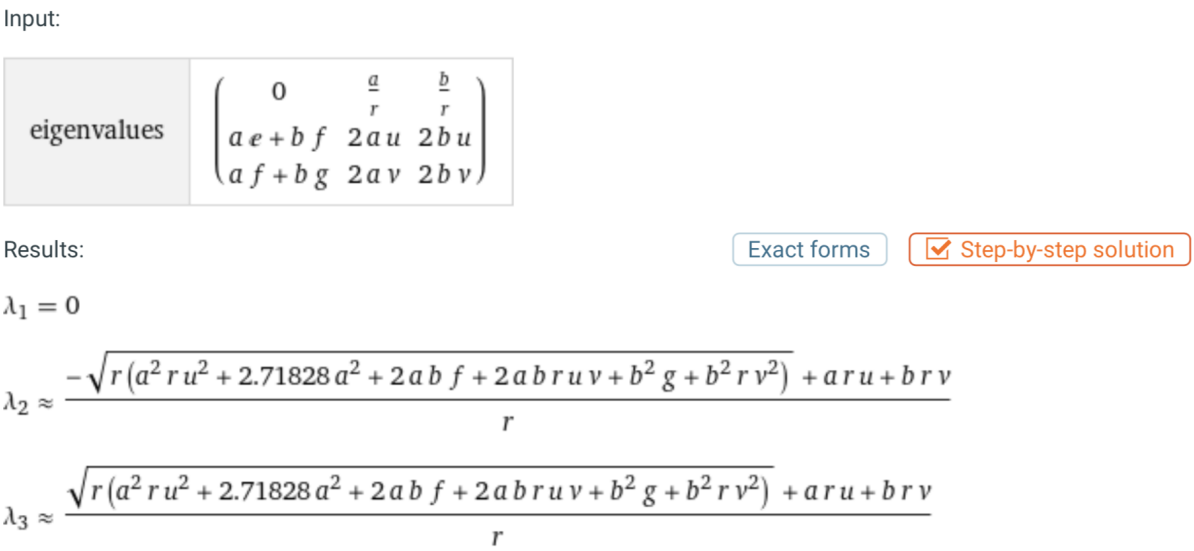


Figure 3.4: Computation of the eigenvalues of $\mathbb{A}_{\alpha,\beta}$ with the symbolic computation engine WolframAlpha

Riemann solver: To compute the upwind penalization parameters, we consider a vertical interface located at $x = 0$ and we assume that the background flow is uniform.
We will solve the problem (3.25) with the following initial condition

$$\begin{aligned} \mathbf{U}(x, y, 0) &= \mathbf{U}^+, && \text{if } x > 0, \\ \mathbf{U}(x, y, 0) &= \mathbf{U}^-, && \text{if } x < 0. \end{aligned}$$

With this choice of initial condition, we obtain a well-posed problem which is invariant with respect to y.
Our goal is to compute $\mathbf{U}$ at $x = 0$.
Due to the invariance with respect to y, we can rewrite (3.25) as

$$\frac{\partial \mathbf{U}}{\partial t} = \mathbb{A}_x \frac{\partial \mathbf{U}}{\partial x}.$$

Furthermore, we can obtain the following system for $[p, \widetilde{\sigma}_x]^T$ only

$$\frac{\partial}{\partial t}\begin{bmatrix} p \\ \widetilde{\sigma}_x \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & -\frac{1}{\rho_0} \\ -M_{0,xx} & -2v_{0,x} \end{bmatrix}}_{=:\mathbb{A}} \frac{\partial}{\partial x}\begin{bmatrix} p \\ \widetilde{\sigma}_x \end{bmatrix},$$

as the DG method is only written in terms of $\widetilde{\boldsymbol{\sigma}} \cdot \boldsymbol{n} = \widetilde{\sigma}_x$.
To compute the eigenvalues of $\mathbb{A}$, we need to solve

$$\begin{vmatrix} -\lambda & -\frac{1}{\rho_0} \\ -M_{0,xx} & -2v_{0,x} - \lambda \end{vmatrix} = 0 \iff \lambda^2 + 2v_{0,x}\lambda - \frac{M_{0,xx}}{\rho_0} = 0.$$

Recalling that

$$M_{0,xx} = \rho_0 c_0^2 - \rho_0 v_{0,x}^2,$$

we obtain the two following eigenvalues

$$\begin{aligned} \lambda_1 &= -\left(c_0 + v_{0,x}\right), \\ \lambda_2 &= c_0 - v_{0,x}, \end{aligned}$$

and the associated eigenvectors are

$$\boldsymbol{w}_1 := \begin{bmatrix} 1 \\ \rho_0(c_0 + v_{0,x}) \end{bmatrix} \quad \text{and} \quad \boldsymbol{w}_2 := \begin{bmatrix} 1 \\ \rho_0(v_{0,x} - c_0) \end{bmatrix}.$$

We can now define

$$\boldsymbol{W} := \begin{bmatrix} 1 & 1 \\ \rho_0(c_0 + v_{0,x}) & \rho_0(v_{0,x} - c_0) \end{bmatrix},$$

and therefore

$$\boldsymbol{W}^{-1} = \frac{1}{2\rho_0 c_0}\begin{bmatrix} \rho_0(c_0 - v_{0,x}) & 1 \\ \rho_0(c_0 + v_{0,x}) & -1 \end{bmatrix} = \begin{bmatrix} \ell_1 \\ \ell_2 \end{bmatrix}.$$

We have

$$\begin{aligned} \begin{bmatrix} p \\ \widetilde{\sigma}_x \end{bmatrix}(0, t) &= \ell_1 \begin{bmatrix} p^+ \\ \widetilde{\sigma}_x^+ \end{bmatrix} \boldsymbol{w}_1 + \ell_2 \begin{bmatrix} p^- \\ \widetilde{\sigma}_x^- \end{bmatrix} \boldsymbol{w}_2 \\ &= \frac{\rho_0(c_0 - v_{0,x})p^+ + \widetilde{\sigma}_x^+}{2\rho_0 c_0}\begin{bmatrix} 1 \\ \rho_0(c_0 + v_{0,x}) \end{bmatrix} + \frac{\rho_0(c_0 + v_{0,x})p^- - \widetilde{\sigma}_x^-}{2\rho_0 c_0}\begin{bmatrix} 1 \\ \rho_0(v_{0,x} - c_0) \end{bmatrix}, \end{aligned}$$

therefore

$$\widehat{p} = \frac{1}{2}\left(p^+ + p^-\right) - \frac{v_{0,x}}{2c_0}\left(p^+ - p^-\right) + \frac{1}{2\rho_0 c_0}\left(\widetilde{\sigma}_x^+ - \widetilde{\sigma}_x^-\right),$$
$$\widehat{\widetilde{\sigma}_x} = \frac{1}{2}\left(\widetilde{\sigma}_x^+ + \widetilde{\sigma}_x^-\right) + \frac{v_{0,x}}{2c_0}\left(\widetilde{\sigma}_x^+ - \widetilde{\sigma}_x^-\right) + \rho_0\frac{c_0^2 - v_{0,x}^2}{2c_0}\left(p^+ - p^-\right).$$

Finally, we can infer the form of the DG flux for a generic interface

$$\widehat{p} = \{\!\{p\}\!\} - \frac{\boldsymbol{v_0}\cdot\boldsymbol{n}^-}{2c_0}[\![p]\!] + \frac{1}{2\rho_0 c_0}[\![\widetilde{\boldsymbol{\sigma}}]\!], \tag{3.27a}$$
$$\widehat{\widetilde{\boldsymbol{\sigma}}}\cdot\boldsymbol{n}^- = \{\!\{\widetilde{\boldsymbol{\sigma}}\}\!\}\cdot\boldsymbol{n}^- + \frac{\boldsymbol{v_0}\cdot\boldsymbol{n}^-}{2c_0}[\![\widetilde{\boldsymbol{\sigma}}]\!] + \rho_0\frac{c_0^2 - (\boldsymbol{v_0}\cdot\boldsymbol{n})^2}{2c_0}[\![p]\!]. \tag{3.27b}$$

Notice that we had to chose an orientation of the normal vector. Following our convention, we have chosen to use $\boldsymbol{n}^-$ as it has the same orientation as $\boldsymbol{v_0}$.
Rewriting (3.27a)–(3.27b) in terms of $\boldsymbol{\sigma}$ instead of $\widetilde{\boldsymbol{\sigma}}$ leads to

$$\widehat{p_h} = \{\!\{p\}\!\} - \frac{\boldsymbol{v_0}\cdot\boldsymbol{n}^-}{2c_0}[\![p]\!] + \frac{1}{2i\omega\rho_0 c_0}[\![\boldsymbol{\sigma}]\!] \tag{3.28a}$$
$$\widehat{\boldsymbol{\sigma}}\cdot\boldsymbol{n}^- = \{\!\{\widetilde{\boldsymbol{\sigma}}\}\!\}\cdot\boldsymbol{n}^- + \frac{\boldsymbol{v_0}\cdot\boldsymbol{n}^-}{2c_0}[\![\widetilde{\boldsymbol{\sigma}}]\!] + i\omega\rho_0\frac{c_0^2 - (\boldsymbol{v_0}\cdot\boldsymbol{n}^-)^2}{2c_0}[\![p]\!]. \tag{3.28b}$$

Comparing (3.28a)–(3.28b) with (3.19a)–(3.19b) , we see that

$$\tau^+ + \tau^- = 2\rho_0 c_0, \tag{3.29a}$$
$$\frac{\tau^+\tau^-}{\tau^+ + \tau^-} = \rho_0\frac{c_0^2 - (\boldsymbol{v_0}\cdot\boldsymbol{n}^-)^2}{2c_0}. \tag{3.29b}$$

The system (3.29a)–(3.29b) leads to the following second-order equation

$$(\tau^+)^2 - 2\rho_0 c_0\tau^+ + \rho_0^2\left(c_0^2 - (\boldsymbol{v_0}\cdot\boldsymbol{n}^-)^2\right) = 0,$$

and to the two following families for $\tau^\pm$

$$\tau_1^+ = \rho_0(c_0 + \boldsymbol{v_0}\cdot\boldsymbol{n}^-), \qquad \tau_1^- = \rho_0(c_0 - \boldsymbol{v_0}\cdot\boldsymbol{n}^-),$$
$$\tau_2^+ = \rho_0(c_0 - \boldsymbol{v_0}\cdot\boldsymbol{n}^-), \qquad \tau_2^- = \rho_0(c_0 + \boldsymbol{v_0}\cdot\boldsymbol{n}^-).$$

To discriminate between $\tau_1^\pm$ and $\tau_2^\pm$ we once again go back to (3.19a)–(3.19b) and we see that the solution must satisfy

$$\frac{\tau^+ - \tau^-}{2(\tau^+ + \tau^-)} = -\frac{\boldsymbol{v_0}\cdot\boldsymbol{n}^-}{2c_0}.$$

We can therefore conclude that the upwind fluxes are obtained by using the $\tau_2^\pm$ solution. We can make this choice independent of the orientation convention by noticing that $\boldsymbol{n}^+ = -\boldsymbol{n}^-$, leading to

$$\tau_2^\pm = \rho_0(c_0 + \boldsymbol{v_0}\cdot\boldsymbol{n}^\pm).$$

Remark 3.3.3: To keep polynomial fluxes on the interfaces, the background quantities will be approximated by their value at the center of the interface.

Remark 3.3.4: In the context of DG and HDG methods, τ is usually chosen to be of the «order of unity» to ensure optimal convergence rate. In the error analysis of the method, we allow the dependency to the background coefficient to be hidden in the constants, so the choice (3.22) is actually possible.

3.3.3 Local solvability

We will now show the local solvability for the *total flux* formulation. Proving the well-posedness of the local problems is always very important when working with HDG methods. For the strongly coercive problems, for which HDG methods were initially designed, this property usually comes directly from the continuous problem. However for harmonic wave equations, which are only weakly coercive, things are more complicated : indeed solving the local problem amounts to solving a wave problem with Dirichlet boundary conditions. We therefore need to ensure that the local problem does not introduce resonance into the method, which is the case when the elements are small enough. In this section, we will prove that the static condensation process is well-defined when the mesh is fine enough.

Notice that in this case the proof relies on an *absorption technique* and is therefore very technical. Readers who are not familiar with HDG theory should probably begin with the proof for the *diffusive flux* formulation which is easier. It will be detailed in Subsection 3.4.2.

First, we need to show the

Lemma 3.3.1:
For $p_h \in \mathcal{P}_k(K)$ with $k > 0$, the following inverse inequality holds

$$\|\nabla p_h \cdot \boldsymbol{n}\|_{\partial K} \lesssim \|\nabla p_h\|_{\partial K} \lesssim h_K^{-\frac{1}{2}} \|\nabla p_h\|_K .$$

Proof:
First, we notice that if p_h is constant the desired inequality reduces to $0 \lesssim 0$. We therefore only consider non-constant p_h.
Let $\widetilde{K}$ be the reference unit element. We consider the map $F : \widetilde{K} \longrightarrow K$. We use $\tilde{\cdot}$ to denote quantities on the reference element instead of the more standard notation $\hat{\cdot}$ to avoid confusion, as we already used $\hat{\cdot}$ to denote the numerical fluxes.
Let $\tilde{\gamma}^1 : H^2(\widetilde{K}) \longrightarrow L^2(\partial\widetilde{K})$ be the normal derivative operator in the reference element. As $\tilde{\gamma}^1$ is continuous, we have

$$\left\|\tilde{\gamma}^1(\widetilde{p_h})\right\|_{\partial\widetilde{K}} \lesssim \|\widetilde{p_h}\|_{2,\widetilde{K}} \lesssim |\widetilde{p_h}|_{1,\widetilde{K}}$$

The second inequality holds as $\widetilde{p_h} \in \mathcal{P}_k(\widetilde{K})$ which is a finite-dimensional vector-space on which all the norms are equivalent and p_h is not constant.
We now recall the following scaling inequalities, see [DS19, Eq (1.6), (1.7) & (1.8)]

$$|\widetilde{p_h}|_{1,\widetilde{K}} \lesssim |\det \mathrm{Jac}(F)|^{-\frac{1}{2}} \|\mathrm{Jac}(F)\| \, |p_h|_{1,K} \lesssim |p_h|_{1,K};$$
$$h_K^{\frac{1}{2}} \|\mu_h\|_{\partial K} \lesssim \|\widetilde{\mu_h}\|_{\partial\widetilde{K}}$$

Due to the regularity of the mesh, we have

$$h_K^{\frac{1}{2}} \|\nabla p_h \cdot \boldsymbol{n}\|_{\partial K} \lesssim \|\nabla p_h\|_K .$$

■

Theorem 2 : *Local solvability for the total flux HDG method*

If τ is chosen such that

$$\exists \tau_0 > 0, \quad \forall e \in \mathcal{E}(K), \quad 0 < \tau_0 \leqslant \tau + \boldsymbol{b_0} \cdot \boldsymbol{n},$$

then there exists a constant $\alpha_+ > 0$ such that the local problem is well-posed if $\omega h_K < \alpha_+$.

Proof: As (3.12a)–(3.12b) is a finite-dimensional problem, we only need to prove uniqueness of the solution. We therefore assume that $\widehat{p_h} = 0$ and $s = 0$, and we need to show that the system

$$\left(\boldsymbol{W_0}\boldsymbol{\sigma}_h, \boldsymbol{r}_h\right)_K - \left(p_h, \operatorname{div}\left(\boldsymbol{r}_h\right)\right)_K + 2i\omega \left(p_h \boldsymbol{W_0} \boldsymbol{b_0}, \boldsymbol{r}_h\right)_K = 0, \qquad \forall \boldsymbol{r}_h \in \boldsymbol{V}_h(K) \quad (3.31a)$$

$$-\omega^2 \left(\rho_0 p_h, w_h\right)_K + \left(\operatorname{div}\left(\boldsymbol{\sigma}_h\right), w_h\right)_K + i\omega \left\langle \tau p_h, w_h \right\rangle_{\partial K} = 0, \qquad \forall w_h \in W_h(K). \quad (3.31b)$$

has only one solution $(\boldsymbol{\sigma}_h, p_h) = (\boldsymbol{0}, 0)$.
We will prove the theorem by contradiction. We therefore assume that there is a non-zero solution $(\boldsymbol{\sigma}_h, p_h)$ to (3.31a)–(3.31b).
Step 1: Energy-like identity.
We test (3.31a) with $\boldsymbol{r}_h = \boldsymbol{\sigma}_h$ and conjugate the resulting equation, we then test (3.31b) with $w_h = p_h$ and add the two resulting equations leading to

$$\left\|\boldsymbol{\sigma}_h\right\|^2_{\boldsymbol{W_0},K} - \omega^2 \left\|p_h\right\|^2_{\rho_0,K} - 2i\omega \left(\boldsymbol{\sigma}_h, p_h \boldsymbol{W_0} \boldsymbol{b_0}\right)_K + i\omega \left\langle \tau p_h, p_h \right\rangle_{\partial K} = 0.$$

We then focus on the third term, as $\boldsymbol{W_0}$ is real and symmetric we have

$$\left(\boldsymbol{\sigma}_h, p_h \boldsymbol{W_0} \boldsymbol{b_0}\right)_K = \left(\boldsymbol{W_0}\boldsymbol{\sigma}_h, p_h \boldsymbol{b_0}\right)_K.$$

Taking $\boldsymbol{r}_h = p_h \left\{\boldsymbol{b_0}\right\}$, where $\left\{\boldsymbol{b_0}\right\}$ is the average of $\boldsymbol{b_0}$ on K, in (3.31a), we have

$$\left(\boldsymbol{W_0}\boldsymbol{\sigma}_h, p_h \left\{\boldsymbol{b_0}\right\}\right)_K - \left(p_h, \left\{\boldsymbol{b_0}\right\} \cdot \nabla p_h\right)_K + 2i\omega \left(p_h \boldsymbol{W_0} \boldsymbol{b_0}, p_h \left\{\boldsymbol{b_0}\right\}\right)_K = 0,$$

leading to

$$-\left(\boldsymbol{W_0}\boldsymbol{\sigma}_h, p_h \boldsymbol{b_0}\right)_K + \left(p_h, \boldsymbol{b_0} \cdot \nabla p_h\right)_K - 2i\omega \left(p_h \boldsymbol{W_0} \boldsymbol{b_0}, p_h \boldsymbol{b_0}\right)_K = -\varepsilon,$$

where

$$\varepsilon := \left(\boldsymbol{W_0}\boldsymbol{\sigma}_h, p_h(\boldsymbol{b_0} - \left\{\boldsymbol{b_0}\right\})\right)_K - \left(p_h, (\boldsymbol{b_0} - \left\{\boldsymbol{b_0}\right\}) \cdot \nabla p_h\right)_K + 2i\omega \left(p_h \boldsymbol{W_0} \boldsymbol{b_0}, p_h(\boldsymbol{b_0} - \left\{\boldsymbol{b_0}\right\})\right)_K.$$

We chose the notation ε to emphasize that this quantity is small, as it will be discussed in *Step 3*.
So we have

$$\begin{aligned} \left\|\boldsymbol{\sigma}_h\right\|^2_{\boldsymbol{W_0},K} - \omega^2 (\left\|p_h\right\|^2_{\rho_0,K} + 4 \left\|p_h \boldsymbol{b_0}\right\|^2_{\boldsymbol{W_0},K}) \\ - 2i\omega \left(p_h, \boldsymbol{b_0} \cdot \nabla p_h\right)_K - 2i\omega\varepsilon + i\omega \left\langle \tau p_h, p_h \right\rangle_{\partial K} = 0 \end{aligned} \quad (3.32)$$

Step 2: Boundary condition
Taking the imaginary part (3.32) leads to

$$\begin{aligned} &2\omega \mathfrak{Re} \left(p_h \boldsymbol{b_0}, \nabla p_h\right)_K + \omega \left\langle \tau p_h, p_h \right\rangle_{\partial K} + 2\omega \mathfrak{Re}\varepsilon = 0, \\ &\text{(by Lemma 4.1.1)} \quad \left\langle (\tau + \boldsymbol{b_0} \cdot \boldsymbol{n}) p_h, p_h \right\rangle_{\partial K} = -2\mathfrak{Re}\varepsilon. \end{aligned}$$

As we have chosen τ such that there is $0 < \tau_0 \leqslant \tau + \boldsymbol{b_0} \cdot \boldsymbol{n}$ where τ_0 does not depend on h_K, we have

$$\|p_h\|_{\partial K}^2 \lesssim |\varepsilon|. \tag{3.33}$$

As we do not have shown that $p|_{\partial K} = 0$, we cannot use Poincaré's inequality, however according to [EG04, Lemma B.63 & Example B.64] we have

$$\|p_h\|_K \leqslant \|p_h\|_{1,K} \leqslant C_K \|\nabla p_h\|_K + \frac{C_K}{\operatorname{meas}(\partial K)} \|p_h\|_{\partial K},$$

where C_K is the Poincaré constant of K[2]. Using standard scaling inequalities, we have

$$C_K \lesssim h_K \quad \text{and} \quad \frac{C_K}{\operatorname{meas}(\partial K)} \lesssim h_K^{\frac{1}{2}}.$$

Using (3.33) to estimate the boundary term, this leads to

$$\|p_h\|_K \lesssim h_K \|\nabla p_h\|_K + h_K^{1/2} |\varepsilon|^{1/2}. \tag{3.34}$$

Step 3: Estimating $|\varepsilon|$ and $\|p_h\|_K$. We have

$$\begin{aligned}
|\varepsilon| \lesssim & h_K \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K} \|p_h\|_K + h_K \|p_h\|_K \|\nabla p_h\|_K + \omega h_K \|p_h\|_K^2 \\
\text{(by Young)} \quad \lesssim & h_K^2 \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}^2 + (1 + \omega h_K) \|p_h\|_K^2 + h_K^2 \|\nabla p_h\|_K^2 \\
\lesssim & h_K^2 \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}^2 + (1 + \omega h_K) \left(h_K^2 \|\nabla p_h\|_K^2 + h_K |\varepsilon| \right) + h_K^2 \|\nabla p_h\|_K^2
\end{aligned}$$

If h_K is small enough, the term $h_K|\varepsilon|$ in the right-hand side can be absorbed by the left-hand side leading to

$$|\varepsilon| \lesssim h_K^2 \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}^2 + (1 + \omega h_K) h_K^2 \|\nabla p_h\|_K^2 .$$

Assuming that h_K is small enough, we can overestimate $\omega h_K \lesssim 1$, leading to

$$|\varepsilon| \lesssim h_K^2 \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}^2 + h_K^2 \|\nabla p_h\|_K^2 . \tag{3.35}$$

Together with the generalized Poincaré's inequality (3.34) we have

$$\|p_h\|_K^2 \lesssim h_K^2 \|\nabla p_h\|_K^2 + h_K^2 \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}^2 , \tag{3.36}$$

and using that $\sqrt{a^2 + b^2} \leqslant |a| + |b|$ we also have

$$\|p_h\|_K \lesssim h_K \|\nabla p_h\|_K + h_K \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K} .$$

Step 4: Estimating $\|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}$
Taking the real part of the Garding's identity (3.32), we have

$$\begin{aligned}
\|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}^2 \lesssim & \omega^2 \|p_h\|_K^2 + \omega \|p_h\|_K \|\nabla p_h\|_K + \omega |\varepsilon| \\
\text{(by (3.36))} \quad \lesssim & \omega^2 \left(h_K^2 \|\nabla p_h\|_K^2 + h_K^2 \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}^2 \right) \\
& + \omega \left(h_K \|\nabla p_h\|_K + h_K \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K} \right) \|\nabla p_h\|_K + \omega |\varepsilon| \\
\text{(by Young)} \quad \lesssim & \left(\omega^2 h_K^2 + \omega h_K \right) \|\nabla p_h\|_K^2 + \omega^2 h_K^2 \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}^2 + \omega h_K \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}^2 + \omega |\varepsilon| \\
\text{(by (3.35))} \quad \lesssim & \left(\omega^2 h_K^2 + \omega h_K \right) \|\nabla p_h\|_K^2 + \left(\omega^2 h_K^2 + \omega h_K \right) \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K} \\
\lesssim & \omega h_K \|\nabla p_h\|_K^2 + \omega h_K \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W_0},K}
\end{aligned}$$

[2]The constant used by the authors of [EG04] is the inverse of the usual Poincaré constant.

We obtained the last line by assuming that $\omega^2 h_K^2 \lesssim \omega h_K$ which is true if h_K is small enough. If h_K is small enough the term $\omega h_K \|\boldsymbol{\sigma}_h\|^2_{\boldsymbol{W}_0,K}$ in the right-hand side can be absorbed by the left-hand side, leading to

$$\|\boldsymbol{\sigma}_h\|^2_{\boldsymbol{W}_0,K} \lesssim \omega h_K \|\nabla p_h\|^2_K . \tag{3.37}$$

Step 5: Estimating $\|\nabla p_h\|_K$
Taking $\boldsymbol{r}_h = \nabla p_h$ in (3.31a) and reverting the integration by parts, we have

$$\begin{aligned}
\|\nabla p_h\|^2_K &= |(\boldsymbol{W}_0\boldsymbol{\sigma}_h, \nabla p_h)_K + 2i\omega\,(p_h \boldsymbol{W}_0\boldsymbol{b}_0, \nabla p_h)_K + \langle p_h, \nabla p_h \cdot \boldsymbol{n}\rangle_{\partial K}| \\
&\lesssim \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W}_0,K} \|\nabla p_h\|_K + \omega \|p_h\|_K \|\nabla p_h\|_K + \|p_h\|_{\partial K} \|\nabla p_h\|_{\partial K} \\
\text{(by LEMMA 3.3.1)} \quad &\lesssim \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W}_0,K} \|\nabla p_h\|_K + \omega \|p_h\|_K \|\nabla p_h\|_K + h_K^{-1/2} \|p_h\|_{\partial K} \|\nabla p_h\|_K
\end{aligned}$$

So we have

$$\begin{aligned}
\|\nabla p_h\|_K &\lesssim \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W}_0,K} + \omega \|p_h\|_K + h_K^{-1/2} \|p_h\|_{\partial K} \\
&\lesssim \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W}_0,K} + \omega \|p_h\|_K + (1+\omega h_K) h_K \|\nabla p_h\|_K \\
&\lesssim \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W}_0,K} + \omega h_K \|\nabla p_h\|_K + \omega h_K^{3/2} \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W}_0,K} + (1+\omega h_K) h_K \|\nabla p_h\|_K
\end{aligned}$$

Finally we have

$$\|\nabla p_h\|_K \lesssim \|\boldsymbol{\sigma}_h\|_{\boldsymbol{W}_0,K} \lesssim \|\boldsymbol{\sigma}_h\|_K . \tag{3.38}$$

Step 6: Contradiction
Combining (3.37) and (3.38), we have

$$\|\boldsymbol{\sigma}_h\|^2_K \lesssim \omega h_K \|\boldsymbol{\sigma}_h\|^2_K$$

as we assumed that $\boldsymbol{\sigma}_h \neq 0$, we can divide by $\|\boldsymbol{\sigma}_h\|_K$, leading to

$$1 \lesssim \omega h_K,$$

which does not hold if ωh_K is small enough.
This is the desired contradiction, and we can therefore conclude that $(\boldsymbol{\sigma}_h, p_h) = (\boldsymbol{0}, 0)$ is the only solution of the system (3.31a)–(3.31b). ∎

3.3.4 Error analysis

The error analysis can be carried out by following the projection analysis for the Helmholtz equation given in [DS19, Sec. 3.5.1 & 3.5.2] with some minor changes.
This error analysis relies on the tailored HDG projection that fits the structure of the numerical trace. This projection $(\boldsymbol{\Pi}, \Pi)$, with

$$(\boldsymbol{\Pi}, \Pi) : \boldsymbol{H}_{\mathrm{div}}(\mathcal{O}) \times H^1(\mathcal{O}) \longrightarrow \boldsymbol{V}_h \times W_h := \boldsymbol{\mathcal{P}}_k(\mathcal{T}_h) \times \mathcal{P}_k(\mathcal{T}_h)$$

is defined by the following equations

$$\begin{aligned}
(\boldsymbol{\Pi}\boldsymbol{\sigma}, \boldsymbol{r}_h)_K &= (\boldsymbol{\sigma}, \boldsymbol{r}_h)_K , && \forall \boldsymbol{r}_h \in \boldsymbol{\mathcal{P}}_{k-1}(K), \\
(\Pi p, w_h)_K &= (p, w_h)_K , && \forall w_h \in \mathcal{P}_{k-1}(K), \\
\langle \boldsymbol{\Pi}\boldsymbol{\sigma} \cdot \boldsymbol{n} + i\omega\tau \Pi p, \mu_h\rangle_{\partial K} &= \langle \boldsymbol{\sigma} \cdot \boldsymbol{n} + i\omega\tau p, \mu_h\rangle_{\partial K} , && \forall \mu_h \in \mathcal{R}_k(\partial K).
\end{aligned}$$

Notice that denoting the image of $(\boldsymbol{\sigma}, p)$ under $(\boldsymbol{\Pi}, \Pi)$ by $(\boldsymbol{\Pi}\boldsymbol{\sigma}, \Pi p)$ is a slight abuse of notation as both components depend on $\boldsymbol{\sigma}$ and p. However it is very convenient and often found in the literature.

We define the following error quantities

$$\boldsymbol{\delta}_h^{\boldsymbol{\sigma}} := \boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma} \quad ; \quad \delta_h^p := \Pi p - p \quad ; \quad \widehat{\delta}_h^p := p - P_M p$$

and

$$\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}} := \boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}_h \in \boldsymbol{V}_h \quad ; \quad \varepsilon_h^p := \Pi p - p_h \in W_h \quad ; \quad \widehat{\varepsilon}_h^p := P_M p - \widehat{p_h} \in M_h.$$

We will split the errors as

$$\begin{aligned} \|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}\|_{\boldsymbol{W_0},\mathcal{T}_h} &\leqslant \|\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + \|\boldsymbol{\delta}_h^{\boldsymbol{\sigma}}\|_{\boldsymbol{W_0},\mathcal{T}_h}, \\ \|p_h - p\|_{\rho_0,\mathcal{T}_h} &\leqslant \|\varepsilon_h^p\|_{\rho_0,\mathcal{T}_h} + \|\delta_h^p\|_{\rho_0,\mathcal{T}_h}, \end{aligned}$$

Notice that the following estimates hold

$$\begin{aligned} \|\delta_h^p\|_K &\lesssim h_K^{k+1}\left(|p|_{k+1,K} + \tau_{\max}^{-1}|\operatorname{div}\boldsymbol{\sigma}|_{k,K}\right), \\ \|\boldsymbol{\delta}_h^{\boldsymbol{\sigma}}\|_K &\lesssim h_K^{k+1}\left(|\boldsymbol{\sigma}|_{k+1,K} + \tau^\star|p|_{k+1,K}\right), \end{aligned}$$

for some constants $\tau_{\max}$ and $\tau^\star$. So we only need to prove estimates for $\|\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\|_{\boldsymbol{W_0},\mathcal{T}_h}$, $\|\varepsilon_h^p\|_{\rho_0,\mathcal{T}_h}$.
The error analysis can be summarized as follows

1. we derive an estimate for $\|\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\|_{\boldsymbol{W_0},\mathcal{T}_h}$ using the energy-like inequality,
2. we use a *dual problem* to estimate $\|\varepsilon_h^p\|_{\rho_0,\mathcal{T}_h}$ in terms of $\|\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\|_{\boldsymbol{W_0},\mathcal{T}_h}$,
3. those estimates are combined through a *bootstrapping process*.

This analysis is therefore strongly related to the *Aubin-Nitsche method* and only works for regular solutions.
We will now give the main changes needed to adapt the error analysis from [DS19].
The error equations (3.30) become

$$\begin{aligned} (\boldsymbol{W_0}\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}, \boldsymbol{r}_h)_{\mathcal{T}_h} - (\varepsilon_h^p, \operatorname{div}(\boldsymbol{r}_h))_{\mathcal{T}_h} + 2i\omega\,(\varepsilon_h^p \boldsymbol{W_0}\boldsymbol{b_0}, \boldsymbol{r}_h)_{\mathcal{T}_h} & \\ + \langle \widehat{\varepsilon}_h^p, \boldsymbol{r}_h\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= (\boldsymbol{W_0}\boldsymbol{\delta}_h^{\boldsymbol{\sigma}}, \boldsymbol{r}_h)_{\mathcal{T}_h} + 2i\omega\,(\delta_h^p \boldsymbol{W_0}\boldsymbol{b_0}, \boldsymbol{r}_h)_{\mathcal{T}_h} \\ -\omega^2\,(\rho_0\varepsilon_h^p, w_h)_{\mathcal{T}_h} + (\operatorname{div}(\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}), w_h)_{\mathcal{T}_h} &= -i\omega\,\langle\tau(\varepsilon_h^p - \widehat{\varepsilon}_h^p), w_h\rangle_{\partial\mathcal{T}_h} - \omega^2\,(\rho_0\delta_h^p, w_h)_{\mathcal{T}_h} \\ -\langle \boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\cdot\boldsymbol{n} + i\omega\tau(\varepsilon_h^p - \widehat{\varepsilon}_h^p), \mu_h\rangle_{\partial\mathcal{T}_h} &= 0. \end{aligned}$$

The energy-like identity of Prop. 3.7 becomes

$$\begin{aligned} \|\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\|^2_{\boldsymbol{W_0},\mathcal{T}_h} - \omega^2\,\|\varepsilon_h^p\|^2_{\rho_0,\mathcal{T}_h} - 2i\omega\,(\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}, \varepsilon_h^p \boldsymbol{W_0}\boldsymbol{b_0})_{\mathcal{T}_h} &= -i\omega\,\langle\tau(\varepsilon_h^p - \widehat{\varepsilon}_h^p), \varepsilon_h^p - \widehat{\varepsilon}_h^p\rangle_{\partial\mathcal{T}_h}\,(\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}, \boldsymbol{W_0}\boldsymbol{\delta}_h^{\boldsymbol{\sigma}})_{\mathcal{T}_h} \\ &\quad - 2i\omega\,(\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}, \delta_h^p \boldsymbol{W_0}\boldsymbol{b_0})_{\mathcal{T}_h} \\ &\quad - \omega^2\,(\rho_0\delta_h^p, \varepsilon_h^p)_{\mathcal{T}_h}, \end{aligned}$$

leading to the following estimate

$$\begin{aligned} \left|\|\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\|^2_{\boldsymbol{W_0},\mathcal{T}_h} + i\omega\,\langle\tau(\varepsilon_h^p - \widehat{\varepsilon}_h^p), \varepsilon_h^p - \widehat{\varepsilon}_h^p\rangle_{\partial\mathcal{T}_h}\right| \lesssim & \omega^2\,\|\varepsilon_h^p\|^2_{\rho_0,\mathcal{T}_h} + \omega\,\|\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\|_{\boldsymbol{W_0},\mathcal{T}_h}\,\|\delta_h^p\|_{\rho_0,\mathcal{T}_h} \\ & + \omega\,\|\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\|_{\boldsymbol{W_0},\mathcal{T}_h}\,\|\delta_h^p\|_{\rho_0,\mathcal{T}_h} + \omega^2\,\|\varepsilon_h^p\|_{\rho_0,\mathcal{T}_h}\,\|\delta_h^p\|_{\rho_0,\mathcal{T}_h} \\ & + \|\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\|_{\boldsymbol{W_0},\mathcal{T}_h}\,\|\boldsymbol{\delta}_h^{\boldsymbol{\sigma}}\|_{\boldsymbol{W_0},\mathcal{T}_h}. \end{aligned}$$

The adjoint problem (3.31) becomes

$$\begin{aligned} \boldsymbol{W_0}\boldsymbol{\xi} - \nabla\theta - 2i\omega\theta\boldsymbol{W_0}\boldsymbol{b_0} &= 0 \\ -\rho_0\omega^2\theta - \operatorname{div}(\boldsymbol{\xi}) &= \varepsilon_h^p. \end{aligned}$$

We state the *elliptic regularity* assumption which is a key ingredient in the error analysis

$$\|\theta\|_{2,\mathcal{O}} + \|\boldsymbol{\xi}\|_{1,\mathcal{O}} \leqslant C_{\text{reg}} \|\varepsilon_h^p\|_{\mathcal{O}}$$

The identity of Prop. 3.8 becomes

$$\begin{aligned}\|\varepsilon_h^p\|_{\rho_0,\mathcal{T}_h}^2 =& \omega^2 \left(\rho_0(\Pi\theta - \theta), \varepsilon_h^p - \delta_h^p\right)_{\mathcal{T}_h} - \omega^2 \left(\rho_0\theta - \{\rho_0\theta\}, \delta_h^p\right)_{\mathcal{T}_h} \\ & - \left(\boldsymbol{W_0}(\boldsymbol{\Pi\xi} - \boldsymbol{\xi}), \boldsymbol{\varepsilon_h^\sigma} - \boldsymbol{\delta_h^\sigma}\right)_{\mathcal{T}_h} + \left(\boldsymbol{W_0\xi} - \{\boldsymbol{W_0\xi}\}, \boldsymbol{\delta_h^\sigma}\right)_{\mathcal{T}_h} \\ & - 2i\omega \left(\theta \boldsymbol{b_0}, \boldsymbol{W_0}\boldsymbol{\varepsilon_h^\sigma}\right)_{\mathcal{T}_h} + 2i\omega \left(\boldsymbol{W_0\xi}, (\varepsilon_h^p - \delta_h^p)\boldsymbol{b_0}\right)_{\mathcal{T}_h} \\ & + 2i\omega \left(\boldsymbol{W_0}(\boldsymbol{\Pi\xi} - \boldsymbol{\xi}), (\varepsilon_h^p - \delta_h^p)\boldsymbol{b_0}\right)_{\mathcal{T}_h},\end{aligned}$$

leading to the following estimate

$$\begin{aligned}\|\varepsilon_h^p\|_{\rho_0,\mathcal{T}_h} \lesssim & h\left(\|\varepsilon_h^p\|_{\rho_0,\mathcal{T}_h} + \|\delta_h^p\|_{\rho_0,\mathcal{T}_h}\right) + h\left(\|\boldsymbol{\varepsilon_h^\sigma}\|_{\boldsymbol{W_0},\mathcal{T}_h} + \|\boldsymbol{\delta_h^\sigma}\|_{\boldsymbol{W_0},\mathcal{T}_h}\right) \\ & + \|\boldsymbol{\varepsilon_h^\sigma}\|_{\boldsymbol{W_0},K} + \|\varepsilon_h^p\|_{\rho_0,K} + \|\delta_h^p\|_{\rho_0,K},\end{aligned}$$

where the dependance to ω has been hidden in the $\lesssim$ symbol.
Notice that in contrast to the HDG method for the standard Helmholtz equation, both p_h and $\boldsymbol{\sigma}_h$ have the *same* convergence rate.
It is now straightforward to follow the bootstrapping argument of Sec. 3.5.2 to obtain the

Theorem 3 : *Convergence of the HDG method with total flux*

Assuming that ωh is small enough and under the elliptic regularity assumption, we have

$$\|p_h - \Pi p\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+1}) \quad ; \quad \|\boldsymbol{\sigma}_h - \boldsymbol{\Pi\sigma}\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+1}).$$

Remark 3.3.5: Some HDG methods are known to achieve *super-convergence*, *ie* taking $p_h \in \mathcal{P}_k$ leads to the following error estimate

$$\|\Pi p - p_h\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+2}).$$

Superconvergence is an attractive property for a numerical scheme, indeed using a postprocessing scheme it is possible to use the solution $(\boldsymbol{\sigma}_h, p_h)$ to construct a new approximation $\widetilde{p_h}$ which converges with order $\mathcal{O}(h^{k+2})$, see [Ste91], [CGS10, Sec. 5] for more details.
Here we were only able to prove *optimal convergence* and the use of post-processing schemes to improve the convergence rate is therefore not possible.

3.3.5 Global solvability

The analysis that we have carried out in the previous subsection works for any solution $(\boldsymbol{\sigma}_h, p_h, \widehat{p_h})$ of the discrete system (3.16a)–(3.16b)–(3.16c) provided that such solution exists. We already discussed the well-posedness of the local problems in Theorem 2, but we have not yet proved that the global problem (3.17) for $\widehat{p_h}$ was well-posed.
To do that we can either directly show the well-posedness of the global problem (3.17). Or we can take advantage of the error estimates of Theorem 3 as we will describe below[3].

Resonant frequencies: We recall that the convected Helmholtz equation is a problem of Fredholm type. It is therefore uniquely solvable except on a set of *resonant frequencies*. For those frequencies, there exist non-zero solutions to the homogenous equation and unique solvability cannot be guaranteed.

[3] In [DS19] this idea is attributed to B. Cockburn.

Main result: We can now state and prove the main result of this section.

Theorem 4 : *Global solvability*

Under the assumptions of THEOREM 2 and THEOREM 3 and if ω is not a resonant frequency of the convected Helmholtz equation (3.1) then the global problem is well-posed, *ie* $\widehat{p_h}$ is uniquely defined by (3.17).

Proof: First we recall that (3.16a)–(3.16b)–(3.16c), or equivalently (3.17), is a square system of linear equations, we therefore only need to show the uniqueness of the solution of the homogenous system (when $g_N = g_D = s = 0$).
Assuming that ω is not a resonant frequency of (3.1), the exact solution is $p = 0$ and $\boldsymbol{\sigma} = \mathbf{0}$, and therefore

$$\|p\|_{s,\mathcal{O}} = 0 \quad \text{and} \quad \|\boldsymbol{\sigma}\|_{t,\mathcal{O}} = 0$$

and

$$\varepsilon_h^p = -p_h \;\; ; \;\; \boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}} = -\boldsymbol{\sigma}_h \;\; ; \;\; \widehat{\varepsilon}_h^p = -\widehat{p_h}.$$

The aim of the error analysis was to prove the following inequalities when h is small enough :

$$\|\varepsilon_h^p\|_{\mathcal{T}_h} \lesssim \|p\|_{s,\mathcal{O}} + \|\boldsymbol{\sigma}_h\|_{t,\mathcal{O}} = 0 \tag{3.39a}$$

$$\|\boldsymbol{\varepsilon}_h^{\boldsymbol{\sigma}}\|_{\mathcal{T}_h} \lesssim \|p\|_{s,\mathcal{O}} + \|\boldsymbol{\sigma}_h\|_{t,\mathcal{O}} = 0 \tag{3.39b}$$

Notice that we have hidden the powers of h in $\lesssim$ as they do not play a role here.
Therefore using (3.39a) and (3.39b) we have shown that

$$p_h \equiv 0 \quad \text{and} \quad \boldsymbol{\sigma}_h \equiv \mathbf{0}$$

when h is small enough.
For all $K \in \mathcal{T}_h$, we can now rewrite (3.12a) as

$$\langle \widehat{p_h}, \boldsymbol{r}_h \cdot \boldsymbol{n} \rangle_{\partial K} = 0, \quad \forall \boldsymbol{r}_h \in \boldsymbol{V}_h(K),$$

which leads to

$$\widehat{p_h} \equiv 0.$$

■

3.4 HDG(+) methods for the diffusive flux formulation

In this section, we will construct HDG methods based on the *diffusive flux* formulation, where $\boldsymbol{q}$ is used instead of $\boldsymbol{\sigma}$. We will mostly describe the HDG+ method where different polynomial degrees are used for the different unknowns, as it the most important novelty of this paper. Adaptation of the formulation construction and theoretical results to a more standard HDG method with the same polynomial interpolation for all the unknowns is straightforward. The main differences between the HDG and HDG+ methods are stated but the details are left out.
The *global solvability* will not be included in this section as the adaptation of the result from the previous section is immediate.

3.4.1 Construction of the method

On an element $K \in \mathcal{T}_h$, the construction of the HDG method relies on the following integration by parts formula, which is related to the weak formulation (3.7a)–(3.7b),

$$\int_K \boldsymbol{W_0 q} \cdot \boldsymbol{r}^* \mathrm{d}\boldsymbol{x} - \int_K p \mathrm{div}\,(\boldsymbol{r}^*)\,\mathrm{d}\boldsymbol{x} + \langle p, \boldsymbol{r}\cdot\boldsymbol{n}\rangle_{\partial K} = 0, \tag{3.40a}$$

$$-\omega^2 \int_K \rho_0 p w^* \mathrm{d}\boldsymbol{x} - 2i\omega \int_K \boldsymbol{b_0}\cdot\nabla p w^* \mathrm{d}\boldsymbol{x} + \int_K \mathrm{div}\,(\boldsymbol{q})\, w^* \mathrm{d}\boldsymbol{x} = \int_K s w^* \mathrm{d}\boldsymbol{x}, \tag{3.40b}$$

for all $(\boldsymbol{r}, w) \in \boldsymbol{H}_{\mathrm{div}}(K) \times H^1(K)$, and where $(\boldsymbol{q}, p) \in \boldsymbol{H}_{\mathrm{div}}(\mathcal{O}) \times H^1(\mathcal{O})$ is the unknown.

Choice of approximation spaces: For the HDG+ method, the choice of approximation spaces is different from the choice made for the previous HDG method. We consider the following *local approximation spaces*

$$\begin{aligned} \boldsymbol{V}_h(K) &= \boldsymbol{\mathcal{P}}_k(K), && \text{for the flux } \boldsymbol{q}_h, \\ W_h(K) &= \mathcal{P}_{k+1}(K), && \text{for the potential } p_h, \end{aligned}$$

where $k \geqslant 2$ is the degree of the method. The use of a higher polynomial degree for p_h is the distinctive feature of the HDG+ method.

Introduction of the hybrid unknown: As we did before, we introduce the *numerical trace* $\widehat{p_h}$ which approximates p on the skeleton $\mathcal{E}_h$ of the mesh. As before the boundary integral in (3.40a) will be discretized as

$$\int_{\partial K} p\boldsymbol{r}^*\cdot\boldsymbol{n}\mathrm{d}\sigma \quad \text{becomes} \quad \int_{\partial K} \widehat{p_h}\boldsymbol{r}_h^*\cdot\boldsymbol{n}\mathrm{d}\sigma.$$

For the HDG+ method, we use the following approximation space for $\widehat{p_h}$

$$M_h(e) = \mathcal{P}_k(e), \quad \forall e \in \mathcal{E}(K).$$

With this choice, p_h and $\widehat{p_h}$ do not have the same polynomial degree and we therefore have two approximations of p with different polynomial degrees on the skeleton of the mesh. We therefore need to change the *penalization term* to

$$\tau(P_M p_h - \widehat{p_h}), \tag{3.41}$$

where P_M is the L^2-orthogonal projection onto M_h. This is called the *reduced stabilization* and it was introduced in [Leh10]. It allows to get convergence rate of $k+2$ for p_h for the cost of a method of degree k. A large penalization parameter $\tau \sim {h_K}^{-1}$ is needed to obtain optimal convergence as it will be detailed in Subsection 3.4.3.

Local problem: Using the integration by parts formula (3.40a)–(3.40b) on an element $K \in \mathcal{T}_h$, we define the so-called *local problem* : seek $(\boldsymbol{q}_h, p_h) \in \boldsymbol{V}_h(K) \times W_h(K)$ such that

$$(\boldsymbol{W_0 q}_h, \boldsymbol{r}_h)_K - (p_h, \mathrm{div}\,(\boldsymbol{r}_h))_K + \langle \widehat{p_h}, \boldsymbol{r}_h\cdot\boldsymbol{n}\rangle_{\partial K} = 0, \tag{3.42a}$$

$$\begin{aligned} -\omega^2 (\rho_0 p_h, w_h)_K - 2i\omega\,(\boldsymbol{b_0}\cdot\nabla p_h, w_h)_K + (\mathrm{div}\,(\boldsymbol{q}_h)\,, w_h)_K \\ +2i\omega\,\langle \tau(P_M p_h - \widehat{p_h}) - \tau_{\mathrm{upw}}(p_h - \widehat{p_h}), w_h\rangle_{\partial K} = (s, w_h)_K\,, \end{aligned} \tag{3.42b}$$

for all $(\boldsymbol{r}_h, w_h) \in \boldsymbol{V}_h(K) \times W_h(K)$.
Following [QS16a], we have introduced a second penalization parameter τ_{upw} defined by

$$\tau_{\mathrm{upw}} := \max(\boldsymbol{b_0}\cdot\boldsymbol{n}, 0).$$

To understand why this second parameter is required, we recall that in HDG methods the penalization serves two purposes:

1. it enforces the Dirichlet boundary condition for the local problems,
2. it controls the stability of the method.

Here as $\boldsymbol{q}_h$ does not take the convection into account, the penalization term (3.41) with τ only stabilizes the diffusion. We therefore need to add a second penalization to stabilize the convection. We denoted it τ_{upw} as it leads to an upwinding behavior that will be detailed in the next paragraph.

Transmission condition: Following the previous example, we introduce the *following numerical flux*

$$\widehat{\boldsymbol{q}_h} \cdot \boldsymbol{n} := \boldsymbol{q}_h \cdot \boldsymbol{n} + 2i\omega\tau \left(P_M p_h - \widehat{p_h}\right), \tag{3.43}$$

where $\tau = \mathcal{O}({h_K}^{-1})$. As discussed before, we need to require the normal continuity of the *total flux* on the interface between two elements, ands the quantity $\widehat{\boldsymbol{q}_h} \cdot \boldsymbol{n}$ only takes the diffusion into account. To deal with convection we add a second numerical flux

$$2i\omega\widehat{p_h\boldsymbol{b_0}} \cdot \boldsymbol{n} := 2i\omega(\boldsymbol{b_0} \cdot \boldsymbol{n})\widehat{p_h} + 2i\omega\tau_{\text{upw}}(p_h - \widehat{p_h}). \tag{3.44}$$

It is important to notice that this flux has an upwind behavior. Let $e = \partial K_+ \cap \partial K_-$ be an interior edge with $\boldsymbol{b_0} \cdot \boldsymbol{n}_- > 0$ on ∂K_-. We have

$$\text{On } \partial K_-: \quad \tau_{\text{upw}} := \max(\boldsymbol{b_0} \cdot \boldsymbol{n}, 0) = \boldsymbol{b_0} \cdot \boldsymbol{n}, \quad \text{so} \quad 2i\omega\widehat{p_h\boldsymbol{b_0}} \cdot \boldsymbol{n} = 2i\omega(\boldsymbol{b_0} \cdot \boldsymbol{n})p_h, \tag{3.45a}$$

$$\text{On } \partial K_+: \quad \tau_{\text{upw}} := \max(\boldsymbol{b_0} \cdot \boldsymbol{n}, 0) = 0, \quad \text{so} \quad 2i\omega\widehat{p_h\boldsymbol{b_0}} \cdot \boldsymbol{n} = 2i\omega(\boldsymbol{b_0} \cdot \boldsymbol{n})\widehat{p_h}. \tag{3.45b}$$

So on the outflow boundary we use the interior value p_h, whereas on the inflow boundary we use the trace value $\widehat{p_h}$.
Finally we write the *transmission condition* as

$$\left\langle \left(\widehat{\boldsymbol{q}_h} - 2i\omega\widehat{p_h\boldsymbol{b_0}}\right) \cdot \boldsymbol{n}, \mu_h \right\rangle_{\partial\mathcal{T}_h \setminus \Gamma_D} + \langle \widehat{p_h} - g_D, \mu_h \rangle_{\Gamma_D} = \langle g_N, \mu_h \rangle_{\Gamma_N}. \tag{3.46}$$

This formulation enforces normal continuity of the total flux between the elements and the boundary conditions on Γ_D and Γ_N.

Remark 3.4.1: To ensure the well-posedness of the *local problems*, the second penalization must be

$$\tau_{\text{upw}}(p_h - \widehat{p_h}),$$

and not

$$\tau_{\text{upw}}(P_M p_h - \widehat{p_h}),$$

see Theorem 5.

Remark 3.4.2: We would like to point out the main theoretical difficulty of this method : when the background flow is not constant, the second flux (3.44) leads to non-polynomial terms on the skeleton. This is usually avoided as much as possible in HDG methods.

Adaptation to a standard HDG method: With this formulation, it is also possible to consider a standard HDG method by using the same polynomial degree for p_h and $\boldsymbol{q}_h$, *i.e.* by using the following local approximation spaces

$$\begin{aligned} \boldsymbol{V}_h(K) &= \boldsymbol{\mathcal{P}}_k(K), && \text{for the flux } \boldsymbol{q}_h, \\ W_h(K) &= \mathcal{P}_k(K), && \text{for the pressure } p_h. \end{aligned}$$

In this case, as W_h and M_h have the same polynomial degree, the projection term becomes simpler, indeed

$$P_M p_h = p_h.$$

For this formulation, we do not require a large penalization parameter anymore, and we only need $\tau = \mathcal{O}(1)$.

Compact formulation of the methods:

HDG methods are usually stated in a compact way that can be obtained by summing the local problems (3.42a)–(3.42b) over the mesh elements and by adding the transmission condition (3.46). This formulation reads : seek $(\boldsymbol{q}_h, p_h, \widehat{p_h}) \in \boldsymbol{V}_h \times W_h \times M_h$, such that

$$(\boldsymbol{W_0} \boldsymbol{q}_h, \boldsymbol{r}_h)_{\mathcal{T}_h} - (p_h, \operatorname{div}(\boldsymbol{r}_h))_{\mathcal{T}_h} + \langle \widehat{p_h}, \boldsymbol{r}_h \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} = 0, \tag{3.47a}$$

$$-\omega^2 (\rho_0 p_h, w_h)_{\mathcal{T}_h} - 2i\omega (\boldsymbol{b_0} \cdot \nabla p_h, w_h)_{\mathcal{T}_h} + (\operatorname{div}(\boldsymbol{q}_h), w_h)_{\mathcal{T}_h} \tag{3.47b}$$

$$+2i\omega \langle \tau(P_M p_h - \widehat{p_h}) - \tau_{\text{upw}}(p_h - \widehat{p_h}), w_h \rangle_{\partial \mathcal{T}_h} = (s, w_h)_{\mathcal{T}_h}$$

$$\left\langle \left(\widehat{\boldsymbol{q}_h} - 2i\omega \widehat{p_h \boldsymbol{b_0}} \right) \cdot \boldsymbol{n}, \mu \right\rangle_{\partial \mathcal{T}_h \setminus \Gamma_D} + \langle \widehat{p_h} - g_D, \mu_h \rangle_{\Gamma_D} = \langle g_N, \mu_h \rangle_{\Gamma_N}, \tag{3.47c}$$

for all $(\boldsymbol{r}_h, w_h, \mu_h) \in \boldsymbol{V}_h \times W_h \times M_h$.

Condensed variational formulation

The compact formulation (3.47a)–(3.47b)–(3.47c) cannot be used to efficiently implement the HDG method, indeed with this formulation it is not clear how the *global problem* for $\widehat{p_h}$ only can be obtained. To describe this process , we will now write a *condensed* variational formulation for $\widehat{p_h}$ only.
We introduce the so-called *local solvers*

$$\mathsf{P}^K : (\widehat{p_h}, s) \longmapsto p_h^K,$$
$$\mathsf{Q}^K : (\widehat{p_h}, s) \longmapsto \boldsymbol{q}_h^K,$$
$$\widehat{\mathsf{Q}}^K : (\widehat{p_h}, s) \longmapsto \widehat{\boldsymbol{q}_h}^K \cdot \boldsymbol{n}^K := \mathsf{Q}^K(\widehat{p_h}, s) - 2i\omega\tau \left(P_M \mathsf{P}^K(\widehat{p_h}, s) - \widehat{p_h} \right) - 2i\omega\tau_{\text{upw}} \left(\mathsf{P}^K(\widehat{p_h}, s) - \widehat{p_h} \right),$$

where $(\boldsymbol{q}_h^K, p_h^K)$ is the solution of (3.42a)–(3.42b) and $\widehat{\boldsymbol{q}_h}^K$ is defined by (3.43).
We can therefore rewrite the transmission condition (3.47c) as

$$a_h(\widehat{p_h}, \mu) = \ell_h(\mu), \tag{3.48}$$

where

$$a_h(\widehat{p_h}, \mu_h) := \left\langle \mathsf{Q}^K(\widehat{p_h}, s) \cdot \boldsymbol{n} + 2i\omega\tau(P_M \mathsf{P}^K(\widehat{p_h}, s) - \widehat{p_h}) + 2i\omega\tau_{\text{upw}}(\mathsf{P}^K(\widehat{p_h}, s) - \widehat{p_h}), \mu_h \right\rangle_{\partial \mathcal{T}_h \setminus \Gamma_D}$$
$$+ \langle \widehat{p_h}, \mu_h \rangle_{\Gamma_D},$$
$$\ell_h(\mu_h) := \langle g_N, \mu_h \rangle_{\Gamma_N} + \langle g_D, \mu_h \rangle_{\Gamma_D}.$$

Equation (3.48) is the so-called *global problem* and is the main equation of the HDG method.

From a computational point of view, we proceed as described in Algorithm 2.

Algorithm 2: Solving HDG+

for $K \in \mathcal{T}_h$ **do**
  Construct the local solvers P^K, Q^K, $\widehat{\mathsf{Q}}^K$
  Add local contribution to the global problem (3.17)
Solve the global problem (3.17) for $\widehat{p_h}$ `// This is the most expensive step`
for $K \in \mathcal{T}_h$ **do**
  Reconstruct the local unknowns $p_h^K = \mathsf{P}^k(\widehat{p_h}, s)$ and $\boldsymbol{q}_h^K = \mathsf{Q}^K(\widehat{p_h}, s)$

3.4.2 Local solvability

It is worth remembering that HDG methods were originally developed for elliptic problems and that harmonic wave equations are only coercive. It is well-known that solving those equations with Dirichlet boundary conditions[4] leads to numerical pollution due to the resonance phenomenon. In this section we will show that the static condensation process is well-defined when the mesh is fine enough, *ie.* the local problem does not produce resonance. Before actually showing the local solvability, we need to prove the following lemma.

Lemma 3.4.1:
If $p \in H^1(K)$ and $\boldsymbol{b_0} \in \boldsymbol{L}^\infty(K) \cap \boldsymbol{\mathcal{C}}(\mathcal{O})$, where $\boldsymbol{\mathcal{C}}(\mathcal{O})$ is the space of vector functions continuous in the domain $\mathcal{O}$, then the following identity holds

$$\mathfrak{Re}\left(p\boldsymbol{b_0}, \nabla p\right)_K = \frac{1}{2}\left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})p, p\right\rangle_{\partial K}.$$

Proof: We use an integration by parts to obtain a relationship between $\left(p\boldsymbol{b_0}, \nabla p\right)_K$ and its complex conjugate :

$$\begin{aligned}
2\mathfrak{Re}\left(p\boldsymbol{b_0}, \nabla p\right)_K &= \left(p\boldsymbol{b_0}, \nabla p\right)_K + \left(p\boldsymbol{b_0}, \nabla p\right)_K^* \\
&= \left(p\boldsymbol{b_0}, \nabla p\right)_K + \left(\nabla p, p\boldsymbol{b_0}\right)_K \\
&= -\left(\operatorname{div}\left(p\boldsymbol{b_0}\right), p\right)_K + \left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})p, p\right\rangle_{\partial K} + \left(\nabla p, p\boldsymbol{b_0}\right)_K \\
(\operatorname{div}(\boldsymbol{b_0}) = 0) \quad &= -\left(\nabla p, p\boldsymbol{b_0}\right)_K + \left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})p, p\right\rangle_{\partial K} + \left(\nabla p, p\boldsymbol{b_0}\right)_K \\
&= \left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})p, p\right\rangle_{\partial K}.
\end{aligned}$$

■

We can now state and prove the main result of this section.

Theorem 5 : *Local solvability for the HDG+ methods*

If

$$\forall e \in \mathcal{E}(K), \quad \tau|_e < 0 \tag{3.49}$$

and if

$$\omega h_K < \frac{-C_{\boldsymbol{W_0},K}\left\|\boldsymbol{b_0}\right\|_{L^\infty(K)} + \left(C^2_{\boldsymbol{W_0},K}\left\|\boldsymbol{b_0}\right\|^2_{L^\infty(K)} + \left\|\rho_0\right\|_{L^\infty(K)}\right)^{\frac{1}{2}}}{C_{\boldsymbol{W_0},K} C \left\|\rho_0\right\|_{L^\infty(K)}} \tag{3.50}$$

where $C > 0$ is a constant that depends only on the shape regularity of K, then the local solver

$$(\widehat{p_h}, s) \longmapsto (p_h, \boldsymbol{q}_h)$$

is well-posed.

[4] Which is what the local solver does.

Proof:
As the local problems have a finite dimension, we only need to prove uniqueness of the solution. We therefore assume that $\widehat{p_h} = s = 0$ and we need to prove that the system

$$(\boldsymbol{W_0 q}_h, \boldsymbol{r}_h)_K - (p_h, \operatorname{div}(\boldsymbol{r}_h))_K = 0, \tag{3.51a}$$

$$-\omega^2 (\rho_0 p_h, w_h)_K - 2i\omega (\boldsymbol{b_0} \cdot \nabla p_h, w_h)_K + (\operatorname{div}(\boldsymbol{q}_h), w_h)_K + 2i\omega \langle \tau P_M p_h - \tau_{\text{upw}} p_h, w_h \rangle_{\partial K} = 0, \tag{3.51b}$$

has only one solution : $(p_h, \boldsymbol{q}_h) = (0, \boldsymbol{0})$.
We will prove the theorem by contradiction. We therefore assume that the system (3.51a)–(3.51b) has a non-zero solution $(p_h, \boldsymbol{q}_h)$.
Step 1: An energy-like system
We begin by testing (3.51b) with $w_h = p_h$

$$-\omega^2 \|p_h\|^2_{\rho_0,K} + 2i\omega (p_h \boldsymbol{b_0}, \nabla p_h)_K + (\operatorname{div}(\boldsymbol{q}_h), p_h)_K + \langle 2i\omega\tau P_M p_h - 2i\omega\tau_{\text{upw}} p_h, p_h \rangle_{\partial K} = 0 \tag{3.52}$$

Then, (3.51a) is tested with $\boldsymbol{r}_h = \boldsymbol{q}_h$ and conjugated :

$$\|\boldsymbol{q}_h\|^2_{\boldsymbol{W_0},K} - (\operatorname{div}(\boldsymbol{q}_h), p_h)_K = 0 \tag{3.53}$$

We now add (3.52) and (3.53) leading to

$$\|\boldsymbol{q}_h\|^2_{\boldsymbol{W_0},K} - \omega^2 \|p_h\|^2_{\rho_0,K} + 2i\omega (p_h \boldsymbol{b_0}, \nabla p_h)_K + \langle 2i\omega\tau P_M p_h - 2i\omega\tau_{\text{upw}} p_h, p_h \rangle_{\partial K} = 0 \tag{3.54}$$

We now obtain the following system by taking the real and imaginary parts of (3.54)

$$\mathfrak{Re} : \quad \|\boldsymbol{q}_h\|^2_{\boldsymbol{W_0},K} - \omega^2 \|p_h\|^2_{\rho_0,K} - 2\omega \mathfrak{Im} (p_h \boldsymbol{b_0}, \nabla p_h)_K = 0 \tag{3.55}$$

$$\mathfrak{Im} : \quad \mathfrak{Re} (p_h \boldsymbol{b_0}, \nabla p_h)_K + \langle \tau P_M p_h, p_h \rangle_{\partial K} = \langle \tau_{\text{upw}} p_h, p_h \rangle_{\partial K} \tag{3.56}$$

Indeed, as $P_M p_h \in M_h$ and τ is constant on each edge, one has

$$\langle \tau P_M p_h, p_h \rangle_{\partial K} = \langle \tau P_M p_h, P_M p_h \rangle_{\partial K} \in \mathbb{R}$$

Step 2: We focus on (3.56) to express $p_h|_{\partial K}$.
By the Lemma 4.1.1 we have

$$\mathfrak{Re} (p_h \boldsymbol{b_0}, \nabla p_h)_K = \frac{1}{2} \langle (\boldsymbol{b_0} \cdot \boldsymbol{n}) p_h, p_h \rangle_{\partial K}$$

and (3.56) becomes

$$\frac{1}{2} \langle (\boldsymbol{b_0} \cdot \boldsymbol{n}) p_h, p_h \rangle_{\partial K} + \langle \tau P_M p_h, p_h \rangle_{\partial K} = \langle \tau_{\text{upw}} p_h, p_h \rangle_{\partial K}$$

For the sake of simplicity, we assume that the sign of $\boldsymbol{b_0} \cdot \boldsymbol{n}$ is constant on each edge. It amounts to assuming that h_K is small enough. For a given edge $e \in \mathcal{E}(K)$, the three following cases are exhaustive:

- *Case 1:* $\boldsymbol{b_0} \cdot \boldsymbol{n} < 0$: therefore $\tau_{\text{upw}} := \max(\boldsymbol{b_0} \cdot \boldsymbol{n}, 0) = 0$ and

$$\underbrace{\langle \tau P_M p_h, P_M p_h \rangle_{\partial K}}_{\leqslant 0 \text{ by (3.49) as } \tau < 0} \overbrace{- \frac{1}{2} \langle |\boldsymbol{b_0} \cdot \boldsymbol{n}| p_h, p_h \rangle_{\partial K}}^{\leqslant 0} = 0$$

- *Case 2:* $\boldsymbol{b_0}\cdot\boldsymbol{n}>0$: therefore $\tau_{\text{upw}}:=\max(\boldsymbol{b_0}\cdot\boldsymbol{n},0)=\boldsymbol{b_0}\cdot\boldsymbol{n}$ and

$$\frac{1}{2}\langle|\boldsymbol{b_0}\cdot\boldsymbol{n}|p_h,p_h\rangle_{\partial K}+\langle\tau P_M p_h,P_M p_h\rangle_{\partial K}=\langle|\boldsymbol{b_0}\cdot\boldsymbol{n}|p_h,p_h\rangle_{\partial K}$$

$$\iff \underbrace{\langle\tau P_M p_h,P_M p_h\rangle_{\partial K}}_{\leqslant 0 \text{ by (3.49) as } \tau<0}\overbrace{-\frac{1}{2}\langle|\boldsymbol{b_0}\cdot\boldsymbol{n}|p_h,p_h\rangle_{\partial K}}^{\leqslant 0}=0$$

- *Case 3:* $\boldsymbol{b_0}\cdot\boldsymbol{n}=0$: in this case we only have

$$\langle\tau P_M p_h,P_M p_h\rangle_{\partial K}=0.$$

In the first two cases we have $p_h|_{\partial K}=P_M p_h=0$ and in the third one we only have $P_M p_h=0$. In particular, the following identity holds for all the three previous cases

$$\int_{\partial K}p_h\mathrm{d}\sigma=0, \tag{3.57}$$

indeed as $P_M p_h$ is the L^2-orthogonal projection of p_h onto M_h, we have

$$\int_{\partial K}P_M p_h\mu_h^*\mathrm{d}\sigma=\int_{\partial K}p_h\mu_h^*\mathrm{d}\sigma,\quad\forall\mu_h\in M_h:=\prod_{e\in\mathcal{E}(K)}\mathcal{P}_k(e),$$

and the previous identity is obtained by taking $\mu_h=1$.

Step 3: Contradiction

As $p_h\in\mathcal{P}_{k+1}(K)$, we have $p_h\in H^1(K)$ and the following Poincaré-Friedrichs inequality holds[5]

$$\|p_h\|_K\leqslant Ch_K\|\nabla p_h\|_K+\int_{\partial K}p_h\mathrm{d}\sigma, \tag{3.58}$$

see [EG04, Lemma B.66] with $f(v)=\int_{\partial K}v\mathrm{d}\sigma$. The constant C is the same one as in (3.50). Using (3.57), the inequality (3.58) becomes

$$\|p_h\|_K\leqslant Ch_K\|\nabla p_h\|_K.$$

Going back to (3.51a), integrating by parts and testing it with $\boldsymbol{r}_h=\nabla p_h$ we have

$$\begin{aligned}\|\nabla p_h\|_K^2&=|(\boldsymbol{W_0}\boldsymbol{q}_h,\nabla p_h)_K|\\&\leqslant C_{\boldsymbol{W_0},K}\|\boldsymbol{q}_h\|_{\boldsymbol{W_0},K}\|\nabla p_h\|_K\\\|\nabla p_h\|_K&\leqslant C_{\boldsymbol{W_0},K}\|\boldsymbol{q}_h\|_{\boldsymbol{W_0},K}\end{aligned} \tag{3.59}$$

On the other hand, from (3.55) we see that

$$\begin{aligned}\|\boldsymbol{q}_h\|_{\boldsymbol{W_0},K}^2&=\omega^2\|p_h\|_{\rho_0,K}^2+2\omega\mathfrak{Im}\,(p_h\boldsymbol{b_0},\nabla p_h)_K\\&\leqslant\omega^2\|\rho_0\|_{L^\infty(K)}\|p_h\|_K^2+2\omega\|\boldsymbol{b_0}\|_{L^\infty(K)}\|p_h\|_K\|\nabla p_h\|_K\\\|\boldsymbol{q}_h\|_{\boldsymbol{W_0},K}^2&\leqslant C^2\|\rho_0\|_{L^\infty(K)}\omega^2h_K^2\|\nabla p_h\|_K^2+2C\|\boldsymbol{b_0}\|_{L^\infty(K)}\omega h_K\|\nabla p_h\|_K^2\end{aligned} \tag{3.60}$$

Combining (3.59) and (3.60) we have

$$\|\nabla p_h\|_K^2\leqslant C_{\boldsymbol{W_0},K}^2\left[C^2\|\rho_0\|_{L^\infty(K)}\omega^2h_K^2\|\nabla p_h\|_K^2+2C\|\boldsymbol{b_0}\|_{L^\infty(K)}\omega h_K\|\nabla p_h\|_K^2\right],$$

[5]When $\boldsymbol{b_0}\cdot\boldsymbol{n}\neq 0$ we can use the standard Poincaré inequality instead.-

as we assumed $(\boldsymbol{q}_h, p_h) \neq (\mathbf{0}, 0)$ we can divide by $\|\nabla p_h\|_K$ to obtain

$$1 \leqslant C^2_{\boldsymbol{W_0},K} \left[C^2 \|\rho_0\|_{L^\infty(K)} \omega^2 h_K^2 + 2C \|\boldsymbol{b_0}\|_{L^\infty(K)} \omega h_K \right]. \tag{3.61}$$

We now define the function

$$f : \alpha \longmapsto C^2_{\boldsymbol{W_0},K} C^2 \|\rho_0\|_{L^\infty(K)} \alpha^2 + 2C^2_{\boldsymbol{W_0},K} C \|\boldsymbol{b_0}\|_{L^\infty(K)} \alpha - 1$$

Rewriting (3.61) in terms of f gives

$$f(\omega h_K) \geqslant 0.$$

We notice that f is a second-order polynomial whose roots are

$$\alpha_\pm = \frac{-C_{\boldsymbol{W_0},K} \|\boldsymbol{b_0}\|_{L^\infty(K)} \pm \left(C^2_{\boldsymbol{W_0},K} \|\boldsymbol{b_0}\|^2_{L^\infty(K)} + \|\rho_0\|_{L^\infty(K)} \right)^{\frac{1}{2}}}{C_{\boldsymbol{W_0},K} C \|\rho_0\|_{L^\infty(K)}}$$

As the leading coefficient of f is positive, we know that

$$\forall \alpha \in (\alpha_-, \alpha_+), \quad f(\alpha) < 0$$

and it is obvious that $\alpha_- < 0$ and $\alpha_+ > 0$.
Finally, we can see that the assumption on ωh_K (3.50) is exactly

$$0 < \omega h_K < \alpha_+,$$

which means

$$f(\omega h_K) < 0.$$

This is the desired contradiction and concludes the proof, as we necessarily have $p_h \equiv 0$ and $\boldsymbol{q}_h \equiv \mathbf{0}$. ∎

Remark 3.4.3: For triangular elements, the constant C satisfies

$$C < \frac{1}{\pi}.$$

Remark 3.4.4: when $\boldsymbol{b_0} = \mathbf{0}$, the solvability assumption (3.49) becomes

$$\omega h_K < \frac{1}{C C_{\boldsymbol{W_0},K} \|\rho_0\|^{\frac{1}{2}}_{L^\infty(K)}}$$

which is similar to the ones given in [DS19, Prop. 3.9] and [Hun19, Prop. 3.4.2].

Remark 3.4.5: this proof is written for the HDG+ method, for the more standard HDG method only minor changes are needed : in *Step 2*, $P_M p$ should be replaced by p. Assumption (3.49) can therefore be replaced with

$$\forall e \in \mathcal{E}(K), \quad \tau|_e < 0 \quad \text{or} \quad \tau|_e > \max_e \left(\frac{1}{2} |\boldsymbol{b_0} \cdot \boldsymbol{n}| \right).$$

3.4.3 Error analysis of the HDG+ method

In this section we will carry out a detailed error analysis of the HDG+ method. The adaptation of this process to the HDG method is straightforward, see Subsection 3.4.4.
We chose to use the orthogonal L^2 projections instead of the tailored HDG(+) projections that fit the numerical trace. As we study problems involving convection, the design of a new projection would be required as using the standard HDG(+) projection would not lead to cleaner error system. The design of such a projection seems very difficult when $\boldsymbol{b_0}$ is not constant.

We denote by $\boldsymbol{\pi_V}$, π_W and P_M the L^2-orthogonal projections onto $\boldsymbol{V}_h$, W_h and M_h respectively.
We recall the following estimates due to standard approximation theory for polynomials and trace inequalities which will be useful for our analysis, see *e.g.* [EG04, Prop. 1.135] :

$$\|p - \pi_W p\|_{\mathcal{O}} \lesssim h^s \|p\|_{s,\mathcal{O}}, \qquad 0 \leqslant s \leqslant k+2, \tag{3.62a}$$
$$\|\boldsymbol{q} - \boldsymbol{\pi_V} \boldsymbol{q}\|_{\mathcal{O}} \lesssim h^t \|\boldsymbol{q}\|_{t,\mathcal{O}}, \qquad 0 \leqslant t \leqslant k+1, \tag{3.62b}$$
$$\|p - P_M p\|_{\partial\mathcal{T}_h} \lesssim h^{s-\frac{1}{2}} \|p\|_{s,\mathcal{O}}, \qquad 1 \leqslant s \leqslant k+1, \tag{3.62c}$$
$$\|p - \pi_W p\|_{\partial K} \lesssim h^{s-\frac{1}{2}} \|p\|_{s,K}, \qquad 1 \leqslant s \leqslant k+2, \tag{3.62d}$$
$$\|\boldsymbol{q}\cdot\boldsymbol{n} - \boldsymbol{\pi_V}\boldsymbol{q}\cdot\boldsymbol{n}\|_{\partial K} \lesssim h^{t-\frac{1}{2}} \|\boldsymbol{q}\|_{t,K}, \qquad 1 \leqslant t \leqslant k+1, \tag{3.62e}$$

where $a \lesssim b$ means that there exists a constant $C > 0$ independent of the mesh size and frequency such that $a \leqslant Cb$.
We will also frequently use the following inverse inequality

$$\|w\|_{\partial K} \lesssim h^{-\frac{1}{2}} \|w\|_K, \qquad \forall w \in W_h. \tag{3.63}$$

Error equations

Let $(p, \boldsymbol{q})$ be the solution of the original problem (3.7a)–(3.7b). We define the projection errors

$$\boldsymbol{\delta}_h^{\boldsymbol{q}} := \boldsymbol{\pi_V}\boldsymbol{q} - \boldsymbol{q} \;\; ; \;\; \delta_h^p := \pi_W p - p \;\; ; \;\; \widehat{\delta}_h^p := p - P_M p$$

and

$$\boldsymbol{\varepsilon}_h^{\boldsymbol{q}} := \boldsymbol{\pi_V}\boldsymbol{q} - \boldsymbol{q}_h \in \boldsymbol{V}_h \;\; ; \;\; \varepsilon_h^p := \pi_W p - p_h \in W_h \;\; ; \;\; \widehat{\varepsilon}_h^p := P_M p - \widehat{p_h} \in M_h$$

Lemma 3.4.2:
The error quantities $(\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}, \varepsilon_h^p, \widehat{\varepsilon}_h^p)$ satisfy the following error equations:

$$(\boldsymbol{W_0}\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}, \boldsymbol{r}_h)_{\mathcal{T}_h} - (\varepsilon_h^p, \operatorname{div}(\boldsymbol{r}_h))_{\mathcal{T}_h} + \langle \widehat{\varepsilon}_h^p, \boldsymbol{r}_h \cdot \boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = (\boldsymbol{W_0}\boldsymbol{\delta}_h^{\boldsymbol{q}}, \boldsymbol{r}_h)_{\mathcal{T}_h} \tag{3.64a}$$

$$-\omega^2(\rho_0\varepsilon_h^p, w_h)_{\mathcal{T}_h} + 2i\omega(\varepsilon_h^p\boldsymbol{b_0}, \nabla w_h)_{\mathcal{T}_h} - (\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}, \nabla w_h)_{\mathcal{T}_h} + \left\langle \widehat{\mathcal{Q}}\cdot\boldsymbol{n} - \widehat{\mathcal{Q}_h}\cdot\boldsymbol{n}, w_h\right\rangle_{\partial\mathcal{T}_h} = -\omega^2(\rho_0\delta_h^p, w_h)_{\mathcal{T}_h} + 2i\omega(\delta_h^p\boldsymbol{b_0}, w_h)_{\mathcal{T}_h} \tag{3.64b}$$

$$\left\langle \widehat{\mathcal{Q}}\cdot\boldsymbol{n} - \widehat{\mathcal{Q}_h}\cdot\boldsymbol{n}, \mu_h\right\rangle_{\partial\mathcal{T}_h} = 0 \tag{3.64c}$$

where

$$\widehat{\mathcal{Q}}\cdot\boldsymbol{n} = \boldsymbol{q}\cdot\boldsymbol{n} - 2i\omega(\boldsymbol{b_0}\cdot\boldsymbol{n})p \quad \text{on } \partial\mathcal{T}_h$$

and

$$\widehat{\mathcal{Q}}\cdot\boldsymbol{n} - \widehat{\mathcal{Q}_h}\cdot\boldsymbol{n} = \boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\cdot\boldsymbol{n} - 2i\omega(\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\varepsilon}_h^p + 2i\omega\tau(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p) - 2i\omega\tau_{\text{upw}}(\varepsilon_h^p - \widehat{\varepsilon}_h^p)$$
$$-\boldsymbol{\delta}_h^{\boldsymbol{q}}\cdot\boldsymbol{n} - 2i\omega(\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\delta}_h^p - 2i\omega\tau P_M\delta_h^p + 2i\omega\tau_{\text{upw}}\left(\delta_h^p - \widehat{\delta}_h^p\right) \tag{3.65}$$

Proof: Notice that $(p, \boldsymbol{q})$ satisfy the equations (3.47a)–(3.47b)–(3.47c), introduce the projections wherever possible and subtract the actual discrete equations. ∎

A useful estimate: We will need to use the following estimate for $\|\nabla\varepsilon_h^p\|_{\partial\mathcal{T}_h}$ to carry out our analysis

Lemma 3.4.3:
The following estimate holds

$$\|\nabla\varepsilon_h^p\|_{\mathcal{T}_h} \leqslant C_{\boldsymbol{W_0},\mathcal{T}_h}\left(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + \|\boldsymbol{\delta}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h}\right) + C\left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}$$

Proof: Going back to (3.64a), testing with $\boldsymbol{r}_h = \nabla\varepsilon_h^p$ and integrating by parts leads to

$$(\boldsymbol{W_0}\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}, \nabla\varepsilon_h^p)_{\mathcal{T}_h} + \|\nabla\varepsilon_h^p\|_{\mathcal{T}_h}^2 + \langle\widehat{\varepsilon}_h^p - \varepsilon_h^p, \nabla\varepsilon_h^p\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = (\boldsymbol{W_0}\boldsymbol{\delta}_h^{\boldsymbol{q}}, \nabla\varepsilon_h^p)_{\mathcal{T}_h}$$

As $\varepsilon_h^p \in W_h$, $\nabla\varepsilon_h^p\cdot\boldsymbol{n} \in \mathcal{P}_k$ and we can use the following property of the projection P_M :

$$\langle P_M\varepsilon_h^p, \nabla\varepsilon_h^p\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = \langle\varepsilon_h^p, \nabla\varepsilon_h^p\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}$$

Using the Cauchy-Schwartz inequality we get

$$\left|\langle\widehat{\varepsilon}_h^p - \varepsilon_h^p, \nabla\varepsilon_h^p\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\right| = \left|\langle\widehat{\varepsilon}_h^p - P_M\varepsilon_h^p, \nabla\varepsilon_h^p\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\right| \leqslant C\|P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p\|_{\partial\mathcal{T}_h}\|\nabla\varepsilon_h^p\|_{\partial\mathcal{T}_h}$$

for some constant $C > 0$.
Using the following trace inequality (3.63)

$$\forall w \in W_h, \quad \|w\|_{\partial K} \leqslant Ch_K^{-\frac{1}{2}}\|w\|_K$$

we have

$$\|\nabla\varepsilon_h^p\|_{\mathcal{T}_h}^2 \leqslant C_{\boldsymbol{W_0},K}\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},K}\|\nabla\varepsilon_h^p\|_{\mathcal{T}_h} + \|P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p\|_{\partial\mathcal{T}_h}Ch_K^{-\frac{1}{2}}\|\nabla\varepsilon_h^p\|_{\mathcal{T}_h} + C_{\boldsymbol{W_0},K}\|\boldsymbol{\delta}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},K}\|\nabla\varepsilon_h^p\|_{\mathcal{T}_h}$$

which is the desired estimate as $\tau|_K = \mathcal{O}(h_K^{-1})$. ∎

Using the Poincaré-Wirtinger inequality: We denote by $\{u\}$ the L^2-projection of u on $\mathcal{P}_0$, *ie*

$$\forall K \in \mathcal{T}_h, \quad \{u\}\,|_K = \frac{1}{|K|}\int_K u \mathrm{d}\boldsymbol{x}$$

We will need to subtract $\{u\}$ from the equations to apply the Poincaré-Wirtinger inequality :

$$\|u - \{u\}\|_{\mathcal{T}_h} \leqslant Ch\,\|\nabla u\|_{\mathcal{T}_h}\,. \tag{3.66}$$

We can do that thanks to the following property of the projections π_W and $\boldsymbol{\pi_V}$, indeed we have for $\boldsymbol{\xi}$ and $\boldsymbol{q}$

$$(\boldsymbol{\pi_V q}, \{\boldsymbol{W_0 \xi}\})_{\mathcal{T}_h} = (\boldsymbol{q}, \{\boldsymbol{W_0 \xi}\})_{\mathcal{T}_h} \quad \text{as} \quad \{\boldsymbol{W_0 \xi}\} \in \boldsymbol{\mathcal{P}}_0 \subset \boldsymbol{\mathcal{P}}_k$$

therefore

$$(\boldsymbol{\delta}_h^{\boldsymbol{q}}, \boldsymbol{W_0 \xi})_{\mathcal{T}_h} = (\boldsymbol{q} - \boldsymbol{\pi_V q}, \boldsymbol{W_0 \xi})_{\mathcal{T}_h} = (\boldsymbol{q} - \boldsymbol{\pi_V q}, \boldsymbol{W_0 \xi} - \{\boldsymbol{W_0 \xi}\})_{\mathcal{T}_h}\,. \tag{3.67}$$

Similar results can be obtained in the same way for the other quantities.

Best approximation property of P_M: During the analysis, we will often need to compare quantities like $\|u - P_M u\|_{\partial K}$ and $\|u - \{u\}\|_{\partial K}$.

Lemma 3.4.4:
For $u \in \mathcal{P}_{k+1}(\mathcal{T}_h)$, the following inequality holds

$$\|u - P_M u\|_{\partial K} \lesssim \|u - \{u\}\|_{\partial K}$$

Proof:
We recall that $M_h := \prod_{e\in\mathcal{E}_h} \mathcal{P}_k(e)$ is a finite-dimensional vector subspace of $L^2(\partial\mathcal{T}_h)$. We recalled that functions in M_h are bi-valued piecewise polynomials of degree up to k on the skeleton of the mesh.
On an internal edge $e = \partial K_- \cap \partial K_+$, we define

$$\{u\}_e := \begin{cases} \{u^-\} \text{ on } K_- \\ \{u^+\} \text{ on } K_+ \end{cases}, \quad \text{where } \; u^\pm = u|_{K^\pm}.$$

With this definition $\{u\}_e$ is a bi-valued piecewise constant on the skeleton of the mesh, and therefore $\{u\}_e \in M_h$.
As P_M is the orthogonal projection onto M_h, we can use the Hilbert projection theorem to obtain

$$\|u - P_M u\|_{\partial\mathcal{T}_h} \leqslant \inf_{v\in M_h} \|u - v\|_{\partial\mathcal{T}_h}\,.$$

We can therefore conclude that

$$\|u - P_M u\|_{\partial\mathcal{T}_h} \leqslant \|u - \{u\}_e\|_{\partial\mathcal{T}_h}\,.$$

When no confusions are possible, we will denote $\{u\}_e$ by $\{u\}$. ∎
This property will often be referred to as the *best approximation property* of P_M.

Discrete energy-like equality: We will now establish a discrete energy-like equality which will be one of the key ingredients to study the convergence of our method.

Lemma 3.4.5:
The following discrete energy-like equality holds

$$\begin{aligned}\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|^2_{\boldsymbol{W_0},\mathcal{T}_h} - \omega^2\|\varepsilon_h^p\|^2_{\rho_0,\mathcal{T}_h} - 2i\omega\left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|^2_{\partial\mathcal{T}_h} - 2i\omega\left\|\left(\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac{1}{2}}(\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|^2_{\partial\mathcal{T}_h}&\\ -2\omega\mathfrak{Im}\left(\varepsilon_h^p\boldsymbol{b_0},\nabla\varepsilon_h^p\right)_K&\\ = -\omega^2\left(\rho_0\delta_h^p,\varepsilon_h^p\right)_K + 2i\omega\left(\delta_h^p\boldsymbol{b_0},\nabla\varepsilon_h^p\right)_K + \left(\boldsymbol{\varepsilon}_h^{\boldsymbol{q}},\boldsymbol{W_0}\boldsymbol{\delta}_h^{\boldsymbol{q}}\right)_K&\\ + \left\langle\boldsymbol{\delta}_h^{\boldsymbol{q}}\cdot\boldsymbol{n} + 2i\omega(\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\delta}_h^p + 2i\omega\tau P_M\delta_h^p - 2i\omega\tau_{\text{upw}}\left(\delta_h^p - \widehat{\delta}_h^p\right),\varepsilon_h^p - \widehat{\varepsilon}_h^p\right\rangle_{\partial\mathcal{T}_h}&\end{aligned} \tag{3.68}$$

Furthermore if $p \in H^s(\mathcal{O})$ and $\boldsymbol{q} \in \boldsymbol{H}^{\boldsymbol{t}}(\mathcal{O})$ where $s \in [1, k+2]$ and $t \in [1, k+1]$ then the following estimate holds

$$\begin{aligned}&\left|\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|^2_{\boldsymbol{W_0},\mathcal{T}_h} - 2i\omega\left(\left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|^2_{\partial\mathcal{T}_h} + \left\|\left(\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac{1}{2}}(\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|^2_{\partial\mathcal{T}_h}\right)\right|\\ \lesssim&\omega^2\|\varepsilon_h^p\|^2_{\mathcal{T}_h} + \omega\|\varepsilon_h^p\|_{\mathcal{T}_h}\left(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + h^t\|\boldsymbol{q}\|_{t,\mathcal{O}} + \left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} + \omega h^{s-1}\|p\|_{s,\mathcal{O}}\right)\\ &+ \|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h}\left(h^t\|\boldsymbol{q}\|_{t,\mathcal{O}} + \omega h^s\|p\|_{s,\mathcal{O}}\right) + h^{2t}\|\boldsymbol{q}\|^2_{t,\mathcal{O}} + \omega h^{s-1}\|p\|_{s,\mathcal{O}}h^t\|\boldsymbol{q}\|_{t,\mathcal{O}}\\ &+ \left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}\left(\omega h^{s-1}\|p\|_{s,\mathcal{O}} + h^t\|\boldsymbol{q}\|_{t,\mathcal{O}}\right)\end{aligned} \tag{3.69}$$

where

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

Proof: Test (3.64a)–(3.64b)–(3.64c) with $(\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}, \varepsilon_h^p, \widehat{\varepsilon}_h^p)$ and sum the resulting equations to obtain

$$\begin{aligned}\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|^2_{\boldsymbol{W_0},K} - \omega^2\|\varepsilon_h^p\|^2_{\rho_0,K} + 2i\omega\left(\varepsilon_h^p\boldsymbol{b_0},\nabla\varepsilon_h^p\right)_K + \left\langle\widehat{\mathcal{Q}}\cdot\boldsymbol{n} - \widehat{\mathcal{Q}_h}\cdot\boldsymbol{n} - \boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, \varepsilon_h^p - \widehat{\varepsilon}_h^p\right\rangle_{\partial K} =&\\ -\omega^2\left(\rho_0\delta_h^p,\varepsilon_h^p\right)_K + 2i\omega\left(\delta_h^p\boldsymbol{b_0},\nabla\varepsilon_h^p\right)_K + \left(\boldsymbol{\varepsilon}_h^{\boldsymbol{q}},\boldsymbol{W_0}\boldsymbol{\delta}_h^{\boldsymbol{q}}\right)_K&\end{aligned}$$

We will now compute the boundary terms using (3.65).
Boundary terms involving P_M:
Notice that, as $P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p \in M_h(K)$ and $\tau \in \mathcal{R}_0$, we have

$$\langle\tau(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p), P_M\varepsilon_h^p\rangle_{\partial K} = \langle\tau(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p), \varepsilon_h^p\rangle_{\partial K}$$

and therefore

$$\langle\tau(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p), \varepsilon_h^p - \widehat{\varepsilon}_h^p\rangle_{\partial K} = \langle\tau(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p), P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p\rangle_{\partial K} = \left\|\tau^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|^2_{\partial K}$$

.
We also have the following estimate

$$\begin{aligned}2\omega\left|\langle\tau P_M\delta_h^p, \varepsilon_h^p - \widehat{\varepsilon}_h^p\rangle_{\partial\mathcal{T}_h}\right| &\leqslant 2\omega\left|\langle|\tau|P_M\delta_h^p, P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p\rangle_{\partial\mathcal{T}_h}\right|\\ &\leqslant 2\omega\left|\left\langle|\tau|^{\frac{1}{2}}\delta_h^p, |\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\rangle_{\partial\mathcal{T}_h}\right|\\ &\lesssim \omega\left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}\||\tau|^{\frac{1}{2}}\|\delta_h^p\|_{\partial\mathcal{T}_h}\\ (\tau = \mathcal{O}(h^{-1}) \text{ and by (3.62d))} \quad &\lesssim \omega h^{s-1}\left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}\|p\|_{s,\mathcal{O}}\end{aligned}$$

Boundary terms involving convection:
As in *Step 2* of the proof of THEOREM 5, we will separate the volumetric term involving $\boldsymbol{b_0}$ into its real and imaginary parts. By the LEMMA 4.1.1 we have

$$\mathfrak{Re}\left(\varepsilon_h^p \boldsymbol{b_0}, \nabla \varepsilon_h^p\right)_K = \frac{1}{2}\left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})\varepsilon_h^p, \varepsilon_h^p\right\rangle_{\partial K}$$

and we can now obtain the second boundary norm

$$\begin{aligned}
&\left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\varepsilon}_h^p, \varepsilon_h^p - \widehat{\varepsilon}_h^p\right\rangle_{\partial\mathcal{T}_h} - \frac{1}{2}\left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})\varepsilon_h^p, \varepsilon_h^p\right\rangle_{\partial\mathcal{T}_h} + \left\langle \tau_{\text{upw}}(\varepsilon_h^p - \widehat{\varepsilon}_h^p), \varepsilon_h^p - \widehat{\varepsilon}_h^p\right\rangle_{\partial\mathcal{T}_h} \\
&= \left\langle -\frac{1}{2}(\boldsymbol{b_0}\cdot\boldsymbol{n})(\varepsilon_h^p - \widehat{\varepsilon}_h^p), \varepsilon_h^p - \widehat{\varepsilon}_h^p\right\rangle_{\partial\mathcal{T}_h} - \frac{1}{2}\left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\varepsilon}_h^p, \widehat{\varepsilon}_h^p\right\rangle_{\partial\mathcal{T}_h} + \left\langle \tau_{\text{upw}}(\varepsilon_h^p - \widehat{\varepsilon}_h^p), \varepsilon_h^p - \widehat{\varepsilon}_h^p\right\rangle_{\partial\mathcal{T}_h} \\
&= \left\langle \left(\tau_{\text{upw}} - \frac{1}{2}(\boldsymbol{b_0}\cdot\boldsymbol{n})\right)(\varepsilon_h^p - \widehat{\varepsilon}_h^p), \varepsilon_h^p - \widehat{\varepsilon}_h^p\right\rangle_{\partial\mathcal{T}_h} \\
&= \left\| \left(\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac{1}{2}}(\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}^2
\end{aligned}$$

indeed as $\widehat{\varepsilon}_h^p$ is single-valued across the skeleton of the mesh we have

$$\left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\varepsilon}_h^p, \widehat{\varepsilon}_h^p\right\rangle_{\partial\mathcal{T}_h} = \left\langle [\![\widehat{\varepsilon}_h^p \boldsymbol{b_0}]\!], \widehat{\varepsilon}_h^p\right\rangle_{\partial\mathcal{T}_h} = 0$$

and we also use that

$$\tau_{\text{upw}} - \frac{1}{2}(\boldsymbol{b_0}\cdot\boldsymbol{n}) = \frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}| \geqslant 0 \tag{3.70}$$

to use the square root.
We will now eliminate the terms involving τ_{upw} from the right-hand side.

$$\begin{aligned}
\left\| \left(\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac{1}{2}}(\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} &\lesssim \|\varepsilon_h^p - \widehat{\varepsilon}_h^p\|_{\partial\mathcal{T}_h} \\
&\lesssim \|\varepsilon_h^p - P_M\varepsilon_h^p\|_{\partial\mathcal{T}_h} + \|P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p\|_{\partial\mathcal{T}_h} \\
\text{(by LEMMA 3.4.4 and } \tau = \mathcal{O}(h^{-1})) \quad &\lesssim \|\varepsilon_h^p - \{\varepsilon_h^p\}\|_{\partial\mathcal{T}_h} + h^{\frac{1}{2}}\left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} \\
\text{(by (3.63))} \quad &\lesssim h^{-\frac{1}{2}}\|\varepsilon_h^p - \{\varepsilon_h^p\}\|_{\mathcal{T}_h} + h^{\frac{1}{2}}\left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} \\
\text{(by (3.66))} \quad &\lesssim h^{\frac{1}{2}}\|\nabla\varepsilon_h^p\|_{\mathcal{T}_h} + h^{\frac{1}{2}}\left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} \\
\text{(by LEMMA 3.4.3)} \quad &\lesssim h^{\frac{1}{2}}\left(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + \|\boldsymbol{\delta}_h^{\boldsymbol{q}}\|_{\mathcal{T}_h} + \left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}\right) \\
\text{(by (3.62e))} \quad &\lesssim h^{\frac{1}{2}}\left(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + h^t\|\boldsymbol{q}\|_{t,\mathcal{O}} + \left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}\right)
\end{aligned} \tag{3.71}$$

Using the following inequalities that can be derived from (3.70)

$$\begin{aligned}
\tau_{\text{upw}} &\lesssim \frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}| \\
\boldsymbol{b_0}\cdot\boldsymbol{n} &\lesssim \frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|
\end{aligned}$$

and using (3.62d) with $s-1$ instead of s to keep $s \leqslant k+2$, we deduce that

$$\begin{aligned}
2\omega \left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\delta}_h^p, \varepsilon_h^p - \widehat{\varepsilon}_h^p \right\rangle_{\partial\mathcal{T}_h} &= 2\omega \left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})^{\frac{1}{2}}\delta_h^p, (\boldsymbol{b_0}\cdot\boldsymbol{n})^{\frac{1}{2}}(\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\rangle_{\partial\mathcal{T}_h} \\
&\lesssim 2\omega \left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})^{\frac{1}{2}}\delta_h^p, (\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|)^{\frac{1}{2}}(\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\rangle_{\partial\mathcal{T}_h} \\
&\lesssim \omega \left\|\widehat{\delta}_h^p\right\|_{\partial\mathcal{T}_h} \left\| \left(\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac{1}{2}} (\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial\mathcal{T}_h} \\
&\lesssim \omega h^{s-\frac{3}{2}} \left\| \left(\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac{1}{2}} (\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial\mathcal{T}_h} \|p\|_{s,\mathcal{O}} \\
&\lesssim \omega h^{s-1} \Big(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} \\
&\qquad + \left\| |\tau|^{\frac{1}{2}} (P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial\mathcal{T}_h} \Big) \|p\|_{s,\mathcal{O}}
\end{aligned}$$

and

$$\begin{aligned}
2\omega \left\langle \tau_{\mathrm{upw}}(\delta_h^p - \widehat{\delta}_h^p), \varepsilon_h^p - \widehat{\varepsilon}_h^p \right\rangle_{\partial\mathcal{T}_h} &\lesssim \omega \left\|\delta_h^p - \widehat{\delta}_h^p\right\|_{\partial\mathcal{T}_h} \left\| \left(\tau_{\mathrm{upw}} - \frac{1}{2}\boldsymbol{b_0}\cdot\boldsymbol{n}\right)^{\frac{1}{2}} (\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial\mathcal{T}_h} \\
&\lesssim \omega h^{s-\frac{3}{2}} \left\| \left(\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac{1}{2}} (\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial\mathcal{T}_h} \|p\|_{s,\mathcal{O}} \\
&\lesssim \omega h^{s-1} \Big(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} \\
&\qquad + \left\| |\tau|^{\frac{1}{2}} (P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial\mathcal{T}_h} \Big) \|p\|_{s,\mathcal{O}}
\end{aligned}$$

Boundary term involving $\boldsymbol{\delta}_h^{\boldsymbol{q}}$:

$$\begin{aligned}
\langle \boldsymbol{\delta}_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, \varepsilon_h^p - \widehat{\varepsilon}_h^p\rangle_{\partial\mathcal{T}_h} &= \langle \boldsymbol{\delta}_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, \varepsilon_h^p - P_M\varepsilon_h^p + P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p\rangle_{\partial\mathcal{T}_h} \\
&= \underbrace{\langle \boldsymbol{\delta}_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, \varepsilon_h^p - P_M\varepsilon_h^p\rangle_{\partial\mathcal{T}_h}}_{=:T_1} + \underbrace{\langle \boldsymbol{\delta}_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p\rangle_{\partial\mathcal{T}_h}}_{=:T_2}
\end{aligned}$$

Using a weighted Cauchy-Schwartz inequality and recalling that $\tau = \mathcal{O}(h^{-1})$ we have

$$\begin{aligned}
T_2 &\lesssim \left\| |\tau|^{\frac{1}{2}} (P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial\mathcal{T}_h} h^{\frac{1}{2}} \|\boldsymbol{\delta}_h^{\boldsymbol{q}}\|_{\partial\mathcal{T}_h} \\
\text{(by (3.62e))} \quad &\lesssim h^t \left\| |\tau|^{\frac{1}{2}} (P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial\mathcal{T}_h} \|\boldsymbol{q}\|_{t,\mathcal{O}}
\end{aligned}$$

$$\begin{aligned}
T_1 &= \langle \boldsymbol{\delta}_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, \varepsilon_h^p - P_M\varepsilon_h^p\rangle_{\partial\mathcal{T}_h} \\
&\lesssim \|\boldsymbol{\delta}_h^{\boldsymbol{q}}\|_{\partial\mathcal{T}_h} \|\varepsilon_h^p - P_M\varepsilon_h^p\|_{\partial\mathcal{T}_h} \\
\text{(by Lemma 3.4.4)} \quad &\lesssim \|\boldsymbol{\delta}_h^{\boldsymbol{q}}\|_{\partial\mathcal{T}_h} \|\varepsilon_h^p - \{\varepsilon_h^p\}\|_{\partial\mathcal{T}_h} \\
\text{(by (3.62e) and (3.62d))} \quad &\lesssim h^{t-\frac{1}{2}} \|\boldsymbol{q}\|_{t,\mathcal{O}} h^{-\frac{1}{2}} \|\varepsilon_h^p - \{\varepsilon_h^p\}\|_{\mathcal{T}_h} \\
\text{(by (3.66))} \quad &\lesssim h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} \|\nabla\varepsilon_h^p\|_{\mathcal{T}_h} \\
\text{(by Lemma 3.4.3)} \quad &\lesssim h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} \left(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + \|\boldsymbol{\delta}_h^{\boldsymbol{q}}\|_{\mathcal{T}_h} + \left\| |\tau|^{\frac{1}{2}} (P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial\mathcal{T}_h} \right) \\
\text{by (3.62b))} \quad &\lesssim h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} \left(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + \left\| |\tau|^{\frac{1}{2}} (P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial\mathcal{T}_h} \right) + h^{2t} \|\boldsymbol{q}\|_{t,\mathcal{O}}^2
\end{aligned}$$

Therefore

$$\langle \boldsymbol{\delta}_h^{\boldsymbol{q}} \cdot \boldsymbol{n}, \varepsilon_h^p - \widehat{\varepsilon}_h^p \rangle_{\partial \mathcal{T}_h} \lesssim h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} \left(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + \left\| |\tau|^{\frac{1}{2}} (P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial \mathcal{T}_h} \right) + h^{2t} \|\boldsymbol{q}\|_{t,\mathcal{O}}^2$$

Volumetric terms: By similar computations using the Cauchy-Schwartz inequality, the projection estimates in (3.62) and LEMMA 3.4.3 we can show that

$$\begin{aligned}
\omega^2 (\rho_0 \delta_h^p, \varepsilon_h^p)_{\mathcal{T}_h} &\lesssim \omega^2 h^s \|p\|_{s,\mathcal{O}} \|\varepsilon_h^p\|_{\mathcal{T}_h} \\
2\omega (\delta_h^p \boldsymbol{b_0}, \nabla \varepsilon_h^p)_{\mathcal{T}_h} &\lesssim \omega h^s \|p\|_{s,\mathcal{O}} \left(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} + \left\| |\tau|^{\frac{1}{2}} (P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial \mathcal{T}_h} \right) \\
(\boldsymbol{W_0} \boldsymbol{\varepsilon}_h^{\boldsymbol{q}}, \boldsymbol{\delta}_h^{\boldsymbol{q}})_{\mathcal{T}_h} &\lesssim h^t \|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} \|\boldsymbol{q}\|_{t,\mathcal{O}} \\
2\omega (\varepsilon_h^p \boldsymbol{b_0}, \nabla \varepsilon_h^p)_{\mathcal{T}_h} &\lesssim \|\varepsilon_h^p\|_{\mathcal{T}_h} \left(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} + \left\| |\tau|^{\frac{1}{2}} (P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial \mathcal{T}_h} \right)
\end{aligned}$$

■

We can rewrite estimate (3.69) in a more readable form :

Corollary 3.4.1:
The following estimate holds

$$\begin{aligned}
&\left| \|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h}^2 - 2i\omega \left(\left\| |\tau|^{\frac{1}{2}} (P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial \mathcal{T}_h}^2 + \left\| \left(\frac{1}{2} |\boldsymbol{b_0} \cdot \boldsymbol{n}| \right)^{\frac{1}{2}} (\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial \mathcal{T}_h}^2 \right) \right| \\
\lesssim &\varepsilon \left(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h}^2 + \left\| |\tau|^{\frac{1}{2}} (P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial \mathcal{T}_h}^2 \right) + \frac{1}{\varepsilon} \left(\omega^2 \|\varepsilon_h^p\|_{\mathcal{T}_h}^2 + h^{2t} \|\boldsymbol{q}\|_{t,\mathcal{O}}^2 + \omega^2 h^{2s-2} \|p\|_{s,\mathcal{O}}^2 \right)
\end{aligned}$$

Proof: apply the weighted Young's inequality to the right-hand side of (3.69). The value of ε will be discussed later. ■

Adjoint problem

As the identity (3.68) does not allow us to directly obtain any error estimate, we need to use a duality argument. For an introduction to the *Aubin-Nitsche method* we refer to [EG04, Sec. 2.3.4], similar processes for HDG(+) methods in the context of wave equations have been carried out in [QSS16], [Hun19, Sec. 3.5] and [DS19, Sec. 3.5.2] and for coercive problems with convection in [QS16a].
The adjoint problem is

$$\begin{aligned}
\boldsymbol{W_0} \boldsymbol{\xi} - \nabla \theta &= 0 && \text{in } \mathcal{O} \\
-\rho_0 \omega^2 \theta - 2i\omega \boldsymbol{b_0} \cdot \nabla \theta - \operatorname{div}(\boldsymbol{\xi}) &= \varepsilon_h^p && \text{in } \mathcal{O} \\
\theta &= 0 && \text{on } \Gamma_D \\
\boldsymbol{\xi} \cdot \boldsymbol{n} - 2i\omega (\boldsymbol{b_0} \cdot \boldsymbol{n}) \theta &= 0 && \text{on } \Gamma_N
\end{aligned}$$

and $(\boldsymbol{\xi}, \theta) \in \boldsymbol{H^1}(\mathcal{O}) \times H^2(\mathcal{O})$ satisfy the following discrete problem for all $(\boldsymbol{r}_h, w_h, \mu_h) \in \boldsymbol{V}_h \times W_h \times M_h$

$$(\boldsymbol{W_0} \boldsymbol{\xi}, \boldsymbol{r}_h)_{\mathcal{T}_h} + (\theta, \operatorname{div}(\boldsymbol{r}_h))_{\mathcal{T}_h} = -\langle \theta, \boldsymbol{r}_h \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \tag{3.73a}$$

$$-\omega^2 (\rho_0 \theta, w_h)_{\mathcal{T}_h} + (2i\omega \theta \boldsymbol{b_0} + \boldsymbol{\xi}, \nabla w_h)_{\mathcal{T}_h} = -\langle \boldsymbol{\xi} \cdot \boldsymbol{n} - 2i\omega (\boldsymbol{b_0} \cdot \boldsymbol{n}) \theta, w_h \rangle_{\partial \mathcal{T}_h} + (\varepsilon_h^p, w_h)_{\mathcal{T}_h} \tag{3.73b}$$

$$\langle \boldsymbol{\xi} \cdot \boldsymbol{n}, \mu_h \rangle_{\partial \mathcal{T}_h \setminus \Gamma_D} = 0 \tag{3.73c}$$

The last equation (3.73c) translates the continuity of $\boldsymbol{\xi} \cdot \boldsymbol{n}$ between the elements and should be interpreted as a jump term. Indeed by the same argument as when we discussed weak continuity of $\boldsymbol{q}_h \cdot \boldsymbol{n}$ in SUBSECTION 3.4.1, we can show that (3.73c) is equivalent to

$$\sum_{K \in \mathcal{T}_h} \int_{\partial K} \boldsymbol{\xi} \cdot \boldsymbol{n} \mu_h^* \mathrm{d}\sigma = \sum_{e \in \mathcal{E}_h^i} \int_e [\![\boldsymbol{\xi}]\!] \mu_h^* \mathrm{d}\sigma - \int_{\Gamma_D} \boldsymbol{\xi} \cdot \boldsymbol{n} \mu_h^* \mathrm{d}\sigma. \tag{3.74}$$

Remark 3.4.6: The functional framework for (3.73c) is a bit complicated. Indeed the interior integrals should formally be interpreted as duality brackets between $H^{-\frac{1}{2}}$ and $H^{\frac{1}{2}}$ and the restriction of those distributions to a segment is not defined. Notice however that the right-hand side of (3.74) is well-defined, therefore giving meaning to the left-hand side. Moreover, as we assume additional regularity $(\boldsymbol{\xi}, \theta) \in \boldsymbol{H^1}(\mathcal{O}) \times H^2(\mathcal{O})$ for the solution of the adjoint problem and as we will work with polynomial quantities at the discrete level, this is not problematic.

In our analysis we will need to use the following elliptic regularity estimate for the dual problem

$$\|\theta\|_{2,\mathcal{O}} + \|\boldsymbol{\xi}\|_{1,\mathcal{O}} \leqslant C_{\text{reg}} \|\varepsilon_h^p\|_{\mathcal{O}}\,. \tag{3.75}$$

This estimate holds when $\boldsymbol{\xi}$ and θ are regular enough, which amounts to requiring enough regularity on the background quantities ρ_0, c_0 and $\boldsymbol{b_0}$, and the convexity of the domain $\mathcal{O}$.

Lemma 3.4.6:
We have the following dual identity :

$$\begin{aligned}
\|\varepsilon_h^p\|_{\mathcal{T}_h}^2 = &- (\boldsymbol{W_0}\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}, \boldsymbol{\pi_V}\boldsymbol{\xi} - \boldsymbol{\xi})_{\mathcal{T}_h} + \omega^2 (\rho_0 \varepsilon_h^p, \pi_W\theta - \theta)_{\mathcal{T}_h} + 2i\omega (\nabla\varepsilon_h^p, (\pi_W\theta - \theta)\boldsymbol{b_0})_{\mathcal{T}_h} \\
&- 2i\omega \langle (\boldsymbol{b_0}\cdot\boldsymbol{n})\varepsilon_h^p, \pi_W\theta - \theta\rangle_{\partial\mathcal{T}_h} \\
&+ (\boldsymbol{W_0}\boldsymbol{\delta}_h^{\boldsymbol{q}}, \boldsymbol{\pi_V}\boldsymbol{\xi})_{\mathcal{T}_h} - \omega^2 (\rho_0\delta_h^p, \pi_W\theta)_{\mathcal{T}_h} + 2i\omega (\delta_h^p\boldsymbol{b_0}, \nabla(\pi_W\theta))_{\mathcal{T}_h} \\
&+ 2i\omega \langle (\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\varepsilon}_h^p - \tau (P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p) + \tau_{\text{upw}} (\varepsilon_h^p - \widehat{\varepsilon}_h^p), \pi_W\theta - P_M\theta\rangle_{\partial\mathcal{T}_h} \\
&- 2i\omega \left\langle (\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\delta}_h^p - \tau P_M\delta_h^p + \tau_{\text{upw}} \left(\delta_h^p - \widehat{\delta}_h^p\right), \pi_W\theta - P_M\theta\right\rangle_{\partial\mathcal{T}_h} \\
&+ \langle \boldsymbol{\delta}_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, \pi_W\theta - P_M\theta\rangle_{\partial\mathcal{T}_h}
\end{aligned}$$

Proof: Introducing the projections in (3.73a)–(3.73b)–(3.73c) and testing with $(\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}, \varepsilon_h^p, \widehat{\varepsilon}_h^p)$

$$(\boldsymbol{W_0}\boldsymbol{\pi_V}\boldsymbol{\xi}, \boldsymbol{\varepsilon}_h^{\boldsymbol{q}})_{\mathcal{T}_h} + (\pi_W\theta, \text{div}\,(\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}))_{\mathcal{T}_h} - \langle P_M\theta, \boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = (\boldsymbol{W_0}\,(\boldsymbol{\pi_V}\boldsymbol{\xi} - \boldsymbol{\xi})\,, \boldsymbol{\varepsilon}_h^{\boldsymbol{q}})_{\mathcal{T}_h} \tag{3.76a}$$

$$\begin{aligned}
&-\omega^2 (\rho_0\pi_W\theta, \varepsilon_h^p)_{\mathcal{T}_h} + 2i\omega ((\pi_W\theta)\boldsymbol{b_0}, \nabla\varepsilon_h^p)_{\mathcal{T}_h} - (\text{div}\,(\boldsymbol{\pi_V}\boldsymbol{\xi})\,, \varepsilon_h^p)_{\mathcal{T}_h} - 2i\omega \langle (\boldsymbol{b_0}\cdot\boldsymbol{n})\pi_W\theta, \varepsilon_h^p\rangle_{\partial\mathcal{T}_h} = \\
&(\varepsilon_h^p, \varepsilon_h^p)_{\mathcal{T}_h} - \omega^2 (\rho_0(\pi_W\theta - \theta), \varepsilon_h^p)_{\mathcal{T}_h} + 2i\omega ((\pi_W\theta - \theta)\boldsymbol{b_0}, \nabla\varepsilon_h^p)_{\mathcal{T}_h} - 2i\omega \langle (\boldsymbol{b_0}\cdot\boldsymbol{n})(\pi_W\theta - \theta), \varepsilon_h^p\rangle_{\partial\mathcal{T}_h}
\end{aligned} \tag{3.76b}$$

$$\langle \boldsymbol{\pi_V}\boldsymbol{\xi}\cdot\boldsymbol{n}, \widehat{\varepsilon}_h^p\rangle_{\partial\mathcal{T}_h} = 0 \tag{3.76c}$$

Now conjugate and sum those equations and compare with the sum of (3.64a)–(3.64b)–(3.64c) tested with $(\boldsymbol{\pi_V}\boldsymbol{\xi}, \pi_W\theta, P_M\theta)$.

$$(\boldsymbol{W_0}\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}, \boldsymbol{\pi_V}\boldsymbol{\xi})_{\mathcal{T}_h} - (\varepsilon_h^p, \text{div}\,(\boldsymbol{\pi_V}\boldsymbol{\xi}))_{\mathcal{T}_h} + \langle \widehat{\varepsilon}_h^p, \boldsymbol{\pi_V}\boldsymbol{\xi}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = \ell_1(\boldsymbol{\pi_V}\boldsymbol{\xi}) \tag{3.77a}$$

$$-\omega^2 (\rho_0\varepsilon_h^p, \pi_W\theta)_{\mathcal{T}_h} - 2i\omega (\nabla\varepsilon_h^p, (\pi_W\theta)\boldsymbol{b_0})_{\mathcal{T}_h} \tag{3.77b}$$

$$+ (\text{div}\,(\boldsymbol{\varepsilon}_h^{\boldsymbol{q}})\,, \pi_W\theta)_{\mathcal{T}_h} + 2i\omega \langle (\boldsymbol{b_0}\cdot\boldsymbol{n})\varepsilon_h^p, \pi_W\theta\rangle_{\partial\mathcal{T}_h} = \ell_2(\pi_W\theta)$$

$$- \langle \boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, P_M\theta\rangle_{\partial\mathcal{T}_h} = \ell_3(P_M\theta) \tag{3.77c}$$

where

$$\begin{aligned}
\ell_1(\boldsymbol{\pi_V \xi}) &:= (\boldsymbol{W_0 \delta_h^q}, \boldsymbol{\pi_V \xi})_{\mathcal{T}_h} \\
\ell_2(\pi_W \theta) &:= -\omega^2 (\rho_0 \delta_h^p, \pi_W \theta)_{\mathcal{T}_h} + 2i\omega (\delta_h^p \boldsymbol{b_0}, \nabla \pi_W \theta)_{\mathcal{T}_h} \\
&\quad + 2i\omega \langle (\boldsymbol{b_0} \cdot \boldsymbol{n}) \widehat{\varepsilon}_h^p - \tau (P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p) + \tau_{\text{upw}} (\varepsilon_h^p - \widehat{\varepsilon}_h^p), \pi_W \theta \rangle_{\partial \mathcal{T}_h} \\
&\quad + \left\langle \boldsymbol{\delta_h^q} \cdot \boldsymbol{n} + 2i\omega (\boldsymbol{b_0} \cdot \boldsymbol{n}) \widehat{\delta}_h^p + 2i\omega \tau P_M \delta_h^p - 2i\omega \tau_{\text{upw}} (\delta_h^p - \widehat{\delta}_h^p), \pi_W \theta \right\rangle_{\partial \mathcal{T}_h} \\
\ell_3(P_M \theta) &:= -2i\omega \langle (\boldsymbol{b_0} \cdot \boldsymbol{n}) \widehat{\varepsilon}_h^p - \tau (P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p) + \tau_{\text{upw}} (\varepsilon_h^p - \widehat{\varepsilon}_h^p), P_M \theta \rangle_{\partial \mathcal{T}_h} \\
&\quad - \left\langle \boldsymbol{\delta_h^q} \cdot \boldsymbol{n} + 2i\omega (\boldsymbol{b_0} \cdot \boldsymbol{n}) \widehat{\delta}_h^p + 2i\omega \tau P_M \delta_h^p - 2i\omega \tau_{\text{upw}} \left(\delta_h^p - \widehat{\delta}_h^p \right), P_M \theta \right\rangle_{\partial \mathcal{T}_h}.
\end{aligned}$$

Notice that an integration by parts has been carried out in ℓ_2.
As $\boldsymbol{W_0}$ is real and symmetric, we have

$$(\boldsymbol{W_0 \pi_V \xi}, \boldsymbol{\varepsilon_h^q})_{\mathcal{T}_h} = (\boldsymbol{\pi_V \xi}, \boldsymbol{W_0 \varepsilon_h^q})_{\mathcal{T}_h}$$

and we can therefore notice that the left-hand sides of (3.76a)–(3.76b)–(3.76c) and (3.77a)–(3.77b)–(3.77c) (after being conjugated) are the same, leading to

$$\begin{aligned}
&(\boldsymbol{W_0 \varepsilon_h^q}, \boldsymbol{\pi_V \xi} - \boldsymbol{\xi})_{\mathcal{T}_h} + (\varepsilon_h^p, \varepsilon_h^p)_{\mathcal{T}_h} - \omega^2 (\rho_0 \varepsilon_h^p, \pi_W \theta - \theta)_{\mathcal{T}_h} \\
&-2i\omega (\nabla \varepsilon_h^p, (\pi_W \theta - \theta) \boldsymbol{b_0})_{\mathcal{T}_h} + 2i\omega \langle (\boldsymbol{b_0} \cdot \boldsymbol{n}) \varepsilon_h^p, \pi_W \theta - \theta \rangle_{\partial \mathcal{T}_h} \\
&\qquad = \ell_1(\boldsymbol{\pi_V \xi}) + \ell_2(\pi_W \theta) + \ell_3(P_M \theta)
\end{aligned}$$

And the identity is obtained by a reorganisation of the different terms. ∎

Remark 3.4.7: conjugating gives the good sign for the convection term, indeed : $[iz]^* = -iz^*$, therefore

$$\left[2i\omega ((\pi_W \theta) \boldsymbol{b_0}, \nabla \varepsilon_h^p)_{\mathcal{T}_h} \right]^* = -2i\omega (\nabla \varepsilon_h^p, (\pi_W \theta) \boldsymbol{b_0})_{\mathcal{T}_h}$$

Lemma 3.4.7:
Assuming that the regularity assumption (3.75) holds and that $\omega^2 h^2 \|\rho_0\|_\infty C_{\text{reg}} C$ (where C is the constant of THEOREM 5) is small enough, if $p \in H^s(\mathcal{O})$ and $\boldsymbol{q} \in \boldsymbol{H^t}(\mathcal{O})$ where $s \in [1, k+2]$ and $t \in [1, k+1]$ then

$$\|\varepsilon_h^p\|_{\mathcal{T}_h} \lesssim h^{t+1}(1+\omega) \|\boldsymbol{q}\|_{t,\mathcal{O}} + h^s(1+\omega+\omega^2) \|p\|_{s,\mathcal{O}} + \omega h \left\| \tau^{\frac{1}{2}} (P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|_{\partial \mathcal{T}_h} + h(1+\omega) \|\boldsymbol{\varepsilon_h^q}\|_{\boldsymbol{W_0}, \mathcal{T}_h}$$

where

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

Proof: we are going to estimate the terms in the right hand side of the LEMMA 3.4.6.
Volumetric terms involving $\boldsymbol{\varepsilon_h^q}$:

$$(\boldsymbol{W_0 \varepsilon_h^q}, \boldsymbol{\xi} - \boldsymbol{\pi_V \xi})_{\mathcal{T}_h} \lesssim h \|\boldsymbol{\varepsilon_h^q}\|_{\boldsymbol{W_0}, K} \|\boldsymbol{\xi}\|_{1,\mathcal{O}} \lesssim h \|\boldsymbol{\varepsilon_h^q}\|_{\boldsymbol{W_0}, \mathcal{T}_h} \|\varepsilon_h^p\|_{\mathcal{T}_h} \tag{3.78a}$$

and

$$\begin{aligned}
\left| (\boldsymbol{W_0 \delta_h^q}, \boldsymbol{\pi_V \xi})_{\mathcal{T}_h} \right| &= \left| (\boldsymbol{W_0 \delta_h^q}, \boldsymbol{\xi})_{\mathcal{T}_h} - (\boldsymbol{W_0 \delta_h^q}, \boldsymbol{\xi} - \boldsymbol{\pi_V \xi})_{\mathcal{T}_h} \right| \\
\text{(by (3.67))} \quad &= \left| (\boldsymbol{\delta_h^q}, \boldsymbol{W_0 \xi} - \{\boldsymbol{W_0 \xi}\})_{\mathcal{T}_h} - (\boldsymbol{W_0 \delta_h^q}, \boldsymbol{\xi} - \boldsymbol{\pi_V \xi})_{\mathcal{T}_h} \right| \\
&\lesssim \|\boldsymbol{q} - \boldsymbol{\pi_V q}\|_{\mathcal{T}_h} \|\boldsymbol{W_0 \xi} - \{\boldsymbol{W_0 \xi}\}\|_{\mathcal{T}_h} + \|\boldsymbol{q} - \boldsymbol{\pi_V q}\|_{\mathcal{T}_h} \|\boldsymbol{\xi} - \boldsymbol{\pi_V \xi}\|_{\mathcal{T}_h} \\
\text{(by (3.66) and (3.62b))} \quad &\lesssim h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} h |\boldsymbol{\xi}|_{1,\mathcal{O}} + h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} h \|\boldsymbol{\xi}\|_{1,\mathcal{O}} \\
\text{(by regularity (3.75))} \quad &\lesssim h^{t+1} \|\boldsymbol{q}\|_{t,\mathcal{O}} \|\varepsilon_h^p\|_{\mathcal{T}_h}
\end{aligned} \tag{3.78b}$$

Volumetric term involving ε_h^p :

$$\omega^2 \left(\rho_0 \varepsilon_h^p, \pi_W \theta - \theta\right)_{\mathcal{T}_h} \lesssim \omega^2 h^2 \left\|\varepsilon_h^p\right\|_{\rho_0,K} \|\theta\|_{2,\mathcal{O}} \lesssim \omega^2 h^2 \left\|\varepsilon_h^p\right\|_{\mathcal{O}} \tag{3.78c}$$

and

$$\begin{aligned}
\left|\omega^2 \left(\rho_0 \delta_h^p, \pi_W \theta\right)_{\mathcal{T}_h}\right| &= \left|\left(\rho_0 \delta_h^p, \theta\right)_{\mathcal{T}_h} - \left(\rho_0 \delta_h^p, \theta - \pi_W \theta\right)_{\mathcal{T}_h}\right| \\
\text{(by (3.67))} \quad &= \omega^2 \left|\left(\delta_h^p, \rho_0 \theta - \{\rho_0 \theta\}\right)_{\mathcal{T}_h} - \left(\rho_0 \delta_h^p, \theta - \pi_W \theta\right)_{\mathcal{T}_h}\right| \\
&\lesssim \omega^2 \left\|p - \pi_W p\right\|_{\mathcal{T}_h} \left\|\rho_0 \theta - \{\rho_0 \theta\}\right\|_{\mathcal{T}_h} + \omega^2 \left\|p - \pi_W p\right\|_{\mathcal{T}_h} \left\|\theta - \pi_W \theta\right\|_{\mathcal{T}_h} \\
\text{(by (3.66) and (3.62a))} \quad &\lesssim \omega^2 h^s \|p\|_{s,\mathcal{O}} h |\theta|_{1,\mathcal{O}} + \omega^2 h^s \|p\|_{s,\mathcal{O}} h \|\theta\|_{1,\mathcal{O}} \\
\text{(by regularity (3.75))} \quad &\lesssim \omega^2 h^{s+1} \|p\|_{s,\mathcal{O}} \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h}
\end{aligned} \tag{3.78d}$$

Volumetric convection term :

$$\begin{aligned}
2\omega \left(\nabla \varepsilon_h^p, (\pi_W \theta - \theta) \boldsymbol{b_0}\right)_{\mathcal{T}_h} &\lesssim \omega \left\|\nabla \varepsilon_h^p\right\|_{\mathcal{T}_h} \left\|\pi_W \theta - \theta\right\|_{\mathcal{T}_h} \\
\text{(by Lemma 3.4.3)} \quad &\lesssim \omega \left(\left\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\right\|_{\boldsymbol{W_0},\mathcal{T}_h} + \left\|\boldsymbol{\delta}_h^{\boldsymbol{q}}\right\|_{\mathcal{T}_h} + \left\|\tau^{\frac{1}{2}} (P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial \mathcal{T}_h} \right) \left\|\pi_W \theta - \theta\right\|_{\mathcal{T}_h}
\end{aligned}$$

using (3.62b), (3.62a), (3.66) and (3.75), we have

$$\begin{aligned}
&\lesssim \omega \left(\left\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\right\|_{\boldsymbol{W_0},\mathcal{T}_h} + h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} + \left\|\tau^{\frac{1}{2}} (P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial \mathcal{T}_h} \right) h \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h} \\
&\lesssim \omega \Big(h \left\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\right\|_{\boldsymbol{W_0},\mathcal{T}_h} + h^{t+1} \|\boldsymbol{q}\|_{t,\mathcal{O}} \\
&\qquad + h \left\|\tau^{\frac{1}{2}} (P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial \mathcal{T}_h} \Big) \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h}
\end{aligned} \tag{3.78e}$$

and

$$\begin{aligned}
2\omega \left|\left(\delta_h^p \boldsymbol{b_0}, \nabla(\pi_W \theta)\right)_{\mathcal{T}_h}\right| &= 2\omega \left|\left(\delta_h^p \boldsymbol{b_0}, \nabla \theta\right)_{\mathcal{T}_h} - \left(\delta_h^p \boldsymbol{b_0}, \nabla(\theta - \pi_W \theta)\right)_{\mathcal{T}_h}\right| \\
\text{(by (3.62a))} \quad &\lesssim \omega \left\|p - \pi_W p\right\|_{\mathcal{T}_h} \|\theta\|_{1,\mathcal{O}} \\
\text{(by (3.62a) and (3.75))} \quad &\lesssim \omega h^s \|p\|_{s,\mathcal{O}} \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h}
\end{aligned} \tag{3.78f}$$

Boundary terms involving $\boldsymbol{b_0} \cdot \boldsymbol{n}$: As we want to keep $s \leqslant k+2$, we will use (3.62c) with $s-1$ instead of s.
As

$$\left\langle (\boldsymbol{b_0} \cdot \boldsymbol{n}) \widehat{\varepsilon}_h^p, \pi_W \theta - P_M \theta \right\rangle_{\partial \mathcal{T}_h} = \left\langle (\boldsymbol{b_0} \cdot \boldsymbol{n}) \widehat{\varepsilon}_h^p, \pi_W \theta - \theta \right\rangle_{\partial \mathcal{T}_h}$$

we focus on

$$\begin{aligned}
2\omega \left\langle (\boldsymbol{b_0} \cdot \boldsymbol{n})(\varepsilon_h^p - \widehat{\varepsilon}_h^p), \pi_W \theta - \theta \right\rangle_{\partial \mathcal{T}_h} &\lesssim \omega \left\|\varepsilon_h^p - \widehat{\varepsilon}_h^p\right\|_{\partial \mathcal{T}_h} \left\|\pi_W \theta - \theta\right\|_{\partial \mathcal{T}_h} \\
\text{(by (3.71))} \quad &\lesssim \omega h^2 \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h} \Big[\left\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\right\|_{\boldsymbol{W_0},\mathcal{T}_h} + h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} \\
&\qquad + \left\||\tau|^{\frac{1}{2}} (P_M \varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial \mathcal{T}_h} \Big]
\end{aligned} \tag{3.78g}$$

and

$$\begin{aligned}
2\omega \left\langle (\boldsymbol{b_0} \cdot \boldsymbol{n}) \widehat{\delta}_h^p, \pi_W \theta - P_M \theta \right\rangle_{\partial \mathcal{T}_h} &\lesssim \omega \left\|\widehat{\delta}_h^p\right\|_{\partial \mathcal{T}_h} \left\|\pi_W \theta - P_M \theta\right\|_{\partial \mathcal{T}_h} \\
\text{(by (3.62d) and (3.62c))} \quad &\lesssim \omega h^{s-\frac{1}{2}} \|p\|_{s,\mathcal{O}} h^{\frac{3}{2}} \|\theta\|_{2,\mathcal{O}} \\
\text{(by (3.75))} \quad &\lesssim \omega h^s \|p\|_{s,\mathcal{O}} \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h}
\end{aligned} \tag{3.78h}$$

Boundary terms involving τ:

$$\begin{aligned} 2\omega \left\langle \tau \left(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p\right), \pi_W\theta - P_M\theta\right\rangle_{\partial\mathcal{T}_h} &\lesssim \omega \left\|\tau(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} \left\|\pi_W\theta - P_M\theta\right\|_{\partial\mathcal{T}_h} \\ (\tau = \mathcal{O}(h^{-1})) \quad &\lesssim \omega h^{-\frac{1}{2}} \left\|\tau^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} h^{\frac{3}{2}} \left\|\theta\right\|_{2,\mathcal{O}} \\ &\lesssim \omega h \left\|\tau^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h} \end{aligned} \tag{3.78i}$$

and

$$\begin{aligned} 2\omega \left\langle \tau P_M\delta_h^p, \pi_W\theta - P_M\theta\right\rangle_{\partial\mathcal{T}_h} &= 2\omega \left\langle \tau P_M\delta_h^p, P_M(\pi_W\theta - \theta)\right\rangle_{\partial\mathcal{T}_h} \\ &\lesssim \omega \left\|\tau P_M\delta_h^p\right\|_{\partial\mathcal{T}_h} \left\|P_M(\theta - \pi_W\theta)\right\|_{\partial\mathcal{T}_h} \\ (\tau = \mathcal{O}(h^{-1})) \quad &\lesssim \omega h^{-1} \left\|p - \pi_W p\right\|_{\partial\mathcal{T}_h} \left\|\theta - \pi_W\theta\right\|_{\partial\mathcal{T}_h} \\ (\text{by (3.62d)}) \quad &\lesssim \omega h^{-1} h^{s-\frac{1}{2}} \left\|p\right\|_{s,\mathcal{O}} h^{2-\frac{1}{2}} \left\|\theta\right\|_{2,\mathcal{O}} \\ &\lesssim \omega h^s \left\|p\right\|_{s,\mathcal{O}} \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h} \end{aligned} \tag{3.78j}$$

Boundary terms involving τ_{upw}: As we want to keep $s \leqslant k+2$, we will use (3.62c) with $s-1$ instead of s.

$$\begin{aligned} 2\omega \left\langle \tau_{\text{upw}} \left(\varepsilon_h^p - \widehat{\varepsilon}_h^p\right), \pi_W\theta - P_M\theta\right\rangle_{\partial\mathcal{T}_h} &\lesssim \omega \left\|\varepsilon_h^p - \widehat{\varepsilon}_h^p\right\|_{\partial\mathcal{T}_h} \left\|\pi_W\theta - P_M\theta\right\|_{\partial\mathcal{T}_h} \\ &\lesssim \omega h^{s-\frac{3}{2}} \left\|p\right\|_{s,\mathcal{O}} h^{\frac{3}{2}} \left\|\theta\right\|_{2,\mathcal{O}} \\ &\lesssim \omega h^s \left\|p\right\|_{s,\mathcal{O}} \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h} \end{aligned} \tag{3.78k}$$

and

$$\begin{aligned} 2\omega \left\langle \tau_{\text{upw}} \left(\delta_h^p - \widehat{\delta}_h^p\right), \pi_W\theta - P_M\theta\right\rangle_{\partial\mathcal{T}_h} &\lesssim \omega h^s \left\|p\right\|_{s,\mathcal{O}} \left\|\theta\right\|_{2,\mathcal{O}} \\ &\lesssim \omega h^s \left\|p\right\|_{s,\mathcal{O}} \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h} \end{aligned} \tag{3.78l}$$

Boundary term involving $\boldsymbol{\delta}_h^{\boldsymbol{q}}$:

$$\begin{aligned} \left\langle \boldsymbol{\delta}_h^{\boldsymbol{q}} \cdot \boldsymbol{n}, \pi_W\theta - P_M\theta\right\rangle_{\partial\mathcal{T}_h} &\lesssim h^{t-\frac{1}{2}} \left\|\boldsymbol{q}\right\|_{t,\mathcal{O}} h^{\frac{3}{2}} \left\|\theta\right\|_{2,\mathcal{O}} \quad (\text{by (3.62e) and (3.62d)}) \\ &\lesssim h^{t+1} \left\|\boldsymbol{q}\right\|_{t,\mathcal{O}} \left\|\varepsilon_h^p\right\|_{\mathcal{T}_h} \end{aligned} \tag{3.78m}$$

The desired estimate can now be obtained by collecting the estimates (3.78a), (3.78b), (3.78c), (3.78d), (3.78e), (3.78f), (3.78g), (3.78h), (3.78i), (3.78j), (3.78k), (3.78l) and (3.78m), and using that $h \longrightarrow 0$. ∎

Bootstraping process

We will now combine the results of Corollary 3.4.1 of Lemma 3.4.5 and Lemma 3.4.7 through a bootstrapping process to obtain a convergence result.

Theorem 6 : *Convergence of the HDG+ method*

Assuming that the regularity assumption (3.75) holds and that $\omega^2 h^2 \|\rho_0\|_\infty C_{\text{reg}} C$ (where C is the constant of THEOREM 5) is small enough, if $p \in H^s(\mathcal{O})$ and $\boldsymbol{q} \in \boldsymbol{H}^{\boldsymbol{t}}(\mathcal{O})$ where $s \in [1, k+2]$ and $t \in [1, k+1]$ then

$$\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + \sqrt{2\omega}\left(\left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} + \left\|\left(\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac{1}{2}}(\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}\right) \lesssim h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} + h^{s-1}\|p\|_{s,\mathcal{O}}$$

and

$$\|\varepsilon_h^p\|_{\mathcal{T}_h} \lesssim (1+\omega)\left(h^{t+1}\|\boldsymbol{q}\|_{t,\mathcal{O}} + h^s\|p\|_{s,\mathcal{O}}\right)$$

Optimal error estimates are

$$\|\pi_W p - p_h\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+2}) \quad \text{and} \quad \|\boldsymbol{\pi_V q} - \boldsymbol{q}_h\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+1}),$$

and

$$\|p - p_h\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+2}) \quad \text{and} \quad \|\boldsymbol{q} - \boldsymbol{q}_h\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+1}).$$

Proof: to make the computations easier we introduce the following notations

$$P := \omega\|\varepsilon_h^p\|_{\mathcal{T}_h} \quad ; \quad Q := \|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} \quad ; \quad T := \left\|\tau^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}$$

and

$$P_h := \omega h^{s-1}\|p\|_{s,\mathcal{O}} \quad ; \quad Q_h := h^t\|\boldsymbol{q}\|_{t,\mathcal{O}}$$

$$B := \left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} + \left\|\left(\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac{1}{2}}(\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h}$$

$$\alpha := (\omega+\omega^2)h \quad ; \quad \beta := \left(1+\omega+\omega^2\right)h \quad ; \quad \gamma := \omega^2 h$$

With those shorthands the estimate of LEMMA 3.4.7 can be rewritten as

$$\begin{aligned} P &\lesssim (\omega+\omega^2)hQ_h + (1+\omega+\omega^2)hP_h + \omega^2 hT + \left(\omega+\omega^2\right)hQ \\ &\lesssim \alpha Q_h + \beta P_h + \gamma B + \alpha Q \end{aligned} \tag{3.79}$$

and the estimate of COROLLARY 3.4.1 can be rewritten as

$$\begin{aligned} Q^2 + 2\omega B^2 \lesssim \left|Q^2 - 2i\omega B^2\right| &\lesssim \varepsilon\left(Q^2+T^2\right) + \frac{1}{\varepsilon}\left(P^2+Q_h^2+P_h^2\right) \\ &\lesssim \varepsilon\left(Q^2+B^2\right) + \frac{1}{\varepsilon}\left(P^2+Q_h^2+P_h^2\right) \end{aligned} \tag{3.80}$$

Taking the square of (3.79) and using Young's inequality leads to

$$P^2 \lesssim \alpha^2 Q_h^2 + \beta^2 P_h^2 + \gamma^2 B^2 + \alpha^2 Q^2. \tag{3.81}$$

Using (3.81) in (3.80) gives

$$\begin{aligned} Q^2 + 2\omega B^2 \lesssim \left(1+\frac{1}{\varepsilon}\right)\left[\left(1+\alpha^2\right)Q_h^2 + \left(1+\beta^2\right)P_h^2\right] + \left(\varepsilon + \gamma^2\left(1+\frac{1}{\varepsilon}\right)\right)B^2 \\ + \left(\varepsilon + \alpha^2\left(1+\frac{1}{\varepsilon}\right)\right)Q^2. \end{aligned}$$

Let C_1 denote the constant hidden in $\lesssim$. Choosing ε so that $C_1\varepsilon < 1$ and assuming that h is small enough the last two terms of the right hand side can be absorbed by the left hand side, leading to

$$Q^2 + 2\omega B^2 \lesssim \left(1 + \frac{1}{\varepsilon}\right)\left[\left(1+\alpha^2\right)Q_h^2 + \left(1+\beta^2\right)P_h^2\right].$$

As ε does not depend on ω and h, we can hide the first factor of the right-hand side into $\lesssim$ leading to

$$Q^2 + 2\omega B^2 \lesssim \left(1+\alpha^2\right)Q_h^2 + \left(1+\beta^2\right)P_h^2.$$

As $\alpha, \beta = \mathcal{O}(h)$, we can overestimate $\alpha, \beta \lesssim 1$ leading to

$$Q^2 + 2\omega B^2 \lesssim Q_h^2 + P_h^2.$$

And finally we have

$$Q + \sqrt{2\omega}B \lesssim P_h + Q_h. \tag{3.82}$$

Now by taking $s = k+2$ and $t = k+1$ in (3.62b) and (3.62d), we can see that

$$Q = \mathcal{O}(h^{k+1}) \quad \text{and} \quad B = \mathcal{O}(h^{k+1}) \tag{3.83}$$

and finally by using (3.82) in (3.79), we have

$$P = \mathcal{O}(h^{k+2}). \tag{3.84}$$

■

It is also possible to obtain a convergence result for the trace $\widehat{p_h}$ which is the main unknown of the method :

Corollary 3.4.2:
Under the assumptions of THEOREM 6, the following error estimates for $\widehat{p_h}$ hold

$$\|\widehat{\varepsilon}_h^p\|_{\partial\mathcal{T}_h} = \mathcal{O}(h^{k+\frac{3}{2}}) \quad \text{and} \quad \|p - \widehat{p_h}\|_{\partial\mathcal{T}_h} = \mathcal{O}(h^{k+\frac{1}{2}}).$$

Proof: We have

$$\begin{aligned}
\|\varepsilon_h^p\|_{\partial\mathcal{T}_h} &\leqslant \|P_M\varepsilon_h^p\|_{\partial\mathcal{T}_h} + \|P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p\|_{\partial\mathcal{T}_h} \\
(\tau = \mathcal{O}(h^{-1})) \quad &\lesssim \|P_M\varepsilon_h^p\|_{\partial\mathcal{T}_h} + h^{\frac{1}{2}}\left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} \\
(P_M \text{ continuous}) \quad &\lesssim \|\varepsilon_h^p\|_{\partial\mathcal{T}_h} + h^{\frac{1}{2}}\left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} \\
(\text{by (3.63)}) \quad &\lesssim h^{-\frac{1}{2}}\|\varepsilon_h^p\|_{\mathcal{T}_h} + h^{\frac{1}{2}}\left\||\tau|^{\frac{1}{2}}(P_M\varepsilon_h^p - \widehat{\varepsilon}_h^p)\right\|_{\partial\mathcal{T}_h} \\
(\text{by (3.83) and (3.84)}) \quad &= \mathcal{O}(h^{-\frac{1}{2}}h^{k+2} + h^{\frac{1}{2}}h^{k+1}) \\
&= \mathcal{O}(h^{k+\frac{3}{2}}),
\end{aligned}$$

which is the first estimate. The second one comes from (3.62c) with $s = k+1$. ■

3.4.4 Error analysis of the HDG method with diffusive flux

As the error analysis for the HDG method is very similar to the one for the HDG+ method we only state the main theorem. The intermediate results (error equations, Garding's equality, dual estimate, ...) are stated without proof in Appendix 3.A.

Theorem 7 : *Convergence of the HDG method with diffusive flux*

Assuming that the regularity assumption (3.75) holds and that $\omega^2 h^2 \|\rho_0\|_\infty C_{\text{reg}} C$ (where C is the constant of Theorem 5) is small enough, if $p \in H^s(\mathcal{O})$ and $\boldsymbol{q} \in \boldsymbol{H}^t(\mathcal{O})$ where $s, t \in [1, k+1]$ then

$$\|\pi_W p - p_h\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+\frac{3}{2}}) \;\; ; \;\; \|P_M p - \widehat{p_h}\|_{\partial\mathcal{T}_h} = \mathcal{O}(h^{k+\frac{1}{2}}) \;\; ; \;\; \text{and} \;\; \|\boldsymbol{\pi_V q} - \boldsymbol{q}_h\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+\frac{1}{2}}),$$

and

$$\|p - p_h\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+1}) \;\; ; \;\; \|p - \widehat{p_h}\|_{\partial\mathcal{T}_h} = \mathcal{O}(h^{k+\frac{1}{2}}) \;\; ; \;\; \text{and} \;\; \|\boldsymbol{q} - \boldsymbol{q}_h\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+\frac{1}{2}}).$$

Remark 3.4.8: Some HDG methods are known to achieve superconvergence, *ie* taking $p_h \in \mathcal{P}_k$ leads to the following error estimate

$$\|\pi_W p - p_h\|_{\mathcal{T}_h} = \mathcal{O}(h^{k+2}),$$

this is not possible for this method because of the convection term

$$2\omega \left| (\delta_h^p \boldsymbol{b_0}, \nabla(\pi_W \theta))_{\mathcal{T}_h} \right| \lesssim \omega h^s \|p\|_{s,\mathcal{O}} \|\varepsilon_h^p\|_{\mathcal{T}_h}$$

in the dual estimate which locks the convergence rate to $\mathcal{O}(h^{k+1})$ for the scalar variable p_h. Superconvergence is an attractive property for a numerical scheme, indeed using a postprocessing scheme it is possible to use the solution $(p_h, \boldsymbol{q}_h)$ to construct a new approximation $\widetilde{p_h}$ which converges with order $\mathcal{O}(h^{k+2})$, see [Ste91], [CGS10, Sec. 5] for more details.

3.5 Implementation

In this section, we will give details on how we have implements the HDG(+) methods. For the sake of simplicity, we will assume that the physical coefficients are constant on each element. In this case, the integrals may be computed using

- an analytic integration procedure that relies on the decomposition of the basis functions in the monomial basis,
- high-order quadrature rules.

Taking into account the variations of the physical parameters inside the elements only requires a straightforward generalization of the material presented here. However this implementation is only possible if quadrature rules are used to evaluate the integrals.

To make the notations lighter, we will drop the subscript h in this section. The quantities $(\boldsymbol{q}, p, \widehat{p})$ will denote the solutions of (3.47a)–(3.47b)–(3.47c) and $(\boldsymbol{q}^K, p^K, \widehat{p}^e)$ their restrictions to an element $K \in \mathcal{T}_h$ and an edge $e \in \mathcal{E}_h$ respectively.
In Table 3.2, we recall the main differences between the three variants of the HDG methods considered in this work.

Variable	Space	HDG $(p_h, \boldsymbol{\sigma}_h)$	HDG $(p_h, \boldsymbol{q}_h)$	HDG+ $(p_h, \boldsymbol{q}_h)$
Pressure p_h	$W_h(K)$	$\mathcal{P}_k(K)$		$\mathcal{P}_{k+1}(K)$
Flux $\boldsymbol{q}_h$ or $\boldsymbol{\sigma}_h$	$\boldsymbol{V}_h(K)$	$\boldsymbol{\mathcal{P}}_k(K)$		
Trace $\widehat{p_h}$	$M_h(e)$	$\mathcal{P}_k(e)$		
Penalization $\tau\|_K$		$\mathcal{O}(1)$		$\mathcal{O}(h_K^{-1})$

Table 3.2: Choice of local spaces and penalization parameter for the different methods

3.5.1 Framework and notations

In this section we present the notations used in [BCDL15, FS20] which are very close to the ones used in `hawen`[6]. We will then give details on how the matrices are assembled.

Local problem: The discretization of the local problem (3.47a)–(3.47b) leads to the following system

$$\mathbb{A}^K \underline{W^K} + \mathbb{C}^K \underline{\Lambda^K} = \mathcal{S}^K, \tag{3.85}$$

where

$$\underline{W^K} := \begin{bmatrix} \underline{p^K} & \underline{q_x^K} & \underline{q_y^K} & \underline{q_z^K} \end{bmatrix}^T \quad \text{and} \quad \underline{\Lambda^K} := \begin{bmatrix} \underline{\widehat{p}^{g(K,1)}} & \underline{\widehat{p}^{g(K,2)}} & \underline{\widehat{p}^{g(K,3)}} & \underline{\widehat{p}^{g(K,4)}} \end{bmatrix}^T,$$

and $\underline{p^K}$ denotes the vector of the coefficients of p_h^K in the basis of $W_h(K)$.

Global problem: The discretization of the transmission condition (3.47c) leads to

$$\sum_{K \in \mathcal{T}_h} \mathbb{B}^K \underline{W^K} + \mathbb{L}^K \underline{\Lambda^K} = 0 \tag{3.86}$$

We denote by $\underline{\Lambda}$ the global trace approximation and for each element $K \in \mathcal{T}_h$ we introduce the *connectivity map* which is the operator $\mathcal{A}_K$ such that[7]

$$\mathcal{A}_K \underline{\Lambda} = \underline{\Lambda^K}.$$

Using (3.85), we can express $\underline{W^K}$ in terms of $\underline{\Lambda}$

$$\underline{W^K} = \left(\mathbb{A}^K\right)^{-1} \mathcal{S}^K - \left(\mathbb{A}^K\right)^{-1} \mathbb{C}^K \mathcal{A}_K \underline{\Lambda}$$

and we can finally construct the global problem using (3.86)

$$\sum_{K \in \mathcal{T}_h} \mathcal{A}_K^T \left[-\mathbb{B}^K \left(\mathbb{A}^K\right)^{-1} \mathbb{C}^K + \mathbb{L}^K \right] \mathcal{A}_K \underline{\Lambda} = - \sum_{K \in \mathcal{T}_h} \mathcal{A}_K^T \mathbb{B}^K \left(\mathbb{A}^K\right)^{-1} \mathcal{S}^K. \tag{3.87}$$

For conciseness, we will denote

$$\mathbb{K}^K := -\mathbb{B}^K \left(\mathbb{A}^K\right)^{-1} \mathbb{C}^K + \mathbb{L}^K, \quad \mathbb{K} := \sum_{K \in \mathcal{T}_h} \mathcal{A}_K^T \mathbb{K}^K \mathcal{A}_K \quad \text{and} \quad \mathcal{S} := - \sum_{K \in \mathcal{T}_h} \mathcal{A}_K^T \mathbb{B}^K \left(\mathbb{A}^K\right)^{-1} \mathcal{S}^K.$$

Remark 3.5.1: the invertibility of matrix $\mathbb{A}^K$ was the subject of Theorem 2 (HDG-$\boldsymbol{\sigma}_h$) and Theorem 5 (HDG+ and HDG-$\boldsymbol{q}_h$) and the invertibility of the global matrix $\mathbb{K}$ was the subject of Theorem 4.

[6] See `https://ffaucher.gitlab.io/hawen-website/`.

[7] The aim of operator $\mathcal{A}_K$ is to copy the global informations to the local solver, depending on the element it may be a simple copy or also involve reordering.

Remark 3.5.2: to efficiently implement the HDG(+) methods, we will need to compute the *inverse* of $\mathbb{A}^K$. As this matrix is not too large, it is possible to perform this operation using `lapack`. However the discretization of the global problem will lead to a large sparse system which will be solved using `mumps`.

Solving HDG: In Algorithm 3, we describe how the HDG method can be efficiently implemented. Details on the construction of the local matrices $\mathbb{A}^K$, $\mathbb{C}^K$, $\mathbb{B}^K$ and $\mathbb{L}^K$ and on the implementation of boundary conditions will be given in the next sections.

Algorithm 3: Solving HDG

`/* Step 1: Construction of the local and global problems */`

for $K \in \mathcal{T}_h$ **do**
- Construct the local matrices $\mathbb{A}^K$, $\mathbb{C}^K$, $\mathbb{B}^K$ and $\mathbb{L}^K$
- Compute $(\mathbb{A}^K)^{-1}$ using `lapack`
- Compute $\mathbb{K}^K = \mathbb{B}^K\left(\mathbb{A}^K\right)^{-1}\mathbb{C}^K + \mathbb{L}^K$
- Modify $\mathbb{K}^K$ to enforce the boundary conditions
- Use the connectivity operator $\mathcal{A}_K$ to add the local contribution to the global matrix $\mathbb{K}$

`/* Step 2: Construction of the source term */`

Localize the source

for $K \in \mathcal{T}_h$ where $s \neq 0$ **do**
- Construct the local source term $\mathcal{S}^K$
- Compute $-\mathbb{B}^K(\mathbb{A}^K)^{-1}\mathcal{S}^K$
- Use the connectivity operator $\mathcal{A}_K$ to add the local contribution to the global source $\mathcal{S}$

`/* Step 3: Resolution of the global system */`

Solve $\mathbb{K}\underline{\Lambda} = \mathcal{S}$ with `mumps`

`/* Step 4: Reconstruction of the solution */`

for $K \in \mathcal{T}_h$ **do**
- Compute the local unknowns $\underline{W^K} = \left(\mathbb{A}^K\right)^{-1}\mathcal{S}^K - \left(\mathbb{A}^K\right)^{-1}\mathbb{C}^K\mathcal{A}_K\underline{\Lambda}$

Convention for the indices: we will always use the following convention for the indices :

- index of basis functions : i (test), j (trial), r (other)
- component of a vector : $u, v \in \{x, y, z\}$
- global number of an edge : m
- local number of an edge : ℓ
- polynomial degree : k

The global number of the ℓ-th face of element K will be denoted by $g(K, \ell)$.

Dimension of polynomial spaces: We recall that the dimension of the space of polynomials of n variables and of degree up to k is given by

$$\dim\left(\mathcal{P}_k(K)\right) = \binom{k+n}{n} =: d_n(k), \qquad \text{for } K \subset \mathbb{R}^n.$$

3.5.2 Implementation of the diffusive flux HDG method

In this section, we focus on the implementation of the HDG-$\boldsymbol{q}_h$ method.

Definition of the elementary matrices

We introduce the basis $(\Phi_j^K)_{j=1}^{d_n(k)}$ of $\mathcal{P}_k(K)$ and $(\Psi_j^e)_{j=1}^{d_{n-1}(k)}$ of $\mathcal{P}_k(e)$ and we decompose the unknowns as

$$p^K = \sum_{j=1}^{d_n(k)} p_j^K \Phi_j^K; \quad q_u^K = \sum_{j=1}^{d_n(k)} q_{u,j}^K \Phi_j^K; \quad \widehat{p}^e(x) = \sum_{j=1}^{d_{n-1}(k)} \widehat{p}_j^e \Psi_j^e$$

for $u \in \{x, y, z\}$, $K \in \mathcal{T}_h$ and $e \in \mathcal{E}_h$.
We introduce the following matrices

$$\begin{aligned} \mathbb{M}_{ij}^K &= \int_K \Phi_i^K \Phi_j^K \mathrm{d}\boldsymbol{x} \qquad \mathbb{D}_{u,ij}^K = \int_K \Phi_j^K \partial_u \Phi_i^K \mathrm{d}\boldsymbol{x} \\ \mathbb{E}_{\ell,ij}^K &= \int_{\partial K^\ell} \Phi_i^K \Phi_j^K \mathrm{d}\sigma \quad \mathbb{F}_{\ell,ij}^K = \int_{\partial K^\ell} \Psi_j^{g(K,\ell)} \Phi_i^K \mathrm{d}\sigma \\ \mathbb{G}_{ij}^m &= \int_{e_m} \Psi_i^m \Psi_j^m \mathrm{d}\sigma \end{aligned} \tag{3.88}$$

where $g(K, \ell)$ is the global number of the ℓ-th face of element K and m is also the global number of the edge e_m.
Notice that all those matrices except $\mathbb{F}_\ell^K$ (which has dimension $d_n(k) \times d_{n-1}(k)$) are square matrices and that all the elementary matrices are real.

Local problem

We introduce the following shorthand to make the algebra easier

$$\alpha_\ell = \min_{\partial K^\ell}(\boldsymbol{b_0} \cdot \boldsymbol{n}, 0) \quad \text{and} \quad t_{\mathrm{upw},\ell} = \max_{\partial K^\ell}(\boldsymbol{b_0} \cdot \boldsymbol{n}, 0)$$

it corresponds to using the expression (3.45a)–(3.45b) of the numerical flux. From an implementation point of view, this can be evaluated using array slicing or a ternary operator.
Using the elementary matrices introduced in (3.88) we can write (3.47a) as

$$\sum_u W_{0,vu} \mathbb{M}^K \underline{q_u^K} - \mathbb{D}_v^K \underline{p^K} + \sum_\ell n_v^{K,\ell} \mathbb{F}_\ell^K \underline{\widehat{p}^{g(K,\ell)}} = 0, \qquad \forall v \in \{x, y, z\} \tag{3.89}$$

and (3.47b) as

$$\begin{aligned} -\omega^2 \rho_0 \mathbb{M}^K \underline{p^K} + 2i\omega \sum_u b_{0,u} \mathbb{D}_u^K \underline{p^K} + \sum_u (\mathbb{D}_u^K)^T \underline{q_u} \\ + 2i\omega \sum_\ell \left[-\alpha_\ell \mathbb{F}_\ell^K \underline{\widehat{p}^{g(K,\ell)}} - t_{\mathrm{upw},\ell} \mathbb{E}_\ell^K \underline{p^K} + \tau_\ell (\mathbb{E}_\ell^K \underline{p^K} - \mathbb{F}_\ell^K \underline{\widehat{p}^{g(K,\ell)}}) \right] = \mathbb{S}^K \end{aligned} \tag{3.90}$$

Matrix form: To construct the global problem (3.87) we need to construct the matrices $\mathbb{A}^K$ and $\mathbb{C}^K$ of the local problem (3.85) using (3.89) and (3.90).

Matrix $\mathbb{A}^K$:

$\mathbb{A}^K$	$\underline{p^K}$	$\underline{q_x^K}$	$\underline{q_y^K}$	$\underline{q_z^K}$
$\underline{p^K}$	$-\rho_0\omega^2\mathbb{M}^K + 2i\omega\sum_u b_{0,u}\mathbb{D}_u^K$ $+2i\omega\sum_\ell(\tau_\ell - t_{\text{upw},\ell})\mathbb{E}_\ell^K$	$(\mathbb{D}_x^K)^T$ $-\mathbb{D}_x^K + \sum_\ell n_x^{K,\ell}\mathbb{E}_\ell^K$	$(\mathbb{D}_y^K)^T$ $-\mathbb{D}_y^K + \sum_\ell n_y^{K,\ell}\mathbb{E}_\ell^K$	$(\mathbb{D}_z^K)^T$ $-\mathbb{D}_z^K + \sum_\ell n_y^{K,\ell}\mathbb{E}_\ell^K$
$\underline{q_x^K}$	$-\mathbb{D}_x^K$	$W_{0,11}^K\mathbb{M}^K$	$W_{0,12}^K\mathbb{M}^K$	$W_{0,13}^K\mathbb{M}^K$
$\underline{q_y^K}$	$-\mathbb{D}_y^K$	$W_{0,21}^K\mathbb{M}^K$	$W_{0,22}^K\mathbb{M}^K$	$W_{0,23}^K\mathbb{M}^K$
$\underline{q_z^K}$	$-\mathbb{D}_z^K$	$W_{0,31}^K\mathbb{M}^K$	$W_{0,32}^K\mathbb{M}^K$	$W_{0,33}^K\mathbb{M}^K$

The red terms can be used instead of the black ones, depending on wether or not an integration by parts is performed.
Matrix $\mathbb{C}^K$:

$\mathbb{C}^K$	$\underline{\widehat{p}^{g(K,1)}}$	$\underline{\widehat{p}^{g(K,2)}}$	$\underline{\widehat{p}^{g(K,3)}}$	$\underline{\widehat{p}^{g(K,4)}}$
$\underline{p^K}$	$-2i\omega(\tau_1+\alpha_1)\mathbb{F}_1^K$	$-2i\omega(\tau_2+\alpha_2)\mathbb{F}_2^K$	$-2i\omega(\tau_3+\alpha_3)\mathbb{F}_3^K$	$-2i\omega(\tau_4+\alpha 4)\mathbb{F}_4^K$
$\underline{q_x^K}$	$n_x^{K,1}\mathbb{F}_1^K$	$n_x^{K,2}\mathbb{F}_2^K$	$n_x^{K,3}\mathbb{F}_3^K$	$n_x^{K,4}\mathbb{F}_4^K$
$\underline{q_y^K}$	$n_y^{K,1}\mathbb{F}_1^K$	$n_y^{K,2}\mathbb{F}_2^K$	$n_y^{K,3}\mathbb{F}_3^K$	$n_y^{K,4}\mathbb{F}_4^K$
$\underline{q_z^K}$	$n_z^{K,1}\mathbb{F}_1^K$	$n_z^{K,2}\mathbb{F}_2^K$	$n_z^{K,3}\mathbb{F}_3^K$	$n_z^{K,4}\mathbb{F}_4^K$

Remark 3.5.3: We chose to write the matrix form of the method with the equation for p^K first which is the opposite of the system (3.47a)–(3.47b)–(3.47c). This convention is interesting because the matrices for the method in dimension n are submatrices of the ones for dimension $n+1$.

Global problem

The transmission condition (3.47c) can be written as

$$\sum_{K,\ell}\Big(\sum_u n_u^{K,\ell}(\mathbb{F}_\ell^K)^T\underline{q_u^K}$$
$$+2i\omega\Big[-\Big(\alpha_\ell\mathbb{G}^{g(K,\ell)}\underline{\widehat{p}}^{g(K,\ell)} + t_{\text{upw},\ell}(\mathbb{F}_\ell^K)^T\underline{p^K}\Big) + \tau\Big((\mathbb{F}_\ell^K)^T\underline{p^K} - \mathbb{G}^{g(K,\ell)}\underline{\widehat{p}}^{g(K,\ell)}\Big)\Big]\Big) = 0 \quad (3.91)$$

Matrix form: We will now use (3.91) to construct the matrices $\mathbb{B}^K$ and $\mathbb{L}^K$ of (3.86).

Matrix $\mathbb{B}^K$:

$\mathbb{B}^K$	$\underline{p^K}$	$\underline{q_x^K}$	$\underline{q_y^K}$	$\underline{q_z^K}$
e_1	$2i\omega(\tau_1 - t_{\text{upw},1})(\mathbb{F}_1^K)^T$	$n_x^{K,1}(\mathbb{F}_1^K)^T$	$n_y^{K,1}(\mathbb{F}_1^K)^T$	$n_z^{K,1}(\mathbb{F}_1^K)^T$
e_2	$2i\omega(\tau_2 - t_{\text{upw},2})(\mathbb{F}_2^K)^T$	$n_x^{K,2}(\mathbb{F}_2^K)^T$	$n_y^{K,2}(\mathbb{F}_2^K)^T$	$n_z^{K,2}(\mathbb{F}_2^K)^T$
e_3	$2i\omega(\tau_3 - t_{\text{upw},3})(\mathbb{F}_3^K)^T$	$n_x^{K,3}(\mathbb{F}_3^K)^T$	$n_y^{K,3}(\mathbb{F}_3^K)^T$	$n_z^{K,3}(\mathbb{F}_3^K)^T$
e_4	$2i\omega(\tau_4 - t_{\text{upw},4})(\mathbb{F}_4^K)^T$	$n_x^{K,4}(\mathbb{F}_4^K)^T$	$n_y^{K,4}(\mathbb{F}_4^K)^T$	$n_z^{K,4}(\mathbb{F}_4^K)^T$

Matrix $\mathbb{L}^K$:

$\mathbb{L}^K$	$\underline{\widehat{p}^{g(K,1)}}$	$\underline{\widehat{p}^{g(K,2)}}$	$\underline{\widehat{p}^{g(K,3)}}$	$\underline{\widehat{p}^{g(K,4)}}$
e_1	$-2i\omega(\tau_1 + \alpha_1)\mathbb{G}^{g(K,1)}$			
e_2		$-2i\omega(\tau_2 + \alpha_2)\mathbb{G}^{g(K,2)}$		
e_3			$-2i\omega(\tau_3 + \alpha_3)\mathbb{G}^{g(K,3)}$	
e_4				$-2i\omega(\tau_4 + \alpha_4)\mathbb{G}^{g(K,4}$

Boundary conditions:

The transmission condition (3.47c) was also used to weakly enforce the boundary conditions

$$\begin{aligned} \boldsymbol{q}\cdot\boldsymbol{n} - 2i\omega(\boldsymbol{b_0}\cdot\boldsymbol{n})p &= g_N && \text{on } \Gamma_N \\ p &= g_D && \text{on } \Gamma_D \end{aligned}$$

Neumann boundary condition: If we want to enforce the Neumann boundary condition on edge ℓ of element K we only need to add a right-hand side to (3.91) :

$$\mathbb{G}^{g(K,\ell)}\underline{g_N}$$

where we assumed that

$$g_N = \sum_i g_{N,i}\Psi_i.$$

Weak Dirichlet boundary conditions: However if we want to enforce the Dirichlet boundary condition on edge ℓ of element K we need to change the matrices $\mathbb{B}^K$ and $\mathbb{L}^K$. This first method of implementation corresponds to weakly enforcing the boundary conditions, *ie* using the expression

$$\langle \widehat{p_h} - g_D, \mu\rangle_{\Gamma_D} = 0.$$

This method is used for a *modal* implementation of the HDG method. It can be implemented as follows described in Algorithm 4.

Algorithm 4: Implementation of weak Dirichlet BC

1 The corresponding row of $\mathbb{B}^K$ is set to 0 : $\mathbb{B}^K[e_\ell, :] = 0$
2 The corresponding entry of $\mathbb{L}^K$ is changed : $\mathbb{L}^K[e_\ell, \underline{\widehat{p}^{g(K,\ell)}}] = \mathbb{G}^{g(K,\ell)}$
3 The following right-hand side is added : $\mathbb{G}^{g(K,\ell)}\underline{g_D}$, where $g_D = \sum_i g_{D,i}\Psi_i$.

Strong Dirichlet boundary conditions: For a nodal implementation of the HDG method, it is also possible to strongly enforce the boundary condition on Γ_D :

$$\widehat{p_h} = g_D,$$

on edge ℓ of element K. We need to change the local contribution

$$\mathbb{K}^K = \mathbb{L}^K - \mathbb{B}^K\left(\mathbb{A}^K\right)^{-1}\mathbb{C}^K,$$

to the global problem as described in Algorithm 5. We recall that $\mathbb{K}^K$ has the same shape as $\mathbb{L}^K$.

Algorithm 5: Implementation of strong Dirichlet BC

1 The corresponding entries of $\mathbb{K}^K$ are replaced with an identity block : $\mathbb{K}^K[e_\ell, \underline{\widehat{p}^{g(K,\ell)}}] = \mathbf{Id}$
2 The following right-hand side is added : $[g_D(\boldsymbol{x}_r)]_r^T$, where $\boldsymbol{x}_r \in \text{dof}\left(e^{g(K,\ell)}\right)$

3.5.3 Implementation of the total flux HDG method

Using the same notations as in the previous section, we can write the discrete form of the *total flux* HDG method (3.16a)–(3.16b)–(3.16c).

Local problem

The local problem (3.12a)–(3.12b) can be written as

$$\sum_u W_{0,vu}\underline{\sigma_u^K} - \mathbb{D}_v^K\underline{p^K} + 2i\omega\left(\sum_u W_{0,vu}b_{0,u}\right)\mathbb{M}^K\underline{p^K} + \sum_\ell n_v^\ell\mathbb{F}_\ell^K\underline{\widehat{p}^{g(K,\ell)}} = 0, \quad \forall v \in \{x, y, e\}$$
$$-\omega^2\rho_0\mathbb{M}^K\underline{p^K} + \sum_u(\mathbb{D}_u^K)^T\underline{\sigma_u^K} + i\omega\sum_\ell\tau^\ell\left(\mathbb{E}_\ell^K\underline{p^K} - \mathbb{F}_\ell^K\underline{\widehat{p}^{g(K,\ell)}}\right) = \mathbb{S}^K$$

Matrix form: Matrix $\mathbb{A}^K$: We introduce $\beta_v := \sum_u W_{0,vu}b_{0,u}$.

$\mathbb{A}^K$	$\underline{p^K}$	$\underline{\sigma_x^K}$	$\underline{\sigma_y^K}$	$\underline{\sigma_z^K}$
$\underline{p^K}$	$-\rho_0\omega^2\mathbb{M}^K + i\omega\sum_\ell\tau_\ell\mathbb{E}_\ell^K$	$(\mathbb{D}_x^K)^T$	$(\mathbb{D}_y^K)^T$	$(\mathbb{D}_z^K)^T$
$\underline{\sigma_x^K}$	$-\mathbb{D}_x^K + 2i\omega\beta_x\mathbb{M}^K$	$W_{0,11}^K\mathbb{M}^K$	$W_{0,12}^K\mathbb{M}^K$	$W_{0,13}^K\mathbb{M}^K$
$\underline{\sigma_y^K}$	$-\mathbb{D}_y^K + 2i\omega\beta_y\mathbb{M}^K$	$W_{0,21}^K\mathbb{M}^K$	$W_{0,22}^K\mathbb{M}^K$	$W_{0,23}^K\mathbb{M}^K$
$\underline{\sigma_z^K}$	$-\mathbb{D}_z^K + 2i\omega\beta_z\mathbb{M}^K$	$W_{0,31}^K\mathbb{M}^K$	$W_{0,32}^K\mathbb{M}^K$	$W_{0,33}^K\mathbb{M}^K$

Matrix $\mathbb{C}^K$:

$\mathbb{C}^K$	$\widehat{p}^{g(K,1)}$	$\widehat{p}^{g(K,2)}$	$\widehat{p}^{g(K,3)}$	$\widehat{p}^{g(K,4)}$
p^K	$-i\omega\tau_1\mathbb{F}_1^K$	$-i\omega\tau_2\mathbb{F}_2^K$	$-i\omega\tau_3\mathbb{F}_3^K$	$-i\omega\tau_4\mathbb{F}_4^K$
σ_x^K	$n_x^{K,1}\mathbb{F}_1^K$	$n_x^{K,2}\mathbb{F}_2^K$	$n_x^{K,3}\mathbb{F}_3^K$	$n_x^{K,4}\mathbb{F}_4^K$
σ_y^K	$n_y^{K,1}\mathbb{F}_1^K$	$n_y^{K,2}\mathbb{F}_2^K$	$n_y^{K,3}\mathbb{F}_3^K$	$n_y^{K,4}\mathbb{F}_4^K$
σ_z^K	$n_z^{K,1}\mathbb{F}_1^K$	$n_z^{K,2}\mathbb{F}_2^K$	$n_z^{K,3}\mathbb{F}_3^K$	$n_z^{K,4}\mathbb{F}_4^K$

Global problem

The discrete transmission condition (3.15) can be written

$$\sum_{K,\ell}\left[n_u^{K,\ell}(\mathbb{F}_\ell^K)^T\underline{\sigma_u^K}+i\omega\tau^\ell\left((\mathbb{F}_\ell^K)\underline{p^K}-\mathbb{G}^{g(K,\ell)}\underline{\widehat{p}^{g(K,\ell)}}\right)\right]=0$$

Matrix form: Matrix $\mathbb{B}^K$:

$\mathbb{B}^K$	$\underline{p^K}$	$\underline{\sigma_x^K}$	$\underline{\sigma_y^K}$	$\underline{\sigma_z^K}$
e_1	$i\omega\tau_1(\mathbb{F}_1^K)^T$	$n_x^{K,1}(\mathbb{F}_1^K)^T$	$n_y^{K,1}(\mathbb{F}_1^K)^T$	$n_z^{K,1}(\mathbb{F}_1^K)^T$
e_2	$i\omega\tau_2(\mathbb{F}_2^K)^T$	$n_x^{K,2}(\mathbb{F}_2^K)^T$	$n_y^{K,2}(\mathbb{F}_2^K)^T$	$n_z^{K,2}(\mathbb{F}_2^K)^T$
e_3	$i\omega\tau_3(\mathbb{F}_3^K)^T$	$n_x^{K,3}(\mathbb{F}_3^K)^T$	$n_y^{K,3}(\mathbb{F}_3^K)^T$	$n_z^{K,3}(\mathbb{F}_3^K)^T$
e_4	$i\omega\tau_4(\mathbb{F}_4^K)^T$	$n_x^{K,4}(\mathbb{F}_4^K)^T$	$n_y^{K,4}(\mathbb{F}_4^K)^T$	$n_z^{K,4}(\mathbb{F}_4^K)^T$

Matrix $\mathbb{L}^K$:

$\mathbb{L}^K$	$\widehat{p}^{g(K,1)}$	$\widehat{p}^{g(K,2)}$	$\widehat{p}^{g(K,3)}$	$\widehat{p}^{g(K,4)}$
e_1	$-i\omega\tau_1\mathbb{G}^{g(K,1)}$			
e_2		$-i\omega\tau_2\mathbb{G}^{g(K,2)}$		
e_3			$-i\omega\tau_3\mathbb{G}^{g(K,3)}$	
e_4				$-i\omega\tau_4\mathbb{G}^{g(K,4}$

3.5.4 Implementation of the HDG+ method

As $p^K\in\mathcal{P}_{k+1}(K)$, $\boldsymbol{q}^K\in\boldsymbol{\mathcal{P}}_k(K)$ and $\widehat{p}^e\in\mathcal{P}_k(e)$, one should be very careful while writing the matrix form of the system, indeed we have

$$\underline{p^K}\in\mathbb{C}^{d_n(k+1)}\quad\text{and}\quad\underline{\boldsymbol{q}^K}\in\mathbb{C}^{n\times d_n(k)}\ \ (\Longleftrightarrow\underline{q_u^K}\in\mathbb{C}^{d_n(k)},u\in\{x,y,z\})$$

so some of the elementary matrices will be rectangular.

Definition of the elementary matrices

Let $(\Phi_j^{K,k})_{j=1}^{d_n(k)}$ be the basis for $\mathcal{P}_k(K)$ and $(\Psi_j^{e,k})_{j=1}^{d_{n-1}(k)}$ be the basis for $\mathcal{P}_k(e)$.
The unknowns are therefore decomposed in the following way

$$p^K = \sum_{j=1}^{d_n(k+1)} p_j^K \Phi_j^{K,k+1} \quad ; \quad q_u^K = \sum_{j=1}^{d_n(k)} q_{u,j}^K \Phi_j^{K,k} \quad ; \quad \widehat{p}^e = \sum_{j=1}^{d_{n-1}(k)} \widehat{p}_j^e \Psi_j^{e,k}$$

for $u \in \{x, y, z\}$.
We define the following elementary matrices :

$$\begin{aligned}
\mathbb{M}_{ij}^{K,k} &= \int_K \Phi_i^{K,k} \Phi_j^{K,k} \mathrm{d}\boldsymbol{x} \\
\mathbb{D}_{u,ij}^{K,k} &= \int_K \Phi_j^{K,k} \partial_u \Phi_i^{K,k} \mathrm{d}\boldsymbol{x} & \mathbb{D}_{u,ij}^{K,k,k+1} &= \int_K \Phi_j^{K,k+1} \partial_u \Phi_i^{K,k} \mathrm{d}\boldsymbol{x} \\
\mathbb{F}_{\ell,ij}^{K,k} &= \int_{\partial K^\ell} \Phi_i^{K,k} \Psi_j^{g(K,\ell),k} \mathrm{d}\sigma & \mathbb{F}_{\ell,ij}^{K,k+1,k} &= \int_{\partial K^\ell} \Phi_i^{K,k+1} \Psi_j^{g(K,\ell),k} \mathrm{d}\sigma \\
\mathbb{E}_{\ell,ij}^{K,k} &= \int_{\partial K^\ell} \Phi_i^{K,k} \Phi_j^{K,k} \mathrm{d}\sigma \\
\mathbb{G}_{i,j}^{m} &= \int_{e_m} \Psi_i^{m,k} \Psi_j^{m,k} \mathrm{d}\sigma
\end{aligned} \tag{3.92}$$

where $g(K, \ell)$ is the global number of the ℓ-th face of element K and m is also the global number of the edge e_m.
Most of those matrices are not square matrices and their sizes are recalled in Table 3.3.

Matrix	Rows	Columns
$\mathbb{M}^{K,k+1}$	$d_n(k+1)$	$d_n(k+1)$
$\mathbb{D}_u^{K,k+1}$	$d_n(k+1)$	$d_n(k+1)$
$\mathbb{D}_u^{K,k,k+1}$	$d_n(k)$	$d_n(k+1)$
$\mathbb{F}_\ell^{K,k+1,k}$	$d_n(k+1)$	$d_{n-1}(k)$
$\mathbb{T}_\ell^{K,k+1}$	$d_n(k+1)$	$d_n(k+1)$
$\mathbb{E}_\ell^{K,k+1}$	$d_n(k+1)$	$d_n(k+1)$

(a) Matrices needed for (3.47b)

Matrix	Rows	Columns
$\mathbb{M}^{K,k}$	$d_n(k)$	$d_n(k)$
$\mathbb{D}_u^{K,k,k+1}$	$d_n(k)$	$d_n(k+1)$
$\mathbb{F}_\ell^{K,k}$	$d_n(k)$	$d_{n-1}(k)$

(b) Matrices needed for (3.47a)

Matrix	Rows	Columns
$\mathbb{F}_\ell^{K,k}$	$d_n(k)$	$d_{n-1}(k)$
$\mathbb{F}_\ell^{K,k+1,k}$	$d_n(k+1)$	$d_{n-1}(k)$
$\mathbb{G}_\ell^{m,k}$	$d_{n-1}(k)$	$d_{n-1}(k)$

(c) Matrices needed for (3.47c)

Table 3.3: Summary of the dimensions of the local matrices

The matrix $\mathbb{T}_\ell^{K,k+1}$ will be used to evaluate the projection $P_M p$ and will be defined in (3.96). It was added to Table 3.3 for completeness.

Remark 3.5.4: To compute the integrals involving polynomials of different degrees in a nodal framework, it is possible to directly use the expressions given in (3.92) or to use the following trick. Let $(\boldsymbol{x}_j^{k+1})_{j=1}^{d_n(k+1)}$ be the degrees of freedom associated with $\mathcal{P}_{k+1}(K)$. We define the *projection matrix* $\mathbb{P}$ by

$$\mathbb{P}_{ij}^{K,k,k+1} = \Phi_i^{K,k}(\boldsymbol{x}_j^{k+1}).$$

Seeing $\Phi_i^{K,k}$ as as polynomial of degree $k+1$, we can express it in the basis $(\Phi_j^{K,k+1})_{j=1}^{d_n(k+1)}$. Due to the nodal nature of the basis functions, we have

$$\Phi_i^{K,k} = \sum_{r=1}^{d_n(k+1)} \Phi_i^{K,k}(\boldsymbol{x}_j^{k+1})\Phi_r^{K,k+1} = \sum_{r=1}^{d_n(k+1)} \mathbb{P}_{ir}^{K,k,k+1}\Phi_r^{K,k+1}.$$

The matrix $\mathbb{D}^{K,k,k+1}$ is therefore given by

$$\mathbb{D}^{K,k,k+1} = \mathbb{P}^{K,k,k+1}\mathbb{D}^{K,k+1}.$$

Evaluating the projection:

The evaluation of the term

$$\langle \tau P_M p_h, w_h \rangle_{\partial K} \tag{3.93}$$

where $w_h \in \mathcal{P}_{k+1}(K)$ is a difficult part of the implementation of the HDG+ method. Notice however that the implementation of the term

$$\langle P_M p_h, \mu_h \rangle_{\partial K}, \quad \mu \in M_h(\partial K)$$

arising in the discretization of the transmission condition (3.47c) is straightforward. Indeed as $\mu_h \in \mathcal{P}_k(e)$, by definition of P_M we have

$$\langle P_M p_h, \mu_h \rangle_{\partial K} = \langle p_h, \mu_h \rangle_{\partial K}.$$

In this section we present three techniques to compute (3.93).

Efficient implementation using a hierarchical basis: The easiest way to compute (3.93) is to use a hierarchical and orthogonal basis for $\mathcal{P}_{k+1}(K)$. Let $(\Phi_j^K)_{j=1}^{d_n(k+1)}$ be such a basis, then $(\Phi_j^K)_{j=1}^{d_n(k)}$ is a basis for $\mathcal{P}_k(K)$. Therefore if

$$p^K = \sum_{j=1}^{d_n(k+1)} p_j^K \Phi_j^K$$

then

$$P_M p^K = \sum_{j=1}^{d_n(k)} p_j^K \Phi_j^K.$$

A good choice for such a basis would probably be Dubiner's one which is L^2-orthogonal if the reference element is

- *in 2D:* the triangle with vertices

$$\boldsymbol{x}_1 = (-1,-1) \;\; ; \;\; \boldsymbol{x}_2 = (1,-1) \;\; ; \;\; \boldsymbol{x}_3 = (-1,1),$$

- *in 3D:* the tetrahedron with vertices

$$\boldsymbol{x}_1 = (-1,-1,-1) \;\; ; \;\; \boldsymbol{x}_2 = (1,-1,-1) \;\; ; \;\; \boldsymbol{x}_3 = (-1,1,-1) \;\; ; \;\; \boldsymbol{x}_4 = (-1,-1,1).$$

See [Dub91, Kir04] for more details.
As our solver is developed in the framework of nodal discontinuous Galerkin methods, we will not use this method for our implementation.

Efficient implementation through Gauss-Legendre quadrature: As τ is constant on each edge, we have

$$\langle \tau P_M p_h, u\rangle_{\partial K} = \langle \tau P_M p_h, P_M u\rangle_{\partial K}\,.$$

This integral can be efficiently computed in two dimensions by using a Gauss-Legendre quadrature. This trick is described in [Oik14, Sec. 3.4]. In two dimensions, the edges of a triangle is a one-dimensional interval I. For simplicity, we assume that $I = [-1, 1]$.
Denote by $\mathcal{G}_k$ the k points Gauss-Legendre quadrature rule. For a function f, it is given by

$$\mathcal{G}_k[f] = \sum_{i=1}^{k} w_i f(a_i)$$

where $(w_i)_{i\in[\![1,k]\!]}$ are the quadrature weights and $(a_i)_{i\in[\![1,k]\!]}$ are the quadrature points. If the function f is regular enough, then $\mathcal{G}_k$ approximates the integral of f over I. In particular $\mathcal{G}_{k+1}$ is exact for polynomials of degree up to $2k+1$. The usual stabilization term with $p, u \in \mathcal{P}_{k+1}(K)$, can be exactly computed with the $k+2$ quadrature rule $\mathcal{G}_{k+2}$ (which is exact for polynomials of degree up to $2k+3$), *i.e.*

$$\mathcal{G}_{k+2}[\tau p u] = \langle \tau p, u\rangle_{\partial K}\,.$$

If the $k+1$ points rule $\mathcal{G}_{k+1}$ is used instead, the HDG+ stabilization term is obtained

$$\mathcal{G}_{k+1}[\tau p u] = \langle \tau P_M p, P_M u\rangle_{\partial K}\,.$$

Lemma 3.5.1:

$$\forall u, v \in \mathcal{P}_{k+1}(I), \quad \mathcal{G}_{k+1}[uv] = \int_I P_M u P_M v \mathrm{d}x$$

Proof: Let φ_m be the Legendre polynomial of order $m \geqslant 0$. On I, we can write

$$u = \sum_{i=0}^{k+1} u_i \varphi_i \quad \text{and} \quad v = \sum_{i=0}^{k+1} v_i \varphi_i.$$

Due to the hierarchical nature of the Legendre polynomials, we know that

$$P_M u = \sum_{i=0}^{k} u_i \varphi_i \quad \text{and} \quad P_M v = \sum_{i=0}^{k} v_i \varphi_i.$$

We recall that $\mathcal{G}_{k+1}$ is exact for polynomials of degree up to $2k+1$ and that φ_{k+1} vanishes at the quadrature points, *ie.*

$$\forall i \in [\![1, k+1]\!], \quad \varphi_{k+1}(a_i) = 0.$$

Then we have

$$\begin{aligned}
\mathcal{G}_{k+1}[uv] &= \sum_{i=1}^{k+1}\left[w_i \sum_{\ell=0}^{k+1} u_\ell \varphi_\ell(a_i) \sum_{m=0}^{k+1} v_m \varphi_m(a_i)\right] \\
&= \sum_{i=1}^{k+1}\left[w_i \sum_{\ell=0}^{k} u_\ell \varphi_\ell(a_i) \sum_{m=0}^{k} v_m \varphi_m(a_i)\right] \\
&= \sum_{i=1}^{k+1} w_i (P_M u)(a_i)(P_M v)(a_i) \\
&= \mathcal{G}_{k+1}[P_M u P_M v] \\
&= \int_I P_M u P_M v \mathrm{d}x,
\end{aligned}$$

as $\deg(P_M u P_M v) = 2k$.

Remark 3.5.5: the decomposition onto the Legendre basis is needed to prove the result, but another basis (*eg.* Lagrange's nodal basis) can be used in the implementation of the method.

Remark 3.5.6: unfortunately this result is not yet generalized for three-dimensional problems.

Generic implementation: We will now present a way to compute (3.93) that works for every choice of polynomial basis in every space dimension. As this technique requires to precompute the projections for all the basis functions of order $k+1$, the nodal HDG+ method may be less attractive when p-adaptivity is needed.
We assume that

$$p^K = \sum_{j=1}^{d_n(k+1)} p_j^K \Phi_j^{K,k+1} \quad \text{and} \quad w_h = \Phi_i^{K,k+1}$$

(3.93) therefore becomes

$$\left\langle P_M p_h^K, w_h \right\rangle_{\partial K^\ell} = \sum_{j=1}^{d_n(k+1)} p_j^K \left\langle P_M \Phi_j^{K,k+1}, \Phi_i^{K,k+1} \right\rangle_{\partial K^\ell}. \tag{3.94}$$

The next step is to decompose $P_M \Phi_j^{K,k+1}$ onto the basis of $M_h(\partial K)$

$$P_M \Phi_j^{K,k+1} = \sum_{r=1}^{d_{n-1}(k)} \theta_{j,r}^{K,\ell} \Psi_r^{e,k},$$

and we can construct a linear system for $\underline{\theta_j^{K,\ell}}$ using the definition of P_M

$$\sum_{r=1}^{d_{n-1}(k)} \left\langle \Psi_r^{e,k}, \Psi_i^{e,k} \right\rangle_{\partial K^\ell} \theta_{j,r}^{K,\ell} = \left\langle \Phi_j^{K,k+1}, \Psi_i^{e,k} \right\rangle_{\partial K^\ell}, \qquad \forall i \in [\![1, d_{n-1}(k)]\!]$$

or in matrix form

$$\mathbb{G}^{g(K,\ell)} \underline{\theta_j^{K,\ell}} = \underline{s_j^{K,\ell}}$$

where

$$s_{j,i}^{K,\ell} := \left\langle \Phi_j^{K,k+1}, \Psi_i^{e,k} \right\rangle_{\partial K^\ell}.$$

Plugging this result into (3.94), we have

$$\left\langle P_M p_h^K, \Phi_i^{K,k+1} \right\rangle_{\partial K^\ell} = \sum_{j=1}^{d_n(k+1)} \sum_{r=1}^{d_{n-1}(k)} p_j^K \theta_{j,r}^{K,\ell} \left\langle \Psi_r^{e,k}, \Phi_i^{K,k+1} \right\rangle_{\partial K^\ell}. \tag{3.95}$$

We define the following matrix

$$\mathbb{T}_{\ell,ij}^{K,k+1} := \sum_{r=1}^{d_{n-1}(k)} \mathbb{F}_{\ell,ir}^{K,k+1,k} \theta_{j,r}^{K,\ell}, \tag{3.96}$$

where we recall that

$$\mathbb{F}_{\ell,ij}^{K,k+1,k} := \int_{\partial K^\ell} \Phi_i^{K,k+1} \Psi_j^{g(K,\ell),k} \mathrm{d}\sigma,$$

and we can rewrite (3.95) in matrix form

$$\left\langle P_M p_h^K, \Phi_i^{K,k+1} \right\rangle_{\partial K} = \sum_{\ell=1}^{4} \sum_{j=1}^{d_n(k+1)} \mathbb{T}_{\ell,ij}^{K,k+1} p_j^K.$$

Local problem

Using the elementary matrices introduced in (3.92) we can write (3.47b) as

$$-\rho_0\omega^2\mathbb{M}^{K,k+1}\underline{p^K} + 2i\omega\sum_u b_0^u\mathbb{D}_u^{K,k+1}\underline{p^K} + \sum_u(\mathbb{D}_u^{K,k,k+1})^T\underline{q_u^K}$$
$$+2i\omega\sum_\ell \tau_\ell\left(\mathbb{T}_\ell^{K,k+1}\underline{p^K} - \mathbb{F}^{K,k+1,k}\underline{\widehat{p}}^{g(K,\ell)}\right) - 2i\omega\sum_\ell\left[\alpha_\ell\mathbb{F}^{K,k+1,k}\underline{\widehat{p}}^{g(K,\ell)} + t_{\text{upw},\ell}\mathbb{E}_\ell^{K,k+1}\underline{p^K}\right] = \mathbb{S}^K \tag{3.97}$$

and (3.47a) as

$$\sum_u W_{0,vu}\mathbb{M}^{K,k}\underline{q_u^K} - \mathbb{D}_v^{K,k,k+1}\underline{p^K} + \sum_\ell n_v^{K,\ell}\mathbb{F}_\ell^{K,k}\underline{\widehat{p}}^{g(K,\ell)} = 0 \qquad \forall v\in\{x,y,z\} \tag{3.98}$$

Using TABLE 3.3 it is easy (and important) to check that (3.97) has $d_n(k+1)$ equations and (3.98) has $d_n(k)$ equations.

Matrix form: We can now construct the matrices $\mathbb{A}^K$ and $\mathbb{C}^K$ for the local problem (3.85).
Matrix $\mathbb{A}^K$: (size : $(d_n(k+1)+3d_n(k))^2$)

$\mathbb{A}^K$	$\underline{p^K}$	$\underline{q_x^K}$	$\underline{q_y^K}$	$\underline{q_z^K}$
$\underline{p^K}$	$-\rho_0\omega^2\mathbb{M}^{K,k+1} + 2i\omega\sum_u b_{0,u}\mathbb{D}_u^{K,k+1}$ $-2i\omega\sum_\ell t_{\text{upw},\ell}\mathbb{E}_\ell^{K,k+1}$ $+2i\omega\sum_\ell \tau_\ell\mathbb{T}_\ell^{K,k+1}$	$(\mathbb{D}_x^{K,k,k+1})^T$	$(\mathbb{D}_y^{K,k,k+1})^T$	$(\mathbb{D}_z^{K,k,k+1})^T$
$\underline{q_x^K}$	$-\mathbb{D}_x^{K,k,k+1}$	$W_{0,11}^K\mathbb{M}^{K,k}$	$W_{0,12}^K\mathbb{M}^{K,k}$	$W_{0,13}^K\mathbb{M}^{K,k}$
$\underline{q_y^K}$	$-\mathbb{D}_y^{K,k,k+1}$	$W_{0,21}^K\mathbb{M}^{K,k}$	$W_{0,22}^K\mathbb{M}^{K,k}$	$W_{0,23}^K\mathbb{M}^{K,k}$
$\underline{q_z^K}$	$-\mathbb{D}_z^{K,k,k+1}$	$W_{0,31}^K\mathbb{M}^{K,k}$	$W_{0,32}^K\mathbb{M}^{K,k}$	$W_{0,33}^K\mathbb{M}^{K,k}$

Matrix $\mathbb{C}^K$: (size: $(d_n(k+1)+3d_n(k))\times(4d_{n-1}(k))$)

$\mathbb{C}^K$	$\underline{\widehat{p}^{g(K,1)}}$	$\underline{\widehat{p}^{g(K,2)}}$	$\underline{\widehat{p}^{g(K,3)}}$	$\underline{\widehat{p}^{g(K,4)}}$
$\underline{p^K}$	$-2i\omega(\tau_1+\alpha_1)\mathbb{F}_1^{K,k+1,k}$	$-2i\omega(\tau_2+\alpha_2)\mathbb{F}_2^{K,k+1,k}$	$-2i\omega(\tau_3+\alpha_3)\mathbb{F}_3^{K,k+1,k}$	$-2i\omega(\tau_4+\alpha_4)\mathbb{F}_4^{K,k+1,k}$
$\underline{q_x^K}$	$n_x^{K,1}\mathbb{F}_1^{K,k}$	$n_x^{K,2}\mathbb{F}_2^{K,k}$	$n_x^{K,3}\mathbb{F}_3^{K,k}$	$n_x^{K,4}\mathbb{F}_4^{K,k}$
$\underline{q_y^K}$	$n_y^{K,1}\mathbb{F}_1^{K,k}$	$n_y^{K,2}\mathbb{F}_2^{K,k}$	$n_y^{K,3}\mathbb{F}_3^{K,k}$	$n_y^{K,4}\mathbb{F}_4^{K,k}$
$\underline{q_z^K}$	$n_z^{K,1}\mathbb{F}_1^{K,k}$	$n_z^{K,2}\mathbb{F}_2^{K,k}$	$n_z^{K,3}\mathbb{F}_3^{K,k}$	$n_z^{K,4}\mathbb{F}_4^{K,k}$

Global problem

Using the elementary matrices introduced in (3.92) we can write (3.47c) as

$$\sum_{K,\ell}\left[\sum_u n_u^{K,\ell}(\mathbb{F}_\ell^{K,k})^T\underline{q_u^K}+2i\omega\sum_\ell\tau_\ell\left((\mathbb{F}_\ell^{K,k+1,k})^T\underline{p^K}-\mathbb{G}^{g(K,\ell),k}\underline{\widehat{p}^{g(K,\ell)}}\right)\right]$$
$$-2i\omega\sum_{K,\ell}\left[\alpha_\ell\mathbb{G}^{g(K,\ell),k}\underline{\widehat{p}^{g(K,\ell)}}+t_{\mathrm{upw},\ell}(\mathbb{F}_\ell^{K,k+1,k})^T\underline{p^K}\right]=0 \tag{3.99}$$

Once again, TABLE 3.3 should be used to check the dimensions of the matrices involved in (3.99).

Matrix form: We will now use (3.99) to construct the matrices $\mathbb{B}^K$ and $\mathbb{L}^K$ of (3.86).
Matrix $\mathbb{B}^K$: (size: $4d_{n-1}(k)\times(d_n(k+1)+3d_n(k))$)

$\mathbb{B}^K$	$\underline{p^K}$	$\underline{q_x^K}$	$\underline{q_y^K}$	$\underline{q_z^K}$
e_1	$2i\omega(\tau_1-t_{\mathrm{upw},1})(\mathbb{F}_1^{K,k+1,k})^T$	$n_x^{K,1}(\mathbb{F}_1^{K,k})^T$	$n_y^{K,1}(\mathbb{F}_1^{K,k})^T$	$n_z^{K,1}(\mathbb{F}_1^{K,k})^T$
e_2	$2i\omega(\tau_2-t_{\mathrm{upw},2})(\mathbb{F}_2^{K,k+1,k})^T$	$n_x^{K,2}(\mathbb{F}_2^{K,k})^T$	$n_y^{K,2}(\mathbb{F}_2^{K,k})^T$	$n_z^{K,2}(\mathbb{F}_2^{K,k})^T$
e_3	$2i\omega(\tau_3-t_{\mathrm{upw},3})(\mathbb{F}_3^{K,k+1,k})^T$	$n_x^{K,3}(\mathbb{F}_3^{K,k})^T$	$n_y^{K,3}(\mathbb{F}_3^{K,k})^T$	$n_z^{K,3}(\mathbb{F}_3^{K,k})^T$
e_4	$2i\omega(\tau_4-t_{\mathrm{upw},4})(\mathbb{F}_4^{K,k+1,k})^T$	$n_x^{K,4}(\mathbb{F}_4^{K,k})^T$	$n_y^{K,4}(\mathbb{F}_4^{K,k})^T$	$n_z^{K,4}(\mathbb{F}_4^{K,k})^T$

Matrix $\mathbb{L}^K$: (size: $(4d_{n-1}(k))^2$)

$\mathbb{L}^K$	$\underline{\widehat{p}^{g(K,1)}}$	$\underline{\widehat{p}^{g(K,2)}}$	$\underline{\widehat{p}^{g(K,3)}}$	$\underline{\widehat{p}^{g(K,4)}}$
e_1	$-2i\omega(\tau_1+\alpha_1)\mathbb{G}^{g(K,1),k}$			
e_2		$-2i\omega(\tau_2+\alpha_2)\mathbb{G}^{g(K,2),k}$		
e_3			$-2i\omega(\tau_3+\alpha_3)\mathbb{G}^{g(K,3),k}$	
e_4				$-2i\omega(\tau_4+\alpha_4)\mathbb{G}^{g(K,4),k}$

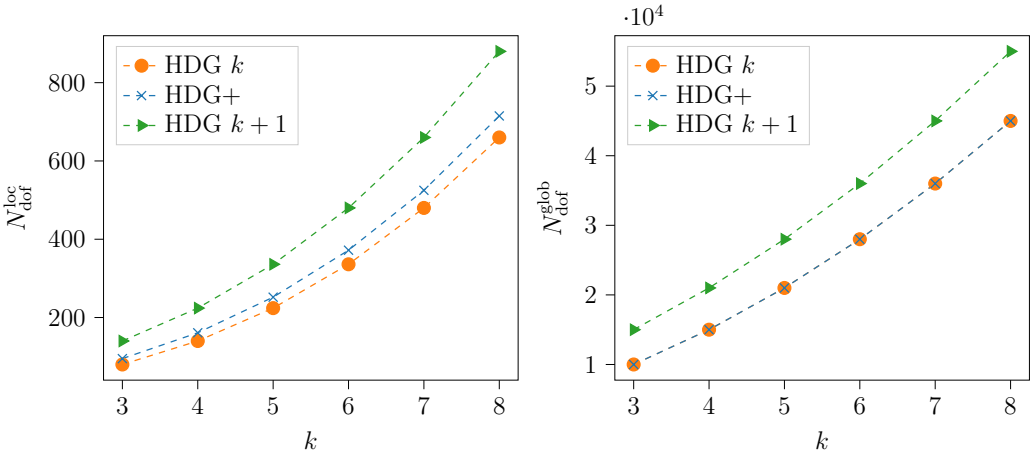
3.5.5 Comparison of the cost of the HDG and HDG+ methods

Now that we have written down the discrete systems for the HDG and HDG+ methods, it is interesting to compare the sizes of systems that we need to solve.
In TABLE 3.4 we have written down the sizes of the matrices for both methods. We can see that the global problem $\mathbb{K}$ of the HDG+ has the same dimension as the one of the HDG method of degree k. As the resolution of the global problem is the most expensive step in the resolution of the method, the cost of the HDG+ method is therefore similar to the cost of the HDG method of degree k, while yielding to an order of convergence of $k+2$ instead of $k+1$. We can also notice that the cost of the local problems of the HDG+ method is intermediate between the cost of the local problems of the HDG methods of degree k and $k+1$. However this has a really limited impact on the computational cost.
We denote by N_{elt} the number of elements in the mesh, *ie* $N_{\mathrm{elt}}=\mathrm{card}(\mathcal{T}_h)$, and by N_{face} the number of faces in the mesh.

Matrix	HDG k	HDG+	HDG $k+1$
$\mathbb{A}^K$	$(4d_n(k))^2$	$(d_n(k+1)+3d_n(k))^2$	$(4d_n(k+1))^2$
$\mathbb{C}^K$	$4d_n(k)\times 4d_{n-1}(k)$	$(d_n(k+1)+3d_n(k))\times(4d_{n-1}(k))$	$4d_n(k+1)\times 4d_{n-1}(k+1)$
$\mathbb{B}^K$	$4d_{n-1}(k)\times 4d_n(k)$	$4d_{n-1}(k)\times(d_n(k+1)+3d_n(k))$	$4d_{n-1}(k+1)\times 4d_n(k+1)$
$\mathbb{L}^K$	$(4d_{n-1}(k))^2$	$(4d_{n-1}(k))^2$	$(4d_{n-1}(k+1))^2$
$\mathbb{K}$	$\sim (N_{\text{face}}d_{n-1}(k))^2$	$\sim (N_{\text{face}}d_{n-1}(k))^2$	$\sim (N_{\text{face}}d_{n-1}(k+1))^2$

Table 3.4: Size of the matrices for the HDG and HDG+ methods in 3D

In FIGURE 3.5, we have plotted the number of degrees of freedom for local problems ($N_{\text{dof}}^{\text{loc}}$) and for the global problem ($N_{\text{dof}}^{\text{glob}}$) with $N_{\text{elt}} = 10^3$ for several polynomial degrees in 3D. Notice that $\mathbb{A}^K$ has dimension $(N_{\text{dof}}^{\text{loc}})^2$ and $\mathbb{K}$ has dimension $(N_{\text{dof}}^{\text{glob}})^2$.



(a) Local number of degrees of freedom

(b) Global number of degrees of freedom with $N_{\text{elt}} = 10^3$

Figure 3.5: Local and global number of degrees of freedom for the three HDG methods in 3D

To put these different costs in perspective, we can also compare them to the cost of two standard DG methods: the SIPDG method which uses a primal formulation with only one scalar unknown, and the LDG method which is a mixed DG method with four unknowns on each elements. First we recall that each element has four faces in 3D and that a face is common to two elements, *i.e.*

$$N_{\text{face}} \sim \frac{4}{2} N_{\text{elt}} \sim 2N_{\text{elt}}.$$

Using TABLE 3.4, we know that the number of degrees of freedom of the HDG method of degree k is

$$N_{\text{dof}}^{\text{HDG-}k} \sim N_{\text{face}} d_{n-1}(k) = N_{\text{face}} \binom{k+n-1}{n-1} \sim 2k^{n-1} N_{\text{elt}},$$

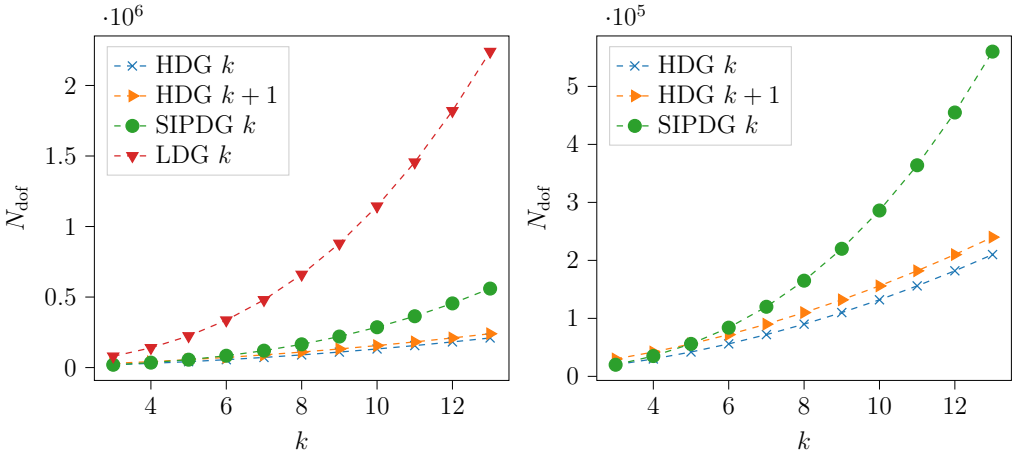
whereas the number of degrees of freedom of a SIPDG method of degree k is

$$N_{\text{dof}}^{\text{SIPDG-}k} \sim N_{\text{elt}} d_n(k) = N_{\text{elt}} \binom{k+n}{n} \sim k^n N_{\text{elt}},$$

and the number of degrees of freedom of a LDG method of degree k is

$$N_{\text{dof}}^{\text{LDG-}k} \sim 4N_{\text{elt}} d_n(k) = 4N_{\text{elt}} \binom{k+n}{n} \sim 4k^n N_{\text{elt}},$$

which becomes much larger as k increases. Those sizes have been depicted in FIGURE 3.6, with and without the LDG method to see the difference between HDG and SIPDG more clearly. We have depicted both the HDG methods with interpolation degrees k and $k+1$ as the HDG+ method has an intermediate cost between those two standard HDG methods. For orders up to 6, the HDG and SIPDG methods have similar costs, while the LDG method is clearly more expensive. Using the HDG method instead of the SIPDG one has one major advantage: the flux can be evaluated without any loss of order as it is an unknown of the HDG method, and the HDG method is therefore a mixed method with the cost of a primal one. We can also see that for orders higher than 7, the HDG method is cheaper than both LDG and SIPDG methods, and we can clearly see the advantage of using a HDG method over a DG one.



(a) With LDG (b) Without LDG

Figure 3.6: $N_{\text{dof}}^{\text{HDG-}k}$, $N_{\text{dof}}^{\text{HDG-}k+1}$, $N_{\text{dof}}^{\text{SIPDG-}k}$ and $N_{\text{dof}}^{\text{LDG-}k}$ in 3D with $N_{\text{elt}} = 10^3$

3.6 Numerical experiments

In this section we will provide some numerical experiments for the three HDG methods that were described and analysed in this paper. We will first focus on the simple case of duct modes propagating in a waveguide to obtain convergence curves and validate the theoretical results of the previous sections. We will then provide some illustrative examples to show that those methods can be used in more realistic cases.

3.6.1 Convergence rate

In this subsection, we will present some numerical experiments to illustrate our theoretical results. As most of the estimates obtained in our analysis involve projection errors of the form

$$\|p_h - \pi_W p\|_{\mathcal{T}_h} \quad \text{or} \quad \|\boldsymbol{q}_h - \boldsymbol{\pi}_{\boldsymbol{V}} \boldsymbol{q}\|_{\mathcal{T}_h},$$

we will need to evaluate those projections before actually computing errors. In TABLE 3.5 we recall the different projections used for the analysis of the three different variants of the HDG method that we considered.

Method	π_W	$\boldsymbol{\pi_V}$	P_M
HDG $\boldsymbol{q}$	L^2	L^2	L^2
HDG +	L^2	L^2	L^2
HDG $\boldsymbol{\sigma}$	HDG	HDG	L^2

Table 3.5: Summary of the different projections used for the analysis of the HDG methods

Geometric settings: As depicted on FIGURE 3.7 we consider a uniform directional flow $\boldsymbol{v_0} = Mc_0\boldsymbol{e}_x$, where M is the *Mach number*.

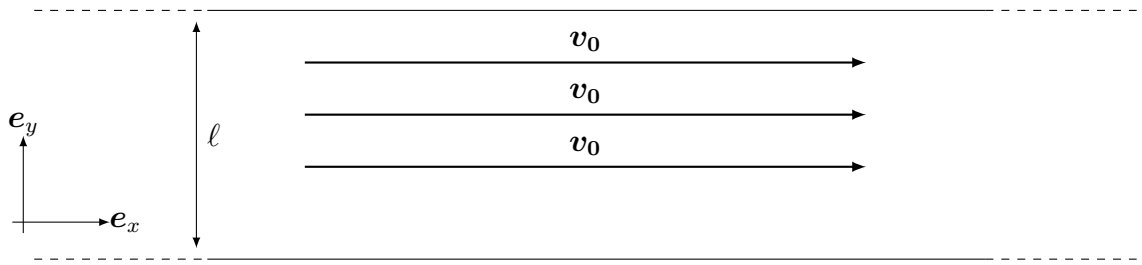


Figure 3.7: Sketch of the geometric configuration

Unless stated otherwise, we will always use the following parameters for the convergence tests

$$\mathcal{O} = (0,2)\times(0,1) \quad ; \quad \rho_0, c_0 \equiv 1 \quad ; \quad \omega = 5.55\pi,$$

and the choice of M will be specified for each numerical experiment.

Analytic solution: The duct modes are a family of analytic solutions of (3.1) in a waveguide, see [BBL03]. They are given by

$$p_n^{\pm}(x,y) = e^{i\beta_n^{\pm}x}\varphi_n(y)$$

where

$$n < N_0 : \qquad \beta_n^{\pm} = \frac{-\kappa M \pm \sqrt{\kappa^2 - \frac{n^2\pi^2}{\ell^2}(1-M^2)}}{1-M^2}$$

$$n > N_0 : \qquad \beta_n^{\pm} = \frac{-\kappa M \pm i\sqrt{\frac{n^2\pi^2}{\ell^2}(1-M^2) - \kappa^2}}{1-M^2}$$

with

$$\kappa = \frac{\omega}{c_0} \quad \text{and} \quad M = \frac{v_0}{c_0}$$

$$N_0 = \left\lfloor \frac{\kappa\ell}{\pi\sqrt{1-M^2}} \right\rfloor$$

and

$$\varphi_0(y) := \sqrt{\ell^{-1}}$$
$$\varphi_n(y) := \sqrt{2\ell^{-1}}\cos\left(\frac{n\pi y}{\ell}\right), \quad n \in \mathbb{N}^*$$

The choice of n will be specified for each numerical experiment.

Evaluating the projections: Here we give the details for evaluating the L^2 projection onto W_h, the process is very similar for the other projections.
We recall that the dimension of the polynomial spaces is given by

$$\dim \mathcal{P}_k(K) = \binom{n+k}{n} =: d_n(k), \quad \text{for } K \subset \mathbb{R}^n.$$

All the projections considered are local to an element K, so evaluating them amounts to solving a linear system on each K. Indeed, the definition of the L^2 projection onto $W_h(K) = \mathcal{P}_k(K)$ gives

$$\forall i \in [\![1, d_n(k)]\!], \quad \left(\pi_W p, \Phi_i^{K,k}\right)_K = \left(p, \Phi_i^{K,k}\right)_K.$$

As $\pi_W p \in W_h(K)$, we can write

$$\pi_W p = \sum_{j=1}^{d_n(k)} \pi_j \Phi_j^{K,k},$$

where $\underline{\pi} = (\pi_j)_j$ is the vector of the coordinates of $\pi_W p$ in the basis $(\Phi_j^K)_j$ of $\mathcal{P}_k(K)$. We therefore obtain the following system

$$\forall i \in [\![1, d_n(k)]\!], \quad \sum_{j=1}^{d_n(k)} \pi_j \left(\Phi_j^{K,k}, \Phi_i^{K,k}\right)_K = \left(p, \Phi_i^{K,k}\right)_K,$$

or in matrix form

$$\mathbb{M}^{K,k}\underline{\pi} = \underline{s}, \quad \text{where } s_i = \left(p, \Phi_i^{K,k}\right)_K.$$

The integral in the right-hand side is evaluated using a 91 points Gauss-Lobatto quadrature formula and the linear system can be solved using `lapack`.

Evaluating the L^2-error: Now that the projections can be evaluated, it remains to compute the L^2 norms. As the numerical solution and the projections are polynomial quantities, this can be done using the mass matrix. Indeed, for $u \in \mathcal{P}_k(K)$ we have

$$\begin{aligned}
\|u\|_K^2 &= \left(\sum_{j=1}^{d_n(k)} u_j \Phi_j^{K,k}, \sum_{i=1}^{d_n(k)} u_i \Phi_i^{K,k}\right)_K \\
&= \sum_{i,j=1}^{d_n(k)} u_j u_i^* \left(\Phi_j^{K,k}, \Phi_i^{K,k}\right)_K \\
&= \sum_{i=1}^{d_n(k)} u_i^* \sum_{j=1}^{d_n(k)} \mathbb{M}_{ij}^{K,k} u_j \\
&= \underline{u}^* \mathbb{M}^{K,k} \underline{u}.
\end{aligned}$$

To obtain more meaningful results, we will use relative errors instead of the standard L^2-error. This choice allows us to compare the errors computed on different meshes without any pollution coming from the different number of elements. The relative error for p_h is given by

$$\mathcal{E}_p := \frac{\|p_h - \pi_W p\|_{\mathcal{T}_h}}{\|\pi_W p\|_{\mathcal{T}_h}},$$

and similar expressions will be used for the other volumetric quantities.

Some notation: To allow the comparison between the HDG+ method and the two HDG methods, we would like to emphasize that k always denotes the polynomial degree used to approximate the *trace unknown* $\widehat{p_h}$. We have chosen to plot the relative errors against the quantity $\frac{k}{h}$ which is proportional to the number of degrees of freedom per wavelength. We would like to point out that all the plots in the next sections will use a log-log scale.

Hardware configuration: Those numerical experiments were carried out on a `miriel` node on the `plafrim` cluster[8]. This node is equipped with a 2 dodeca-core Haswell Intel Xeon E5-2680 v3 with a clock rate of 2.5 GHz and 128 Go of memory.

Acoustic case without flow

A first important test to validate our implementation in the *acoustic case without flow*. In this case, the convected Helmholtz equation reduces to the standard Helmholtz equation and the HDG methods, either with the diffusive or total flux, are the same when there is no convection. We should therefore be able to reproduce the *super-convergence* of the HDG method for the Helmholtz equation , *i.e.*

$$\mathcal{E}_p = \mathcal{O}(h^{k+2}),$$

when an approximation of degree k is used for all the unknowns.
Here we have chosen to use the following parameters

$$n = 0, \quad \text{and} \quad M = 0,$$

which correspond to a plane-wave.
The resulting convergence curve is depicted on FIGURE 3.8 and we can clearly see that the super-convergence is obtained.

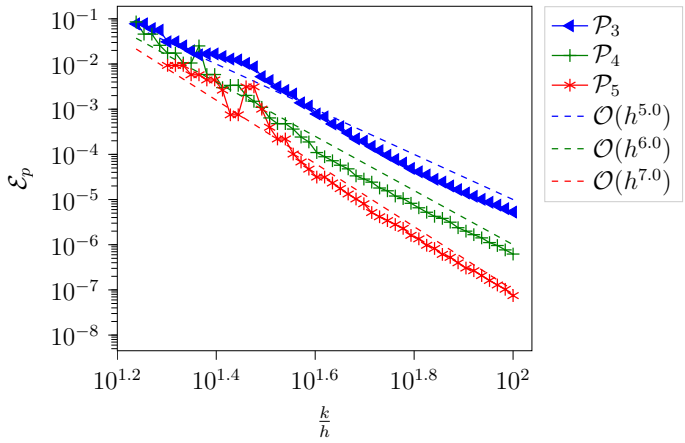


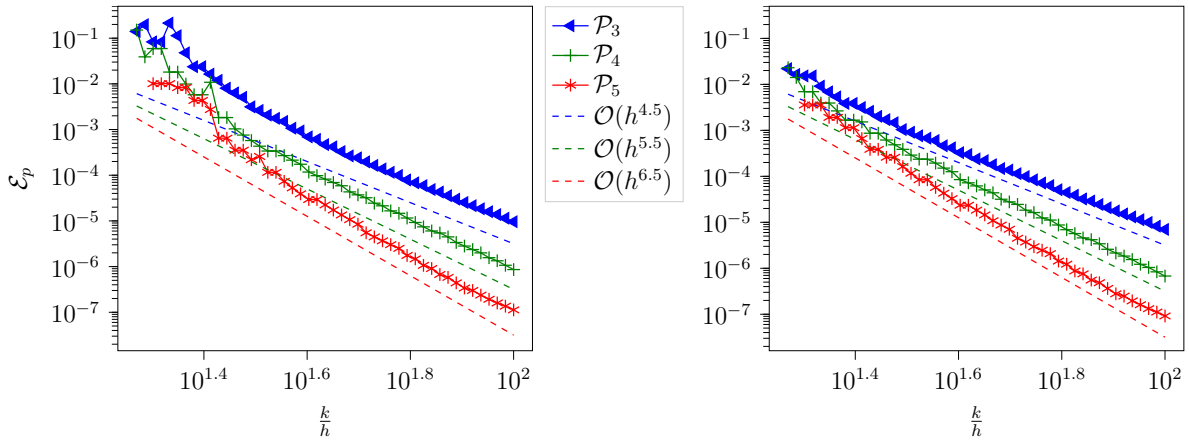
Figure 3.8: Convergence history for the HDG method without flow for the volumetric unknown p_h

Low Mach

We then move to a flow with a low Mach number. In this case we have used the following parameters

$$n = 3, \quad \text{and} \quad M = 0.2.$$

[8] See `http://www.plafrim.fr`.



(a) HDG with diffusive flux $\boldsymbol{q}_h$ (b) HDG with total flux $\boldsymbol{\sigma}_h$

Figure 3.9: Low Mach convergence history for the volumetric unknown p_h for both the diffusive and total flux HDG method

The convergence history for the volumetric unknown p_h is displayed on FIGURE 3.9 for both of the HDG methods. We can see that the *diffusive flux* formulation achieves an order of convergence of $k + 3/2$ as expected. On the other hand the *total flux* formulation also achieves an order of convergence of $k + 3/2$ which is better than the expected order $k + 1$. However for uniform flows and upon well choosing the penalization parameters both of those methods are algebraically equivalent and we therefore expect to obtain the same order of convergence for p_h. On FIGURE 3.10 the convergence rate for the volumetric unknown p_h for the HDG+ method is displayed. As expected the optimal convergence rate of $k + 2$ is obtained. As discussed in TABLE 3.6, it is clear that the use of the HDG+ method is less expensive than the use of the HDG methods.

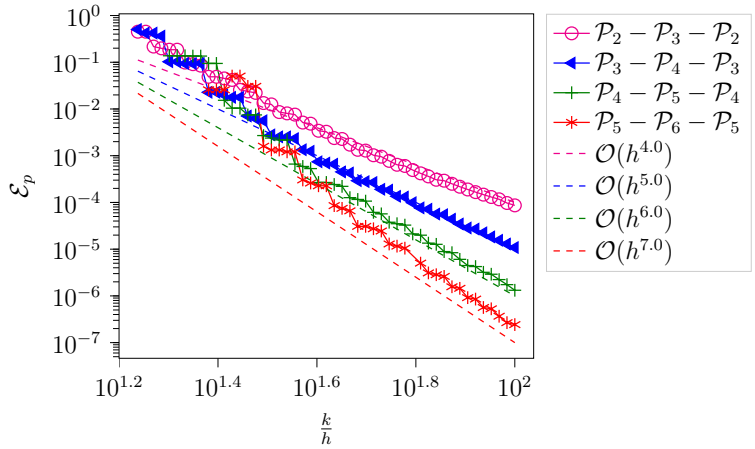


Figure 3.10: Low Mach convergence history for the volumetric unknown p_h for the HDG+ method

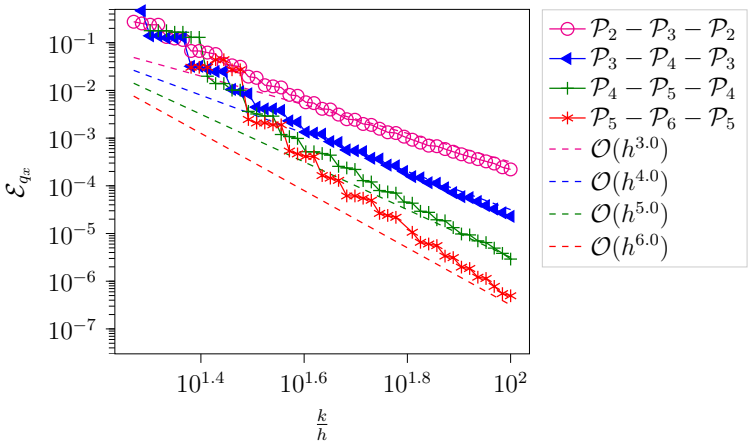
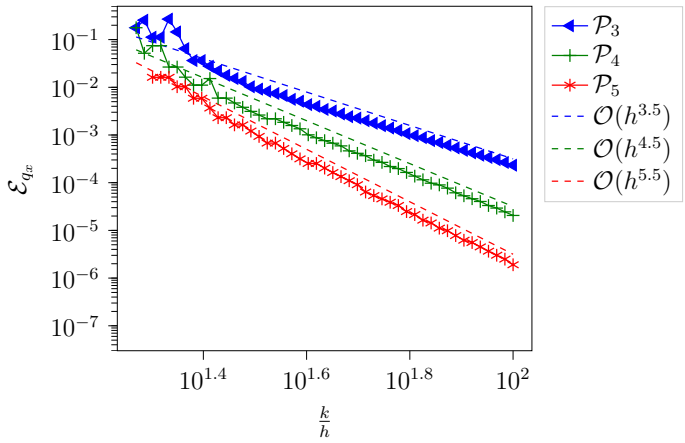
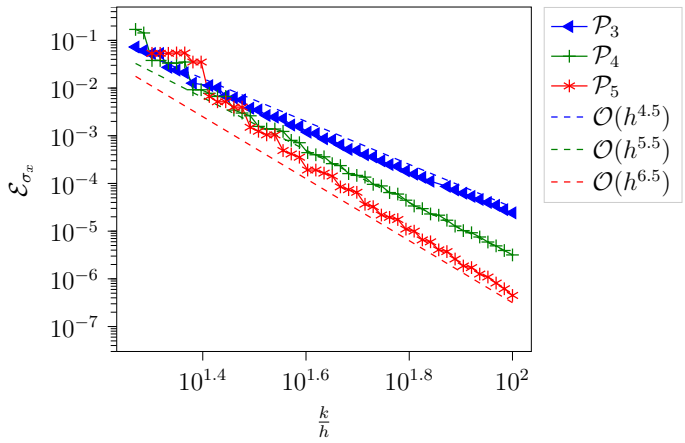


Figure 3.11: Low Mach convergence history for the first component of volumetric unknown $\boldsymbol{q}_h$ for the HDG+ method



(a) HDG with diffusive flux $\boldsymbol{q}_h$



(b) HDG with total flux $\boldsymbol{\sigma}_h$

Figure 3.12: Low Mach convergence history for the first component of the volumetric flux unknown

If we now move to the flux unknown, we can see on FIGURE 3.11 that $\boldsymbol{q}_h$ converges with the optimal order $k+1$ as expected. For the HDG methods, the convergence histories for $\boldsymbol{q}_h$ and $\boldsymbol{\sigma}_h$ are depicted on FIGURE 3.12. The *diffusive flux* formulation achieves a convergence order of $k+1/2$ as expected, whereas the *total flux* formulation outperforms the theoretical results and achieves an order of $k+3/2$. For this formulation it was however expected that both p_h

and $\boldsymbol{\sigma}_h$ converge with the same order, which is actually the case. Finally we would like to point out that the choice between the formulations with $\boldsymbol{q}_h$ or $\boldsymbol{\sigma}_h$ is application dependent. Finally information regarding the size of the global problems for a fixed number of degrees of freedom per wavelength are given in TABLE 3.6. We have included the HDG+ method with $(k, k-1)$ and the HDG+ method with $(k+1, k)$. When using the HDG+ method with $(k, k-1)$, a smaller global problem is solved to obtain the same convergence rate as the HDG methods. When using the HDG+ method with $(k+1, k)$, larger elements can be used to obtain the same number of degrees of freedom per wavelength because of the higher polynomial interpolation degree. We can clearly see that the HDG+ methods are computationally less expensive than both of the HDG methods.

k	h/k		HDG-$\boldsymbol{q}_h$	HDG-$\boldsymbol{\sigma}_h$	HDG+ $(k, k-1)$	HDG+ $(k+1, k)$
3	10^{-1}	nnz	20 384	20 384	10 968	8 720
		nnz LU	36 872	36 872	15 216	9 800
		MUMPS time	$2.5 \cdot 10^{-2}$	$1.6 \cdot 10^{-2}$	$1.8 \cdot 10^{-2}$	$1.7 \cdot 10^{-2}$
3	10^{-2}	nnz	872 216	872 216	481 528	467 280
		nnz LU	3 572 752	3 572 752	1 835 148	1 673 560
		MUMPS time	0.27	0.23	0.15	0.10
4	10^{-1}	nnz	17 130	17 130	8 720	4 980
		nnz LU	24 125	24 125	9 800	4 644
		MUMPS time	$1.9 \cdot 10^{-2}$	$2.0 \cdot 10^{-2}$	$1.7 \cdot 10^{-2}$	$9.5 \cdot 10^{-3}$
4	10^{-2}	nnz	752 330	752 330	467 280	423 600
		nnz LU	2 874 075	2 874 075	1 673 560	1 397 418
		MUMPS time	0.19	0.17	0.10	$8.6 \cdot 10^{-2}$
5	10^{-1}	nnz	12 552	12 552	4 980	4 332
		nnz LU	14 112	14 112	4 644	4 068
		MUMPS time	$1.7 \cdot 10^{-2}$	$2.7 \cdot 10^{-2}$	$9.5 \cdot 10^{-3}$	$9.9 \cdot 10^{-3}$
5	10^{-2}	nnz	672 840	672 840	432 600	414 456
		nnz LU	2 409 006	2 409 006	1 397 418	1 333 368
		MUMPS time	0.12	0.11	$8.6 \cdot 10^{-2}$	$6.8 \cdot 10^{-2}$

Table 3.6: Size of the global systems and `mumps` elapsed time for different interpolation degree k for the low Mach case with a fixed number of degrees of freedom per wavelength

Notice that the numerical experiments performed here lead to relatively small linear system. Indeed our goal was to validate the numerical method rather than showing the ability of `hawen` to handle large numerical simulations.

In TABLE 3.7, we review the size of the global system for various interpolation degrees k with a fixed error threshold for p_h. As we did before, we use the relative L^2-error defined by

$$\mathcal{E}_p := \frac{\|p_h - \pi_W p\|_{\mathcal{T}_h}}{\|\pi_W p\|_{\mathcal{T}_h}}.$$

We can clearly see that, when the desired error is smaller than 10^{-2}, increasing the interpolation degree and thus the order of the method leads to solving a smaller linear system to obtain the same error level. This less visible for an error threshold of 10^{-2}, as this accuracy can be obtained with relatively large elements even with a low interpolation degree. For this error level, we can also see that there is no real difference in the size of the linear systems of HDG and HDG+ methods. This is expected as there is no real difference between the

different interpolation degrees for this error threshold. However for smaller error thresholds, we can see that the HDG-$\boldsymbol{q}_h$, HDG-$\boldsymbol{\sigma}_h$ and HDG+ with $(k, k-1)$ lead to linear systems with similar sizes, which is expected as they share the same order. The HDG+ method with $(k+1, k)$ also has a higher convergence rate and therefore lead to significantly smaller linear system for the same error level.

k	$\mathcal{E}_p$		HDG-$\boldsymbol{q}_h$	HDG-$\boldsymbol{\sigma}_h$	HDG+ $(k, k-1)$	HDG+ $(k+1, k)$
3	10^{-2}	nnz	49 784	39 216	37 200	26 920
		nnz LU	116 208	85 416	97 371	53 088
		MUMPS time	$3.1 \cdot 10^{-2}$	$1.8 \cdot 10^{-2}$	$2.5 \cdot 10^{-2}$	$1.3 \cdot 10^{-2}$
3	10^{-3}	nnz	99 920	102 800	92 040	49 784
		nnz LU	293 328	293 328	283 155	116 272
		MUMPS time	$3.0 \cdot 10^{-2}$	$2.8 \cdot 10^{-2}$	$3.2 \cdot 10^{-2}$	$1.9 \cdot 10^{-2}$
3	10^{-4}	nnz	210 360	267 784	255 552	140 640
		nnz LU	689 568	926 352	944 826	440 584
		MUMPS time	$4.9 \cdot 10^{-2}$	$6.2 \cdot 10^{-2}$	$8.1 \cdot 10^{-2}$	$3.63 \cdot 10^{-2}$
4	10^{-2}	nnz	28 240	28 240	26 920	28 240
		nnz LU	51 050	51 150	53 088	51 050
		MUMPS time	$2.3 \cdot 10^{-2}$	$1.4 \cdot 10^{-2}$	$1.3 \cdot 10^{-2}$	$1.4 \cdot 10^{-2}$
4	10^{-3}	nnz	58 560	58 560	49 784	42 050
		nnz LU	126 350	126 450	116 272	84 075
		MUMPS time	$1.9 \cdot 10^{-2}$	$2.0 \cdot 10^{-2}$	$1.9 \cdot 10^{-2}$	$2.2 \cdot 10^{-2}$
4	10^{-4}	nnz	106 880	151 600	140 640	77 770
		nnz LU	284 050	433 450	440 584	182 875
		MUMPS time	$2.5 \cdot 10^{-2}$	$3.2 \cdot 10^{-2}$	$3.63 \cdot 10^{-2}$	$2.1 \cdot 10^{-2}$
4	10^{-5}	nnz	255 580	255 580	372 640	156 100
		nnz LU	787 925	787 925	1 383 296	446 855
		MUMPS time	$3.9 \cdot 10^{-2}$	$4.1 \cdot 10^{-2}$	0.10	$3.4 \cdot 10^{-2}$
5	10^{-2}	nnz	24 660	24 660	28 240	24 660
		nnz LU	34 740	34 740	51 050	34 740
		MUMPS time	$1.1 \cdot 10^{-2}$	$1.4 \cdot 10^{-2}$	$1.4 \cdot 10^{-2}$	$1.9 \cdot 10^{-2}$
5	10^{-3}	nnz	40 656	40 656	42 050	40 656
		nnz LU	73 440	73 296	84 075	73 512
		MUMPS time	$1.6 \cdot 10^{-2}$	$1.4 \cdot 10^{-2}$	$2.2 \cdot 10^{-2}$	$1.3 \cdot 10^{-2}$
5	10^{-4}	nnz	84 312	88 200	77 770	60 540
		nnz LU	181 944	192 996	182 875	120 636
		MUMPS time	$2.6 \cdot 10^{-2}$	$1.9 \cdot 10^{-2}$	$2.1 \cdot 10^{-2}$	$1.5 \cdot 10^{-2}$
5	10^{-5}	nnz	148 704	153 888	156 100	111 972
		nnz LU	392 580	409 752	446 855	263 340
		MUMPS time	$3.1 \cdot 10^{-2}$	$2.9 \cdot 10^{-2}$	$3.4 \cdot 10^{-2}$	$2.5 \cdot 10^{-2}$

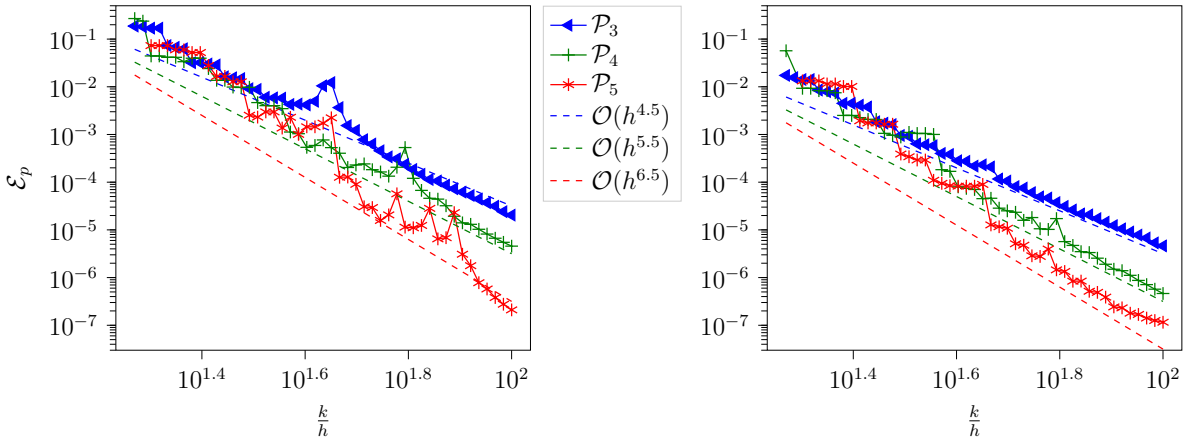
Table 3.7: Size of the global systems and `mumps` elapsed time for different interpolation degree k for the low Mach case with a fixed error threshold $\mathcal{E}_p$

Large Mach

Finally we also considered a flow with a large March number. In this case, we used the following parameters

$$n = 3, \quad \text{and} \quad M = 0.8.$$

As the simulations of acoustic wave propagation in flows with large Mach numbers is known to be more challenging, we expect to see worse performances than in the previous subsection.



(a) HDG with diffusive flux $\boldsymbol{q}$ (b) HDG with total flux $\boldsymbol{\sigma}$

Figure 3.13: Large Mach convergence history for the volumetric unknown p_h for both of the HDG methods

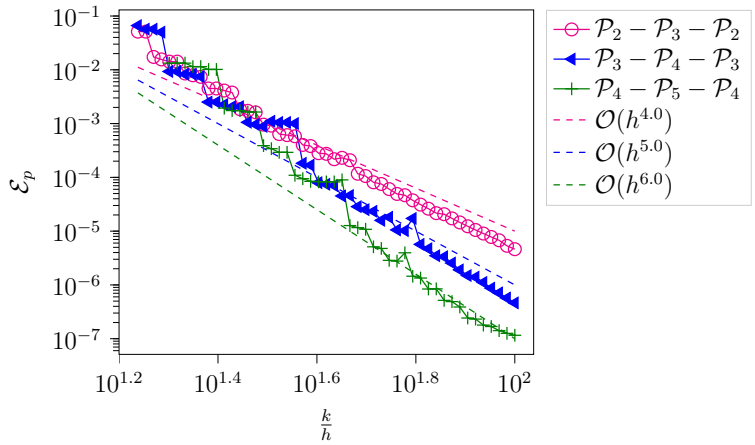
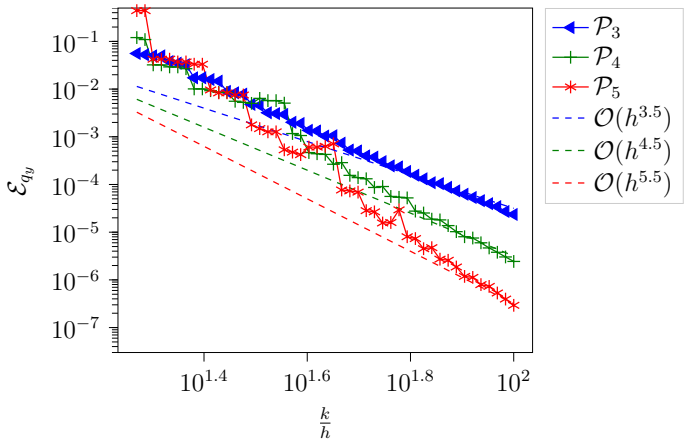
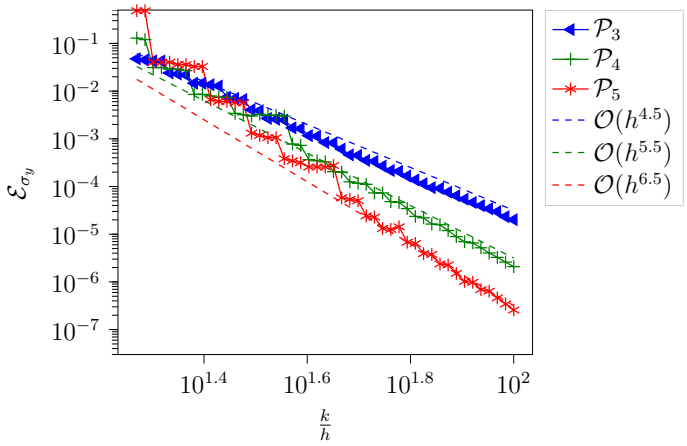


Figure 3.14: Large Mach convergence history for the volumetric unknown p_h for the HDG+ method

The convergence history for the volumetric unknown is depicted in FIGURE 3.13 for both of the HDG methods. We can see that the *total flux* formulation still achieves the convergence order of $k + 3/2$ whereas the behaviour of the *diffusive flux* formulation seems less robust. The same convergence history for HDG+ method is displayed on FIGURE 3.14 and we can see that it still achieves the optimal convergence rate of $k + 2$.



(a) HDG with diffusive flux $\boldsymbol{q}_h$



(b) HDG with total flux $\boldsymbol{\sigma}_h$

Figure 3.15: Large Mach convergence history for the second component of the volumetric flux unknown

The convergence history for the volumetric flux unknown $\boldsymbol{q}_h$ or $\boldsymbol{\sigma}_h$ is depicted in FIGURE 3.15 for the HDG methods and in FIGURE 3.16 for the HDG+ method. As in the low-Mach case, the HDG methods have a convergence rate of $k + 3/2$ and the HDG+ method has a convergence rate of $k + 1$. Notice that the HDG-$\boldsymbol{\sigma}_h$ method seems to be the most robust method for the approximation of the flux unknown for high Mach numbers.

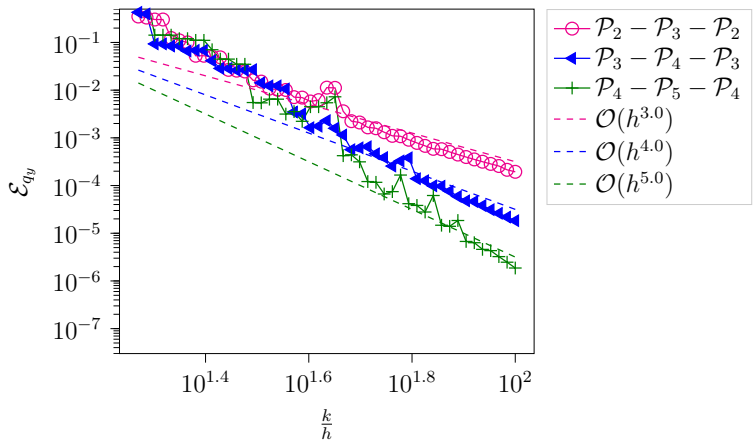


Figure 3.16: Large Mach convergence history for the second component of the volumetric unknown $\boldsymbol{q}_h$ for the HDG+ method

3.6.2 A posteriori error estimate

In this section, we will show that it is possible to compute a simple *a posteriori* error indicator. For more complete approach to *a posteriori* error analysis for HDG methods, we refer to [CZ12, CZ13]. We introduce the *relative jump error*

$$\mathcal{E}_{\text{jump}} = \frac{\sqrt{\sum_{K,\ell} \|\widehat{p_h} - p_h\|^2_{\partial K^\ell}}}{\sqrt{\sum_{K,\ell} \|\widehat{p_h}\|^2_{\partial K^\ell}}},$$

which measures the jump between p_h and $\widehat{p_h}$. For the HDG+ method, p_h should be replaced with $P_M p_h$.

Jump error and residuals for the HDG methods: To understand the importance of this quantity, we introduce the residuals

$$\delta_h := -\omega^2 \rho_0 p_h + \operatorname{div}(\boldsymbol{\sigma}_h) - s, \quad \text{and} \quad \boldsymbol{\Delta}_h := \boldsymbol{W_0}\boldsymbol{\sigma}_h + \nabla p_h + 2i\omega p_h \boldsymbol{W_0 b_0}.$$

We have chosen to work with the HDG-$\boldsymbol{\sigma}_h$ formulation as we recommend it over the HDG-$\boldsymbol{q}_h$ one, but the adaptation to the HDG-$\boldsymbol{q}_h$ formulation is immediate.
Reverting the integrations by parts in (3.12a)–(3.12b), we have

$$\begin{aligned}(\boldsymbol{\Delta}_h, \boldsymbol{r}_h)_K &= \langle p_h - \widehat{p_h}, \boldsymbol{r}_h \cdot \boldsymbol{n}\rangle_{\partial K}, \\ (\delta_h, w_h)_K &= -i\omega \langle \tau(p_h - \widehat{p_h}), w_h\rangle_{\partial K},\end{aligned}$$

as δ_h and $\boldsymbol{\Delta}_h$ are usually not polynomial quantities this means that

$$\begin{aligned}(\boldsymbol{P_V}\boldsymbol{\Delta}_h, \boldsymbol{r}_h)_K &= \langle p_h - \widehat{p_h}, \boldsymbol{r}_h \cdot \boldsymbol{n}\rangle_{\partial K}, \\ (P_W\delta_h, w_h)_K &= -i\omega \langle \tau(p_h - \widehat{p_h}), w_h\rangle_{\partial K},\end{aligned}$$

where $\boldsymbol{P_V}$ and P_W are the L^2-orthogonal projections onto $\boldsymbol{V}_h(K)$ and $W_h(K)$ respectively. Notice that if δ_h and $\boldsymbol{\Delta}_h$ are actually polynomials things are even simpler as $P_W\delta_h = \delta_h$ and $\boldsymbol{P_V}\boldsymbol{\Delta}_h = \boldsymbol{\Delta}_h$. Taking $\boldsymbol{r}_h = \boldsymbol{P_V}\boldsymbol{\Delta}_h$ and $w_h = P_W\delta_h$, and using the following inverse inequality

$$\|w_h\|_{\partial K} \leqslant C h_K^{-\frac{1}{2}} \|w_h\|_K, \quad \forall w_h \in W_h,$$

leads to

$$\begin{aligned}\|\boldsymbol{P_V}\boldsymbol{\Delta}_h\|_K &\leqslant C_{\boldsymbol{\Delta}} h_K^{-\frac{1}{2}} \|p_h - \widehat{p_h}\|_{\partial K}, \\ \|P_W\delta_h\|_K &\leqslant C_\delta h_K^{-\frac{1}{2}} \|p_h - \widehat{p_h}\|_{\partial K}.\end{aligned}$$

As

$$\|\delta_h\|_K \leqslant \|P_W\delta_h\|_K + \|(\mathrm{Id} - P_W)\delta_h\|_K, \tag{3.100}$$

we can see that the size of the residuals and hence the quality of the approximation depends only on the size of the jump $p_h - \widehat{p_h}$ (first term of the rhs) and the approximation properties of $W_h(K)$ (second term of the rhs).
When $p \in H^s(K)$ with $s \in [\![0, k+1]\!]$, we have

$$\|p - P_W p\|_K \leqslant C h_K^s \|p\|_{s,K},$$

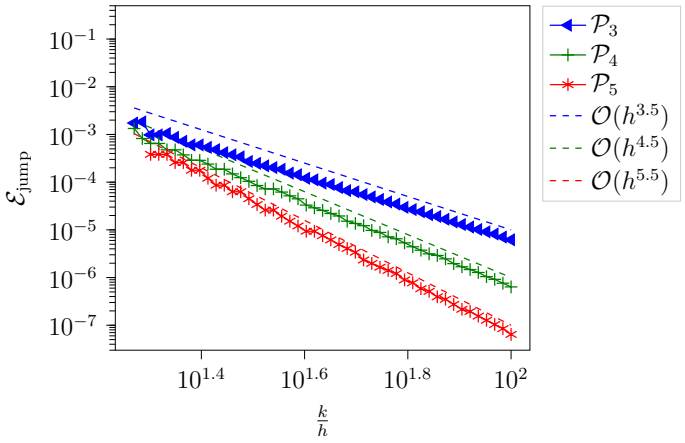
so the penalization term $\tau(p_h - \widehat{p_h})$ will ensure the stability of the method. Indeed (3.100) will therefore lead to

$$\|\delta_h\|_K \leqslant C_\delta h_K^{-\frac{1}{2}} \|p_h - \widehat{p_h}\|_{\partial K} + C h_K^s \|\delta_h\|_{s,K},$$

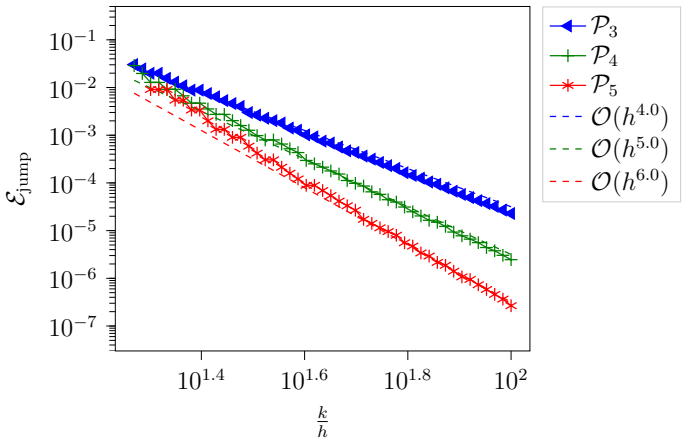
and assuming that h_K is small enough, we have

$$\|\delta_h\|_K \leqslant C h_K^{-\frac{1}{2}} \|p_h - \widehat{p_h}\|_{\partial K},$$

as the last term of the right-hand side can be absorbed by the left-hand side. A similar result can be obtained for $\mathbf{\Delta}_h$.
To illustrate this, we have depicted $\mathcal{E}_{\text{jump}}$ for the HDG methods for a low-Mach flow in FIGURE 3.17 and for a large-Mach flow in FIGURE 3.18.

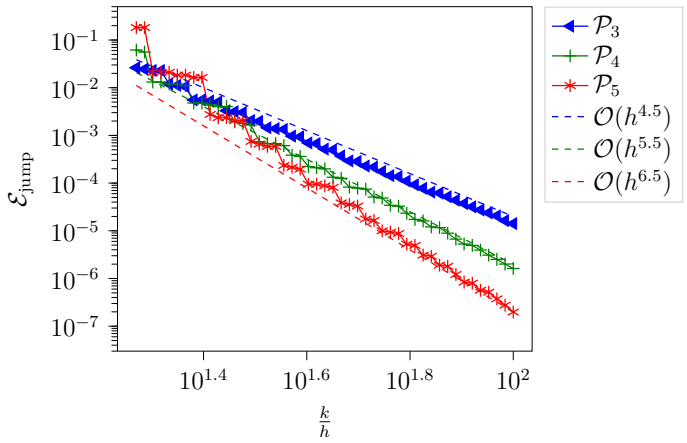


(a) HDG with diffusive flux $\boldsymbol{q}_h$

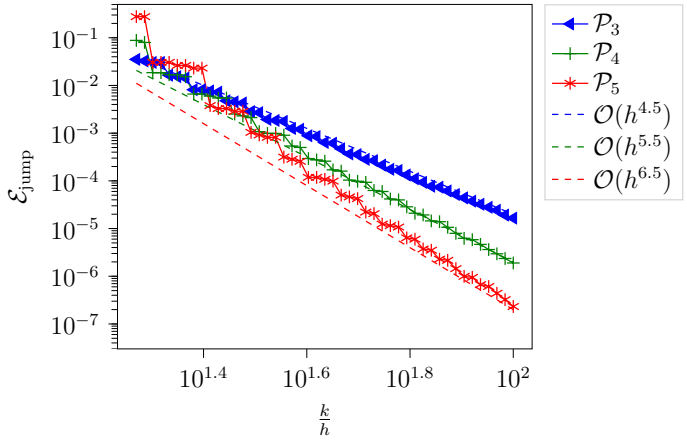


(b) HDG with total flux $\boldsymbol{\sigma}_h$

Figure 3.17: Low Mach convergence history for the jump error



(a) HDG with diffusive flux $\boldsymbol{q}_h$



(b) HDG with total flux $\boldsymbol{\sigma}_h$

Figure 3.18: Large Mach convergence history for the jump error

Adaptation to the HDG+ method: For the HDG+ method, nothing changes for the vectorial residual $\boldsymbol{\Delta}_h$ but things are a little different for the scalar residual δ_h. Indeed this time we have

$$\begin{aligned} \|P_W \delta_h\|_K &\lesssim h_K^{-\frac{1}{2}} \left(\| |\tau|(P_M p_h - \widehat{p_h}) \|_{\partial K} + \| |\tau_{\text{upw}}|(p_h - \widehat{p_h} \|_{\partial K} \right), \\ &\lesssim h_K^{-\frac{1}{2}} \| |\tau|(P_M p_h - \widehat{p_h}) \|_{\partial K}, \\ (\text{as } \tau = \mathcal{O}({h_K}^{-1})) \quad &\lesssim h_K^{-\frac{3}{2}} \| P_M p_h - \widehat{p_h} \|_{\partial K}. \end{aligned}$$

To illustrate this, we have depicted $\mathcal{E}_{\text{jump}}$ for the HDG+ method for a low-Mach flow in FIGURE 3.19 and for a large-Mach flow in FIGURE 3.20. Once again, we can see that these quantities are clearly decreasing.

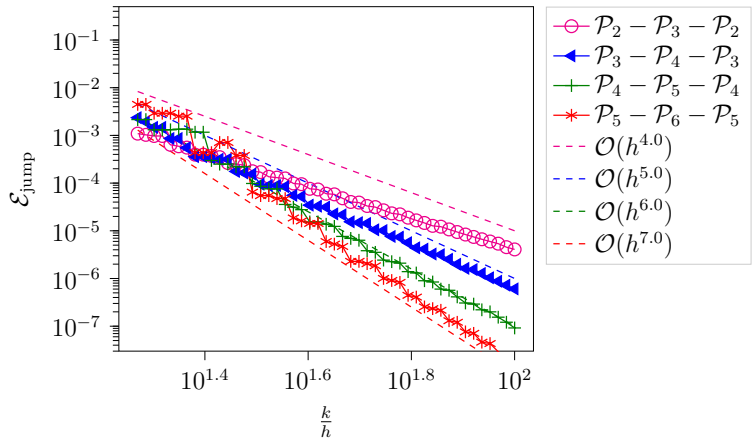


Figure 3.19: Low Mach convergence history for the jump error for the HDG+ method

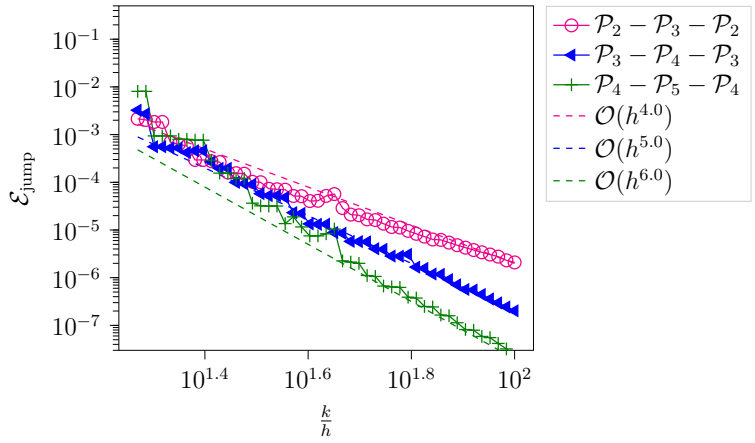


Figure 3.20: Large Mach convergence history for the jump error for the HDG+ method

3.6.3 Is the upwinding mechanism necessary ?

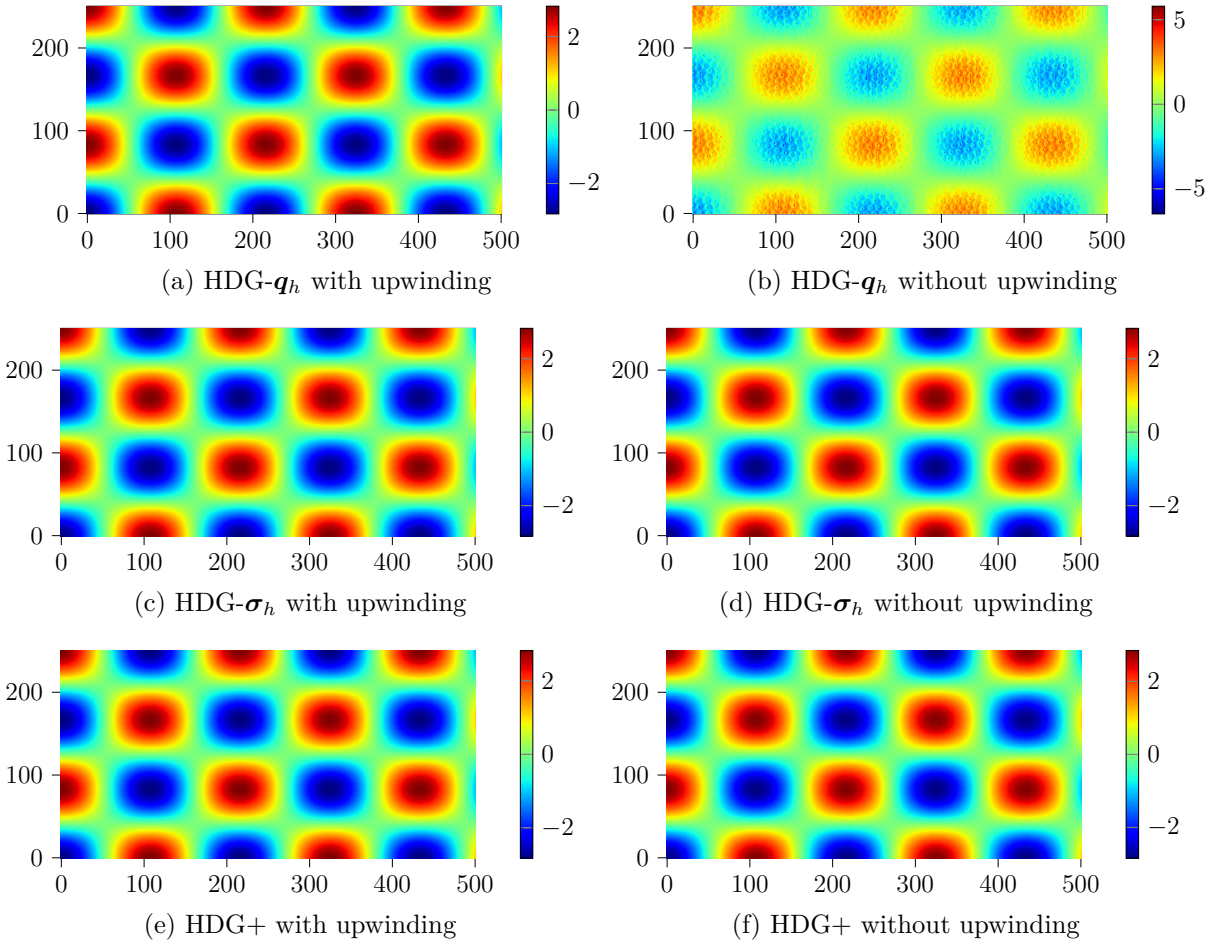
In this section we investigate the numerical impact of upwinding mechanisms. We have performed the same numerical simulations with the three methods with and without those mechanisms. For the HDG-$\boldsymbol{q}_h$ and the HDG+ methods, deactivating the upwinding mechanism corresponds to taking $\tau_{\text{upw}} \equiv 0$. For the HDG-$\boldsymbol{\sigma}_h$ it corresponds to taking $\tau = 1$ instead of the value given by the Riemann solver.
As it can be seen on FIGURE 3.21 the HDG-$\boldsymbol{q}_h$ without upwinding leads to poor numerical results, whereas the two other methods seem to perform well.

Method	With upwinding	Without upwinding
HDG-$\boldsymbol{q}_h$	$4.3 \cdot 10^{-7}$	0.53
HDG-$\boldsymbol{\sigma}_h$	$4.9 \cdot 10^{-7}$	$5.8 \cdot 10^{-7}$
HDG+	$6.3 \cdot 10^{-8}$	$5.77 \cdot 10^{-7}$

Table 3.8: Jump error $\mathcal{E}_{\text{jump}}$ for the different method

We have also computed the jump error $\mathcal{E}_{\text{jump}}$ as quality indicator of the numerical solution. The values are given in TABLE 3.8. Even if the error is higher without upwinding for the HDG+ and HDG-$\boldsymbol{\sigma}_h$, it seems to remain at a reasonable level. For the HDG+ method, we can conclude that the τ_{upw} penalization seems optional. This can be understood as we

need to choose $\tau = \mathcal{O}({h_K}^{-1})$ which is large, so the method seems less sensitive to changes in the penalization. On the other hand, this τ_{upw} penalization is mandatory to make the HDG-$\boldsymbol{q}_h$ formulation work. Finally, for the HDG-$\boldsymbol{\sigma}_h$ we still recommend to use the upwind penalization parameter τ as it leads to a method with no arbitrary choice to make.



(a) HDG-$\boldsymbol{q}_h$ with upwinding

(b) HDG-$\boldsymbol{q}_h$ without upwinding

(c) HDG-$\boldsymbol{\sigma}_h$ with upwinding

(d) HDG-$\boldsymbol{\sigma}_h$ without upwinding

(e) HDG+ with upwinding

(f) HDG+ without upwinding

Figure 3.21: p_h for the different HDG methods with and without upwind

3.6.4 Point-sources in a uniform flow

For many practical applications it is necessary to consider point-sources. In this section, we show that our method can be used for such computations.

Settings: We consider a uniform flow in an infinite plane. We use a Dirac point source $s = \delta_{(0,0)}$ located at the origin. We write $\boldsymbol{b_0}$ as

$$\boldsymbol{b_0} = M\left(\cos\alpha \boldsymbol{e}_x + \sin\alpha \boldsymbol{e}_y\right),$$

where M is the Mach number, α the angle between $\boldsymbol{b_0}$ and the horizontal axis, and we normalize $\rho_0, c_0 \equiv 1$.
The physical domain $\mathcal{O}_{\text{phys}}$ is surrounded with PMLs, see FIGURE 3.22 for a sketch of the geometric settings.

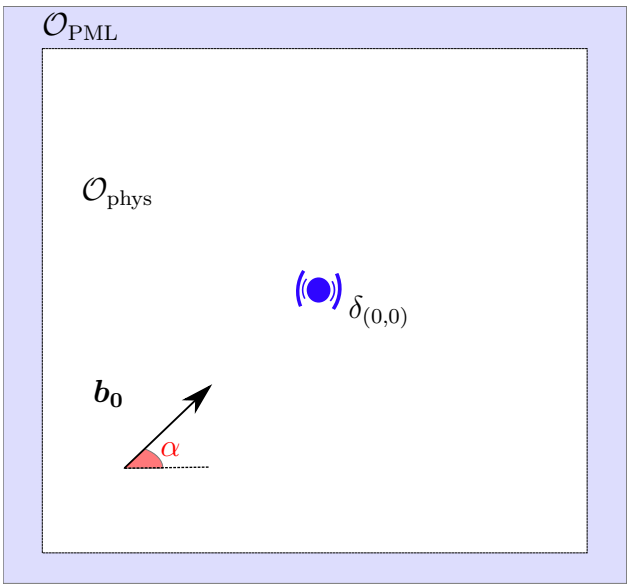
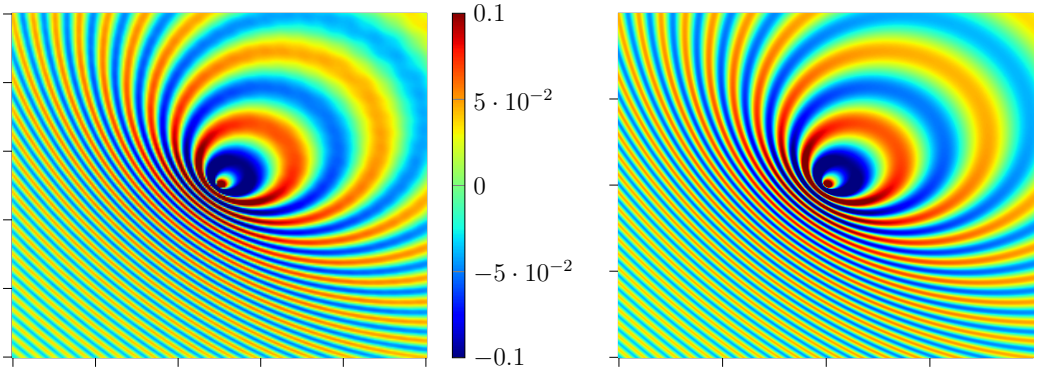


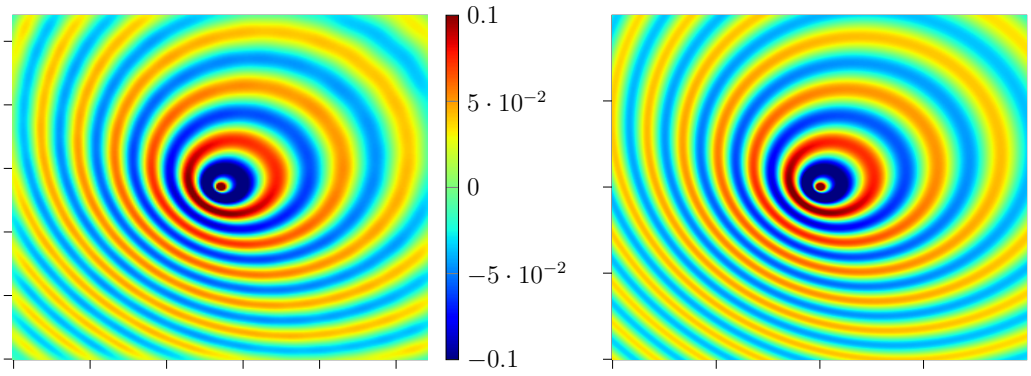
Figure 3.22: Geometric settings

We have depicted the solution for a large Mach number ($M = 0.8$) in FIGURE 3.23 and for a low Mach number ($M = 0.4$) in FIGURE 3.24. The presence of convection leads to a clearly visible Doppler effect : indeed because of the convection, the apparent frequency changes in the domain. The apparent frequency is higher in the bottom-left part of the domain than in the top-right part. The artifacts in the top-right part of the domain are due to the PMLs. Notice that the artifacts are more visible for the large Mach flow.



(a) Numerical result (b) Analytic solution

Figure 3.23: p_h computed with the HDG-$\boldsymbol{\sigma}_h$ method for $\omega = 6\pi$, $M = 0.8$ and $\alpha = \pi/4$. PMLs are not displayed. The colorbar is the same for the two pictures.



(a) Numerical result (b) Analytic solution

Figure 3.24: p_h computed with the HDG-$\boldsymbol{\sigma}_h$ method for $\omega = 6\pi$, $M = 0.4$ and $\alpha = \pi/4$. PMLs are not displayed. The colorbar is the same for the two pictures.

Analytic solution: In this context it is possible to write an analytic solution for (3.1). Following [HPN19], we first need to introduce the so-called *Prandlt-Glauert-Lorentz transformation* in the frequency domain. This transformation maps $\boldsymbol{x} = (x, y)$ to $\widetilde{\boldsymbol{x}} = (\widetilde{x}, \widetilde{y})$ and ω to $\widetilde{\omega}$, where

$$\begin{cases} \widetilde{x} = \left(1 + M^2 \dfrac{\cos^2\alpha}{\beta(1+\beta)}\right) x + M^2 \dfrac{\cos\alpha \sin\alpha}{\beta(1+\beta)} y \\ \widetilde{y} = M^2 \dfrac{\cos\alpha \sin\alpha}{\beta(1+\beta)} x + \left(1 + M^2 \dfrac{\sin^2\alpha}{\beta(1+\beta)}\right) y \,, \\ \widetilde{\omega} = \dfrac{\omega}{\beta} \end{cases}$$

with $\beta := \sqrt{1 - M^2}$. As a shorthand, we write this transformation

$$\widetilde{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x}, \quad \text{with } \boldsymbol{A} := \begin{bmatrix} \left(1 + M^2 \frac{\cos^2\alpha}{\beta(1+\beta)}\right) & M^2 \frac{\cos\alpha\sin\alpha}{\beta(1+\beta)} \\ M^2 \frac{\cos\alpha\sin\alpha}{\beta(1+\beta)} & \left(1 + M^2 \frac{\sin^2\alpha}{\beta(1+\beta)}\right) \end{bmatrix}.$$

This Prandtl-Glauert-Lorentz transformation is closely related to the Lorentz transform arising in special relativity. It is well-known that, when the background coefficients are uniform, these transformation maps the convected Helmholtz equation to a standard Helmholtz equation, see *eg.* [MBAG20, HPN19]. This can be understood as the Lorentz transformation was introduced as the transformation between two inertial frames that preserves wave equations, for a deeper insight of the connection between flow acoustics and Lorentzian geometry we refer to [Vis98]. The analytic solution is given by

$$p_{\text{exact}}(\boldsymbol{x}, \omega) = -\frac{i}{4\beta} H_0^{(1)}\left(\frac{\omega}{\beta}|\boldsymbol{A}\boldsymbol{x}|\right) \exp\left[\frac{i\omega}{\beta} \boldsymbol{A}\boldsymbol{x} \cdot \boldsymbol{b_0}\right]$$

where $H_0^{(1)}$ is the Hankel function of the first kind of order 0 and

$$\widetilde{r} := \sqrt{\widetilde{x}^2 + \widetilde{y}^2}.$$

Even if there is an analytic solution in this case, we were not able to obtain meaningful convergence plots due to the bad quality of the PMLs.

Computational cost: In Table 3.9 we have written down the sizes of the linear systems to solve for the different HDG methods using an interpolation of order 5 for the trace

variable. To give a reference, we also added the size of the system obtained when solving the convected Helmholtz equation (3.1) with a continuous finite element method (CG) with same interpolation degree using the `montjoie` solver[9]. We have also added the size of the system obtained when solving the standard Helmholtz equation with a Local Discontinuous Galerkin method (LDG), which is a first-order DG method, using the `montjoie` solver (as LDG methods are not implemented for the convected Helmholtz equation in `montjoie`).

Method	HDG-$\boldsymbol{q}_h$	HDG-$\boldsymbol{\sigma}_h$	HDG+	CG	LDG
k	5	5	4	5	5
nnz	12 237 750	12 237 750	6 243 750	1 353 750	46 862 808
nnz LU	67 515 161	67 515 161	34 244 845	39 008 186	259 272 979
$\mathcal{E}_{\text{jump}}$	$1.4 \cdot 10^{-3}$	$1.0 \cdot 10^{-3}$	$8.9 \cdot 10^{-4}$		

Table 3.9: Size of the linear system to solve for the Dirac in a uniform flow

We can see that using HDG instead of LDG leads to significantly smaller linear systems. For the convected Helmholtz equation, it is possible to use the CG method which less expensive than the HDG ones, see [CD16]. However, we would like to point out that the CG method is known to give bad numerical results for more realistic aeroacoustic models such as Galbrun's equation, see [CD18].

Local refinement: To handle the point-sources, a mesh with a local refinement around the source should be used, otherwise artifacts could be present in the numerical solution. With the HDG method, those artifacts seem really limited when no local refinement is used, see FIGURE 3.25.

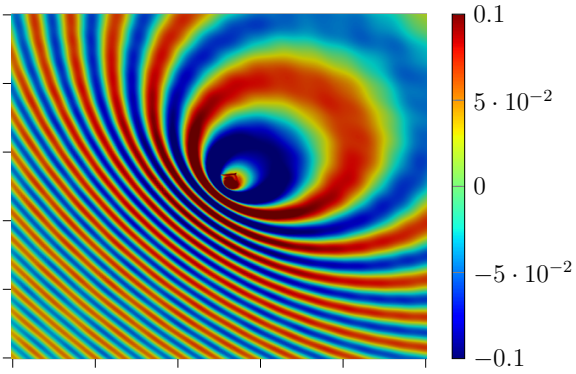


Figure 3.25: Point-source without local refinement

Illustrative example: To show that our method can handle more complex simulations, we consider the same test-case as before but with two point-sources located near the origin. On FIGURE 3.26, we can still see the changes in the apparent frequency due to the Doppler effect, but also interference patterns due to the interactions between the two sources.

[9] `montjoie` is a versatile and well-tested high-order finite element solver. For more informations about the numerical method used to solve the convected Helmholtz equation, see `http://montjoie.gforge.inria.fr/helmholtz.php`.

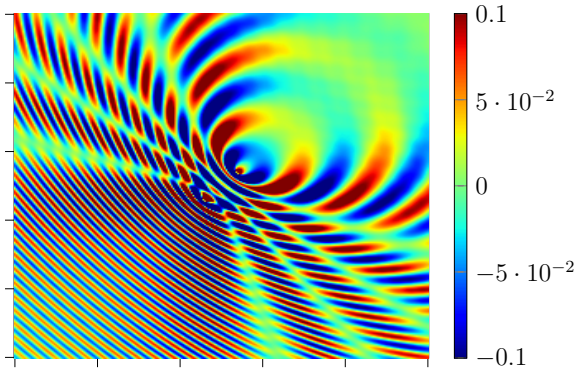


Figure 3.26: Interference between two point-sources

3.6.5 Gaussian jet

In this section we focus on a space-varying flow. We will use a *gaussian jet*, which is a common test-case in the literature, see *e.g.* [MMMP17, Sec. 3]. To work with parameters and unknowns without dimension, we choose

$$\rho_0 = c_0 \equiv 1,$$

and we consider the following gaussian jet flow

$$\boldsymbol{v_0} = M_0(y)\boldsymbol{e}_x, \quad \text{where} \quad M_0(y) := M_\infty + \mu \exp\left(-\frac{y^2}{R^2}\right),$$

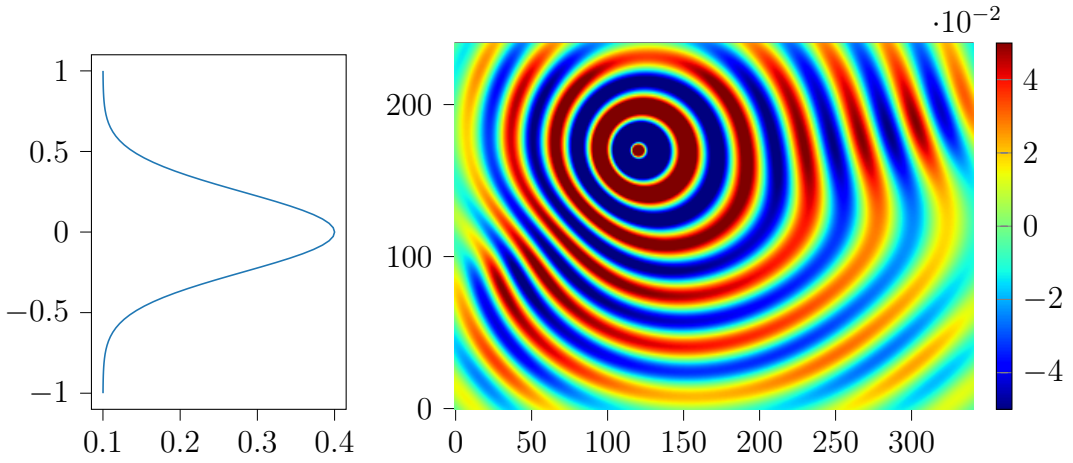
which is a gaussian perturbation of the uniform flow $\boldsymbol{v_0} = M_\infty \boldsymbol{e}_x$.
For this simulation, we work with the following values for the parameters

$$\mathcal{O}_{\text{phys}} = (0,3) \times (-1,1) \;\; ; \;\; \omega = 6\pi \;\; ; \;\; s = \delta_{(1,0.5)} \;\; ;$$

and

$$M_\infty = 0.1 \;\; ; \;\; \mu = 0.3 \;\; ; \;\; R = 0.35.$$

The physical domain $\mathcal{O}_{\text{phys}}$ is surrounded by PMLs.
Notice that these kind of jet-flows are not potential flows and should therefore not be used with the convected Helmholtz equation. However, as discussed in [MMMP17], with this choice of parameters the vorticity of the flow stays small and can be neglected. The convected Helmholtz equation is therefore a good approximation of the more realistic models.



(a) $M_0(y)$ (b) Numerical solution

Figure 3.27: $\mathfrak{Re}p_h$ for the gaussian jet, obtained with the HDG+ method

On FIGURE 3.27, we can see that there is a phase-shift inside the jet, as it is expected : this refraction-like effect can be seen in the center-left and center-right parts of the domain at the limit between the jet and uniform flows. There is also a small Doppler effect as the apparent frequency changes in the domain : it is different in the the bottom-left part of the domain and in the center-right one. However this effect is less obvious than in the previous example as the Mach number is significantly lower.

Conclusion

In this chapter, we have introduced three HDG methods to solve the convected Helmholtz equation. Two of them are standard HDG methods that use the same polynomial degree for the approximation of all the unknowns. The third one is less standard and uses a higher polynomial degree for the scalar unknown and a reduced stabilization process.

For all of those methods, detailed theoretical results on convergence and well-posedness are provided. It is important to note that we could not obtain the super-convergence property for the two standard HDG methods because of convection. We also provided numerical experiments that are consistent with the absence of super-convergence. The HDG+ method achieves optimal convergence. Due to the reduced stabilization this leads to a "super-convergence like behaviour" as we obtain a convergence rate of $k+2$ for the cost of a HDG method of degree k without post-processing.

During the numerical experiments, it occurred to us that the HDG+ and HDG-$\boldsymbol{\sigma}_h$ methods seemed more robust than the HDG-$\boldsymbol{q}_h$ method. In particular, they are less sensitive to the choice of penalization parameter.

Appendix

3.A Intermediate results for the error analysis of the HDG method with diffusive flux

In this appendix we state the intermediate results allowing to adapt the analysis conducted in Subsection 3.4.3 to prove the convergence of the HDG method stated in Theorem 7. The proofs are omitted but are very similar to the ones given in Subsection 3.4.3.
The main differences to keep in mind are : $s \in [1, k+1]$ instead of $s \in [1, k+2]$ and $\tau = \mathcal{O}(1)$ instead of $\tau = \mathcal{O}(h^{-1})$.

Gradient estimate: this corresponds to Lemma 3.4.3

Lemma 3.A.1:
The following estimate holds

$$\|\nabla \varepsilon_h^p\|_{\mathcal{T}_h} \lesssim \|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + \|\boldsymbol{\delta}_h^{\boldsymbol{q}}\|_{\mathcal{T}_h} + h^{-\frac{1}{2}} \|\varepsilon_h^p - \widehat{\varepsilon}_h^p\|_{\partial\mathcal{T}_h}$$

Energy-like equality: this corresponds to Lemma 3.4.5.

Lemma 3.A.2:
The following energy-like equality holds

$$\begin{aligned}
&\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|^2_{\boldsymbol{W_0},\mathcal{T}_h} - \omega^2 \|\varepsilon_h^p\|^2_{\rho_0,K} - 2i\omega \left(\|\varepsilon_h^p - \widehat{\varepsilon}_h^p\|^2_{\partial\mathcal{T}_h} + \left\| \left(\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac{1}{2}} (\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|^2_{\partial\mathcal{T}_h} \right) \\
&= -\omega^2 (\rho_0 \delta_h^p, \varepsilon_h^p)_{\mathcal{T}_h} + 2i\omega (\delta_h^p \boldsymbol{b_0}, \nabla\varepsilon_h^p)_{\mathcal{T}_h} + (\boldsymbol{W_0}\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}, \boldsymbol{\delta}_h^{\boldsymbol{q}})_{\mathcal{T}_h} + 2\omega \mathfrak{Im} (\varepsilon_h^p \boldsymbol{b_0}, \nabla\varepsilon_h^p)_{\mathcal{T}_h} \\
&+ \left\langle \boldsymbol{\delta}_h^{\boldsymbol{q}}\cdot\boldsymbol{n} + 2i\omega(\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\delta}_h^p - 2i\omega\tau(\delta_h^p - \widehat{\delta}_h^p) + 2i\omega\tau_{\mathrm{upw}}(\delta_h^p - \widehat{\delta}_h^p), \varepsilon_h^p - \widehat{\varepsilon}_h^p \right\rangle_{\partial\mathcal{T}_h}
\end{aligned}$$

Furthermore if $p \in H^s(\mathcal{O})$ and $\boldsymbol{q} \in \boldsymbol{H^t}(\mathcal{O})$ where $s, t \in [1, k+1]$ then the following estimate holds

$$\begin{aligned}
&\left| \|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|^2_{\boldsymbol{W_0},\mathcal{T}_h} - 2i\omega \left(\|\varepsilon_h^p - \widehat{\varepsilon}_h^p\|^2_{\partial\mathcal{T}_h} + \left\| \left(\frac{1}{2}|\boldsymbol{b_0}\cdot\boldsymbol{n}|\right)^{\frac{1}{2}} (\varepsilon_h^p - \widehat{\varepsilon}_h^p) \right\|^2_{\partial\mathcal{T}_h} \right) \right| \\
\lesssim &\omega^2 \|\varepsilon_h^p\|^2_{\mathcal{T}_h} + \omega \|\varepsilon_h^p\|_{\mathcal{T}_h} \left(\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} + h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} + h^{-\frac{1}{2}} \|\varepsilon_h^p - \widehat{\varepsilon}_h^p\|_{\partial\mathcal{T}_h} + \omega h^s \|p\|_{s,\mathcal{O}} \right) \\
&+ \|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h} \left(h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} + \omega h^s \|p\|_{s,\mathcal{O}} \right) + h^{2t} \|\boldsymbol{q}\|^2_{t,\mathcal{O}} \\
&+ h^{-\frac{1}{2}} \|\varepsilon_h^p - \widehat{\varepsilon}_h^p\|_{\partial\mathcal{T}_h} \left(\omega h^s \|p\|_{s,\mathcal{O}} + h^t \|\boldsymbol{q}\|_{t,\mathcal{O}} \right)
\end{aligned}$$

Dual identity: this corresponds to Lemma 3.4.6.

Lemma 3.A.3:
The following dual identity holds

$$\begin{aligned}\|\varepsilon_h^p\|_{\mathcal{T}_h}^2 =& -(\boldsymbol{W_0}\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}, \boldsymbol{\pi_V}\boldsymbol{\xi}-\boldsymbol{\xi})_{\mathcal{T}_h} + \omega^2(\rho_0\varepsilon_h^p, \pi_W\theta-\theta)_{\mathcal{T}_h} + 2i\omega(\nabla\varepsilon_h^p, (\pi_W\theta-\theta)\boldsymbol{b_0})_{\mathcal{T}_h}\\ &- 2i\omega\langle(\boldsymbol{b_0}\cdot\boldsymbol{n})\varepsilon_h^p, \pi_W\theta-\theta\rangle_{\partial\mathcal{T}_h}\\ &+ (\boldsymbol{W_0}\boldsymbol{\delta}_h^{\boldsymbol{q}}, \boldsymbol{\pi_V}\boldsymbol{\xi})_{\mathcal{T}_h} - \omega^2(\rho_0\delta_h^p, \pi_W\theta)_{\mathcal{T}_h} + 2i\omega(\delta_h^p\boldsymbol{b_0}, \nabla(\pi_W\theta))_{\mathcal{T}_h}\\ &+ 2i\omega\langle(\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\varepsilon}_h^p - \tau(\varepsilon_h^p-\widehat{\varepsilon}_h^p) + \tau_{\text{upw}}(\varepsilon_h^p-\widehat{\varepsilon}_h^p), \pi_W\theta - P_M\theta\rangle_{\partial\mathcal{T}_h}\\ &- 2i\omega\left\langle(\boldsymbol{b_0}\cdot\boldsymbol{n})\widehat{\delta}_h^p - \tau(\delta_h^p-\widehat{\delta}_h^p) + \tau_{\text{upw}}\left(\delta_h^p-\widehat{\delta}_h^p\right), \pi_W\theta - P_M\theta\right\rangle_{\partial\mathcal{T}_h}\\ &+ \langle\boldsymbol{\delta}_h^{\boldsymbol{q}}\cdot\boldsymbol{n}, \pi_W\theta - P_M\theta\rangle_{\partial\mathcal{T}_h}\end{aligned}$$

Dual estimate: this corresponds to Lemma 3.4.7.

Lemma 3.A.4:
Assuming that the regularity assumption (3.75) holds and that $\omega^2h^2\|\rho_0\|_\infty C_{\text{reg}}C$ (where C is the constant of Theorem 5) is small enough, if $p \in H^s(\mathcal{O})$ and $\boldsymbol{q} \in \boldsymbol{H^t}(\mathcal{O})$ where $s,t \in [1, k+1]$ then

$$\|\varepsilon_h^p\|_{\mathcal{T}_h} \lesssim h^{t+1}(1+\omega)\|\boldsymbol{q}\|_{t,\mathcal{O}} + h^s(1+\omega+\omega^2)\|p\|_{s,\mathcal{O}} + h(1+\omega)\|\boldsymbol{\varepsilon}_h^{\boldsymbol{q}}\|_{\boldsymbol{W_0},\mathcal{T}_h}$$

where

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

Bibliography

[ALA13] Natalia C. B. Arruda, Abimael F. D. Loula, and Regina C. Almeida. Locally discontinuous but globally continuous Galerkin methods for elliptic problems. *Computer Methods in Applied Mechanics and Engineering*, 255:104–120, March 2013.

[BBL03] Eliane Bécache, Anne-Sophie Bonnet-Ben Dhia, and Guillaume Legendre. Perfectly matched layers for the convected Helmholtz equation. Technical report, 2003.

[BCDL15] Marie Bonnasse-Gahot, Henri Calandra, Julien Diaz, and Stéphane Lanteri. Hybridizable Discontinuous Galerkin method for the simulation of the propagation of the elastic wave equations in the frequency domain. Research Report RR-8990, INRIA Bordeaux ; INRIA Sophia Antipolis - Méditerranée, June 2015.

[BDMP21] Hélène Barucq, Julien Diaz, Rose-Cloé Meyer, and Ha Pham. Implementation of hybridizable discontinuous Galerkin method for time-harmonic anisotropic poroelasticity in two dimensions. *International Journal for Numerical Methods in Engineering*, 122(12):3015–3043, 2021.

[CC12] Yanlai Chen and Bernardo Cockburn. Analysis of variable-degree HDG methods for convection–diffusion equations. Part I: General nonconforming meshes. *IMA Journal of Numerical Analysis*, 32(4):1267–1293, October 2012.

[CC14] Yanlai Chen and Bernardo Cockburn. Analysis of variable-degree HDG methods for convection-diffusion equations. Part II: Semimatching nonconforming meshes. *Mathematics of Computation*, 83(285):87–111, January 2014.

[CD16] Juliette Chabassier and Marc Duruflé. High Order Finite Element Method for solving Convected Helmholtz equation in radial and axisymmetric domains. Application to Helioseismology. Research Report RR-8893, Inria Bordeaux Sud-Ouest, March 2016.

[CD18] Juliette Chabassier and Marc Duruflé. Solving time-harmonic Galbrun's equation with an arbitrary flow. Application to Helioseismology. *Rapport de recherche Inria*, 2018.

[CDG+09] Bernardo Cockburn, Bo Dong, Johnny Guzmán, Marco Restelli, and Riccardo Sacco. A Hybridizable Discontinuous Galerkin Method for Steady-State Convection-Diffusion-Reaction Problems. *SIAM Journal on Scientific Computing*, 31(5):3827–3846, January 2009.

[CDPE16] Bernardo Cockburn, Daniele A. Di Pietro, and Alexandre Ern. Bridging the hybrid high-order and hybridizable discontinuous Galerkin methods. *ESAIM: Mathematical Modelling and Numerical Analysis*, 50(3):635–650, May 2016.

[CGL09] Bernardo Cockburn, Jayadeep Gopalakrishnan, and Raytcho Lazarov. Unified Hybridization of Discontinuous Galerkin, Mixed, and Continuous Galerkin Methods for Second Order Elliptic Problems. *SIAM J. Numer. Anal*, 47:1319–1365, August 2009.

[CGS10] Bernardo Cockburn, Jayadeep Gopalakrishnan, and Francisco-Javier Sayas. A projection-based error analysis of HDG methods. *Mathematics of Computation*, 79(271):1351–1367, March 2010.

[CLOS20] Liliana Camargo, Bibiana López-Rodríguez, Mauricio Osorio, and Manuel Solano. An HDG method for Maxwell's equations in heterogeneous media. *Computer Methods in Applied Mechanics and Engineering*, 368:113178, August 2020.

[Coc14] Bernardo Cockburn. Static Condensation, Hybridization, and the Devising of the HDG Methods. *springerprofessional.de*, 2014.

[CQS18] Huangxin Chen, Weifeng Qiu, and Ke Shi. A priori and computable a posteriori error estimates for an HDG method for the coercive Maxwell equations. *Computer Methods in Applied Mechanics and Engineering*, 333:287–310, May 2018.

[CQSS17] Huangxin Chen, Weifeng Qiu, Ke Shi, and Manuel Solano. A Superconvergent HDG Method for the Maxwell Equations. *Journal of Scientific Computing*, 70(3):1010–1029, March 2017.

[CS13] Bernardo Cockburn and Ke Shi. Superconvergent HDG methods for linear elasticity with weakly symmetric stresses | IMA Journal of Numerical Analysis | Oxford Academic. *IMA Journal of Numerical Analysis*, 2013.

[CZ12] Bernardo Cockburn and Wujun Zhang. A Posteriori Error Estimates for HDG Methods. *Journal of Scientific Computing*, 51(3):582–607, June 2012.

[CZ13] Bernardo Cockburn and Wujun Zhang. A Posteriori Error Analysis for Hybridizable Discontinuous Galerkin Methods for Second Order Elliptic Problems. *SIAM Journal on Numerical Analysis*, 51(1):676–693, January 2013.

[DS19] Shukai Du and Francisco-Javier Sayas. *An Invitation to the Theory of the Hybridizable Discontinuous Galerkin Method: Projections, Estimates, Tools*. SpringerBriefs in Mathematics. Springer International Publishing, Cham, 2019.

[Dub91] Moshe Dubiner. Spectral methods on triangles and other domains. *Journal of Scientific Computing*, 6(4):345–390, December 1991.

[EG04] Alexandre Ern and Jean-Luc Guermond. *Theory and Practice of Finite Elements*. Applied Mathematical Sciences. Springer-Verlag, New York, 2004.

[Fau21] Florian Faucher. 'hawen': Time-harmonic wave modeling and inversion using hybridizable discontinuous Galerkin discretization. *Journal of Open Source Software*, 6(57):2699, January 2021.

[FCS15] Guosheng Fu, Bernardo Cockburn, and Henryk Stolarski. Analysis of an HDG method for linear elasticity. *International Journal for Numerical Methods in Engineering*, 102(3-4):551–575, 2015.

[FLd14] Cristiane O. Faria, Abimael F. D. Loula, and Antônio J. B. dos Santos. Primal stabilized hybrid and DG finite element methods for the linear elasticity problem. *Computers & Mathematics with Applications*, 68(4):486–507, August 2014.

[FS20] Florian Faucher and Otmar Scherzer. Adjoint-state method for Hybridizable Discontinuous Galerkin discretization, application to the inverse acoustic wave problem. *Computer Methods in Applied Mechanics and Engineering*, 372:113406, December 2020.

[GM11] Roland Griesmaier and Peter Monk. Error Analysis for a Hybridizable Discontinuous Galerkin Method for the Helmholtz Equation. *Journal of Scientific Computing*, 49(3):291–310, December 2011.

[GSV18] Jay Gopalakrishnan, Manuel Solano, and Felipe Vargas. Dispersion Analysis of HDG Methods. *Journal of Scientific Computing*, 77(3):1703–1735, December 2018.

[HPN19] Fang Q. Hu, Michelle E. Pizzo, and Douglas M. Nark. On the use of a Prandtl-Glauert-Lorentz transformation for acoustic scattering by rigid bodies with a uniform flow. *Journal of Sound and Vibration*, 443:198–211, March 2019.

[HPS17] Allan Hungria, Daniele Prada, and Francisco-Javier Sayas. HDG methods for elastodynamics. *Computers & Mathematics with Applications*, 74(11):2671–2690, December 2017.

[Hun19] Allan Hungria. *Using HDG+ to Compute Solutions of the 3D Linear Elastic and Poroelastic Wave Equations*. PhD thesis, University of Delaware, 2019.

[HW08] Jan S. Hesthaven and Tim Warburton. *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications*. Texts in Applied Mathematics. Springer-Verlag, New York, 2008.

[Kir04] Robert C. Kirby. Algorithm 839: FIAT, a new paradigm for computing finite element basis functions. *ACM Transactions on Mathematical Software (TOMS)*, 30(4):502–516, December 2004.

[KSC12] Robert J. Kirby, Spencer J. Sherwin, and Bernardo Cockburn. To CG or to HDG: A Comparative Study. *Journal of Scientific Computing*, 51(1):183–212, April 2012.

[Leh10] Christoph Lehrenfeld. *Hybrid Discontinuous Galerkin Methods for Solving Incompressible Flow Problems*. PhD thesis, RWTH Aachen, 2010.

[MBAG20] Philippe Marchner, Hadrien Beriot, Xavier Antoine, and Christophe Geuzaine. Stable Perfectly Matched Layers with Lorentz transformation for the convected Helmholtz equation. Technical report, 2020.

[MMMP17] Jean-François Mercier, Colin Mietka, Florence Millot, and Vincent Pagneux. Acoustic propagation in a vortical homentropic flow. page 19, 2017.

[NPRC15] Ngoc-Cuong Nguyen, Jaime Peraire, Fernando Reitich, and Bernardo Cockburn. A phase-based hybridizable discontinuous Galerkin method for the numerical solution of the Helmholtz equation. *Journal of Computational Physics*, 290:318–335, June 2015.

[Oik14] Issei Oikawa. A hybridized discontinuous Galerkin method with reduced stabilization. *arXiv:1405.2491 [math]*, November 2014.

[Oik16] Issei Oikawa. Analysis of a Reduced-Order HDG Method for the Stokes Equations. *Journal of Scientific Computing*, 67(2):475–492, May 2016.

[Oik18] Issei Oikawa. An HDG Method with Orthogonal Projections in Facet Integrals. *Journal of Scientific Computing*, 76(2):1044–1054, August 2018.

[PE12] Daniele Antonio Di Pietro and Alexandre Ern. *Mathematical Aspects of Discontinuous Galerkin Methods*. Mathématiques et Applications. Springer-Verlag, Berlin Heidelberg, 2012.

[Pie90] Allan Pierce. Wave equation for sound in fluids with unsteady inhomogeneous flow. *The Journal of the Acoustical Society of America*, 87(6):2292–2299, June 1990.

[QS16a] Weifeng Qiu and Ke Shi. An HDG Method for Convection Diffusion Equation. *Journal of Scientific Computing*, 66(1):346–357, January 2016.

[QS16b] Weifeng Qiu and Ke Shi. A superconvergent HDG method for the incompressible Navier–Stokes equations on general polyhedral meshes. *IMA Journal of Numerical Analysis*, 36(4):1943–1967, October 2016.

[QSS16] Weifeng Qiu, Jiguang Shen, and Ke Shi. An HDG method for linear elasticity with strong symmetric stresses. *arXiv:1312.1407 [math]*, February 2016.

[SM50] Jack Sherman and Winifried Morrison. Adjustment of an Inverse Matrix Corresponding to a Change in One Element of a Given Matrix. *Annals of Mathematical Statistics*, 21(1):124–127, March 1950.

[Ste91] Rolf Stenberg. Postprocessing schemes for some mixed finite elements. *ESAIM: Mathematical Modelling and Numerical Analysis*, 25(1):151–167, 1991.

[Vis98] Matt Visser. Acoustic black holes: Horizons, ergospheres, and Hawking radiation. *Classical and Quantum Gravity*, 15(6):1767–1791, June 1998.

[YMKS16] Sergey Yakovlev, David Moxey, Robert M. Kirby, and Spencer J. Sherwin. To CG or to HDG: A Comparative Study in 3D. *Journal of Scientific Computing*, 67(1):192–220, April 2016.

Chapter 4

Absorbing Boundary Conditions for the convected Helmholtz equation

Contents

Introduction

Domain truncation is an important problem of computational wave dynamics. Indeed in many applications, the waves propagate in an infinite medium which must be truncated to a finite one to allow numerical simulation using finite element methods. The introduction of an artificial boundary should be treated carefully to avoid numerical pollution due to reflections at the boundary. Among others, the two most popular truncation techniques are

- *Perfectly Matched Layers (PMLs)*, introduced in [Ber94], which consist in surrounding the computational domain by a layer in which the waves are highly absorbed,

- *Absorbing Boundary Conditions (ABCs)*, introduced in [EM77], which minimize the reflected waves generated by waves impinging the boundary introduced for truncating the computational domain.

It is worth remembering that ABCs are often expressed in terms of non-local operators and that finding efficient local approximation of those operators is usually a challenging problem.

In this chapter we consider the *convected Helmholtz equation* which is the simplest model for acoustic wave propagation in a flow. It has been shown that standard PMLs are unstable because of the existence of back-propagating modes. In [BBL03], stable PMLs for the convected Helmholtz equation were derived, and this work was generalized to more general configurations in [MBAG20]. The key ingredient of those stable PMLs is a change of variables that transforms the convected Helmholtz equation into the standard one. The so-called *Prandtl-Glauert-Lorentz transformation* (PGL) of [HPN19] is the most common example of such a change of variables. Similar work for time-domain acoustics has been performed in [DJ06]. On the other hand, ABCs for the convected Helmholtz equation in a waveguide were derived in [Kim14] and [MAGB21]. For industrial applications in which an exact radiation condition is required, a FEM-BEM coupling that approximates the non-local Dirichlet-to-Neumann operator has been considered in [CES14, BCD$^+$14].

Herein, we construct a family of local ABCs based on the *Prandtl-Glauert-Lorentz transformation* for the convected Helmholtz equation. Those ABCs are constructed from the ABCs formerly derived for the standard Helmholtz equation in [EM77]. As the *Prandtl-Glauert-Lorentz transformation* is only used to compute the impedance operator on the artificial boundary, the new ABCs are valid for a flow that varies inside the domain but becomes uniform at infinity. Those new ABCs are efficient for low and intermediate Mach numbers and are easy to implement in an existing finite-element solver.

This chapter is organized as follows:

- in Section 4.1, we recall some results on the convected Helmholtz equation and precisely state the assumptions under which the new ABCs are derived,
- in Section 4.2, we use the *Prandtl-Glauert-Lorentz transformation* to construct the new ABCs,
- in Section 4.3, we present numerical results to illustrate the performance of the new ABCs.

4.1 Model problem and geometric settings

In this section, we introduce the *convected Helmholtz equation* and recall some of its properties. We also precisely state the assumptions under which the new Absorbing Boundary Condition are constructed. Finally, we show that on a bounded domain the convected Helmholtz equation with absorbing boundary conditions is a well-posed problem.

4.1.1 Equation and carrier flow

We consider the following *convected Helmholtz equation* written in pressure formulation

$$-\omega^2 p - 2i\omega \boldsymbol{v_0} \cdot \nabla p + \boldsymbol{v_0} \cdot \nabla[\boldsymbol{v_0} \cdot \nabla p] - \operatorname{div}\left(c_0^2 \nabla p\right) = s, \tag{4.1}$$

where p is the acoustic potential, $\boldsymbol{v_0}$ is the velocity of the carrier flow, c_0 is the adiabatic sound speed and s is the acoustic source.
In this chapter, we use the following convention for time-harmonic solutions

$$\mathsf{p}(\boldsymbol{x}, t) = p(\boldsymbol{x}, \omega) e^{-i\omega t}. \tag{4.2}$$

To make the notations lighter, the $e^{-i\omega t}$ factor is omitted.

As there are several ways to write the convected Helmholtz equation, see [Pie90], we have chosen the simple form (4.1) as a model problem for this chapter. The construction of the ABCs presented here can then be adapted to the other forms of the equation without major difficulties.

Assumptions on the carrier flow. In this chapter we will make the following usual assumptions on the carrier flow:

Assumption 5 (Conservation of mass):
The velocity $\boldsymbol{v_0}$ satisfies the following mass conservation equation

$$\operatorname{div}(\boldsymbol{v_0}) = 0.$$

Assumption 6 (Subsonic flow):
The carrier flow is subsonic, *i.e.*

$$\inf_{\mathcal{O}} \left(c_0^2 - |\boldsymbol{v_0}|^2\right) > 0.$$

Second-order operator: We introduce the matrix

$$\boldsymbol{K_0} := c_0^2 \mathbf{Id} - \boldsymbol{v_0}\boldsymbol{v_0}^T,$$

using Assumption 5, (4.1) can be rewritten as

$$-\omega^2 p - \operatorname{div}(\boldsymbol{K_0}\nabla p + 2i\omega p \boldsymbol{v_0}) = s. \tag{4.3}$$

The quantity

$$\boldsymbol{\sigma} := -\boldsymbol{K_0}\nabla p - 2i\omega p \boldsymbol{v_0}$$

is called the *total flux* and will be convenient to derive absorbing boundary conditions. Indeed, let Γ be a curve and $\boldsymbol{n}_\Gamma$ its unitary normal vector, then the acoustic energy flux going through Γ is proportional to $\boldsymbol{\sigma}|_\Gamma \cdot \boldsymbol{n}_\Gamma$.

4.1.2 Absorbing boundary condition, weak formulation and well-posedness

Truncated domain. Equation (4.3) is posed on the free space $\mathbb{R}^n$ for $n = 2$ or 3. Even if we only present numerical examples in two dimensions, the ABCs derived here are also valid in dimension 3. However to perform numerical simulation we need to work on a finite domain. In practice, we introduce a boundary Σ, the so-called artificial boundary, which limits a region including the support of the acoustic source and thus defines a bounded computational domain $\mathcal{O}$. The ABC is then set on the artificial boundary Σ.

Boundary condition. There exists an exact boundary condition which makes the artificial boundary transparent. It can be written as

$$\boldsymbol{\sigma} \cdot \boldsymbol{n} + \mathrm{DtN}(p) = 0, \quad \text{on } \Sigma, \tag{4.4}$$

where DtN is the *Dirichlet-to-Neumann map* which is a non-local operator. In the context of convected acoustics, the quantity $\boldsymbol{\sigma}\cdot\boldsymbol{n}$ should be interpreted as $-\partial_n p$. We have therefore chosen to call DtN the operator involved in the exact radiation condition.
In this chapter, we will approximate (4.4) by the following *local* absorbing boundary condition

$$\boldsymbol{\sigma}\cdot\boldsymbol{n} + \mathcal{Z}p = 0, \qquad \text{on } \Sigma, \tag{4.5}$$

where $\mathcal{Z}\in\mathbb{C}$ is a local approximation of the DtN-operator which will be determined later. The condition (4.5) has the form of an *impedance boundary condition* and we refer to $\mathcal{Z}$ as an *impedance-like coefficient*.
For now we will make the following assumption on $\mathcal{Z}$ and we will later show that it holds.

Assumption 7 (Value of $\mathcal{Z}$):
The value of $\mathcal{Z}$ is chosen such that

$$\mathfrak{Im}\mathcal{Z} \neq \omega\boldsymbol{v_0}\cdot\boldsymbol{n},$$

and the sign of $\mathfrak{Im}\mathcal{Z}$ does not change on Σ.

Weak formulation: Equation (4.3) and the impedance boundary condition (4.5) lead to the following problem:

$$\text{Seek } p\in H^1(\mathcal{O}) \text{ such that } \mathsf{a}(p,w_h)=\ell(w_h), \text{ for all } w_h\in H^1(\mathcal{O}), \tag{4.6}$$

where

$$\mathsf{a}(p,w_h) := \int_{\mathcal{O}} \boldsymbol{K_0}\nabla p\cdot\nabla w_h^* + 2i\omega p\boldsymbol{v_0}\cdot\nabla w_h^* - \omega^2 p w_h^* \mathrm{d}\boldsymbol{x} - \int_{\partial\mathcal{O}} \mathcal{Z}p w_h^* \mathrm{d}\sigma, \tag{4.7a}$$

$$\ell(w_h) := \int_{\mathcal{O}} s w_h^* \mathrm{d}\boldsymbol{x}. \tag{4.7b}$$

Notice that this formulation can only be obtained when Assumption 5 is true.

Well-posedness. Before actually showing the well-posedness of the problem (4.6), we recall the Lemma 4.1.1 from Chapter 3.

Lemma 4.1.1:
Under Assumption 5 and if $p\in H^1(\mathcal{O})$ and $\boldsymbol{v_0}\in\boldsymbol{L}^\infty(\mathcal{O})\cap\boldsymbol{\mathcal{C}}(\mathcal{O})$, the following identity holds

$$\mathfrak{Re}\int_{\mathcal{O}} p\boldsymbol{v_0}\cdot\nabla p^*\mathrm{d}\boldsymbol{x} = \frac{1}{2}\int_{\partial\mathcal{O}}(\boldsymbol{v_0}\cdot\boldsymbol{n})|p|^2\mathrm{d}\sigma.$$

Using this lemma, we can now prove that the convected Helmholtz equation with absorbing boundary conditions is a well-posed problem.

Theorem 8 :

Under Assumption 5, Assumption 6 and Assumption 7, the problem (4.6) has a unique solution $p\in H^1(\mathcal{O})$.

Proof : When Assumption 6 holds, $\boldsymbol{K_0}$ is a symmetric positive-definite matrix and its spectrum is given by

$$\mathrm{Sp}(\boldsymbol{K_0}) = \left\{c_0^2, c_0^2 - |\boldsymbol{v_0}|^2\right\}.$$

The problem (4.6) is therefore of Fredholm type and existence of the solution will follow from its uniqueness.
Taking $w_h = p$ in (4.7a) and $s = 0$ in (4.7b) leads to the following energy-like identity

$$\left\|\boldsymbol{K_0}^{\frac{1}{2}}\nabla p\right\|_{\mathcal{O}}^2 - \omega^2\|p\|_{\mathcal{O}}^2 + 2i\omega\int_{\mathcal{O}} p\boldsymbol{v_0}\cdot\nabla p^*\mathrm{d}\boldsymbol{x} - \int_{\partial\mathcal{O}}\mathcal{Z}|p|^2\mathrm{d}\sigma = 0, \tag{4.8}$$

where $\|\cdot\|_{\mathcal{O}}$ is the norm of $L^2(\mathcal{O})$.
Taking the imaginary part of (4.8) leads to

$$2\omega\mathfrak{Re}\int_{\mathcal{O}} p\boldsymbol{v_0}\cdot\nabla p^*\mathrm{d}\boldsymbol{x} - \mathfrak{Im}\int_{\partial\mathcal{O}}\mathcal{Z}|p|^2\mathrm{d}\sigma = 0,$$

and using Lemma 4.1.1, we have

$$\omega\int_{\partial\mathcal{O}}(\boldsymbol{v_0}\cdot\boldsymbol{n})|p|^2\mathrm{d}\sigma - \mathfrak{Im}\int_{\partial\mathcal{O}}\mathcal{Z}|p|^2\mathrm{d}\sigma = 0,$$

which yields

$$\int_{\partial\mathcal{O}}(\omega\boldsymbol{v_0}\cdot\boldsymbol{n} - \mathfrak{Im}\mathcal{Z})|p|^2\mathrm{d}\sigma = 0.$$

Finally we obtain that $p|_{\partial\mathcal{O}} = 0$ if $\mathfrak{Im}\mathcal{Z} \neq \omega\boldsymbol{v_0}\cdot\boldsymbol{n}$ and if the sign of $\mathfrak{Im}\mathcal{Z}$ does not change.
On $\partial\mathcal{O}$ we have

$$\nabla p = (\partial_{\boldsymbol{n}}p)\boldsymbol{n} + \nabla_{\partial\mathcal{O}}p,$$

where $\nabla_{\partial\mathcal{O}}$ is the tangential gradient on $\partial\mathcal{O}$. As $p|_{\partial\mathcal{O}} = 0$, we have $\nabla_{\partial\mathcal{O}}p = 0$ and therefore

$$\nabla p = (\partial_{\boldsymbol{n}}p)\boldsymbol{n}.$$

Owing to boundary condition (4.5), we also have

$$\boldsymbol{\sigma}\cdot\boldsymbol{n} = 0,$$

and, as $p|_{\partial\mathcal{O}} = 0$,

$$0 = (\boldsymbol{K_0}\nabla p)\cdot\boldsymbol{n} = (\partial_{\boldsymbol{n}}p)\underbrace{\boldsymbol{n}^T\boldsymbol{K_0}\boldsymbol{n}}_{>0},$$

which finally leads to $\partial_{\boldsymbol{n}}p = 0$ as $\boldsymbol{K_0}$ is positive-definite.
To prove the uniqueness of the solution we rely on the *unique continuation principle*: let P be a second-order differential operator

$$P = \sum_{i,j=1}^{n} a_{ij}\frac{\partial^2}{\partial x_i\partial x_j} + \sum_{j=1}^{n} a_j\frac{\partial}{\partial x_j} + a$$

that satisfies the ellipticity condition

$$\sum_{i,j=1}^{n} a_{ij}\xi_i\xi_j \geqslant |\boldsymbol{\xi}|^2, \quad \text{for all } \boldsymbol{\xi} = (\xi_i)_{i=1}^n,$$

and let S be an hypersurface located at x_0 then

$$\begin{cases} Pu = 0 \\ u|_S = 0 \\ \partial_{\boldsymbol{n}}u|_S = 0 \end{cases} \Longrightarrow u = 0 \text{ in a neighborhood of } x_0,$$

where $\partial_{\boldsymbol{n}}$ is the normal derivative on S. For more details, we refer to [Pro59].

■

4.1.3 Geometric assumptions for the background flow

As the absorbing boundary conditions will be constructed using the *Prandtl-Glauert-Lorentz transformation*, the background flow must be locally uniform around the artificial boundary. We therefore make the following assumption.

Assumption 8 (Uniform flow outside $\mathcal{O}$):
There exists a compact $K \subsetneq \mathcal{O}$ such that $\boldsymbol{v_0}$ and c_0 are constant outside K.

ASSUMPTION 8 is a reasonable assumption: in practice we can always assume that the background flow becomes uniform if the artificial boundary is located far away from the acoustic source. It is also worth noticing that a similar assumption is usually used to derive ABCs for the standard Helmholtz equation. A sketch of this configuration is given in FIGURE 4.1.

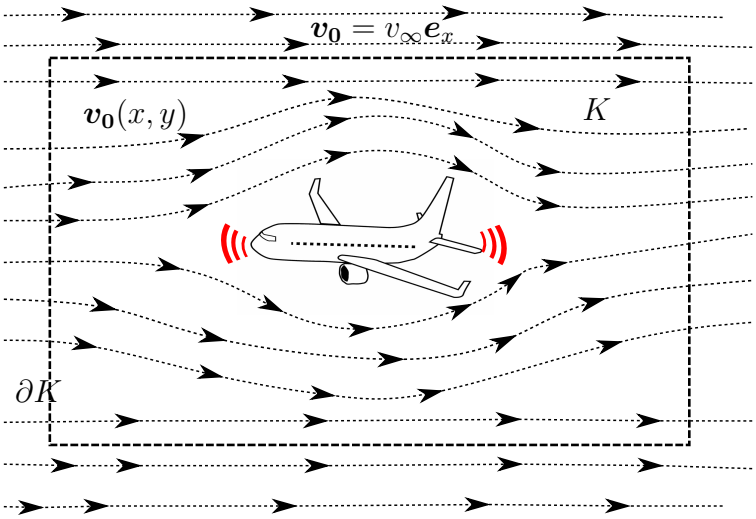


Figure 4.1: Flow around an aircraft, the flow is uniform outside of K

4.2 Prandtl-Glauert-Lorentz transformation and approximate ABCs

In this section we introduce the *Prandtl-Glauert-Lorentz transformation*, which is the main tool for the construction of the new ABCs. We recall that the *Prandtl-Glauert-Lorentz transformation* maps the convected Helmholtz equation into the standard one. As for the Helmholtz equation, the behaviour of the solution of the convected Helmholtz equation at infinity must be specified to ensure its uniqueness, leading to the notion of *outgoing solution*. In this section, we prove that the *Prandtl-Glauert-Lorentz transformation* of an outgoing solution of the standard Helmholtz equation is an outgoing solution of the convected Helmholtz equation. We then show how the boundary conditions are transformed under this change of coordinates.

4.2.1 Prandtl-Glauert-Lorentz transformation

We define the *Lorentz factor*

$$\alpha := \sqrt{1 - M^2},$$

where M is the Mach-number defined as

$$\boldsymbol{M_0} := \frac{1}{c_0}\boldsymbol{v_0}, \quad \text{and} \quad M := |\boldsymbol{M_0}|.$$

Following [HPN19], we use the following transformation

$$\widetilde{t} = \alpha t + \frac{1}{\alpha c_0}(\boldsymbol{M_0}\cdot\boldsymbol{x}) \;\; ; \;\; \widetilde{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} := \left(\mathbf{Id} + \frac{1}{\alpha(1+\alpha)}\boldsymbol{M_0}\boldsymbol{M_0}^T\right)\boldsymbol{x}. \tag{4.9}$$

This transformation called *Lorentz transformation* in [MBAG20], *Prandtl-Glauert transformation* in [Cha00] and *Prandtl-Glauert-Lorentz transformation* in [HPN19]. The term *Lorentz transformation* is used as this transformation is close to the Lorentz transformation arising in special relativity.
Using the *Sherman-Morrisson formula* [SM50], it is easy to show that

$$\boldsymbol{A}^{-1} = \mathbf{Id} - \frac{1}{1+\alpha}\boldsymbol{M_0}\boldsymbol{M_0}^T.$$

We can therefore write the inverse *Prandtl-Glauert-Lorentz transformation*

$$t = \frac{1}{\alpha}\widetilde{t} - \frac{1}{\alpha c_0}(\boldsymbol{M_0}\cdot\widetilde{\boldsymbol{x}}) \;\; ; \;\; \boldsymbol{x} = \boldsymbol{A}^{-1}\widetilde{\boldsymbol{x}} := \left(\mathbf{Id} - \frac{1}{1+\alpha}\boldsymbol{M_0}\boldsymbol{M_0}^T\right)\widetilde{\boldsymbol{x}}. \tag{4.10}$$

4.2.2 Transformation of the convected Helmholtz equation

In this subsection, we prove that the *Prandtl-Glauert-Lorentz transformation* maps the convected Helmholtz equation to the standard Helmholtz equation. The material covered here is standard and can be found *e.g.* in [HPN19]. We have added it here for the sake of completeness. We would like to point out that as the *Prandtl-Glauert-Lorentz transformation* is a spacetime transformation, it is easier to work in time-domain. Then we go back to the frequency domain thanks to a Fourier transform.

Transformation of the time-domain wave equation. Let us consider the time-domain convected wave equation

$$\left(\frac{\partial}{\partial t} + \boldsymbol{v_0}\cdot\nabla\right)^2 p - c_0^2\Delta p = s, \tag{4.11}$$

where s denotes a source term.

Lemma 4.2.1:
If c_0 and $\boldsymbol{v_0}$ are constant, the quantity

$$\widetilde{p}(\widetilde{\boldsymbol{x}},\widetilde{t}) := p(\boldsymbol{x},t), \tag{4.12}$$

satisfies the following wave equation

$$\partial^2_{\widetilde{t}\widetilde{t}}\widetilde{p} - c_0^2\widetilde{\Delta}\widetilde{p} = \widetilde{s}, \tag{4.13}$$

where $\widetilde{\Delta}$ is the Laplace operator with respect to $\widetilde{\boldsymbol{x}}$ and $(\widetilde{\boldsymbol{x}},\widetilde{t})$ is defined by (4.9).

Proof : To make the proof more concise, we only consider the case where $\boldsymbol{M_0} = M\boldsymbol{e}_x$. The other cases reduce to this one by rotation. In this case, (4.11) can be rewritten as

$$\frac{\partial^2 p}{\partial t^2} + 2c_0 M\partial^2_{tx}p - c_0^2\left(1-M^2\right)\frac{\partial^2 p}{\partial x^2} - c_0^2\frac{\partial^2 p}{\partial y^2} = s, \tag{4.14}$$

and the Lorentz factor is still given by

$$\alpha = \sqrt{1 - M^2}.$$

Using the chain rule and (4.10), we have

$$\partial^2_{\widetilde{tt}}\widetilde{p} = \frac{1}{\alpha^2}\frac{\partial^2 p}{\partial t^2}, \tag{4.15a}$$

$$\partial^2_{\widetilde{xx}}\widetilde{p} = \frac{M^2}{\alpha^2 c_0^2}\frac{\partial^2 p}{\partial t^2} - \frac{2M}{\alpha c_0}\left(1 - \frac{M^2}{1+\alpha}\right)\partial^2_{tx} p + \left(1 - \frac{M^2}{1+\alpha}\right)^2 \frac{\partial^2 p}{\partial x^2}, \tag{4.15b}$$

$$\partial^2_{\widetilde{yy}}\widetilde{p} = \frac{\partial^2 p}{\partial y^2}. \tag{4.15c}$$

We notice that

$$1 - \frac{M^2}{1+\alpha} = \frac{1 - M^2 + \alpha}{1+\alpha} = \frac{\alpha(1+\alpha)}{1+\alpha} = \alpha. \tag{4.16}$$

As (4.13) can be rewritten as

$$\partial^2_{\widetilde{tt}}\widetilde{p} - c_0^2\partial^2_{\widetilde{xx}}\widetilde{p} - c_0^2\partial^2_{\widetilde{yy}}\widetilde{p} = \widetilde{s},$$

using (4.15a)–(4.15b)–(4.15c) and (4.16) immediately leads to (4.14). ∎

Transformation of time-harmonic solutions. We will now derive a relationship between the solutions in the frequency domain.

Lemma 4.2.2:
Let p be a solution of (4.3) and let $\widetilde{p}$ be defined by (4.12). If c_0 and $\boldsymbol{v_0}$ are constant, the following identity holds

$$\mathcal{F}[\widetilde{p}](\widetilde{\boldsymbol{x}}, \widetilde{\omega}) = \alpha \exp\left[i\omega \frac{\boldsymbol{M_0}\cdot\boldsymbol{x}}{\alpha^2 c_0}\right] \mathcal{F}[p](\boldsymbol{x}, \omega), \tag{4.17}$$

where $\mathcal{F}$ denotes the Fourier transform with respect to time and the relationship between the frequencies is

$$\widetilde{\omega} = \frac{\omega}{\alpha}. \tag{4.18}$$

The same identity holds for the source-term s, *i.e.*

$$\mathcal{F}[\widetilde{s}](\widetilde{\boldsymbol{x}}, \widetilde{\omega}) = \alpha \exp\left[i\omega \frac{\boldsymbol{M_0}\cdot\boldsymbol{x}}{\alpha^2 c_0}\right] \mathcal{F}[s](\boldsymbol{x}, \omega).$$

Proof : As discussed in (4.2) we chose the $e^{-i\omega t}$ convention for time-harmonic solutions, we shall therefore use the following convention for the Fourier transform with respect to time

$$\mathcal{F}[p](\boldsymbol{x}, \omega) := \int_{\mathbb{R}} p(\boldsymbol{x}, t) e^{i\omega t} \mathrm{d}t.$$

Let us begin with the solution in Lorentz coordinates, the frequency domain solution is given by

$$\mathcal{F}[\widetilde{p}](\widetilde{\boldsymbol{x}}, \widetilde{\omega}) = \int_{\mathbb{R}} \widetilde{p}(\widetilde{\boldsymbol{x}}, \widetilde{t}) e^{i\widetilde{\omega}\widetilde{t}} \mathrm{d}\widetilde{t},$$

and (4.9) implies

$$= \int_{\mathbb{R}} \widetilde{p}(\widetilde{\boldsymbol{x}}, \widetilde{t}) e^{i\widetilde{\omega}\alpha t} \exp\left(i\widetilde{\omega}\frac{\boldsymbol{M_0}\cdot\boldsymbol{x}}{\alpha c_0}\right) \alpha \mathrm{d}t,$$

and we have

$$= \alpha \exp\left(i\widetilde{\omega}\frac{\boldsymbol{M_0}\cdot\boldsymbol{x}}{\alpha c_0}\right) \int_{\mathbb{R}} p(\boldsymbol{x}, t) e^{i\alpha\widetilde{\omega}t} \mathrm{d}t.$$

Finally, we get that

$$\mathcal{F}[\widetilde{p}](\widetilde{\boldsymbol{x}}, \widetilde{\omega}) = \alpha \exp\left(i\widetilde{\omega}\frac{\boldsymbol{M_0}\cdot\boldsymbol{x}}{\alpha c_0}\right) \mathcal{F}[p](\boldsymbol{x}, \alpha\widetilde{\omega}), \tag{4.19}$$

Taking $\omega = \alpha\widetilde{\omega}$ in (4.19) leads to (4.17) and (4.18). ∎

To summarize, when c_0 and $\boldsymbol{v_0}$ are constant, then Lemma 4.2.1 and Lemma 4.2.2 imply that the *Prandtl-Glauert-Lorentz transformation* maps the convected Helmholtz equation (4.3) into the following Helmholtz equation

$$-\left(\frac{\omega}{\alpha}\right)^2 \widetilde{p} - c_0^2 \widetilde{\Delta}\widetilde{p} = \widetilde{s}.$$

4.2.3 Transformation of the boundary condition

We will now focus on how the boundary conditions are changed under the *Prandtl-Glauert-Lorentz transformation*.
We denote by Σ the artificial boundary in physical coordinates and $\widetilde{\Sigma}$ the artificial boundary in Lorentz coordinates. We will choose a circle centered on the origin for $\widetilde{\Sigma}$, Σ will therefore be an ellipse. An example is depicted on Figure 4.2.

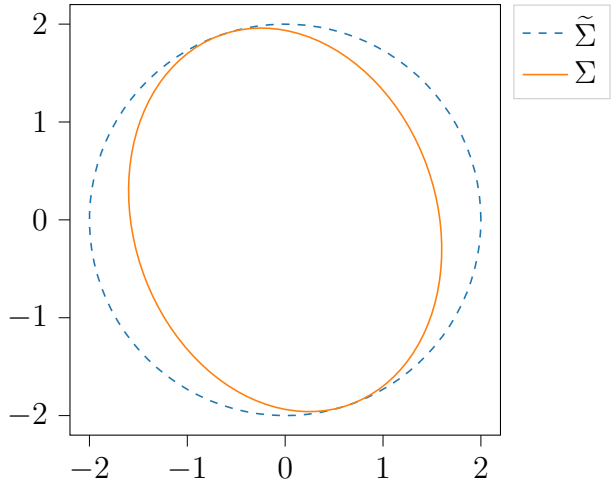


Figure 4.2: Σ and $\widetilde{\Sigma}$ for $\boldsymbol{M_0} = [0.6\ 0.2]^T$ and $R = 2$

More precisely, the following lemma holds.

Lemma 4.2.3:
Let $R > 0$. If the artificial boundary in Lorentz coordinates satisfies

$$\widetilde{\Sigma} = \left\{ \widetilde{\boldsymbol{x}} \;\middle|\; |\widetilde{\boldsymbol{x}}|^2 = R^2 \right\} = \left\{ \widetilde{\boldsymbol{x}} = R\left(\cos(t)\boldsymbol{e}_x + \sin(t)\boldsymbol{e}_y\right) \;\middle|\; t \in \mathbb{S}^1 \right\},$$

then the artificial boundary in physical coordinates satisfies

$$\Sigma = \left\{ \boldsymbol{x} \;\middle|\; |\boldsymbol{A}\boldsymbol{x}| = R^2 \right\} = \left\{ \boldsymbol{x} = R\left(\cos(t)\boldsymbol{A}_1 + \sin(t)\boldsymbol{A}_2\right) \;\middle|\; t \in \mathbb{S}^1 \right\},$$

where $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ are the first and second columns of $\boldsymbol{A}^{-1}$.

Remark 4.2.1: At first glance, the other choice (taking a circle in physical coordinates and an ellipse in Lorentz coordinates) may seem more natural. However ABCs for the Helmholtz equation on an ellipsoid are more complicated to implement. If they are derived using an analytical point of view, those ABCs are expressed using special functions whose numerical evaluation may be unstable, see [BST12, Sai08]. On the other hand, if a geometric point of view is used, the resulting ABCs involve curvature operators that are difficult to approximate, see [BT80, ABB99].

Transformation of the impedance boundary condition. We will now state and prove the main result of this chapter.

Theorem 9 :

Let p be the solution of

$$-\omega^2 p - \operatorname{div}\left(\boldsymbol{K_0}\nabla p + 2i\omega\boldsymbol{v_0}\right) = s, \qquad \text{in } \mathcal{O}, \tag{4.20a}$$
$$\boldsymbol{\sigma}\cdot\boldsymbol{n} + \mathcal{Z}p = 0, \qquad \text{on } \Sigma, \tag{4.20b}$$

then if $\boldsymbol{v_0}$ and c_0 are constant, $\widetilde{p}$ defined by (4.17) satisfies

$$-\widetilde{\omega}^2\widetilde{p} - c_0^2\widetilde{\Delta}\widetilde{p} = \widetilde{s}, \qquad \text{in } \widetilde{\mathcal{O}}, \tag{4.21a}$$
$$\partial_{\widetilde{\boldsymbol{n}}}\widetilde{p} + \widetilde{\mathcal{Z}}\widetilde{p} = 0, \qquad \text{on } \widetilde{\Sigma}, \tag{4.21b}$$

if the impedance operators satisfy

$$\mathcal{Z}(\boldsymbol{x},\omega) = -\frac{c_0^2}{\mu}\widetilde{\mathcal{Z}}(\widetilde{\boldsymbol{x}},\widetilde{\omega}) + i\omega\boldsymbol{v_0}\cdot\boldsymbol{n}, \tag{4.22}$$

and $\mu := |\boldsymbol{A}^{-T}\boldsymbol{n}|^{-1}$ is a scaling factor.

Remark 4.2.2: In (4.22) we can see that the impedance-like coefficient for the convected Helmholtz equation is the impedance coefficient of the standard Helmholtz equation scaled and shifted with a correction term that takes convection into account.

Proof : The equivalence of the volumetric parts (4.20a) and (4.21a) is a direct consequence of Lemma 4.2.1 and Lemma 4.2.2. Let us now focus on the boundary conditions (4.20b) and (4.21b). This is adapted from [HPN19, Sec. 3].

We assume that Σ and $\widetilde{\Sigma}$ have the following cartesian equations

$$\Sigma : \{\Phi(\boldsymbol{x}) = 0\},$$
$$\widetilde{\Sigma} : \left\{\widetilde{\Phi}(\widetilde{\boldsymbol{x}}) = 0\right\}.$$

Step 1:
Because of the chain rule, we have the following relationship between their normal directions

$$\nabla\Phi = \frac{\partial\Phi}{\partial\boldsymbol{x}} = \frac{\partial\widetilde{\boldsymbol{x}}}{\partial\boldsymbol{x}}\frac{\partial\widetilde{\Phi}}{\partial\widetilde{\boldsymbol{x}}} = \boldsymbol{A}^T\widetilde{\nabla}\widetilde{\Phi} \tag{4.23}$$

Let $\boldsymbol{n}$ and $\widetilde{\boldsymbol{n}}$ be the unit normal vectors to Σ and $\widetilde{\Sigma}$ respectively. We have

$$\widetilde{\boldsymbol{n}} = \frac{1}{|\widetilde{\nabla}\widetilde{\Phi}|}\widetilde{\nabla}\widetilde{\Phi},$$

which becomes thanks to (4.23)

$$= \frac{1}{|\boldsymbol{A}^{-T}\nabla\Phi|}\boldsymbol{A}^{-T}\nabla\Phi,$$

then, since $\nabla\Phi = |\nabla\Phi|\boldsymbol{n}$, we have

$$= \frac{|\nabla\Phi|}{|\boldsymbol{A}^{-T}\nabla\Phi|}\boldsymbol{A}^{-T}\boldsymbol{n},$$

which implies

$$\widetilde{\boldsymbol{n}} = \mu\boldsymbol{A}^{-T}\boldsymbol{n},$$

where μ is a scaling factor that ensures that $|\widetilde{\boldsymbol{n}}| = 1$.
Step 2:
We have

$$\boldsymbol{A}^{-1}\boldsymbol{M_0} = \left(\mathbf{Id} - \frac{1}{1+\alpha}\boldsymbol{M_0}\boldsymbol{M_0}^T\right)\boldsymbol{M_0} = \left(1 - \frac{|\boldsymbol{M_0}|^2}{1+\alpha}\right)\boldsymbol{M_0} = \alpha\boldsymbol{M_0}.$$

Starting from [HPN19, Eq. (11)], multiplying by $\boldsymbol{A}^{-1}$ on the left and by $\boldsymbol{A}^{-T}$ on the right, we have

$$c_0^2\boldsymbol{M_0}\boldsymbol{M_0}^T - c_0^2\mathbf{Id} = -c_0^2\boldsymbol{A}^{-1}\boldsymbol{A}^{-T} \iff \boldsymbol{A}^{-1}\boldsymbol{A}^{-T} = \mathbf{Id} - \boldsymbol{M_0}\boldsymbol{M_0}^T.$$

Step 3:
We recall that

$$\widetilde{p}(\widetilde{\boldsymbol{x}}, \widetilde{\omega}) = \alpha\alpha\exp\left[\frac{i\omega}{\alpha^2 c_0}\boldsymbol{M_0}\cdot\boldsymbol{x}\right]p(\boldsymbol{x}, \omega).$$

Using the chain rule, we have

$$\widetilde{\nabla}p = \boldsymbol{A}^{-T}\nabla p, \quad \text{and} \quad \widetilde{\nabla}\left[\alpha\exp\left[\frac{i\omega}{\alpha^2 c_0}\boldsymbol{M_0}\cdot\boldsymbol{x}\right]\right] = \frac{i\omega}{\alpha^2 c_0}\alpha\exp\left[\frac{i\omega}{\alpha^2 c_0}\boldsymbol{M_0}\cdot\boldsymbol{x}\right]\boldsymbol{A}^{-T}\boldsymbol{M_0},$$

and therefore

$$\begin{aligned}
\partial_{\widetilde{\boldsymbol{n}}}\widetilde{p} &= \widetilde{\boldsymbol{n}}\cdot\widetilde{\nabla}\widetilde{p} \\
&= \alpha\exp\left[\frac{i\omega}{\alpha^2 c_0}\boldsymbol{M_0}\cdot\boldsymbol{x}\right]\widetilde{\boldsymbol{n}}^T\left(\boldsymbol{A}^{-T}\nabla p + \frac{i\omega}{\alpha^2 c_0}p\boldsymbol{A}^{-T}\boldsymbol{M_0}\right) \\
&= \alpha\exp\left[\frac{i\omega}{\alpha^2 c_0}\boldsymbol{M_0}\cdot\boldsymbol{x}\right]\mu\boldsymbol{n}^T\boldsymbol{A}^{-1}\left(\boldsymbol{A}^{-T}\nabla p + \frac{i\omega}{\alpha c_0}p\boldsymbol{M_0}\right) \\
&= \alpha\exp\left[\frac{i\omega}{\alpha^2 c_0}\boldsymbol{M_0}\cdot\boldsymbol{x}\right]\mu\boldsymbol{n}^T\left((\mathbf{Id}-\boldsymbol{M_0}\boldsymbol{M_0}^T)\nabla p + i\frac{\omega}{c_0}p\boldsymbol{M_0}\right) \\
&= \alpha\exp\left[\frac{i\omega}{\alpha^2 c_0}\boldsymbol{M_0}\cdot\boldsymbol{x}\right]\frac{\mu}{c_0^2}\boldsymbol{n}^T\left(\underbrace{\boldsymbol{K_0}\nabla p + 2i\omega p\boldsymbol{v_0}}_{=-\boldsymbol{\sigma}} - i\omega p\boldsymbol{v_0}\right) \\
&= \alpha\exp\left[\frac{i\omega}{\alpha^2 c_0}\boldsymbol{M_0}\cdot\boldsymbol{x}\right]\frac{\mu}{c_0^2}\left(\mathcal{Z}p - i\omega p\boldsymbol{v_0}\cdot\boldsymbol{n}\right)
\end{aligned}$$

Using (4.22) and

$$p(\boldsymbol{x},\omega) = \frac{1}{\alpha}\exp\left[\frac{-i\omega}{\alpha^2 c_0}\boldsymbol{M_0}\cdot\boldsymbol{x}\right]\widetilde{p}(\widetilde{\boldsymbol{x}},\widetilde{\omega}),$$

we have

$$\partial_{\widetilde{\boldsymbol{n}}}\widetilde{p} = -\widetilde{\mathcal{Z}}\widetilde{p},$$

which concludes the proof. ∎

Remark 4.2.3: The minus sign in (4.22) comes from the fact that we have chosen to work with $\boldsymbol{\sigma}\cdot\boldsymbol{n}$ which should be interpreted as $-\partial_{\boldsymbol{n}}p$. As described in CHAPTER 3, this choice of unknown is convenient for constructing mixed finite element methods.

Alternative flux. Up to this point, we have worked with the *total flux* $\boldsymbol{\sigma}$ which is a convenient quantity to work with mixed finite-element methods. However when working with continuous primal finite-element methods, it is possible to obtain a symmetric formulation, which would lead to an alternative boundary flux [HPN17, HPN19, LMG+20]. This flux is

$$\boldsymbol{g} := -c_0^2\nabla p + (-i\omega p + \boldsymbol{v_0}\cdot\nabla p)\boldsymbol{v_0} = -\boldsymbol{K_0}\nabla p - i\omega p\boldsymbol{v_0}.$$

This alternative flux is related to $\boldsymbol{\sigma}$ by

$$\boldsymbol{g} = \boldsymbol{\sigma} + i\omega p\boldsymbol{v_0},$$

and we can define the following ABC

$$\boldsymbol{g}\cdot\boldsymbol{n} + \mathcal{Z}_{\boldsymbol{g}}p = 0.$$

Corollary 4.2.1:
The impedance operator $\mathcal{Z}_{\boldsymbol{g}}$ satisfies

$$\mathcal{Z}_{\boldsymbol{g}}(\boldsymbol{x},\omega) = -\frac{c_0^2}{\mu}\widetilde{\mathcal{Z}}(\widetilde{\boldsymbol{x}},\widetilde{\omega}).$$

The ABCs derived here can therefore be easily adapted to other finite-element formulations.

4.2.4 Outgoing solutions of the convected Helmholtz equation

The existence of outgoing solutions for an equation is the key condition for obtaining absorbing conditions that can give well-posed boundary problems. In this section, we will prove that the notion of *outgoing solution* for the convected Helmholtz equation can be inherited from the one for the standard Helmholtz equation through the *Prandtl-Glauert-Lorentz transformation*.

Lemma 4.2.4:
The *Prandtl-Glauert-Lorentz transformation* maps outgoing solutions of the convected Helmholtz equation to outgoing solutions of the Helmholtz equation. More precisely, the following diagram is commutative

$$\begin{array}{ccc} p_{\text{conv}}(\boldsymbol{x},t) & \xrightarrow{t\to\infty} & p_{\text{conv},\omega}(\boldsymbol{x}) \\ \Big\downarrow PGL_t & & PGL_\omega^{-1}\Big\uparrow \\ \widetilde{p}_{\text{std}}(\widetilde{\boldsymbol{x}},\widetilde{t}) & \xrightarrow{\widetilde{t}\to\infty} & \widetilde{p}_{\text{std},\widetilde{\omega}}(\widetilde{\boldsymbol{x}}) \end{array}$$

where p_{conv} and $\widetilde{p}_{\text{std}}$ are the time-domain solutions of the convected and standard wave equations, and $p_{\text{conv},\omega}$ and $\widetilde{p}_{\text{std},\widetilde{\omega}}$ are outgoing solutions of the convected and standard Helmholtz equations. In particular, we have

$$\widetilde{p}_{\text{std},\widetilde{\omega}}(\widetilde{\boldsymbol{x}}) = \alpha \exp\left[i\omega \frac{\boldsymbol{M_0}\cdot\boldsymbol{x}}{\alpha^2 c_0}\right] p_{\text{conv},\omega}(\boldsymbol{x}).$$

This result is proven by using the *limiting amplitude principle* which consists in studying the long-term behaviour of the time-domain solution with the following source term

$$g(\boldsymbol{x},t) = \begin{cases} s(\boldsymbol{x})e^{-i\omega t}, & \text{if } t > 0 \\ 0, & \text{if } t \leqslant 0 \end{cases},$$

where s is a function in $L^2(\mathcal{O})$ with compact support. As the proof of this lemma is rather long and technical, it is given in Appendix 4.A.

4.2.5 New ABCs for the convected Helmholtz equation in 2D

In this section, we derive the practical values used for the impedance-like operators for the convected Helmholtz equation from Absorbing Boundary Conditions formerly derived for the Helmholtz equation.

Lorentz ABCs

In this section, we recall some absorbing boundary conditions on a circle for the Helmholtz equation. The ABCs of this section are specific to the 2D setting, however the extension to 3D is straightforward as the result of Theorem 9 holds for any dimension. Notice that even if Theorem 9 is stated for a uniform flow, the new Absorbing Boundary Conditions can be used in the more general case described by Assumption 8. Indeed as the *Prandtl-Glauert-Lorentz transformation* will only be used to construct the impedance operator, the carrier flow only needs to be locally uniform close to the artificial boundary, which we can always assume if the boundary is located far away from the source. Those conditions will define the $\widetilde{\mathcal{Z}}$ operator applied on $\widetilde{\Sigma}$. The value of $\mathcal{Z}$ will then be obtained using the transformation of

the time-harmonic solutions described in LEMMA 4.2.2 and the transformation of impedance-like coefficients described in THEOREM 9.
We will consider the Engquist-Madja absorbing boundary conditions, see [EM77], in the Lorentz variable.
We assume that the artificial boundary in Lorentz coordinates $\widetilde{\Sigma}$ is a circle of radius R.

Proposition 4.2.1:
The ABCs constructed as the *Prandtl-Glauert-Lorentz transformation* of the Engquist-Madja ABCs are

- ABC of order 0:

$$\boldsymbol{\sigma}\cdot\boldsymbol{n}+\mathcal{Z}p=0,\quad \text{with } \mathcal{Z}:=i\left(\frac{c_0}{\alpha\mu}+\boldsymbol{v_0}\cdot\boldsymbol{n}\right)\omega, \tag{ABC0}$$

- ABC of order 1:

$$\boldsymbol{\sigma}\cdot\boldsymbol{n}+\mathcal{Z}p=0,\quad \text{with } \mathcal{Z}:=\frac{-c_0^2}{2\mu R}+i\left(\frac{c_0}{\alpha\mu}+\boldsymbol{v_0}\cdot\boldsymbol{n}\right)\omega. \tag{ABC1}$$

As a comparison point, the ABC that selects the outgoing planewaves that are locally orthogonal to the boundary is

$$\boldsymbol{\sigma}\cdot\boldsymbol{n}+\mathcal{Z}p=0,\quad \text{with } \mathcal{Z}:=i(c_0+\boldsymbol{v_0}\cdot\boldsymbol{n})\omega. \tag{ABC-PW}$$

For all of those three ABCs, ASSUMPTION 7 holds and (ABC0), (ABC1) and (ABC-PW) all lead to a well-posed problem.

ABC of order 0 The zeroth order ABC reads

$$\left(\partial_{\widetilde{\boldsymbol{n}}}-i\frac{\widetilde{\omega}}{c_0}\right)\widetilde{p}=0.$$

The impedance-like operators are therefore defined by

$$\widetilde{\mathcal{Z}}:=-i\frac{\widetilde{\omega}}{c_0}\quad\text{and}\quad\mathcal{Z}:=i\left(\frac{c_0}{\alpha\mu}+\boldsymbol{v_0}\cdot\boldsymbol{n}\right)\omega. \tag{4.24}$$

It is now clear that ASSUMPTION 7 holds, so the ABC (4.24) leads to a well-posed problem. In the remaining of this chapter, this ABC will be called ABC0.

ABC of order 1 The first order ABC reads

$$\left(\partial_{\widetilde{\boldsymbol{n}}}-i\frac{\widetilde{\omega}}{c_0}+\frac{1}{2R}\right)\widetilde{p}=0.$$

The impedance-like operators are therefore defined by

$$\widetilde{\mathcal{Z}}:=-i\frac{\widetilde{\omega}}{c_0}+\frac{1}{2R}\quad\text{and}\quad\mathcal{Z}:=\frac{-c_0^2}{2\mu R}+i\left(\frac{c_0}{\alpha\mu}+\boldsymbol{v_0}\cdot\boldsymbol{n}\right)\omega. \tag{4.25}$$

Once again it is clear that ASSUMPTION 7 holds, so the ABC (4.25) also leads to a well-posed problem. In the remaining of this chapter, this ABC will be called ABC1.

Plane-wave ABC

To dispose of comparison elements, we will derive an absorbing boundary condition that selects outgoing plane-waves that are locally orthogonal to the artificial boundary Σ. In the remaining of this chapter, it will be called ABC-PW.
We begin by computing the plane waves for the convected Helmholtz equation (4.1). We write p as

$$p_{PW} = p_0 e^{i\boldsymbol{\kappa}\cdot\boldsymbol{x}}$$

where $p_0 \in \mathbb{C}$ and where $\boldsymbol{\kappa}$ is the wave vector.
We can rewrite (4.1) as

$$p_0 \left[-\omega^2 + 2\omega \boldsymbol{v_0}\cdot\boldsymbol{\kappa} + \boldsymbol{\kappa}^T \boldsymbol{K_0}\boldsymbol{\kappa}\right] = 0,$$

and therefore

$$-\rho_0\omega^2 + 2\omega\rho_0 \boldsymbol{v_0}\cdot\boldsymbol{\kappa} + \rho_0 c_0^2|\boldsymbol{\kappa}|^2 - \rho_0(\boldsymbol{v_0}\cdot\boldsymbol{\kappa})^2 = 0. \tag{4.26}$$

As we only have one equation for two unknowns, we make the following assumptions

$$\boldsymbol{\kappa} := \kappa \begin{bmatrix}\cos\theta \\ \sin\theta\end{bmatrix}, \qquad \text{where } \theta \in [0, 2\pi) \text{ is assumed to be known,}$$
$$\boldsymbol{v_0} := Mc_0 \begin{bmatrix}\cos\theta_0 \\ \sin\theta_0\end{bmatrix}, \qquad \text{where } \theta_0 \in [0, 2\pi).$$

With those assumptions, (4.26) becomes

$$-\omega^2 + 2\omega\kappa M c_0 \left(\cos\theta_0\cos\theta + \sin\theta_0\sin\theta\right) + c_0^2\kappa^2 - M^2c_0^2\kappa^2\left(\cos\theta_0\cos\theta + \sin\theta_0\sin\theta\right)^2 = 0,$$

or equivalently

$$-\omega^2 + 2\omega\kappa M c_0\cos(\theta-\theta_0) + c_0^2\kappa^2\left(1 - M^2\cos^2(\theta-\theta_0)\right) = 0.$$

Solving this last equation for κ leads to the two following solutions

$$\kappa_- = \frac{-\omega}{c_0(1 - M\cos(\theta-\theta_0))} \quad \text{and} \quad \kappa_+ = \frac{\omega}{c_0(1 + M\cos(\theta-\theta_0))}.$$

Notice that $\boldsymbol{\kappa}$ can be rewritten as

$$\boldsymbol{\kappa} = \frac{\kappa_\pm}{Mc_0}\left(\cos(\theta-\theta_0)\boldsymbol{v_0} + \sin(\theta-\theta_0)\boldsymbol{v_0^\perp}\right).$$

The plane-wave solutions of (4.3) are therefore

$$p_{PW} = p_0 e^{i\boldsymbol{\kappa}_\pm\cdot\boldsymbol{x}}, \quad \text{with} \quad \kappa_\pm = \frac{\pm\omega}{c_0(1 \pm M\cos(\theta-\theta_0))},$$

and the total flux $\boldsymbol{\sigma}_{PW}$ therefore satisfies

$$\boldsymbol{\sigma}_{PW} = -i\frac{c_0\kappa_\pm}{M}\left((1-M^2)\cos(\theta-\theta_0)\boldsymbol{v_0} + \sin(\theta-\theta_0)\boldsymbol{v_0^\perp}\right)p_0 e^{i\boldsymbol{\kappa}\cdot\boldsymbol{x}}.$$

As we want to select plane-waves that are locally orthogonal to the artificial boundary Σ, we have

$$\boldsymbol{n} = \frac{\boldsymbol{\kappa}_\pm}{|\boldsymbol{\kappa}_\pm|} = \begin{bmatrix}\cos\theta \\ \sin\theta\end{bmatrix},$$

leading to

$$\boldsymbol{v_0}\cdot\boldsymbol{n} = Mc_0\cos(\theta-\theta_0), \quad \text{and} \quad \boldsymbol{v_0^\perp}\cdot\boldsymbol{n} = Mc_0\sin(\theta-\theta_0).$$

We then have

$$\boldsymbol{\sigma}\cdot\boldsymbol{n} = -i\omega c_0\left(1 \pm M\cos(\theta-\theta_0)\right)p = -i\omega\left(c_0 \pm \boldsymbol{v_0}\cdot\boldsymbol{n}\right)p.$$

We can now define the impedance-like operator $\mathcal{Z}$ as

$$\mathcal{Z} = \frac{-\boldsymbol{\sigma}_{PW}\cdot\boldsymbol{n}}{p_{PW}} = i\omega(c_0 + \boldsymbol{v_0}\cdot\boldsymbol{n}),$$

which leads to (ABC-PW)

$$\boldsymbol{\sigma}\cdot\boldsymbol{n} + i\omega(c_0 + \boldsymbol{v_0}\cdot\boldsymbol{n})p = 0.$$

We have chosen the ABC with the + sign as a consequence of the convention $e^{-i\omega t}$ that we use for time-harmonic solutions .
Once again Assumption 7 holds, so using this ABC leads to a well-posed problem.

4.3 Numerical experiments

The numerical experiments of this section will be performed using the HDG-$\boldsymbol{\sigma}$ method of Chapter 3, which uses the following *first-order in space and second-order in frequency* formulation:

$$\begin{aligned}\boldsymbol{\sigma} + \boldsymbol{K_0}\nabla p + 2i\omega p\boldsymbol{v_0} &= 0,\\ -\omega^2 p + \operatorname{div}(\boldsymbol{\sigma}) &= s.\end{aligned}$$

As the *total flux* $\boldsymbol{\sigma}$ is an unknown of the method, it is very natural to work with the absorbing boundary condition (4.5).
This method has been implemented in the `hawen` solver, see [Fau21].

4.3.1 Experiments with a uniform flow

We consider a point-source in a uniform flow

$$\boldsymbol{M_0} := M\begin{bmatrix}\cos\theta_0\\ \sin\theta_0\end{bmatrix}, \quad s = \delta_{\mathbf{0}}, \quad c_0 = 1, \quad \omega = 6\pi.$$

Using Lemma 4.2.2 and standard theory on the Helmholtz equation we can express the reference solution as

$$\widetilde{p}_{\mathrm{ref}}(\widetilde{\boldsymbol{x}},\widetilde{\omega}) = \frac{i}{4}H_0^{(1)}(\widetilde{\omega}\widetilde{r}),$$

which leads to

$$p_{\mathrm{ref}}(\boldsymbol{x},\omega) = \frac{i}{4\alpha}H_0^{(1)}\left(\frac{\omega}{\alpha}|\boldsymbol{A}\boldsymbol{x}|\right)\exp\left(-\frac{i\omega}{\alpha^2 c_0}\boldsymbol{M_0}\cdot\boldsymbol{x}\right). \tag{4.27}$$

Unless stated otherwise, we will use $\theta_0 = \frac{\pi}{4}$ in this section as it leads to a circular artificial boundary in both physical and PGL coordinates. The effect of θ_0 will only be illustrated in the end of this section.

Validation of the ABCs.

We define the relative error as

$$\mathcal{E}_{\mathcal{O}} := \sqrt{\frac{\sum_{K,i}|\mathfrak{Re}(p_h - p_{\mathrm{ref}})(\boldsymbol{x}_i^K)|^2}{\sum_{K,i}|\mathfrak{Re}(p_{\mathrm{ref}})(\boldsymbol{x}_i^K)|^2}},$$

where p_h is the numerical solution and $(\boldsymbol{x}_i^K)_i$ are the degrees of freedom in element $K\in\mathcal{T}_h$. As the solution p_{ref} is singular at $\boldsymbol{x}=(0,0)$, the error will be computed on $\mathcal{O}\backslash B(\mathbf{0},\rho)$ where $B(\mathbf{0},\rho)$ is the open ball centered on $\boldsymbol{x}=\mathbf{0}$ with radius $\rho = 2h$.

Low Mach number. The relative error $\mathcal{E}_{\mathcal{O}}$ for various values of R is given in TABLE 4.1 for a low Mach number $M = 0.4$. In this case, we can see that the three ABCs perform well, even if the two PGL-based ABCs ((ABC0) and (ABC1)) perform better than (ABC-PW). As expected, (ABC1) performs better than (ABC0) and for $M = 0.4$, we can see that a relative error of $\sim 0.1\%$ is obtained even with $R = 0.5$.

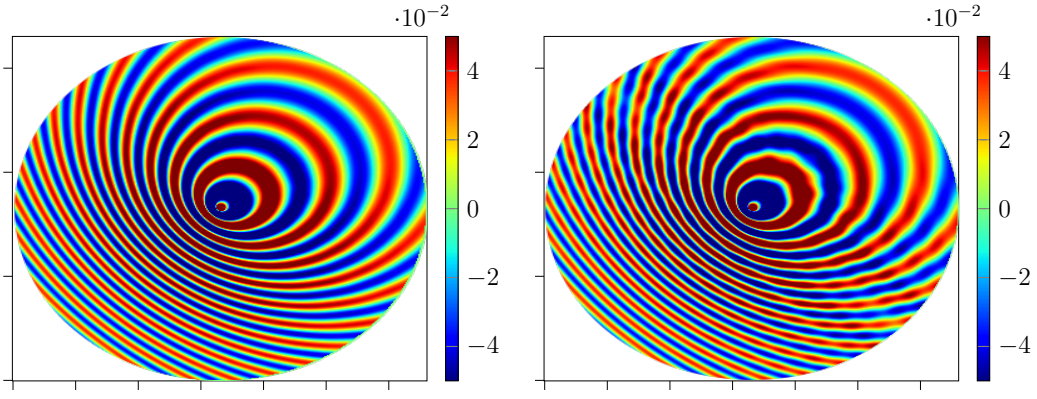
R	(ABC0)	(ABC1)	(ABC-PW)
0.5	3.49%	0.11%	4.69%
1.0	1.73%	0.15%	4.57%
1.5	1.16%	0.15%	3.89%
2.0	0.86%	0.14%	3.56%

Table 4.1: Relative error $\mathcal{E}_{\mathcal{O}}$ in the domain for $M = 0.4$

Intermediate Mach number. In TABLE 4.2, the relative error $\mathcal{E}_{\mathcal{O}}$ is computed for several values of R for $M = 0.6$. In this case (ABC-PW) performs badly, leading to an error level of $\sim 10\%$, whereas the PGL-based ABCs give good numerical results with an error level of $\sim 1\%$. This is also illustrated in FIGURE 4.3. Once again (ABC1) performs better than (ABC0). Indeed (ABC1) leads to an error level below 1% even for very small values of R whereas a larger R is required to obtain a similar error level with (ABC0).

R	(ABC0)	(ABC1)	(ABC-PW)
0.5	3.20%	0.91%	10.14%
1.0	1.58%	0.82%	9.02%
1.5	1.17%	0.83%	9.60%
2.0	0.98%	0.75%	8.23%

Table 4.2: Relative error $\mathcal{E}_{\mathcal{O}}$ in the domain for $M = 0.6$



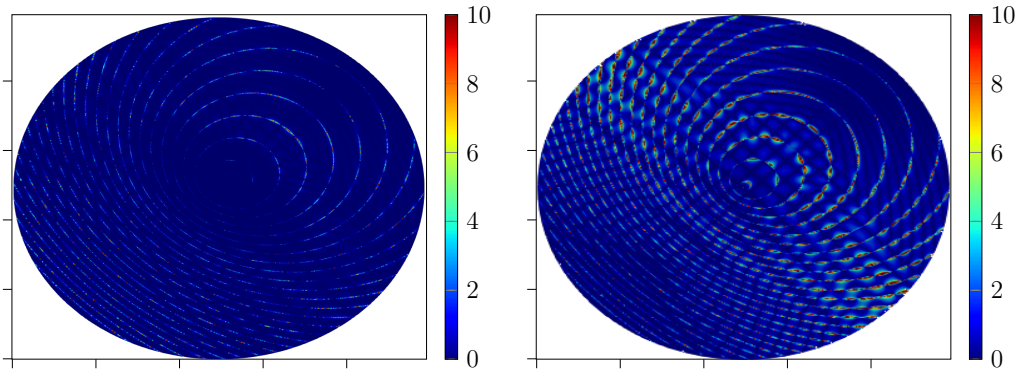
(a) (ABC1) (b) (ABC-PW)

Figure 4.3: Comparison between two ABCs for $M = 0.6$ and $R = 2$

To understand the distribution of the error in the domain, we also consider the error in decibel (dB)

$$\mathcal{E}_{\text{dB}} := 20 \log_{10} \left| \frac{p_h}{p_{\text{ref}} + 10^{-12}} \right|.$$

Notice that we have added absolute values and a small term in the denominator to avoid invalid values due to floating-point arithmetic inside the logarithm.



(a) (ABC1) (b) (ABC-PW)

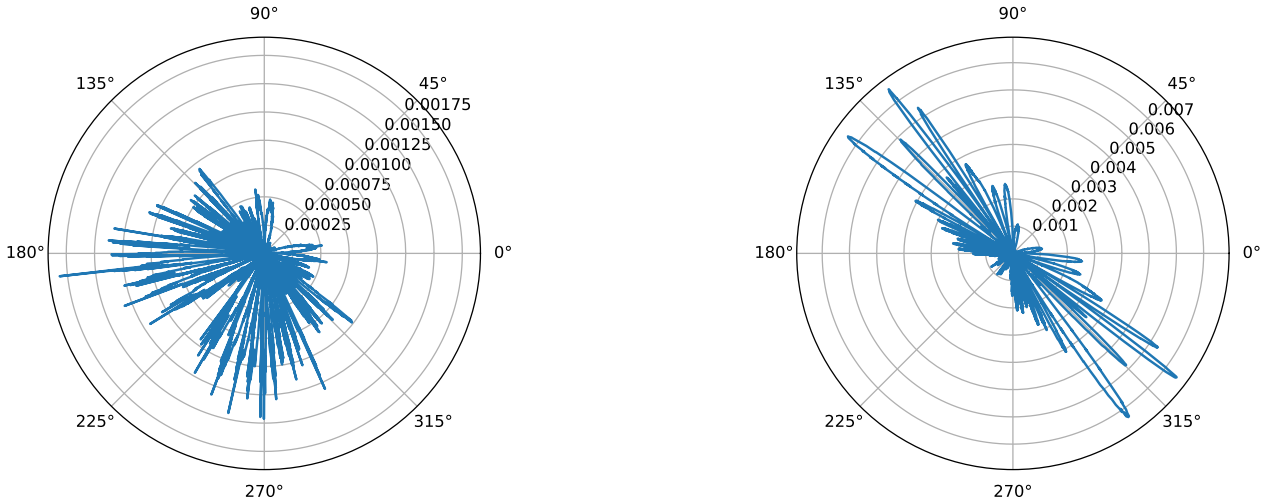
Figure 4.4: Error in dB $|\mathcal{E}_{\text{dB}}|$ in the domain for $M = 0.6$ and $R = 2$

In FIGURE 4.4, the error in decibel in plotted for (ABC1) and (ABC-PW) for $M = 0.6$ and $R = 2$. For (ABC1), we cannot really see a pattern for the error distribution. For (ABC-PW) however there is a clear pattern: the ABC performs better in the top-right and bottom-left parts of the domain. This can be understood as this ABC is constructed to select outgoing plane-waves that are locally orthogonal to the boundary. By looking at FIGURE 4.3, we can see that p is almost orthogonal to the boundary in the top-right and bottom-left parts of the domain where (ABC-PW) performs well. However due to the presence of convection, p is clearly not orthogonal to the boundary in the top-left and bottom-right parts of the domain, where ABC-PW exhibits its worst behavior.

To further illustrate this idea, we also consider the *local error* on the artificial boundary Σ

$$\mathcal{E}_\Sigma := |\mathfrak{Re}(p_h - p_{\text{ref}})|_\Sigma| \,.$$

This error is depicted in FIGURE 4.5 for (ABC1) and (ABC-PW). For (ABC-PW), the same effect as in FIGURE 4.4 can be seen. For (ABC1) however, we notice that most of the error is located in bottom part of the domain. The effects of the flow angle θ_0 on this error distribution will be discussed in TABLE 4.4 at the end of this section. We would like to point out that the error levels for (ABC1) are one order of magnitude lower than those of (ABC-PW) as expected.



(a) (ABC1) (b) (ABC-PW)

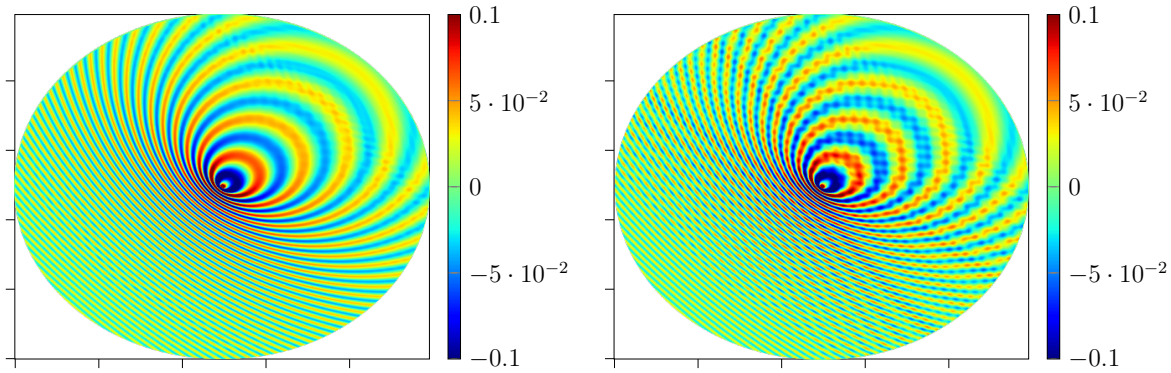
Figure 4.5: Local error $\mathcal{E}_\Sigma$ on Σ for $M = 0.6$ and $R = 2$

Large Mach number. For a large Mach number, the use of the PGL-based ABCs leads to a higher error. It therefore seems natural to consider larger domain, but as it can be seen in TABLE 4.3, even for $R = 10$ the error does not get below 2% with (ABC1). It also seems that in this case, using (ABC1) instead of (ABC0) does not improve the quality of the solution.

R	(ABC0)	(ABC1)	(ABC-PW)
3	2.71%	2.69%	16.92%
10	2.06%	2.05%	14.87%

Table 4.3: Relative error $\mathcal{E}_\mathcal{O}$ in the domain for $M = 0.8$

In FIGURE 4.6 we have plotted the numerical solution p_h obtained with (ABC1) and (ABC-PW) for the smaller domain. Even if we can see some artifacts in the bottom-left part of the domain with (ABC1), the quality of the result is clearly far superior than with (ABC-PW).

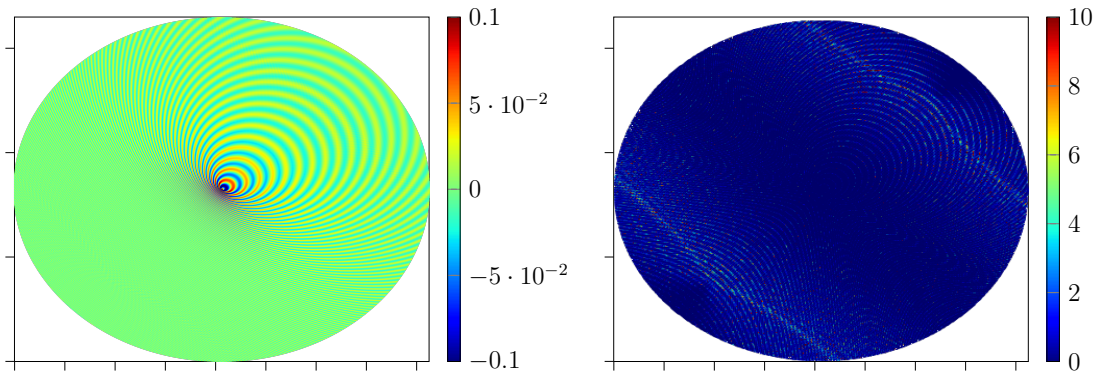


(a) (ABC1) (b) (ABC-PW)

Figure 4.6: Comparison between two ABCs for $M = 0.8$ with $R = 3$

In FIGURE 4.7 we have depicted the numerical solution p_h and the error in decibel $\mathcal{E}_{\text{dB}}$ for

(ABC1) with $R = 10$. Even if the relative error $\mathcal{E}_{\mathcal{O}}$ is of 2.05%, by looking at the plot of $\mathcal{E}_{\text{dB}}$ it seems possible to trust the solution far enough from the artificial boundary.



(a) Numerical solution p_h (b) $\mathcal{E}_{\text{dB}}$

Figure 4.7: Results and error for (ABC1) with $M = 0.8$ and $R = 10$

Finally we would like to point out that considering a larger domain leads to a much more expensive problem from a computational point of view. The numerical experiments were runned on a `miriel` node of the `plafrim` cluster[1] equipped with 2 dodeca-core Haswell Intel Xeon E5-2680 v3 with a clock rate of 2.5 GHz and 128 Go of memory. The case in the small domain with $R = 3$ ran in 1min, whereas for the test with $R = 10$, it took 1h15min to run.

Influence of the flow angle

In this section, we try to understand the influence of the flow angle θ_0 on the performance of the ABCs. In Table 4.4 the errors obtained when using (ABC1) are given for various values of θ_0. We can clearly see that $\mathcal{E}_{\mathcal{O}}$ is smaller when the flow is aligned with one of the axis (for $\theta_0 = 0$ or $\theta_0 = \pi/2$). This can be understood as the change of spatial coordinates $\boldsymbol{A}^{-1}$, which is used to construct the ABCs, is contracting: the domain is smaller in physical coordinates than in PGL ones. When $\theta_0 = 0$ or $\pi/2$, this contraction only occurs in one direction, whereas it occurs in both directions for the other values of θ_0. The error is therefore smaller when the flow is aligned with one of the axis. The maximal error appears to be achieved for $\theta_0 = \pi/4$ which is the angle with the highest contraction, as both directions are contracted in the same way.

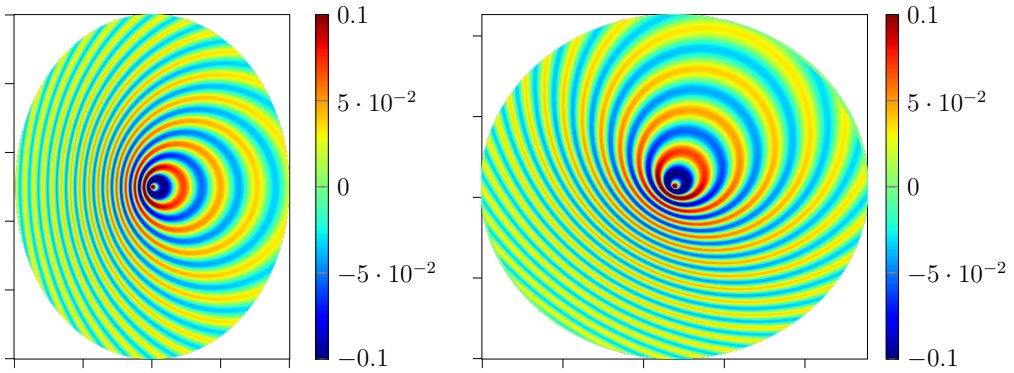
Flow angle θ_0	Error $\mathcal{E}_{\mathcal{O}}$
0	$1.38 \cdot 10^{-3}$%
$\pi/3$	0.40%
$\pi/4$	0.65%
$\pi/6$	0.41%
$\pi/2$	$1.37 \cdot 10^{-3}$%

Table 4.4: Domain error $\mathcal{E}_{\mathcal{O}}$ for $M = 0.6$ with (ABC1) and $R = 2.5$

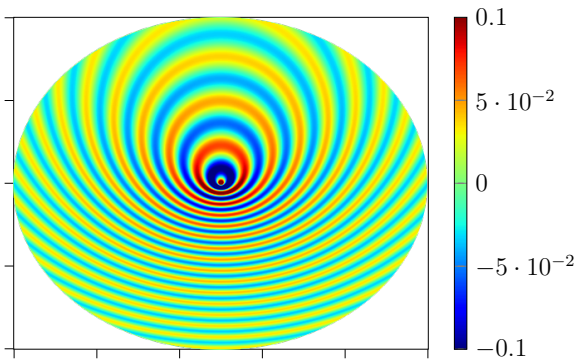
As the error seems minimal for a flow aligned with one of the axis, performing a rotation should therefore be considered before using those ABCs. However as the relative error stays below 1%, results obtained for other values of θ_0 are exploitable.

[1] See `https://www.plafrim.fr`.

Some of the cases of TABLE 4.4 are depicted in FIGURE 4.8. On those figures, the elliptical shapes of the domain can clearly be seen when $\theta_0 \neq \pi/4$, as well as the contraction effect due to the change of coordinates.



(a) $\theta_0 = 0$ (b) $\theta_0 = \frac{\pi}{3}$



(c) $\theta_0 = \frac{\pi}{2}$

Figure 4.8: Numerical examples for various values of θ_0 with $M = 0.6$ and (ABC1)

Illustrative examples with multiple sources

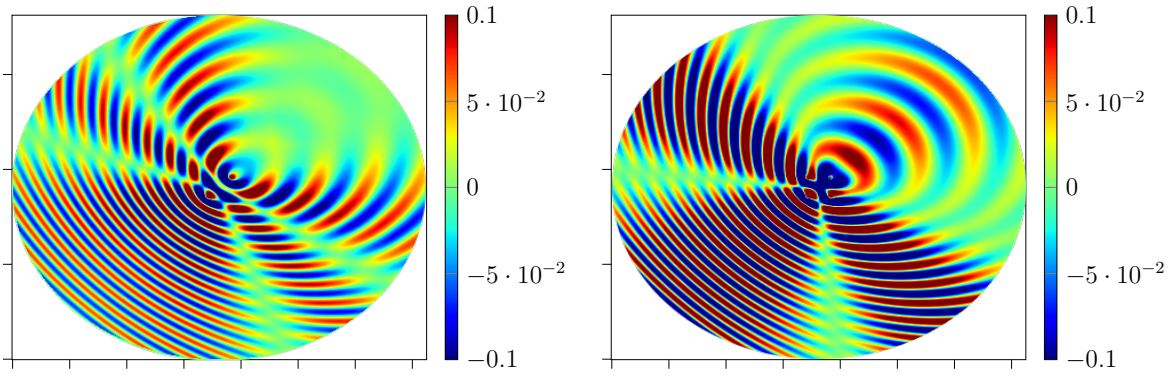
To illustrate the ability of the ABCs to handle more complex cases, we consider multiple point-sources in a uniform flow. For the numerical simulations, the following parameters are used

$$M = 0.6, \quad \theta_0 = \frac{\pi}{4}, \quad \omega = 6\pi, \quad R = 2.$$

The sources will be located at $(\pm 0.1, \pm 0.1)$ for the case with two point-sources and at $(\pm 0.1, \pm 0.1)$ and $(\pm 0.1, \mp 0.1)$ for the case with four point-sources.
In FIGURE 4.9 the results obtained with the above parameters and (ABC1) are depicted. We can clearly see the interference patterns between the sources as well as changes in the apparent frequency due to the Doppler effect. Those physical phenomena seem to be handled well by the ABC and it looks like there is no numerical pollution inside the domain

4.3.2 Experiments with a potential flow

In this section, we give illustrative examples where the PGL-based ABCs are used with a non-uniform flow. As the convected Helmholtz equation is only valid for potential flows, we focus on the case of a potential flow around a circular obstacle.



(a) Two point-sources (b) Four point-sources

Figure 4.9: Interferences between multiple point-sources for $M = 0.6$ with (ABC1)

We have constructed ABCs for an artificial boundary Σ centered on the source. The expression for the potential flow should therefore be translated as it is usually written for a circular obstacle located at $(0, 0)$.
We assume that the obstacle has a radius of R_C and its center is located at $\boldsymbol{x}_C = (x_C, y_C)$. For a point $\boldsymbol{x} = (x, y)$, we define the translated polar coordinates around $\boldsymbol{x}_C$ by

$$r := |\boldsymbol{x} - \boldsymbol{x}_C|, \quad \text{and} \quad \theta := \arctan\left(\frac{y - y_C}{x - x_C}\right).$$

Let M_∞ be the Mach number of the flow at infinity, the potential flow is naturally expressed in the previous polar coordinates by

$$\boldsymbol{v_0} := M_\infty \left[\left(1 - \frac{R_C^2}{r^2}\right)\cos\theta \boldsymbol{e}_r - \left(1 + \frac{R_C^2}{r^2}\right)\sin\theta \boldsymbol{e}_\theta\right],$$

leading to the following expression in cartesian coordinates

$$\begin{aligned}\boldsymbol{v_0} = M_\infty \Bigg[&\left(\frac{x - x_C}{r}\left(1 - \frac{R_C^2}{r^2}\right)\cos\theta + \frac{y - y_C}{r}\left(1 + \frac{R_C^2}{r^2}\right)\sin\theta\right)\boldsymbol{e}_x \\ &+ \left(\frac{y - y_C}{r}\left(1 - \frac{R_C^2}{r^2}\right)\cos\theta - \frac{x - x_C}{r}\left(1 + \frac{R_C^2}{r^2}\right)\sin\theta\right)\boldsymbol{e}_y\Bigg].\end{aligned}$$

The configuration for this case is depicted in FIGURE 4.10.
In FIGURE 4.11, we depicted p_h obtained with the following parameters

$$\boldsymbol{x}_C = (\pm 1, 0), \quad R_C = 0.5, \quad M_\infty = 0.4, \quad \omega = 6\pi.$$

It seems there is no reflection at the artificial boundary Σ and the expected physical phenomena are visible. We can clearly see a change in apparent frequency due to the Doppler effect, phase-shifts due to a refraction-like effect and an interference pattern due to the reflection of the wave on the obstacle. For the upstream example, we can also see a silent zone behind the obstacle as expected. For the downstream case, creeping waves can be seen around the obstacle, and a constructive interference pattern is visible behind the obstacle.

Similar cases to the downstream one have been considered in [LMG+20] and in [BCD+14]. The results obtained by the authors are depicted on FIGURE 4.12, and they are very similar to the ones of this paper, therefore validating the new ABCs that we have constructed.

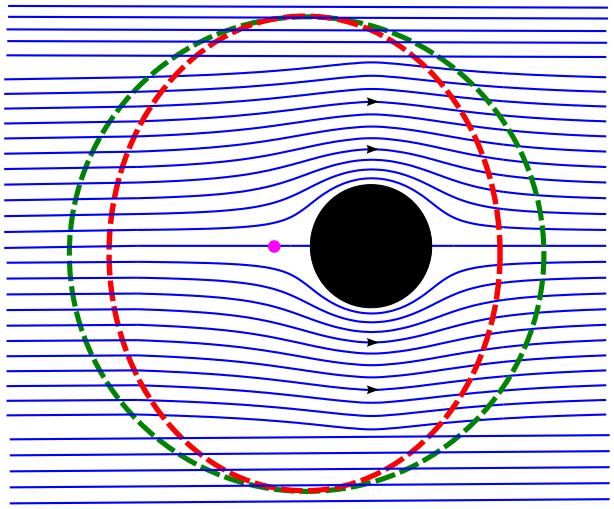
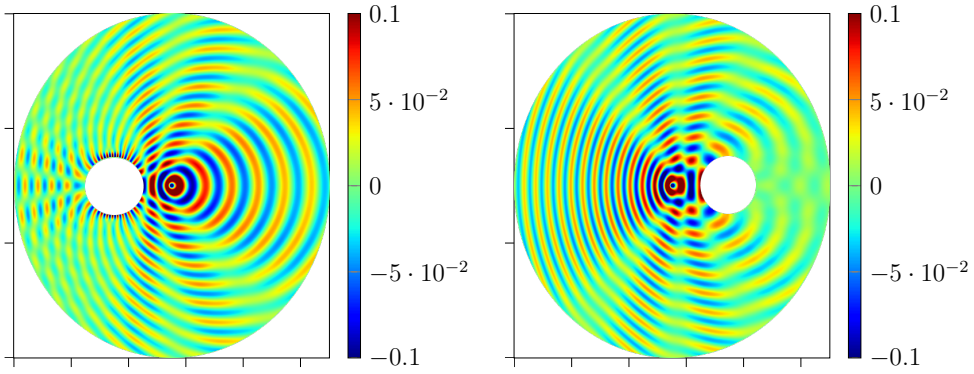
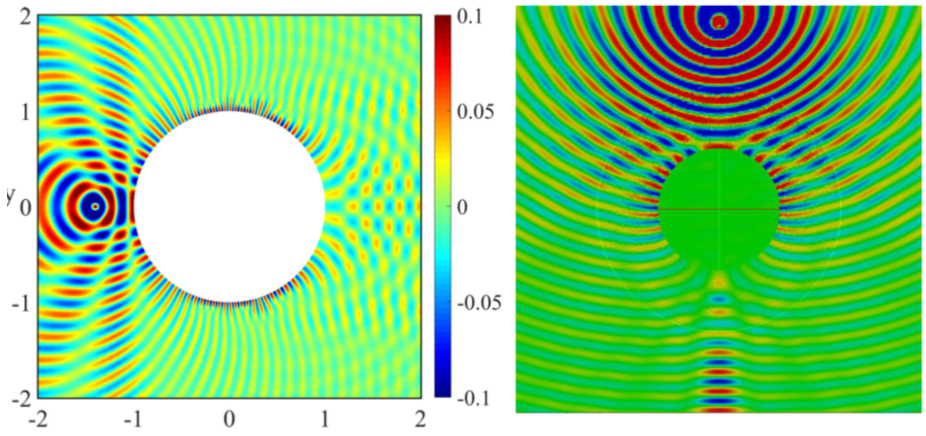


Figure 4.10: Sketch of the settings: in blue: streamlines of $\boldsymbol{v_0}$, in red: artificial boundary Σ, in green: artificial boundary in PGL coordinates, in magenta: point-source, in black: obstacle



(a) Downstream obstacle (b) Upstream obstacle

Figure 4.11: Point-source in a potential flow around a circular obstacle for $M_\infty = 0.4$



(a) Extracted from [LMG+20] (b) Extracted from [BCD+14]

Figure 4.12: Similar cases found in the literature

Conclusion

In this chapter, we have seen that the use of the *Prandtl-Glauert-Lorentz transformation* to construct ABCs for the convected Helmholtz equation leads to very good results for low and intermediate Mach numbers. For higher Mach numbers, a very large domain seems to be required and the ABCs constructed in this chapter should be used carefully. Due to the presence of convection, the usual idea of selecting waves that are locally orthogonal to the boundary leads to bad results, and the PGL-based ABCs should be preferred. Among those ABCs, the use of ABC1 instead of ABC0 gives better results for no additional cost. Finally we would like to point out that the ABCs constructed in this chapter are really easy to implement in an existing finite-element solver for the convected Helmholtz equation as the PGL transformation is only required to compute the impedance-like operator $\mathcal{Z}$.

In [Gab03, Chap. 3] and [Bér08, Chap. 5], ABCs for Galbrun's equation are studied. Galbrun's equation is an aeroacoustic model which is more realistic than the convected Helmholtz equation as it allows for more general background flows. However the ABCs derived in those dissertations were only illustrated with uniform flows for which the two models are equivalent. This can clearly be seen as the reference solution used in [Bér08, p. 129 - last equation] is the same as the one we described in (4.27). Even if various conditions are described, the best performing one is somewhat similar to the ABC-PW condition (ABC-PW) of this chapter, and the authors only obtained an error level of 4.5% for $M = 0.3$ and $\omega = 6\pi$. Extending the results of could therefore lead to better ABCs for Galbrun's equation.

For the convected Helmholtz equation, usual PMLs are known to be unstable. In [BBL03] stable PMLs for the convected Helmholtz equation were derived, but only for propagation in a waveguide. This work was extended to arbitrary geometries in [MBAG20], this formulation is also based on the *Prandtl-Glauert-Lorentz transformation*. For high Mach numbers, the so-called Lorentz PMLs of [MBAG20] seem to perform better than the ABCs of this chapter. Indeed, the authors were able to obtain an error level fo 1.2% using a smaller domain for $M = 0.8$. However this formulation seems difficult to implement in the context of HDG methods and, as stated in [MAGB21], their efficiency for non-uniform flows remains unclear. Finally, we would like to point out that similar PMLs have been derived for time-domain convected acoustic wave equation in [DJ06].

Perspectives Even if we only presented 2D results, the extension of this work to 3D is straightforward as the result of Theorem 9 relating $\mathcal{Z}$ and $\widetilde{\mathcal{Z}}$ holds for any dimension.

Extension of this work to higher order boundary conditions is possible even if the expected gain is not clear. As pointed out in [ABB99] higher order ABCs for the Helmholtz equation are expressed as perturbations of low order ABCS. These perturbations involve curvature operators (Gaussian and mean curvatures essentially) and their derivatives which are canceled for a circle. Hence there is no expected gain in 2D in using higher order ABCs. The question remains open in 3D.

Extension to more realistic aeroacoustic models seem possible. The most straightforward model to consider for this extension is Goldstein's equation [MMMP17] which consists in a coupling between a convected Helmholtz equation and a vectorial transport equation. The ABCs of this chapter could be used for the convected Helmholtz equation and it is possible to obtain an exact outgoing condition for the transport equation. Extension to Galbrun's equation [BMM+12] or to the Linearized Euler's Equations [BBJ02] seems more complicated, mostly due to the lack of a "standard" equation for which ABCs can be derived.

Appendix

4.A Proof of Lemma 4.2.4

In this appendix, we prove the result of LEMMA 4.2.4. As the numerical examples of this chapter are in 2D, we have chosen to prove the result in the same dimension. Extension to 3D is straightforward and would involve the 3D Green functions instead of the Hankel functions.

Proof : This result is proven by using the *limiting amplitude principle* which consists in studying the long-term behaviour of the time-domain solution with the following source term

$$g(\boldsymbol{x},t) = \begin{cases} s(\boldsymbol{x})e^{-i\omega t}, & \text{if } t > 0 \\ 0, & \text{if } t \leqslant 0 \end{cases},$$

where s is a function in $L^2(\mathcal{O})$ with compact support.

Step 1: Green functions in time-domain.

The Green function for the two-dimensional acoustic wave-equation in time-domain reads

$$G_{\text{std}}(\boldsymbol{x},t) = \frac{H\left(t-\frac{r}{c_0}\right)}{2\pi\sqrt{t^2-\frac{r^2}{c_0^2}}},$$

where H is the Heaviside function and $r = |\boldsymbol{x}|$. For the proof, we refer to [DJ06, Th. 1] or [Ver10, Sec. A.4.2].

The Green function for the convected wave equation in time-domain has been computed in the case $\boldsymbol{M_0} = M\boldsymbol{e}_x$ using the *Cagniard-de Hoop method* in [DJ06, Th. 6]. The case of a generic $\boldsymbol{M_0}$ reduces to this one by rotation. In [DJ06], the parameter σ denotes the absorption parameter used in PMLs, so we take $\sigma = 0$ leading to $A = B = 0$ in [DJ06, Th. 6].

With the assumption that $\boldsymbol{M_0} = M\boldsymbol{e}_x$, the *Prandtl-Glauert-Lorentz transformation* reads

$$\widetilde{t} = \alpha t + \frac{1}{\alpha c_0}Mx, \quad \widetilde{x} = \frac{x}{\alpha}, \quad \widetilde{y} = y, \quad \widetilde{r}^2 = \widetilde{x}^2 + \widetilde{y}^2, \quad \widetilde{\theta} = \arctan\frac{\widetilde{y}}{\widetilde{x}}. \tag{4.28}$$

We define the following change of variables in space

$$\hat{x} = \frac{\widetilde{x}}{\alpha}, \quad \text{and} \quad \hat{y} = \frac{\widetilde{y}}{\alpha}. \tag{4.29a}$$

In polar coordinates, we also define

$$\hat{r}^2 = \hat{x}^2 + \hat{y}^2, \quad \text{and} \quad \hat{\theta} = \arctan\frac{\hat{y}}{\hat{x}}. \tag{4.29b}$$

It is straightforward to check that

$$\hat{r} = \frac{\widetilde{r}}{\alpha}, \quad \text{and} \quad \hat{\theta} = \widetilde{\theta}.$$

The Green function corresponding to the convected acoustic wave equation reads

$$G_{\text{conv}}(\boldsymbol{x},t) = \frac{H(t-\frac{\hat{r}}{c_0}(1-M\cos\hat{\theta}))}{2\pi\sqrt{1-M^2}\sqrt{(t+\frac{M\hat{r}}{c_0}\cos\hat{\theta})^2-\frac{\hat{r}^2}{c_0}}}.$$

Using the intermediate change of variables (4.29a)–(4.29b) and expressing in terms of the *Prandt-Glauert-Lorentz coordinates* (4.28), we have

$$\begin{aligned}
G_{\text{conv}}(\boldsymbol{x},t) &= \frac{H\left(t+\frac{M}{c_0}\hat{x}-\frac{\hat{r}}{c_0}\right)}{2\pi\alpha\sqrt{(t+\frac{M}{c_0}\hat{x})^2-\frac{\hat{r}^2}{c_0^2}}},\\
&= \frac{H\left(t+\frac{M}{\alpha c_0}\widetilde{x}-\frac{\widetilde{r}}{\alpha c_0}\right)}{2\pi\alpha\sqrt{(t+\frac{M}{\alpha c_0}\widetilde{x})^2-\frac{\widetilde{r}^2}{\alpha^2 c_0^2}}},\\
&= \frac{H\left(t+\frac{M}{\alpha^2 c_0}x-\frac{\widetilde{r}}{\alpha c_0}\right)}{2\pi\alpha\sqrt{(t+\frac{M}{\alpha^2 c_0}x)^2-\frac{\widetilde{r}^2}{\alpha^2 c_0^2}}},\\
&= \frac{H\left(\frac{1}{\alpha}\left[\alpha t+\frac{M}{\alpha c_0}x-\frac{\widetilde{r}}{c_0}\right]\right)}{2\pi\alpha\frac{1}{\alpha}\sqrt{(\alpha t+\frac{M}{\alpha c_0}x)^2-\frac{\widetilde{r}^2}{c_0^2}}},\\
&= \frac{H\left(\widetilde{t}-\frac{\widetilde{r}}{c_0}\right)}{2\pi\sqrt{\widetilde{t}^2-\frac{\widetilde{r}^2}{c_0^2}}}\\
&= G_{\text{std}}(\widetilde{\boldsymbol{x}},\widetilde{t}).
\end{aligned}$$

As G_{conv} is mapped to G_{std} through the *Prandtl-Glauert-Lorentz transformation*, we can therefore prove that this transform also maps outgoing solutions to the convected Helmholtz equation to outgoing solutions to the standard Helmholtz equation by using standard arguments of the *limiting amplitude principle*.

Step 2: Limiting amplitude principle for the standard Helmholtz equation.

We recall that the following identities hold

$$\int_r^{+\infty}\frac{e^{i\omega t}}{2\pi\sqrt{t^2-r^2}}\mathrm{d}t = \frac{1}{2\pi}\int_0^{+\infty}e^{i\omega r\cosh\theta}\mathrm{d}\theta = \frac{i}{4}H_0^{(1)}(\omega r), \tag{4.30}$$

where we used the following change of variables $t = r\cosh\theta$ and the integral form of the Hankel function, see [OLBC10, Eq. 10.9.10].

Using the integral representation we have

$$p_{\text{std}}(\boldsymbol{x},t) = \int_{\mathbb{R}^2}\int_{\mathbb{R}}G_{\text{std}}(\boldsymbol{x}_1,t_1)g(\boldsymbol{x}-\boldsymbol{x}_1,t-t_1)\mathrm{d}t_1\mathrm{d}\boldsymbol{x}_1,$$

using the integral identity (4.30) for the Hanekl function, we have

$$= \int_{\mathbb{R}^2}\int_{\frac{|\boldsymbol{x}_1|}{c_0}}^{+\infty}\frac{g(\boldsymbol{x}-\boldsymbol{x}_1,t-t_1)}{2\pi\sqrt{t_1^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}}}\mathrm{d}t_1\mathrm{d}\boldsymbol{x}_1$$

and since $g(\boldsymbol{x},t) = s(\boldsymbol{x})e^{-i\omega t}H(t)$, where s is a L^2-function with compact support, we obtain

$$\begin{aligned}
&= \int_{\mathbb{R}^2}s(\boldsymbol{x}-\boldsymbol{x}_1)e^{-i\omega t}\int_{\frac{|\boldsymbol{x}_1|}{c_0}}^{t}\frac{e^{i\omega t_1}}{2\pi\sqrt{t_1^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}}}\mathrm{d}t_1\mathrm{d}\boldsymbol{x}_1\\
&= \int_{\mathbb{R}^2}s(\boldsymbol{x}-\boldsymbol{x}_1)e^{-i\omega t}\left[\frac{i}{4}H_0^{(1)}\left(\frac{\omega}{c_0}|\boldsymbol{x}_1|\right)+\mathfrak{r}_{\text{std}}\right]\mathrm{d}\boldsymbol{x}_1
\end{aligned}$$

where $\mathfrak{r}_{\text{std}}(\boldsymbol{x}_1,t) = -\int_t^{+\infty} \frac{e^{i\omega t_1}}{2\pi\sqrt{t_1^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}}}\mathrm{d}t_1$ with $t > \frac{|\boldsymbol{x}_1|}{c_0}$,

$$= \int_{\mathbb{R}^2} s(\boldsymbol{x}-\boldsymbol{x}_1)e^{-i\omega t}\frac{i}{4}H_0^{(1)}\left(\frac{\omega}{c_0}|\boldsymbol{x}_1|\right)\mathrm{d}\boldsymbol{x}_1 + \mathfrak{R}_{\text{std}},$$

and

$$\mathfrak{R}_{\text{std}}(\boldsymbol{x},t) = e^{-i\omega t}\int_{\mathbb{R}^2} s(\boldsymbol{x}-\boldsymbol{x}_1)\mathfrak{r}_{\text{std}}(\boldsymbol{x}_1,t)\mathrm{d}\boldsymbol{x}_1 = -e^{-i\omega t}\int_{\mathbb{R}^2} s(\boldsymbol{x}-\boldsymbol{x}_1)\int_t^{+\infty}\frac{e^{i\omega t_1}}{2\pi\sqrt{t_1^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}}}\mathrm{d}t_1\mathrm{d}\boldsymbol{x}_1.$$

We now prove that $\mathfrak{R}_{\text{std}}(\boldsymbol{x},t) \xrightarrow[t\to+\infty]{} 0$. The first step is to show that the integral defining $\mathfrak{r}_{\text{std}}$ is actually finite-valued. We have

$$\begin{aligned}\mathfrak{r}_{\text{std}}(\boldsymbol{x}_1,t) &:= -\int_t^{+\infty}\frac{e^{i\omega t_1}}{2\pi\sqrt{t_1^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}}}\mathrm{d}t_1\\ &= -\lim_{A\to+\infty}\left(\left[\frac{e^{i\omega t_1}}{2i\pi\omega\sqrt{t_1^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}}}\right]_t^A + \int_t^A \frac{t_1 e^{i\omega t_1}}{2\pi\left(t_1^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}\right)^{\frac{3}{2}}}\mathrm{d}t_1\right)\\ &< +\infty,\end{aligned}$$

as

$$\lim_{A\to+\infty}\frac{e^{i\omega A}}{2i\pi\omega\sqrt{A^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}}} = 0, \quad \text{and} \quad \left|\frac{t_1 e^{i\omega t_1}}{2\pi\left(t_1^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}\right)^{\frac{3}{2}}}\right| = \frac{t_1}{2\pi\left(t_1^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}\right)^{\frac{3}{2}}} \underset{t_1\to\infty}{\sim} \frac{1}{t_1^2}.$$

We can therefore express $\mathfrak{r}_{\text{std}}$ as

$$\mathfrak{r}_{\text{std}}(\boldsymbol{x}_1,t) = \frac{e^{i\omega t}}{2i\pi\omega\sqrt{t^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}}} - \int_t^{+\infty}\frac{t_1 e^{i\omega t_1}}{2\pi\left(t_1^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}\right)^{\frac{3}{2}}}\mathrm{d}t_1$$

and $\mathfrak{R}_{\text{std}}$ as

$$\mathfrak{R}_{\text{std}}(\boldsymbol{x},t) = \int_{\mathbb{R}^2}\frac{s(\boldsymbol{x}-\boldsymbol{x}_1)}{2\pi\sqrt{t^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}}}\mathrm{d}\boldsymbol{x}_1 + e^{-i\omega t}\int_{\mathbb{R}^2} s(\boldsymbol{x}-\boldsymbol{x}_1)\int_t^{+\infty}\frac{t_1 e^{i\omega t_1}}{2\pi\left(t_1^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}\right)^{\frac{3}{2}}}\mathrm{d}t_1\mathrm{d}\boldsymbol{x}_1.$$

We can now show that both of these quantities vanish at infinity. Let us first focus on the first term in $\mathfrak{R}_{\text{std}}$

$$\int_{\mathbb{R}^2}\frac{s(\boldsymbol{x}-\boldsymbol{x}_1)}{2\pi\sqrt{t^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}}}\mathrm{d}\boldsymbol{x}_1.$$

As we are interested in the limit as $t\to+\infty$ and as the space integral is over a compact domain, we can assume that

$$\frac{3t^2}{4} > \frac{|\boldsymbol{x}_1|^2}{c_0^2},$$

which leads to

$$\frac{|s(\boldsymbol{x}-\boldsymbol{x}_1)|}{2\pi\sqrt{t^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}}} < \frac{|s(\boldsymbol{x}-\boldsymbol{x}_1)|}{\pi|t|},$$

and we have

$$\int_{\mathbb{R}^2} \left| \frac{s(\boldsymbol{x}-\boldsymbol{x}_1)}{2\pi\sqrt{t^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}}} \right| \mathrm{d}\boldsymbol{x}_1 \leqslant \frac{1}{\pi|t|} \int_{\mathbb{R}^2} |s(\boldsymbol{x}-\boldsymbol{x}_1)| \mathrm{d}\boldsymbol{x}_1 \xrightarrow[t\to+\infty]{} 0,$$

as the last integral is over a compact domain. By taking the limit in the previous inequality, we have

$$\lim_{t\to+\infty} \int_{\mathbb{R}^2} \left| \frac{s(\boldsymbol{x}-\boldsymbol{x}_1)}{2\pi\sqrt{t^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}}} \right| \mathrm{d}\boldsymbol{x}_1 = 0.$$

We can now work on the second term.
As we are interested in the limit as $t \to +\infty$ and as the space integral is over a compact domain, we can assume once again that

$$\frac{3t_1^2}{4} > \frac{3t^2}{4} > \frac{|\boldsymbol{x}_1|^2}{c_0^2},$$

which leads to

$$\frac{t_1}{2\pi\left(t_1^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}\right)^{\frac{3}{2}}} < \frac{t_1}{\frac{\pi}{4}t_1^3} = \frac{4}{\pi t_1^2}.$$

We therefore have

$$\int_t^{+\infty} \left| \frac{t_1 e^{i\omega t_1}}{2\pi\left(t_1^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}\right)^{\frac{3}{2}}} \right| \mathrm{d}t_1 < \frac{4}{\pi} \int_t^{+\infty} \frac{1}{t_1^2} \mathrm{d}t_1 = \frac{4}{\pi t} \xrightarrow[t\to+\infty]{} 0.$$

Finally we have

$$\begin{aligned} \int_{\mathbb{R}^2} |s(\boldsymbol{x}-\boldsymbol{x}_1)| \int_t^{+\infty} \left| \frac{t_1 e^{i\omega t_1}}{2\pi\left(t_1^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}\right)^{\frac{3}{2}}} \right| \mathrm{d}t_1 \mathrm{d}\boldsymbol{x}_1 &\leqslant \int_{\mathbb{R}^2} |s(\boldsymbol{x}-\boldsymbol{x}_1)| \frac{4}{\pi t} \mathrm{d}\boldsymbol{x}_1 \\ &\leqslant \frac{4}{\pi t} \int_{\mathbb{R}^2} |s(\boldsymbol{x}-\boldsymbol{x}_1)| \mathrm{d}\boldsymbol{x}_1 \xrightarrow[t\to+\infty]{} 0. \end{aligned}$$

Notice that the last integral is finite as the support of s is compact. By taking the limit in the previous inequalities, we have

$$\lim_{t\to+\infty} \int_{\mathbb{R}^2} |s(\boldsymbol{x}-\boldsymbol{x}_1)| \int_t^{+\infty} \left| \frac{t_1 e^{i\omega t_1}}{2\pi\left(t_1^2-\frac{|\boldsymbol{x}_1|^2}{c_0^2}\right)^{\frac{3}{2}}} \right| \mathrm{d}t_1 \mathrm{d}\boldsymbol{x}_1 = 0.$$

We therefore have

$$\lim_{t\to+\infty} |\mathfrak{R}_{\mathrm{std}}(\boldsymbol{x},t)| = 0, \quad \forall \boldsymbol{x} \in \mathbb{R}^2.$$

Now, we define the time-harmonic Green function by

$$W_{\mathrm{std}}(\boldsymbol{x},\omega) = \frac{i}{4} H_0^{(1)}\left(\frac{\omega}{c_0}|\boldsymbol{x}|\right).$$

We end up with the *limiting amplitude principle*

$$\lim_{t\to\infty}\left\|p_{\mathrm{std}}(\boldsymbol{x},t)-p_{\mathrm{std},\omega}(\boldsymbol{x})e^{-i\omega t}\right\|=0,$$

where $p_{\mathrm{std},\omega}:=W_{\mathrm{std}}(\cdot,\omega)*s$ is the outgoing solution of the standard Helmholtz equation.
Step 3: Limiting amplitude principle for the convected Helmholtz equation. We have

$$\begin{aligned}
p(\boldsymbol{x},t)&=\int_{\mathbb{R}^2}\int_{\mathbb{R}}G_{\mathrm{conv}}(\boldsymbol{x}_1,t_1)g(\boldsymbol{x}-\boldsymbol{x}_1,t-t_1)\mathrm{d}t_1\mathrm{d}\boldsymbol{x}_1\\
&=\int_{\mathbb{R}^2}\int_{\mathbb{R}}G_{\mathrm{std}}(\widetilde{\boldsymbol{x}}_1,\widetilde{t}_1)g\left(\boldsymbol{A}^{-1}(\widetilde{\boldsymbol{x}}-\widetilde{\boldsymbol{x}}_1),\frac{1}{\alpha}(\widetilde{t}-\widetilde{t}_1)+\frac{1}{\alpha c_0}\boldsymbol{M_0}\cdot(\widetilde{\boldsymbol{x}}-\widetilde{\boldsymbol{x}}_1)\right)\mathrm{d}\widetilde{t}_1\mathrm{d}\widetilde{\boldsymbol{x}}_1\\
&=\int_{\mathbb{R}^2}\int_{\frac{|\widetilde{\boldsymbol{x}}_1|}{c_0}}^{+\infty}\frac{g\left(\boldsymbol{A}^{-1}(\widetilde{\boldsymbol{x}}-\widetilde{\boldsymbol{x}}_1),\frac{1}{\alpha}(\widetilde{t}-\widetilde{t}_1)-\frac{1}{\alpha c_0}\boldsymbol{M_0}\cdot(\widetilde{\boldsymbol{x}}-\widetilde{\boldsymbol{x}}_1)\right)}{2\pi\sqrt{\widetilde{t}_1^2-\frac{|\widetilde{\boldsymbol{x}}_1|^2}{c_0^2}}}\mathrm{d}\widetilde{t}_1\mathrm{d}\widetilde{\boldsymbol{x}}_1
\end{aligned}$$

using that $g(\boldsymbol{x},t)=s(\boldsymbol{x})e^{-i\omega t}H(t)$, where s is a function with compact support,

$$\begin{aligned}
&=\int_{\mathbb{R}^2}s(\boldsymbol{A}^{-1}(\widetilde{\boldsymbol{x}}-\widetilde{\boldsymbol{x}}_1))e^{-i\omega(\frac{1}{\alpha}\widetilde{t}-\frac{1}{\alpha c_0}\boldsymbol{M_0}\cdot\widetilde{\boldsymbol{x}})}e^{-i\frac{\omega}{\alpha c_0}\boldsymbol{M_0}\cdot\widetilde{\boldsymbol{x}}_1}\int_{\frac{|\widetilde{\boldsymbol{x}}_1|}{c_0}}^{\widetilde{t}-\frac{\boldsymbol{M_0}}{c_0}\cdot(\widetilde{\boldsymbol{x}}-\widetilde{\boldsymbol{x}}_1)}\frac{e^{i\frac{\omega}{\alpha}\widetilde{t}}}{2\pi\sqrt{\widetilde{t}_1^2-\frac{|\widetilde{\boldsymbol{x}}_1|^2}{c_0^2}}}\mathrm{d}\widetilde{t}_1\mathrm{d}\widetilde{\boldsymbol{x}}_1\\
&=\int_{\mathbb{R}^2}s(\boldsymbol{A}^{-1}(\widetilde{\boldsymbol{x}}-\widetilde{\boldsymbol{x}}_1))e^{-i\omega(\frac{1}{\alpha}\widetilde{t}-\frac{1}{\alpha c_0}\boldsymbol{M_0}\cdot\widetilde{\boldsymbol{x}})}e^{-i\frac{\omega}{\alpha c_0}\boldsymbol{M_0}\cdot\widetilde{\boldsymbol{x}}_1}\left[\frac{i}{4}H_0^{(1)}\left(\frac{\omega}{\alpha c_0}|\widetilde{\boldsymbol{x}}_1|\right)+\mathfrak{r}_{\mathrm{conv}}\right]\mathrm{d}\widetilde{\boldsymbol{x}}_1\\
&=\int_{\mathbb{R}^2}\frac{1}{\alpha}s(\boldsymbol{x}-\boldsymbol{x}_1)e^{-i\omega t}e^{-i\frac{\omega}{\alpha^2c_0}\boldsymbol{M_0}\cdot\boldsymbol{x}_1}\frac{i}{4}H_0^{(1)}\left(\frac{\omega}{\alpha c_0}|\boldsymbol{A}\boldsymbol{x}_1|\right)\mathrm{d}\boldsymbol{x}_1+\mathfrak{R}_{\mathrm{conv}},
\end{aligned}$$

where $\mathfrak{r}_{\mathrm{conv}}$ and $\mathfrak{R}_{\mathrm{conv}}$ are defined as $\mathfrak{r}_{\mathrm{std}}$ and $\mathfrak{R}_{\mathrm{std}}$ in the previous step. By using similar arguments to the ones in the previous step, we can show that

$$\lim_{t\to+\infty}\|\mathfrak{R}_{\mathrm{conv}}(\cdot,t)\|=0.$$

We now define the time-harmonic Green function by

$$W_{\mathrm{conv}}(\boldsymbol{x},\omega)=\frac{1}{\alpha}\exp\left[-\frac{i\omega}{\alpha^2c_0}\boldsymbol{M_0}\cdot\boldsymbol{x}\right]\frac{i}{4}H_0^{(1)}\left(\frac{\omega}{\alpha c_0}|\boldsymbol{A}\boldsymbol{x}|\right),$$

and we can also obtain the limiting amplitude principle for the convected Helmholtz equatio

$$\lim_{t\to\infty}\left\|p(\boldsymbol{x},t)-p_{\mathrm{conv},\omega}(\boldsymbol{x})e^{-i\omega t}\right\|=0,$$

where $p_{\mathrm{conv},\omega}=W_{\mathrm{conv}}(\cdot,\omega)*s$ is the outgoing solution of the convected Helmholtz equation.
Step 4: Equivalence under *Prandtl-Glauert-Lorentz transformation.*
We notice that

$$W_{\mathrm{conv}}(\boldsymbol{x},\omega)=\frac{1}{\alpha}\exp\left[-\frac{i\omega}{\alpha^2c_0}\boldsymbol{M_0}\cdot\boldsymbol{x}\right]\frac{i}{4}H_0^{(1)}\left(\frac{\omega}{\alpha c_0}|\boldsymbol{A}\boldsymbol{x}|\right)=\frac{1}{\alpha}\exp\left[-\frac{i\omega}{\alpha^2c_0}\boldsymbol{M_0}\cdot\boldsymbol{x}\right]W_{\mathrm{std}}\left(\widetilde{\boldsymbol{x}},\frac{\omega}{\alpha}\right),$$

We recall that $\widetilde{s}$ is defined by

$$\widetilde{s}(\widetilde{\boldsymbol{x}},\widetilde{\omega}):=\alpha\exp\left[i\omega\frac{\boldsymbol{M_0}\cdot\boldsymbol{x}}{\alpha^2c_0}\right]s(\boldsymbol{x},\omega),$$

and we define

$$\widetilde{p}_{\mathrm{std},\widetilde{\omega}}:=W_{\mathrm{std}}(\cdot,\widetilde{\omega})*\widetilde{s}=\int_{\mathbb{R}^2}\widetilde{s}(\cdot-\widetilde{\boldsymbol{x}}_1)W_{\mathrm{std}}(\widetilde{\boldsymbol{x}}_1,\widetilde{\omega})\mathrm{d}\widetilde{\boldsymbol{x}}_1.$$

We have

$$\widetilde{p}_{\mathrm{std},\widetilde{\omega}}(\widetilde{\boldsymbol{x}}) = \alpha \exp\left[i\omega \frac{\boldsymbol{M_0}\cdot\boldsymbol{x}}{\alpha^2 c_0}\right] p_{\mathrm{conv},\omega}(\boldsymbol{x}),$$

indeed

$$\begin{aligned}
\widetilde{p}_{\mathrm{std},\widetilde{\omega}}(\widetilde{\boldsymbol{x}}) &= \int_{\mathbb{R}^2} \widetilde{s}(\widetilde{\boldsymbol{x}}-\widetilde{\boldsymbol{x}}_1) W_{\mathrm{std}}(\widetilde{\boldsymbol{x}}_1,\widetilde{\omega})\mathrm{d}\widetilde{\boldsymbol{x}}_1, \\
&= \int_{\mathbb{R}^2} \alpha \exp\left[i\omega \frac{\boldsymbol{M_0}\cdot(\boldsymbol{x}-\boldsymbol{x}_1)}{\alpha^2 c_0}\right] s(\boldsymbol{x}-\boldsymbol{x}_1)\alpha \exp\left[i\omega \frac{\boldsymbol{M_0}\cdot\boldsymbol{x}_1}{\alpha^2 c_0}\right] W_{\mathrm{conv}}(\boldsymbol{x}_1,\omega)\frac{\mathrm{d}\boldsymbol{x}_1}{\alpha}, \\
&= \alpha \exp\left[i\omega \frac{\boldsymbol{M_0}\cdot\boldsymbol{x}}{\alpha^2 c_0}\right] \int_{\mathbb{R}^2} s(\boldsymbol{x}-\boldsymbol{x}_1) W_{\mathrm{conv}}(\boldsymbol{x}_1,\omega)\mathrm{d}\boldsymbol{x}_1.
\end{aligned}$$

So the outgoing solutions to the convected Helmholtz equation are mapped to outgoing solution to the standard Helmholtz equation through the *Prandtl-Glauert-Lorentz transformation.*

■

Bibliography

[ABB99] Xavier Antoine, Helene Barucq, and Abderrahmane Bendali. Bayliss-Turkel-like Radiation Condition on Surfaces of Arbitrary Shape. *Journal of Mathematical Analysis and Applications*, 229(1):184–211, 1999.

[BBJ02] Christophe Bogey, Christophe Bailly, and Daniel Juvé. Computation of Flow Noise Using Source Terms in Linearized Euler's Equations. *Aiaa Journal - AIAA J*, 40:235–243, February 2002.

[BBL03] Eliane Bécache, Anne-Sophie Bonnet-Ben Dhia, and Guillaume Legendre. Perfectly matched layers for the convected Helmholtz equation. Technical report, 2003.

[BCD+14] Nolwenn Balin, Fabien Casenave, Francois Dubois, Eric Duceau, Stefan Duprey, and Isabelle Terrasse. Boundary Element and Finite Element Coupling for Aeroacoustics Simulations. Technical report, February 2014.

[Ber94] Jean-Pierre Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, 114(2):185–200, October 1994.

[Bér08] Hadrien Bériot. *Eléments Finis d'ordre Élévé Pour l'opérateur de Galbrun En Régime Harmonique.* Thesis, Compiègne, January 2008.

[BMM+12] Anne-Sophie Bonnet-Ben Dhia, Jean-François Mercier, Florence Millot, Sebastien Pernet, and Emilie Peynaud. Time-Harmonic Acoustic Scattering in a Complex Flow: A Full Coupling Between Acoustics and Hydrodynamics. *Communications in Computational Physics*, 11(2):555–572, February 2012.

[BST12] Helene Barucq, Anne-Gaelle St-Guirons, and Sebastien Tordeux. Non-reflecting boundary condition on ellipsoidal boundary. *Numerical Analysis and Applications*, 5(2):109–115, April 2012.

[BT80] Alvin Bayliss and Eli Turkel. Radiation boundary conditions for wave-like equations. *Communications on Pure and Applied Mathematics*, 33(6):707–725, 1980.

[CES14] Fabien Casenave, Alexandre Ern, and Guillaume Sylvand. Coupled BEM–FEM for the convected Helmholtz equation with non-uniform flow in a bounded domain. *Journal of Computational Physics*, 257:627–644, January 2014.

[Cha00] C. John Chapman. SIMILARITY VARIABLES FOR SOUND RADIATION IN A UNIFORM FLOW. *Journal of Sound and Vibration*, 233(1):157–164, May 2000.

[DJ06] Julien Diaz and Patrick Joly. A time domain analysis of PML models in acoustics. *Computer Methods in Applied Mechanics and Engineering*, 195(29-32):3820–3853, June 2006.

[EM77] Bjorn Engquist and Andrew Majda. Absorbing Boundary Conditions for the Numerical Simulation of Waves. *Mathematics of Computation*, page 23, 1977.

[Fau21] Florian Faucher. 'hawen': Time-harmonic wave modeling and inversion using hybridizable discontinuous Galerkin discretization. *Journal of Open Source Software*, 6(57):2699, January 2021.

[Gab03] Gwenael Gabard. *Méthodes Numériques et Modèles de Sources Aéroacoustiques Fondées Sur l'équation de Galbrun*. PhD thesis, 2003.

[HPN17] Fang Q. Hu, Michelle E. Pizzo, and Douglas M. Nark. On a time domain boundary integral equation formulation for acoustic scattering by rigid bodies in uniform mean flow. *The Journal of the Acoustical Society of America*, 142(6):3624, December 2017.

[HPN19] Fang Q. Hu, Michelle E. Pizzo, and Douglas M. Nark. On the use of a Prandtl-Glauert-Lorentz transformation for acoustic scattering by rigid bodies with a uniform flow. *Journal of Sound and Vibration*, 443:198–211, March 2019.

[Kim14] Seungil Kim. Analysis of the convected Helmholtz equation with a uniform mean flow in a waveguide with complete radiation boundary conditions. *Journal of Mathematical Analysis and Applications*, 410(1):275–291, February 2014.

[LMG+20] Alice Lieu, Philippe Marchner, Gwenael Gabard, Hadrien Beriot, Xavier Antoine, and Christophe Geuzaine. A Non-Overlapping Schwarz Domain Decomposition Method with High-Order Finite Elements for Flow Acoustics. *Computer Methods in Applied Mechanics and Engineering*, 2020.

[MAGB21] Philippe Marchner, Xavier Antoine, Christophe Geuzaine, and Hadrien Bériot. Construction and Numerical Assessment of Local Absorbing Boundary Conditions for Heterogeneous Time-Harmonic Acoustic Problems. April 2021.

[MBAG20] Philippe Marchner, Hadrien Beriot, Xavier Antoine, and Christophe Geuzaine. Stable Perfectly Matched Layers with Lorentz transformation for the convected Helmholtz equation. Technical report, 2020.

[MMMP17] Jean-François Mercier, Colin Mietka, Florence Millot, and Vincent Pagneux. Acoustic propagation in a vortical homentropic flow. page 19, 2017.

[OLBC10] Frank Olver, Daniel W. Lozier, Ronald F. Boisvert, and Charles W. Clark. NIST Handbook of Mathematical Functions. January 2010.

[Pie90] Allan Pierce. Wave equation for sound in fluids with unsteady inhomogeneous flow. *The Journal of the Acoustical Society of America*, 87(6):2292–2299, June 1990.

[Pro59] Murray H. Protter. Unique continuation for elliptic equations. 1959.

[Sai08] Anne-Gaelle Saint-Guirons. *Construction et Analyse de Conditions Absorbantes de Type Diritchlet-to-Neumann Pour Des Frontières Ellipsoïdales*. PhD thesis, Pau, January 2008.

[SM50] Jack Sherman and Winifried Morrison. Adjustment of an Inverse Matrix Corresponding to a Change in One Element of a Given Matrix. *Annals of Mathematical Statistics*, 21(1):124–127, March 1950.

[Ver10] Arnold Verruijt. *An Introduction to Soil Dynamics*. Theory and Applications of Transport in Porous Media. Springer Netherlands, 2010.

Part III

Per aspera ad astra

Chapter 5

First steps toward the construction of a computational framework for realistic simulations of helioseismic waves

Contents

Introduction

In this chapter, we perform some illustrative numerical simulations to show how the HDG solver introduced in Chapter 3 could be used to perform pure acoustic simulation with solar-like parameters. We will then describe the difficulties encountered when trying to extend this solver to the more realistic Galbrun's equation.

5.1 Solar-like numerical simulations

As discussed in the introduction, the *convected Helmholtz equation* can be used to study the propagation of pure acoustic waves, or p-modes, inside the Sun. This simple model has

already been studied in *e.g.* [GBD+17] using the numerical methods described in [CD16]. In this section, we will show that the HDG solver described in CHAPTER 3 can be used to perform numerical simulation in solar-like settings.

Evaluation of the physical parameters. Let $R_\odot = 696340 \cdot 10^3$m be the solar radius. For this application, we use scaled coordinates so the computational mesh is a circle of radius 1.0007126. The background physical parameters ρ_0 and c_0 come from the Model S of [CDA+96][1]. For $r \leqslant 1$, the Model S describes the solar interior, and for $r > 1$ it describes the solar atmosphere. They are interpolated in the computational domain using a spline representation and converted into SI units instead of the cgs units used in [CDA+96]. As no velocity field is provided by the Model S, we have chosen to use an orthoradial flow

$$\boldsymbol{v_0} = M c_0 \boldsymbol{e}_\theta,$$

where the Mach number M is chosen so that there is no convection in the radiative zone ($0 \leqslant r \leqslant 0.6 R_\odot$) and the velocity $|\boldsymbol{v_0}|$ is 30% of the sound-speed c_0 in the convection zone ($r \geqslant 0.6 R_\odot$). This choice of velocity field also satisfies the mass-conservation property

$$\operatorname{div}(\boldsymbol{v_0}) = 0.$$

Nondimensionalization. As we are interested in waves with a frequency of several mHz, we need to perform a nondimensionalization of the system. This will allow us to work with quantities having the same order of magnitude leading to more stable numerical methods. We denote by $\boldsymbol{x}$ the physical coordinates inside the Sun

$$0 \leqslant |\boldsymbol{x}| \leqslant R_\odot,$$

and by $\widetilde{\boldsymbol{x}} = \boldsymbol{x}/R_\odot$ the normalized coordinates used to perform the numerical simulations

$$0 \leqslant |\widetilde{\boldsymbol{x}}| = \frac{|\boldsymbol{x}|}{R_\odot} \leqslant 1.$$

For a quantity f defined in the Sun, we can evaluate it on the mesh using the normalized coordinates

$$\widetilde{f}(\widetilde{\boldsymbol{x}}) := f(\boldsymbol{x}).$$

The derivatives in scaled coordinates can then be evaluated using the chain rule, and we have

$$\nabla_{\boldsymbol{x}} = \frac{1}{R_\odot} \nabla_{\widetilde{\boldsymbol{x}}}, \quad \text{and} \quad \operatorname{div}_{\boldsymbol{x}} = \frac{1}{R_\odot} \operatorname{div}_{\widetilde{\boldsymbol{x}}}.$$

The convected Helmholtz equation

$$-\rho_0 \omega^2 p - 2i\omega\rho_0 \boldsymbol{v_0} \cdot \nabla_{\boldsymbol{x}} p - \operatorname{div}_{\boldsymbol{x}} (\boldsymbol{K_0} \nabla_{\boldsymbol{x}} p) = s,$$

therefore becomes

$$-\rho_0 \omega^2 \widetilde{p} - \frac{2i\omega}{R_\odot} \rho_0 \boldsymbol{v_0} \cdot \nabla_{\widetilde{\boldsymbol{x}}} \widetilde{p} - \frac{1}{R_\odot^2} \operatorname{div}_{\widetilde{\boldsymbol{x}}} (\boldsymbol{K_0} \nabla_{\widetilde{\boldsymbol{x}}} \widetilde{p}) = s,$$

and we multiply by $R_\odot^2$ leading to

$$-\rho_0 \widetilde{\omega}^2 \widetilde{p} - 2i\widetilde{\omega}\rho_0 \boldsymbol{v_0} \cdot \nabla_{\widetilde{\boldsymbol{x}}} \widetilde{p} - \operatorname{div}_{\widetilde{\boldsymbol{x}}} (\boldsymbol{K_0} \nabla_{\widetilde{\boldsymbol{x}}} \widetilde{p}) = \widetilde{s},$$

where $\widetilde{\omega} = R_\odot \omega$ and $\widetilde{s} = R_\odot^2 s$. This multiplication by $R_\odot^2$ allows us to work with the correct nondimensionalization with minimal changes: we only need to modify the frequency in the input file, and we do not have to make any change inside the code.

[1] Values for this model can be obtained here: https://phys.au.dk/~jcd/solar_models/.

Boundary conditions. For solar applications, the acoustic source must be located in the vicinity of the surface, so the Prandtl-Glauert-Lorentz based ABCs of Chapter 4 do not seem to be well-suited for those applications : indeed to use the PGL-based ABCs, we would need to consider a computational domain centered on the source, and the domain of interest would represent only about half of the computational domain. We will therefore use the ABC-PW that selects outgoing plane-waves that are locally orthogonal to the boundary. Even if the quality of this ABC is lower than the quality of the PGL-based ones, using it here should be sufficient as our goal is only to illustrate the robustness of the HDG solver to solar-like configuration. More realistic ABCs for helioseismology were derived in [BCD+17, BFP19] by carefully studying the solar atmosphere when no velocity field is present. However, it is not clear how those ABCs should be extended in the presence of a background flow.

Numerical results. On Figure 5.1, some numerical results obtained with HDG-$\boldsymbol{\sigma}$ method in those settings are depicted. We have used a complex frequency to model the damping

$$\omega \longleftarrow \omega + i\sigma,$$

and we have considered $\omega = 2.6 \cdot 10^6$ after nondimensionalization, which corresponds to about 3mHz in physical values, and three values for σ: $\sigma_1 = 2.6 \cdot 10^5$, $\sigma_2 = 1.2 \cdot 10^5$ and $\sigma_3 = 6.0 \cdot 10^4$ (corresponding to 10%, 5% and 2.5% of ω respectively) after nondimensionalization. As we can see on Figure 5.1, this is quite effective as a higher value for σ leads to more attenuated waves in the domain, however it is not clear if this is a good model for the physical damping process occurring in the Sun. We can also see on those simulations that the acoustic waves do not propagate through the solar core, as predicted by the asymptotic theory. Finally, we would like to mention that we had to use a mesh refined near the surface to account for the rapid variations of the background parameters in this particular area of the Sun. Some elements on the size of the numerical problem are given in Table 5.1, and we can see that simulation ran fast even if the problem is quite large.

Number of triangles	214 562
nnz LU	449 043 667
MUMPS time	11 min
Memory used	13 GiB

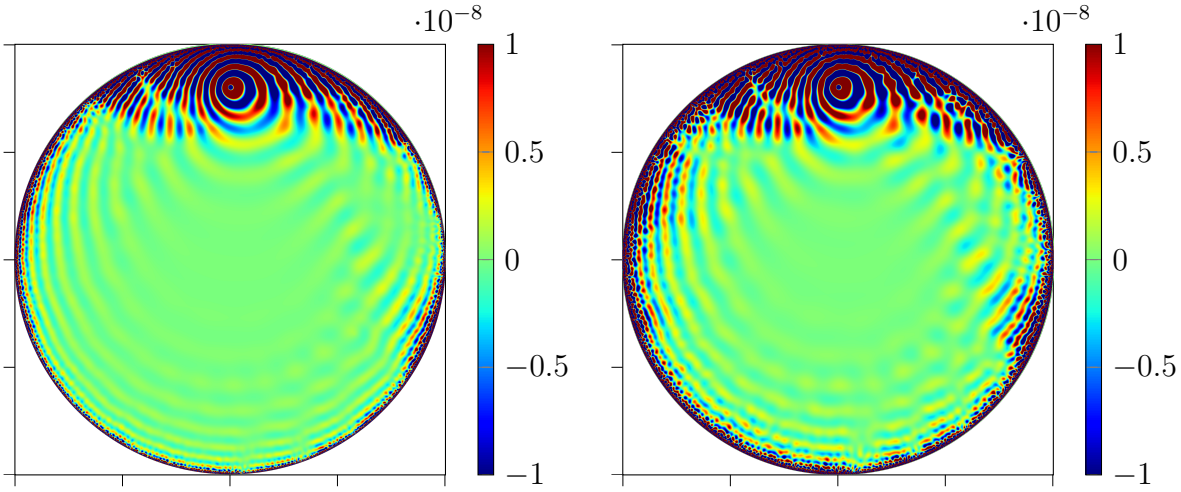
Table 5.1: Size of the numerical problem

5.2 Approximation of Galbrun's equation in the low-regularity settings

As we discussed in Chapter 2, using standard finite-element methods to solve Galbrun's equation in H^1 leads to bad numerical results, due to the presence of spurious modes. In a recent work, [HH20], a well-posedness result was obtained for Galbrun's equation in the low-regularity space

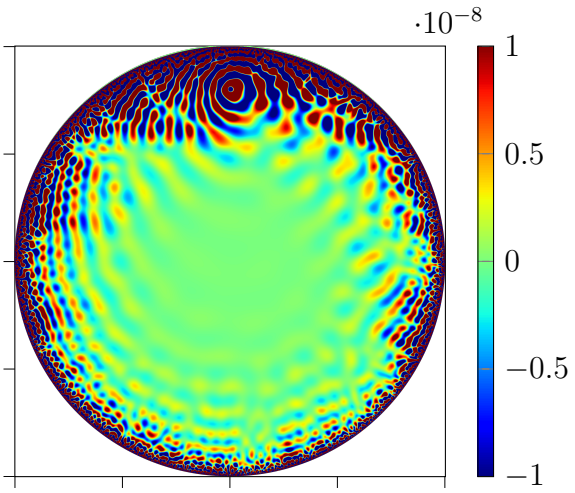
$$\boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O}) := \left\{ \boldsymbol{\xi} \in \boldsymbol{L^2}(\mathcal{O}) \;\middle|\; \operatorname{div}(\boldsymbol{\xi}) \in L^2(\mathcal{O}), \quad \nabla_{\boldsymbol{v_0}} \boldsymbol{\xi} \in \boldsymbol{L^2}(\mathcal{O}) \right\}.$$

As this space does not define the behavior of $\boldsymbol{\xi}$ on the resonant lines that we have computed in Chapter 2, trying to approximate Galbrun's equation in $\boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O})$ instead of $H^1(\mathcal{O})$ could lead to better numerical results. In [HH20] the authors have proven that the variational



(a) Damping σ_1

(b) Damping σ_2



(c) Damping σ_3

Figure 5.1: Numerical results for the Sun.

form associated to Galbrun's equation is *weakly* T*-coercive* on $\boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O})$, which leads to a "Fredholm-like" behavior where existence and unicity are equivalent.
As discussed in Chapter 1, the continuity of a function $\boldsymbol{\xi} \in \boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O})$ on the interface between two elements depends on the interface:

- if $\boldsymbol{v_0} \cdot \boldsymbol{n} = 0$: only the normal part of $\boldsymbol{\xi}$ is continuous (coming from the divergence),
- if $\boldsymbol{v_0} \cdot \boldsymbol{n} \neq 0$: $\boldsymbol{\xi}$ is fully continuous (coming from the directional derivative),

where $\boldsymbol{n}$ is a unitary normal vector to the interface. Because of this non-standard continuity between elements, it seems rather difficult to devise a conforming method for $\boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O})$, leading us to consider HDG methods instead.

5.2.1 Different notions of coercivity

As mentioned in the previous section, the authors of [HH20] used the notion of weak T-coercivity to prove the well-posedness of Galbrun's equation in $\boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O})$. As this notion is not very common, we will investigate it in this section. In particular, we will focus on how weak T-coercivity can be used to devise numerical methods. Let $a : V \times V \longrightarrow \mathbb{C}$ be a continuous sesquilinear form.

Définition 2 : *Coercivity*

We say that

- a is V*-coercive* if
$$\exists \alpha > 0,\ \forall v \in V,\ |a(v,v)| \geqslant \alpha \|v\|_V^2 ,$$
- a is *weakly* V*-coercive* if there exists a compact operator c such that $a + c$ is coercive,
- a is T*-coercive* if
$$\exists \mathsf{T} \in \mathcal{L}(V),\quad \text{bijective},\ \exists \alpha > 0,\ \forall v \in V,\ |a(v,\mathsf{T}v)| \geqslant \alpha \|v\|_V^2 ,$$
- a is *weakly* T*-coercive* if $a(\cdot,\mathsf{T}\cdot)$ is only weakly coercive in the previous definition.

Remark 5.2.1: In the previous definition, we can notice that the weak notions have the same structure: they can be decomposed as the strong one and a compact perturbation.

The notion of T-coercivity was introduced to deal with problems with sign-changing coefficients, see *e.g.* [CC12]. For applications of this notion to more usual wave problems, we refer to [Cia12]. Weak T-coercivity is a less common notion, other examples of weakly T-coercive problems can be found in [Hal19a, Hal19b]. Before further investigating this notion, we recall the Banach-Necas-Babushka (BNB) theorem on the well-posedness of variational problems, for the proof we refer to [EG04, Th. 2.6].

Theorem 10 : *Banach-Necas-Babushka*

Let $a : V \times V \longrightarrow \mathbb{C}$ be a continuous sesquilinear form and $\ell \in V'$, then the problem

$$\text{Seek } v \in V, \text{ such that } a(v,w) = \langle \ell, v \rangle_{V',V} \text{ for all } w \in V,$$

is well-posed if, and only if, a satisfies

- the inf-sup (or stability) condition

$$\exists \alpha > 0,\ \forall v \in V,\ \sup_{w \in V \setminus \{0\}} \frac{|a(v,w)|}{\|w\|_V} \geqslant \alpha \|v\|_V , \tag{5.1}$$

- the unicity condition

$$(\forall w \in V,\ a(v,w) = 0) \implies v = 0.$$

To discretize continuous problems in the framework of Theorem 10, we usually require that the stability condition (5.1) holds uniformly, *i.e.* that there is a sequence of finite-dimensional Hilbert spaces $(V_h)_h$, where the index h is a small parameter that can be interpreted as the mesh size[2] and $V_h \subset V$, such that

$$\exists \alpha_\dagger,\ \exists h_\dagger,\ \forall h \in (0, h_\dagger),\ \forall v_h \in V_h,\quad \sup_{w_h \in V_h \setminus \{0\}} \frac{|a(v_h, w_h)|}{\|w_h\|_V} \geqslant \alpha_\dagger \|v_h\|_V .$$

5.2.2 Numerical approximation of weakly T-coercive problems

In this section, we will focus on the numerical approximation of weakly T-coercive problems using a conforming method where the discrete space V_h is a subset of the continuous space V. Some of the most common examples of conforming numerical methods are *Lagrange finite-elements* used to discretize problems posed in the Sobolev space H^1, and *Hermite finite-elements* that are used for problems in the Sobolev space H^2. However *Discontinuous Galerkin Methods* do not belong to the category of conforming numerical methods.

[2] Even if there is no mesh in this abstract analysis, in practice the space V_h will be associated with a mesh of the domain to perform numerical simulations.

Proposition 5.2.1:
Let V be a Hilbert space endowed with its hermitian product $(\cdot,\cdot)_V$ and associated norm $\|\cdot\|_V$. Let $b : V \times V \to \mathbb{C}$ and $c : V \times V \to \mathbb{C}$ be two continuous sesquilinear forms such that

- b is T-coercive,
- c is compact.

Let $a = b+c$ be a continuous sesquilinear form on V, and $\ell \in V'$. We consider the variational problem

$$\text{Seek } v \in V, \text{ such that } a(v,w) = \langle \ell, v\rangle_{V',V} \text{ for all } w \in V. \tag{5.2}$$

Let $(V_h)_h$ be a family of finite-dimensional subsets of V that satisfy the following *approximation property*

$$\forall v \in V, \quad \lim_{h\to 0}\left(\inf_{v_h\in V_h} \|v - v_h\|_V\right) = 0. \tag{5.3}$$

Assuming that

- the continuous problem (5.2) is well-posed, *i.e.* the operator A associated with the sesquilinear form a is invertible with bounded inverse,
- the sesquilinear form b satisfies the following discrete T_h-coercivity property

$$\exists \alpha^\star, \beta^\star > 0,\ \forall h > 0,\ \exists \mathsf{T}_h \in \mathcal{L}(V_h),\ |||\mathsf{T}_h||| \leqslant \beta^\star \text{ and } \forall v_h \in V_h |b(v_h, \mathsf{T}_h v_h)| \geqslant \alpha^\star \|v_h\|_V^2, \tag{5.4}$$

then a satisfies the following *uniform discrete stability condition*

$$\exists \alpha_\dagger > 0,\ \exists h_\dagger > 0,\ \forall h \in (0, h_\dagger),\ \forall v_h \in V_h,\quad \sup_{w_h\in V_h\setminus\{0\}} \frac{|a(v_h, w_h)|}{\|w_h\|_V} \geqslant \alpha_\dagger \|v_h\|_V. \tag{5.5}$$

Proof : [3] This result will be proven by contradiction. We assume that

$$\forall \alpha_\dagger,\ \forall h_\dagger,\ \exists h \in (0, h_\dagger),\ \exists v_h \in V_h,\quad \sup_{w_h\in V_h\setminus\{0\}} \frac{|a(v_h, w_h)|}{\|w_h\|_V} < \alpha_\dagger \|v_h\|_V,$$

which is the negation of (5.5).
We can therefore extract a subsequence $(V_{h_k})_{k\in\mathbb{N}^*}$ of $(V_h)_h$ with $\lim_{k\to+\infty} h_k = 0$, and construct a sequence $(\mu_{h_k})_{k\in\mathbb{N}^*}$ of strictly positive real numbers that tends to 0 such that

$$\forall k \in \mathbb{N}^*,\ \exists v^0_{h_k} \in V_{h_k},\ \left\|v^0_{h_k}\right\|_V = 1 \quad \text{and} \quad \sup_{w_{h_k}\in V_{h_k}\setminus\{0\}} \frac{|a(v^0_{h_k}, w_{h_k})}{\|w_{h_k}\|_V} \leqslant \mu_{h_k}.$$

In the following, we will denote those quantities by $(V_h)_h$, $(\mu_h)_h$ and $(v^0_h)_h$ to make the notations lighter.
Step 1: For $w \in V$, $h > 0$ and $\forall w_h \in V_h$, we have

$$\begin{aligned}
|a(v^0_h, w)| &= \left|a(v^0_h, w_h) + a(v^0_h, w - w_h)\right|, \\
&\leqslant |a(v^0_h, w_h)| + \left(\mathsf{A}v^0_h, w - w_h\right)_V, \\
&\leqslant |a(v^0_h, w_h)| + |||\mathsf{A}||| \underbrace{\left\|v^0_h\right\|_V}_{=1} \|w - w_h\|_V, \\
&\leqslant \mu_h \|w_h\|_V + |||\mathsf{A}||| \|w - w_h\|_V.
\end{aligned}$$

[3] The blueprint for this proof was found in an exam given at ENSTA Paris by P. Ciarlet and A.-S. Bonnet-BenDhia. This result on the approximation of weakly T-coercive problems seems clearer than the ones usually found in the literature.

As this inequality holds for any $w_h \in V_h$, we can chose $w_h = P^{\perp}_{V_h}(w)$ where $P^{\perp}_{V_h}$ denotes the orthogonal projection onto V_h and we therefore have

$$\|w - w_h\|_V = \inf_{v_h \in V_h} \|w - v_h\|_V .$$

Using this with the approximation property (5.3) and using that $\lim_{h\to 0} \mu_h = 0$, we have

$$\forall w \in V, \quad \lim_{h\to 0} a(v_h^0, w) = 0. \tag{5.6}$$

Step 2: Let $w \in V$, as A is bijective so is its adjoint A^* and there exists $\widetilde{w} \in V$ such that $w = \mathsf{A}^*\widetilde{w}$. We therefore have

$$\lim_{h\to 0} \left(v_h^0, w\right)_V = \lim_{h\to 0} \left(v_h^0, \mathsf{A}^*\widetilde{w}\right)_V = \lim_{h\to 0} \left(\mathsf{A}v_h^0, \widetilde{w}\right)_V = 0,$$

where we used (5.6). We can therefore conclude that

$$v_h^0 \underset{h\to 0}{\rightharpoonup} 0.$$

Step 3: Let $h > 0$ and let T_h be the operator defined in (5.4), we have

$$\begin{aligned} |b(v_h^0, \mathsf{T}_h v_h^0)| &= \left|a(v_h^0, \mathsf{T}_h v_h^0) - c(v_h^0, \mathsf{T}_h v_h^0)\right|, \\ &= \left|a(v_h^0, \mathsf{T}_h v_h^0) - \left(\mathsf{C}v_h^0, \mathsf{T}_h v_h^0\right)_V\right| \\ &\leqslant \left(\mu_h + \left\|\mathsf{C}v_h^0\right\|_V\right) \left\|\mathsf{T}_h v_h^0\right\|_V , \end{aligned}$$

where C is the operator associated with the sesquilinear form c.
As C is compact and v_h^0 converges weakly to 0 in V, we know that

$$\lim_{h\to 0} \left\|\mathsf{C}v_h^0\right\|_V = 0.$$

Using this and the fact that $\lim_{h\to 0} \mu_h = 0$, we have

$$\lim_{h\to 0} |b(v_h^0, \mathsf{T}_h v_h^0)| = 0,$$

which leads to a contradiction. Indeed, owing to the discrete T-coercivity of b, we also have

$$|b(v_h^0, \mathsf{T}_h v_h^0) \geqslant \alpha^\star \left\|v_h^0\right\|_V^2 ,$$

where the term in the right-hand side does not tend to 0. ∎

Proposition 5.2.1 can be summarized by saying that if a continuous problem has the structure "T-coercive + compact", then it can be discretized using conforming numerical methods. We would like to point out that the assumption of the discrete T-coercivity for b is not too restrictive in the framework of conforming methods. Indeed, it is proven in [Cia12, Cor. 1] that for those methods, the discrete T-coercivity can be inherited from the continuous one.

Finally, we would like point out that there are two ways to discretize a T-coercive sesquilinear form:

- the *Petrov-Galerkin approach*, where the coercive form $b(\cdot, \mathsf{T}\cdot)$ is discretized and the space of test-functions is different from the space of trial-functions, see [CC12],
- the T-*conform mesh approach*, where the knowledge of T is translated into constraint on the mesh so that $\mathsf{T}V_h \subset V_h$, see [CC12, Sec. 4].

However, it is not clear how those techniques should be translated for "T-coercive + compact" problems. As the proof presented here to derive an inf-sup condition is not constructive, it also seems difficult to use this condition directly to devise numerical methods.

5.2.3 Theoretical gaps for the non-conforming discretization of (weakly) T-coercive problems

In the previous section, we have seen that the conforming discretization (*i.e.* with $\boldsymbol{V}_h \subset V$) of weakly T-coercive problems is possible. However, as we already discussed, devising conforming approximation spaces for $\boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O})$ turns out to be difficult. It seems natural to consider non-conforming discretization methods instead. Indeed, using a (H)DG method to solve Galbrun's equation in $\boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O})$ would allow us to require weakly the continuity constraints associated with this space, leading to a more natural numerical method. To the best of our knowledge, there is no theory allowing us to do that, and we use this section to state some of the open questions that should be answered before using such a numerical method to solve Galbrun's equation:

- With non-conforming methods, the discrete T-coercivity is not inherited from the continuous one and should be proven by hand in a discrete mesh-dependent norm, how is this norm chosen for (H)DG methods ? What is the relationship between the discrete operator T_h and the continuous one T ?
- As the *Petrov-Galerkin approach* seems difficult for weakly T-coercive problems, how can we build (weak T)-conform meshes for those problems ?
- How can we use the (discrete) weak T-coercivity to prove the well-posedness of the local problems of the HDG methods ?
- Proofs of convergence of HDG methods usually rely on some *Aubin-Nitsche*-like techniques and therefore require very regular solutions, can we prove the convergence of HDG methods in $\boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O})$?

Notice that some of those questions are specific to HDG methods as we prefer them over usual DG methods to reduce the cost of the numerical simulations. Using DG methods instead may lead to simpler answers to some of those questions, in particular results on the convergence of DG methods to low-regularity solutions can be found *e.g.* in [PE12, Sec. 4.2.5]. To the best of our knowledge no such results are available for HDG methods and the extension of the DG results to HDG does not seem straightforward[4].

5.3 Towards the construction of a HDG method for Galbrun's equation

Before actually trying to devise a HDG method for Galbrun's equation, we will consider a simpler vectorial wave problem: the *velocity formulation of the Helmholtz equation*. It is interesting to notice that Galbrun's equation reduces to this model when there is no convection and no hydrodynamic terms. We will then present a way to add the convection terms to the HDG method for this model.

5.3.1 Velocity formulation of the Helmholtz equation

First we consider the following equation

$$\begin{aligned} -\rho_0\omega^2\boldsymbol{\xi} - \nabla\left[\rho_0 c_0^2 \mathrm{div}\,(\boldsymbol{\xi})\right] &= \boldsymbol{s}, && \text{in } \mathcal{O}, \\ c_0^2 \mathrm{div}\,(\boldsymbol{\xi}) - i\omega\boldsymbol{\xi}\cdot\boldsymbol{n} &= 0, && \text{on } \partial\mathcal{O}. \end{aligned} \tag{5.7a}$$

[4]Those results are usually proven for SIPDG-like formulations and extended to other DG flavors by proving that the method can be locally rewritten as a primal formulation, for HDG methods this can only be done globally.

To reach a first-order formulation we introduce the *lagrangian pressure*

$$p = -\rho_0 c_0^2 \mathrm{div}\,(\boldsymbol{\xi})\,,$$

leading to the following system

$$\begin{aligned} -\rho_0\omega^2\boldsymbol{\xi} + \nabla p &= \boldsymbol{s}, && \text{in } \mathcal{O},\\ \frac{p}{\rho_0 c_0^2} + \mathrm{div}\,(\boldsymbol{\xi}) &= 0, && \text{in } \mathcal{O},\\ \frac{p}{\rho_0} - i\omega\boldsymbol{\xi}\cdot\boldsymbol{n} &= 0, && \text{on } \partial\mathcal{O}. \end{aligned}$$

In [Cha19, Th 3.2] it is proven that if $\boldsymbol{s} \in \boldsymbol{L^2}(\mathcal{O})$, then $\boldsymbol{\xi} \in \boldsymbol{H}_{\mathrm{div}}(\mathcal{O})$ and $p \in H^1(\mathcal{O})$.

Weak formulation. By multiplying by a test-function and integrating over an element $K \in \mathcal{T}_h$, we obtain the following weak formulation: seek $(\boldsymbol{\xi}, p) \in \boldsymbol{H}_{\mathrm{div}}(K) \times H^1(K)$ such that

$$\begin{aligned} -\omega^2 \int_K \rho_0 \boldsymbol{\xi}\cdot\boldsymbol{r}_h^* \mathrm{d}\boldsymbol{x} + \int_K \nabla p\cdot\boldsymbol{r}_h^*\mathrm{d}\boldsymbol{x} &= \int_K \boldsymbol{s}\cdot\boldsymbol{r}_h^*\mathrm{d}\boldsymbol{x},\\ \int_K \frac{p}{\rho_0 c_0^2} w_h^*\mathrm{d}\boldsymbol{x} - \int_K \boldsymbol{\xi}\cdot\nabla w_h^*\mathrm{d}\boldsymbol{x} + \int_{\partial K} \boldsymbol{\xi}\cdot\boldsymbol{n} w_h^*\mathrm{d}\sigma &= 0, \end{aligned}$$

for all $(\boldsymbol{r}_h, w_h) \in \boldsymbol{H}_{\mathrm{div}}(K) \times H^1(K)$.

Approximation spaces. To reach a HDG formulation we introduce the following approximation spaces

$$\begin{aligned} \boldsymbol{V}_h(K) &:= \boldsymbol{\mathcal{P}}_k(K) && \text{for } \boldsymbol{\xi}_h,\\ W_h(K) &:= \mathcal{P}_{k-1}(K) && \text{for } p_h. \end{aligned}$$

The interpolation degree for $\boldsymbol{\xi}$ is higher than the one for p as $\boldsymbol{\xi}$ is the main unknown of the model. If we assume that $\boldsymbol{\xi}_h \in \boldsymbol{\mathcal{P}}_{k+1}(K)$ it is therefore natural to assume that $p_h := \mathrm{div}\,(\boldsymbol{\xi}_h) \in \mathcal{P}_k(K)$.

Numerical trace. As usual with HDG methods, we also introduce a *numerical trace*. Although it may look at appealing to use an approximation of $\boldsymbol{\xi}\cdot\boldsymbol{n}$ as it would lead to a scalar global problem, this oriented quantity is dual-valued on interior edges, indeed we have

$$\boldsymbol{\xi}\cdot\boldsymbol{n}_+ = -\boldsymbol{\xi}\cdot\boldsymbol{n}_-, \quad \text{on } \mathcal{E}_h^i \ni e = \partial K_- \cap \partial K_+,$$

where the normal continuity of $\boldsymbol{\xi}$ over e was assumed. However, we can notice that

$$(\boldsymbol{\xi}\cdot\boldsymbol{n}_+)\boldsymbol{n}_+ = (\boldsymbol{\xi}\cdot\boldsymbol{n}_-)\boldsymbol{n}_-, \quad \text{on } \mathcal{E}_h^i \ni e = \partial K_- \cap \partial K_+,$$

we therefore introduce a new unknown $\boldsymbol{\lambda}_h$ which approximates $\boldsymbol{\xi}$ on $\mathcal{E}_h$. We will only require normal continuity leading to the following stabilization term

$$i\omega \left\langle \tau \left(P_M(\boldsymbol{\xi}_h)\cdot\boldsymbol{n} - \boldsymbol{\lambda}_h\cdot\boldsymbol{n}\right)\boldsymbol{n}, \boldsymbol{\mu}\right\rangle_{\partial K}.$$

For this new unknown $\boldsymbol{\lambda}_h$ we introduce the following approximation space

$$\boldsymbol{M}_h(e) = \boldsymbol{\mathcal{P}}_{k-1}(e), \quad \forall e \in \mathcal{E}(K).$$

Construction of the local problems. We can now construct the local problems by approximating the weak formulation on an element $K \in \mathcal{T}_h$: seek $(p_h, \boldsymbol{\xi}_h) \in W_h(K) \times \boldsymbol{V}_h(K)$ such that

$$\left(\frac{p_h}{\rho_0 c_0^2}, w_h\right)_K - (\boldsymbol{\xi}_h, \nabla w_h)_K + \langle \boldsymbol{\lambda}_h \cdot \boldsymbol{n}, w_h \rangle_{\partial K} = 0,$$
$$-\omega^2 (\rho_0 \boldsymbol{\xi}_h, \boldsymbol{r}_h)_K + (\nabla p_h, \boldsymbol{r}_h)_K + i\omega \langle \tau (P_M(\boldsymbol{\xi}_h) \cdot \boldsymbol{n} - \boldsymbol{\lambda}_h \cdot \boldsymbol{n}) \boldsymbol{n}, \boldsymbol{r}_h \rangle_{\partial K} = (\boldsymbol{s}, \boldsymbol{r}_h)_K, \quad (5.10a)$$

for all $(w_h, \boldsymbol{r}_h) \in W_h(K) \times \boldsymbol{V}_h(K)$. The boundary term in (5.10a) corresponds to a weak enforcement of Dirichlet boundary conditions on ∂K. This amounts to requiring the normal continuity of $\boldsymbol{\xi}_h$ between two elements. The projection P_M has been added to deal with the different interpolation degrees of $\boldsymbol{\xi}_h$ and $\boldsymbol{\lambda}_h$.

Transmission condition. As $p \in H^1(\mathcal{O})$, we need to enforce the continuity of p_h between two elements. To this end, we introduce the *numerical flux* for p_h

$$\widehat{p_h \boldsymbol{n}} := p_h \boldsymbol{n} + i\omega\tau (P_M(\boldsymbol{\xi}_h) \cdot \boldsymbol{n} - \boldsymbol{\lambda}_h \cdot \boldsymbol{n}) \boldsymbol{n} \in \boldsymbol{M}_h,$$

where $\tau \neq 0$ is a *stabilization parameter*. We have written the flux using $\widehat{p_h \boldsymbol{n}}$ instead of $\widehat{p_h}$ to make things easier when we will add the convection to the model. If one is only interested in a HDG method for the problem without convection, then a scalar flux can be used instead. This flux $\widehat{p_h \boldsymbol{n}}$ satisfies the following *transmission condition*

$$\langle \widehat{p_h \boldsymbol{n}}, \boldsymbol{\mu} \rangle_{\partial \mathcal{T}_h \setminus \Gamma_D} + \langle p_h \boldsymbol{n} + i\omega [\tau \boldsymbol{\xi}_h \cdot \boldsymbol{n} - (\tau + 1) \boldsymbol{\lambda}_h \cdot \boldsymbol{n}] \boldsymbol{n}, \boldsymbol{\mu} \rangle_{\Gamma_D} = 0, \quad (5.11)$$

for all $\boldsymbol{\mu} \in \boldsymbol{M}_h$. Notice that (5.11) enforces the continuity of p_h on the interior edges, indeed we have

$$\langle \widehat{p_h \boldsymbol{n}}, \boldsymbol{\mu} \rangle_{\partial \mathcal{T}_h \setminus \Gamma_D} = \sum_{e \in \mathcal{E}_h^i} \langle [\![\widehat{p_h}]\!], \boldsymbol{\mu} \rangle_e = 0.$$

As $\widehat{p_h \boldsymbol{n}} \in \boldsymbol{M}_h$, we have $[\![\widehat{p_h}]\!] \in \boldsymbol{M}_h$ and therefore $[\![\widehat{p_h}]\!] = \boldsymbol{0}$.
We can notice that (5.11) also enforces the absorbing boundary condition which has been discretized as

$$\widehat{p_h} - i\omega \boldsymbol{\lambda}_h \cdot \boldsymbol{n} = 0.$$

Local solvability. We can now prove the local problems are well-posed.

Proposition 5.3.1:
There is a constant $C_{\text{solv}} > 0$ such that if $\omega h_K < C_{\text{solv}}$ and $\tau \neq 0$, then the local problem is uniquely solvable.

Proof : As the local problems are finite-dimensional, we only need to prove the uniqueness of the solution. If there are two distinct solutions $(p_h^1, \boldsymbol{\xi}_h^1)$ and $(p_h^2, \boldsymbol{\xi}_h^2)$, then their difference $(p_h, \boldsymbol{\xi}_h)$ with $p_h = p_h^2 - p_h^1$ and $\boldsymbol{\xi}_h = \boldsymbol{\xi}_h^2 - \boldsymbol{\xi}_h^1$ satisfies the following system

$$\left(\frac{p_h}{\rho_0 c_0^2}, w_h\right)_K - (\boldsymbol{\xi}_h, \nabla w_h)_K = 0, \quad (5.12a)$$
$$-\omega^2 (\rho_0 \boldsymbol{\xi}_h, \boldsymbol{r}_h)_K + (\nabla p_h, \boldsymbol{r}_h)_K + i\omega \langle \tau P_M(\boldsymbol{\xi}_h) \cdot \boldsymbol{n}, \boldsymbol{r}_h \cdot \boldsymbol{n} \rangle_{\partial K} = 0, \quad (5.12b)$$

with $\boldsymbol{s} = \boldsymbol{\lambda}_h = \boldsymbol{0}$. We will prove the result by contradiction. We therefore assume that the system (5.12a)–(5.12b) has a non-zero solution $(p_h, \boldsymbol{\xi}_h) \in \mathcal{P}_{k-1}(K) \times \boldsymbol{\mathcal{P}}_k(K)$.

Testing (5.12a)–(5.12b) with $w_h = p_h$ and $\boldsymbol{r}_h = \boldsymbol{\xi}_h$, and adding the resulting identities after conjugating (5.12a) leads to

$$\left\| \frac{p_h}{\sqrt{\rho_0 c_0^2}} \right\|_K^2 - \omega^2 \left\| \sqrt{\rho_0} \boldsymbol{\xi}_h \right\|_K^2 + i\omega \left\langle \tau P_M \boldsymbol{\xi}_h \cdot \boldsymbol{n}, P_M \boldsymbol{\xi}_h \cdot \boldsymbol{n} \right\rangle_{\partial K} = 0, \tag{5.13}$$

where we used that

$$\left\langle \tau P_M \boldsymbol{\xi}_h \cdot \boldsymbol{n}, \boldsymbol{\xi}_h \cdot \boldsymbol{n} \right\rangle_{\partial K} = \left\langle \tau P_M \boldsymbol{\xi}_h \cdot \boldsymbol{n}, P_M \boldsymbol{\xi}_h \cdot \boldsymbol{n} \right\rangle_{\partial K}, \quad \text{as } \deg(P_M \boldsymbol{\xi}_h) = k - 1,$$

to obtain the boundary term.
Splitting the energy-like identity (5.13) into its real and imaginary parts, we obtain

$$\mathfrak{Re}: \qquad \left\| \frac{p_h}{\sqrt{\rho_0 c_0^2}} \right\|_K^2 - \omega^2 \left\| \sqrt{\rho_0} \boldsymbol{\xi}_h \right\|_K^2 = 0, \tag{5.14a}$$

$$\mathfrak{Im}: \qquad \left\langle \tau P_M \boldsymbol{\xi}_h \cdot \boldsymbol{n}, P_M \boldsymbol{\xi}_h \cdot \boldsymbol{n} \right\rangle_{\partial K} = 0. \tag{5.14b}$$

As $\tau \neq 0$, (5.14b) leads to $P_M \boldsymbol{\xi}_h \cdot \boldsymbol{n} = 0$ on ∂K.
Going back to (5.12a) and testing with $w_h = 1$ leads to

$$\{p_h\} = 0,$$

where $\{\cdot\}$ denotes the average value over the element. We can therefore use the following Poincaré-Wirtinger inequality

$$\|p_h\|_K = \|p_h - \{p_h\}\|_K \leqslant C h_K \|\nabla p_h\|_K . \tag{5.15}$$

Testing (5.12b) with $\boldsymbol{r}_h = \nabla p_h$ leads to

$$\|\nabla p_h\|_K^2 = \omega^2 \left| (\rho_0 \boldsymbol{\xi}_h, \nabla p_h)_K \right| \leqslant \omega^2 \|\rho_0 \boldsymbol{\xi}_h\|_K \|\nabla p_h\|_K .$$

As we assumed that $p_h \neq 0$ and we have proven that $\{p_h\} = 0$, we have $\nabla p_h \neq \mathbf{0}$ and we can divide by $\|\nabla p_h\|_K$ to obtain

$$\|\nabla p_h\|_K^2 \lesssim \omega^4 \left\| \sqrt{\rho_0} \boldsymbol{\xi}_h \right\|_K^2 , \tag{5.16}$$

where the constant hidden in $\lesssim$ accounts for the changes involving ρ_0.
Going back to (5.14a) we have

$$\begin{aligned} \omega^2 \left\| \sqrt{\rho_0} \boldsymbol{\xi}_h \right\|_K^2 &= \left\| \frac{p_h}{\sqrt{\rho_0 c_0^2}} \right\|_K^2 \\ &\lesssim \|p_h\|_K^2 \\ \text{(by (5.15)} \quad &\lesssim h_K^2 \|\nabla p_h\|_K^2 \\ \text{(by (5.16))} \quad &\lesssim \omega^4 h_K^2 \left\| \sqrt{\rho_0} \boldsymbol{\xi}_h \right\|_K^2 . \end{aligned}$$

As we have assumed that $\boldsymbol{\xi}_h \neq \mathbf{0}$, we can divide by $\omega^2 \left\| \sqrt{\rho_0} \boldsymbol{\xi}_h \right\|_K^2$ to obtain

$$1 \lesssim \omega^2 h_K^2,$$

if we denote by C_{solv}^2 the constant hidden in $\lesssim$, we have

$$\omega^2 h_K^2 \geqslant C_{\text{solv}}^2,$$

which is in contradiction with the assumption that $\omega h_K < C_{\text{solv}}$. We therefore obtained a contradiction and we can conclude that $(p_h, \boldsymbol{\xi}_h) = (0, \mathbf{0})$. The system (5.12a)–(5.12b) is therefore uniquely solvable. ∎

Characterization of the global problem. As we have proven that the local problems are well-posed, we can introduce the so-called *local solvers* that satisfy

$$p_h^K = \mathsf{P}^K(\boldsymbol{\lambda}_h) + \mathsf{P}^K_{\rm src}(\boldsymbol{s}), \quad \text{and} \quad \boldsymbol{\xi}^K = \boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h) + \boldsymbol{\Xi}^K_{\rm src}(\boldsymbol{s}),$$

where $(\mathsf{P}^K(\boldsymbol{\lambda}_h), \boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h))$ is defined by

$$\left(\frac{\mathsf{P}^K(\boldsymbol{\lambda}_h)}{\rho_0 c_0^2}, w_h\right)_K - \left(\boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h), \nabla w_h\right)_K = -\left\langle \boldsymbol{\lambda}_h \cdot \boldsymbol{n}, w_h \right\rangle_{\partial K}, \tag{5.17a}$$

$$-\omega^2 \left(\rho_0 \boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h), \boldsymbol{r}_h\right)_K + \left(\nabla \mathsf{P}^K(\boldsymbol{\lambda}_h), \boldsymbol{r}_h\right)_K = -i\omega \left\langle \tau (P_M \boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h) - \boldsymbol{\lambda}_h) \cdot \boldsymbol{n}, \boldsymbol{r}_h \cdot \boldsymbol{n} \right\rangle_{\partial K} \tag{5.17b}$$

for all $(w_h, \boldsymbol{r}_h) \in W_h(K) \times \boldsymbol{V}_h(K)$ and $(\mathsf{P}^K_{\rm src}(\boldsymbol{s}), \boldsymbol{\Xi}^K_{\rm src}(\boldsymbol{s}))$ is defined by

$$\left(\frac{\mathsf{P}^K_{\rm src}(\boldsymbol{s})}{\rho_0 c_0^2}, w_h\right)_K - \left(\boldsymbol{\Xi}^K_{\rm src}(\boldsymbol{s}), \nabla w_h\right)_K = 0,$$

$$-\omega^2 \left(\rho_0 \boldsymbol{\Xi}^K_{\rm src}(\boldsymbol{s}), \boldsymbol{r}_h\right)_K + \left(\nabla \mathsf{P}^K_{\rm src}(\boldsymbol{s}), \boldsymbol{r}_h\right)_K + i\omega \left\langle \tau (P_M \boldsymbol{\Xi}^K_{\rm src}(\boldsymbol{s}) \cdot \boldsymbol{n}), \boldsymbol{r}_h \cdot \boldsymbol{n} \right\rangle_{\partial K} = (\boldsymbol{s}, \boldsymbol{r}_h)_K,$$

for all $(w_h, \boldsymbol{r}_h) \in W_h(K) \times \boldsymbol{V}_h(K)$.
We want to obtain a characterization of the global problem

$$a_h(\boldsymbol{\lambda}_h, \boldsymbol{\mu}) = \ell(\boldsymbol{\mu}).$$

As they do not depend on $\boldsymbol{\lambda}_h$ the terms involving $\mathsf{P}^K_{\rm src}(\boldsymbol{s})$ and $\boldsymbol{\Xi}^K_{\rm src}(\boldsymbol{s})$ can be moved into $\ell(\boldsymbol{\mu})$, and we will only focus on the terms involving $\mathsf{P}^K(\boldsymbol{\lambda}_h)$ and $\boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h)$.

Proposition 5.3.2:
The bilinear form a_h of the global problem satisfies

$$\begin{aligned} a_h(\boldsymbol{\lambda}_h, \boldsymbol{\mu}) = &\left(\frac{\mathsf{P}^K(\boldsymbol{\lambda}_h)}{\sqrt{\rho_0 c_0^2}}, \frac{\mathsf{P}^K(\boldsymbol{\mu})}{\sqrt{\rho_0 c_0^2}}\right)_{\mathcal{T}_h} - \omega^2 \left(\sqrt{\rho_0}\boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h), \sqrt{\rho_0}\boldsymbol{\Xi}^K(\boldsymbol{\mu})\right)_{\mathcal{T}_h} \\ &+ i\omega \left\langle \tau \left(P_M \boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h) \cdot \boldsymbol{n} - \boldsymbol{\lambda}_h \cdot \boldsymbol{n}\right), P_M \boldsymbol{\Xi}^K(\boldsymbol{\mu}) \cdot \boldsymbol{n} - \boldsymbol{\mu} \cdot \boldsymbol{n} \right\rangle_{\partial \mathcal{T}_h} \\ &+ i\omega \left\langle \boldsymbol{\lambda}_h \cdot \boldsymbol{n}, \boldsymbol{\mu} \cdot \boldsymbol{n} \right\rangle_{\Gamma_D}, \end{aligned}$$

for all $\boldsymbol{\mu} \in \boldsymbol{M}_h$.

Proof : We recall that the transmission condition (5.11) reads

$$\left\langle p_h + i\omega\tau \left(P_M \boldsymbol{\xi}_h \cdot \boldsymbol{n} - \boldsymbol{\lambda}_h \cdot \boldsymbol{n}\right), \boldsymbol{\mu} \cdot \boldsymbol{n} \right\rangle_{\partial \mathcal{T}_h} - i\omega \left\langle \boldsymbol{\lambda}_h \cdot \boldsymbol{n}, \boldsymbol{\mu} \cdot \boldsymbol{n} \right\rangle_{\Gamma_D} = 0,$$

for all $\boldsymbol{\mu} \in \boldsymbol{M}_h$, and we rewrite it as

$$a_h(\boldsymbol{\lambda}_h, \boldsymbol{\mu}) := -\left\langle \mathsf{P}^K(\boldsymbol{\lambda}_h) + i\omega\tau \left(P_M \boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h) \cdot \boldsymbol{n} - \boldsymbol{\lambda}_h \cdot \boldsymbol{n}\right), \boldsymbol{\mu} \cdot \boldsymbol{n} \right\rangle_{\partial \mathcal{T}_h} + i\omega \left\langle \boldsymbol{\lambda}_h \cdot \boldsymbol{n}, \boldsymbol{\mu} \cdot \boldsymbol{n} \right\rangle_{\Gamma_D} = \ell(\boldsymbol{\mu}).$$

Going back to (5.17a) written for an arbitrary $\boldsymbol{\mu} \in \boldsymbol{M}_h$, conjugating it and summing over the elements, we obtain

$$\left(w_h, \frac{\mathsf{P}^K(\boldsymbol{\mu})}{\rho_0 c_0^2}\right)_{\mathcal{T}_h} - \left(\nabla w_h, \boldsymbol{\Xi}^K(\boldsymbol{\mu})\right)_{\mathcal{T}_h} + \left\langle w_h, \boldsymbol{\mu} \cdot \boldsymbol{n} \right\rangle_{\partial \mathcal{T}_h} = 0,$$

testing this last equation with $w_h = \mathsf{P}^K(\boldsymbol{\lambda}_h)$ we obtain

$$\left(\frac{\mathsf{P}^K(\boldsymbol{\lambda}_h)}{\sqrt{\rho_0 c_0^2}}, \frac{\mathsf{P}^K(\boldsymbol{\mu})}{\sqrt{\rho_0 c_0^2}}\right)_{\mathcal{T}_h} - \left(\nabla \mathsf{P}^K(\boldsymbol{\lambda}_h), \boldsymbol{\Xi}^K(\boldsymbol{\mu})\right)_{\mathcal{T}_h} = -\left\langle \mathsf{P}^K(\boldsymbol{\lambda}_h), \boldsymbol{\mu}\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h}.$$

We therefore have

$$\begin{aligned} a_h(\boldsymbol{\lambda}_h, \boldsymbol{\mu}) = &\left(\frac{\mathsf{P}^K(\boldsymbol{\lambda}_h)}{\sqrt{\rho_0 c_0^2}}, \frac{\mathsf{P}^K(\boldsymbol{\mu})}{\sqrt{\rho_0 c_0^2}}\right)_{\mathcal{T}_h} - \left(\nabla \mathsf{P}^K(\boldsymbol{\lambda}_h), \boldsymbol{\Xi}^K(\boldsymbol{\mu})\right)_{\mathcal{T}_h} \\ &- i\omega \left\langle \tau\left(P_M \boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h)\cdot\boldsymbol{n} - \boldsymbol{\lambda}_h\cdot\boldsymbol{n}\right), \boldsymbol{\mu}\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h} \\ &+ i\omega \left\langle \boldsymbol{\lambda}_h\cdot\boldsymbol{n}, \boldsymbol{\mu}\cdot\boldsymbol{n}\right\rangle_{\Gamma_D}. \end{aligned}$$

Going back to (5.17b), testing with $\boldsymbol{r}_h = \boldsymbol{\Xi}^K(\boldsymbol{\mu})$ and summing over the elements, we have

$$\begin{aligned} -\omega^2\left(\sqrt{\rho_0}\boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h), \sqrt{\rho_0}\boldsymbol{\Xi}^K(\boldsymbol{\mu})\right)_{\mathcal{T}_h} + i\omega\left\langle \tau\left(P_M\boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h)\cdot\boldsymbol{n} - \boldsymbol{\lambda}_h\cdot\boldsymbol{n}\right)\boldsymbol{n}, \boldsymbol{\Xi}^K(\boldsymbol{\mu})\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h} \\ = -\left(\nabla\mathsf{P}^K(\boldsymbol{\lambda}_h), \boldsymbol{\Xi}^K(\boldsymbol{\mu})\right)_{\mathcal{T}_h}, \end{aligned}$$

which leads to

$$\begin{aligned} a_h(\boldsymbol{\lambda}_h, \boldsymbol{\mu}) = &\left(\frac{\mathsf{P}^K(\boldsymbol{\lambda}_h)}{\sqrt{\rho_0 c_0^2}}, \frac{\mathsf{P}^K(\boldsymbol{\mu})}{\sqrt{\rho_0 c_0^2}}\right)_{\mathcal{T}_h} - \omega^2\left(\sqrt{\rho_0}\boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h), \sqrt{\rho_0}\boldsymbol{\Xi}^K(\boldsymbol{\mu})\right)_{\mathcal{T}_h} \\ &+ i\omega\left\langle \tau\left(P_M\boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h)\cdot\boldsymbol{n} - \boldsymbol{\lambda}_h\cdot\boldsymbol{n}\right), \boldsymbol{\Xi}^K(\boldsymbol{\mu})\cdot\boldsymbol{n} - \boldsymbol{\mu}\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h} \\ &+ i\omega\left\langle \boldsymbol{\lambda}_h\cdot\boldsymbol{n}, \boldsymbol{\mu}\cdot\boldsymbol{n}\right\rangle_{\Gamma_D}. \end{aligned}$$

The desired characterization is obtained by noticing that

$$\left\langle \tau\left(P_M\boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h)\cdot\boldsymbol{n} - \boldsymbol{\lambda}_h\cdot\boldsymbol{n}\right), \boldsymbol{\Xi}^K(\boldsymbol{\mu})\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h} = \left\langle \tau\left(P_M\boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h)\cdot\boldsymbol{n} - \boldsymbol{\lambda}_h\cdot\boldsymbol{n}\right), P_M\boldsymbol{\Xi}^K(\boldsymbol{\mu})\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h},$$

as $\deg\left[\tau\left(P_M\boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h)\cdot\boldsymbol{n} - \boldsymbol{\lambda}_h\cdot\boldsymbol{n}\right)\right] = k-1$. ∎

We can clearly see that a_h has the expected structure of the discretization of a harmonic wave equation: it is symmetric but can become indefinite if ω is large enough.
If we take $\boldsymbol{\mu} = \boldsymbol{\lambda}_h$ we can also obtain the following Gårding-like inequality

$$\mathfrak{Re}\, a_h(\boldsymbol{\lambda}_h, \boldsymbol{\lambda}_h) + (\omega^2+1)\left\|\sqrt{\rho_0}\boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h)\right\|_{\mathcal{T}_h}^2 \geqslant \left\|\frac{\mathsf{P}^K(\boldsymbol{\lambda}_h)}{\sqrt{\rho_0 c_0^2}}\right\|_{\mathcal{T}_h}^2 + \left\|\sqrt{\rho_0}\boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h)\right\|_{\mathcal{T}_h}^2,$$

where the right-hand side should be interpreted as a weighted $\boldsymbol{H}_{\mathrm{div}}(\mathcal{O})$-norm. This suggests that the global problem is uniquely solvable when ω is not a resonant frequency.

5.3.2 Adding the convection to the numerical method

We will now show a way to add convection to this method. We consider the following equation

$$-\rho_0\omega^2\boldsymbol{\xi} - 2i\omega\rho_0\nabla_{\boldsymbol{v_0}}\boldsymbol{\xi} + \rho_0\nabla_{\boldsymbol{v_0}}[\nabla_{\boldsymbol{v_0}}\boldsymbol{\xi}] - \nabla\left[\rho_0 c_0^2 \mathrm{div}\,(\boldsymbol{\xi})\right] = \boldsymbol{s},$$

instead of (5.7a). Contrary to the scalar case, we cannot combine the two second-order terms into only one anisotropic Laplace operator. We will therefore use a construction similar to

the *extended scheme* of [DS19, Sec. 4.4]. We therefore add an additional unknown to the problem

$$\boldsymbol{q} := \nabla_{\boldsymbol{v_0}} \boldsymbol{\xi}.$$

With this choice, a boundary term involving $\boldsymbol{v_0} \cdot \boldsymbol{n}$ will be added to the problem. Even if boundary terms involving background coefficient should be avoided as much as possible in the context of HDG methods, this one can be handled in a way that is quite similar to the HDG+ scheme for the convected Helmholtz equation.

We obtain the following first-order formulation

$$\begin{aligned} -\rho_0\omega^2\boldsymbol{\xi} - 2i\omega\rho_0\nabla_{\boldsymbol{v_0}}\boldsymbol{\xi} + \rho_0\nabla_{\boldsymbol{v_0}}\boldsymbol{q} + \nabla p &= \boldsymbol{s}, \\ \frac{p}{\rho_0 c_0^2} + \operatorname{div}(\boldsymbol{\xi}) &= 0, \\ \rho_0\boldsymbol{q} - \rho_0\nabla_{\boldsymbol{v_0}}\boldsymbol{\xi} &= \boldsymbol{0}. \end{aligned}$$

In the previous case, a vectorial unknown was used for the numerical trace to ensure the solvability of the method for the velocity formulation of the Helmholtz equation. We can therefore use this vectorial unknown to enforce the full continuity when $\boldsymbol{v_0} \cdot \boldsymbol{n} \neq 0$ without having to add a second skeleton unknown associated with $\boldsymbol{q}$. So even if the size of the local problems is increased when the convection is added to the model, the size of global problem (which is the main problem of the HDG method) does not change.

To reach a HDG formulation we introduce the following approximation spaces

$$\begin{aligned} \boldsymbol{V}_h(K) &:= \boldsymbol{\mathcal{P}}_k(K) && \text{for } \boldsymbol{\xi}_h, \\ W_h(K) &:= \mathcal{P}_{k-1}(K) && \text{for } p_h, \\ \boldsymbol{Q}_h(K) &:= \boldsymbol{\mathcal{P}}_{k-1}(K) && \text{for } \boldsymbol{q}_h. \end{aligned}$$

Depending on the expected regularity for the solution of Galbrun's equation, multiple choices are possible for the flux $\widehat{\boldsymbol{q}_h}$ which has the form

$$\widehat{\boldsymbol{q}_h} = (\text{continuity of } \boldsymbol{q}_h) + (\text{penalization})(P_M\boldsymbol{\xi}_h - \boldsymbol{\lambda}_h),$$

where the first term can be:

- $(\boldsymbol{v_0} \cdot \boldsymbol{n})\boldsymbol{q}_h$ to translate full continuity of $\boldsymbol{q}_h$ on the interfaces where $\boldsymbol{v_0} \cdot \boldsymbol{n} \neq 0$,
- $(\boldsymbol{v_0} \cdot \boldsymbol{n})(\boldsymbol{q}_h \cdot \boldsymbol{n})\boldsymbol{n}$ to translate the normal continuity of $\boldsymbol{q}_h$ on the interfaces where $\boldsymbol{v_0} \cdot \boldsymbol{n} \neq 0$,

and the penalization term can be chosen among:

- the most natural choice $\boldsymbol{v_0} \cdot \boldsymbol{n}$ to enforce the full continuity of $\boldsymbol{\xi}$ on interfaces where $\boldsymbol{v_0} \cdot \boldsymbol{n} \neq 0$,
- some kind of upwinding parameter such as $\max(\boldsymbol{v_0} \cdot \boldsymbol{n}, 0)$ which mimics the τ_{upw} parameter of the scalar case.

In the weak formulation the term $\nabla_{\boldsymbol{v_0}}\boldsymbol{q}$ will appear either as

$$(\nabla_{\boldsymbol{v_0}}\boldsymbol{q}, \boldsymbol{r}_h)_K, \quad \text{or} \quad -(\boldsymbol{q}, \operatorname{div}(\boldsymbol{v_0} \otimes \boldsymbol{r}_h))_K + \langle(\boldsymbol{v_0} \cdot \boldsymbol{n})\boldsymbol{q}, \boldsymbol{r}_h\rangle_{\partial K},$$

and both of those expressions make sense only if we assume that

$$\boldsymbol{q} \in \left\{ \boldsymbol{\xi} \in \boldsymbol{L^2}(\mathcal{O}) \;\middle|\; \nabla_{\boldsymbol{v_0}}\boldsymbol{\xi} \in \boldsymbol{L^2}(\mathcal{O}) \right\}.$$

This is quite similar to what happened in the case without convection: even if we only had $\boldsymbol{\xi} \in \boldsymbol{H}_{\text{div}}(\mathcal{O})$, it was possible to prove that $p = \rho_0 c_0^2 \operatorname{div}(\boldsymbol{\xi}) \in H^1(\mathcal{O})$. However this additional regularity for $\boldsymbol{q}$, or equivalently for $\nabla_{\boldsymbol{v_0}}\boldsymbol{\xi}$, is not proven and we should remain careful when

using it. Indeed, as we have discussed in Chapter 2, requiring too much regularity may lead to spurious oscillations in the numerical results. It therefore seems natural to consider the following *numerical flux* for $\boldsymbol{q}$

$$\widehat{\boldsymbol{q}_h} = (\boldsymbol{v_0}\cdot\boldsymbol{n})\left[\boldsymbol{q}_h + i\omega\tau_{\boldsymbol{q}}(P_M\boldsymbol{\xi}_h - \boldsymbol{\lambda}_h)\right].$$

Notice that we have multiplied the penalization term by $i\omega\tau_{\boldsymbol{q}}$: the multiplication by $i\omega$ was performed to ensure the solvability of the local problems as it was the case for the previous models, and we added the parameter $\tau_{\boldsymbol{q}}$ as HDG methods with different polynomial orders on one element usually require large penalization values to converge.

We notice that

$$\operatorname{div}(\rho_0\boldsymbol{v_0}\otimes\boldsymbol{v}) = \rho_0\nabla_{\boldsymbol{v_0}}\boldsymbol{v} + \operatorname{div}(\rho_0\boldsymbol{v_0})\,\boldsymbol{v},$$

together with the usual *continuity equation*

$$\operatorname{div}(\rho_0\boldsymbol{v_0}) = 0,$$

this leads to

$$\operatorname{div}(\rho_0\boldsymbol{v_0}\otimes\boldsymbol{v}) = \rho_0\nabla_{\boldsymbol{v_0}}\boldsymbol{v},$$

and we therefore have

$$\begin{aligned}(\rho_0\nabla_{\boldsymbol{v_0}}\boldsymbol{\xi}_h,\boldsymbol{v})_K &= -(\boldsymbol{\xi}_h,\operatorname{div}(\rho_0\boldsymbol{v_0}\otimes\boldsymbol{v}))_K + \langle\rho_0(\boldsymbol{v_0}\cdot\boldsymbol{n})\boldsymbol{\xi}_h,\boldsymbol{v}\rangle_{\partial K}\\ &= -(\boldsymbol{\xi}_h,\nabla_{\boldsymbol{v_0}}\boldsymbol{v})_K + \langle\rho_0(\boldsymbol{v_0}\cdot\boldsymbol{n})\boldsymbol{\xi}_h,\boldsymbol{v}\rangle_{\partial K},\end{aligned}$$

introducing the numerical trace $\boldsymbol{\lambda}_h$ into the boundary term, we obtain

$$= -(\boldsymbol{\xi}_h,\nabla_{\boldsymbol{v_0}}\boldsymbol{v})_K + \langle\rho_0(\boldsymbol{v_0}\cdot\boldsymbol{n})\boldsymbol{\lambda}_h,\boldsymbol{v}\rangle_{\partial K}.$$

We can now construct the *local problems:* seek $(p_h,\boldsymbol{q}_h,\boldsymbol{\xi}_h)\in W_h(K)\times\boldsymbol{Q}_h(K)\times\boldsymbol{V}_h(K)$ such that

$$\begin{aligned}\left(\frac{p_h}{\rho_0c_0^2},w_h\right)_K - (\boldsymbol{\xi}_h,\nabla w_h)_K + \langle\boldsymbol{\lambda}_h\cdot\boldsymbol{n},w_h\rangle_{\partial K} &= 0,\\ (\rho_0\boldsymbol{q}_h,\boldsymbol{v})_K + (\boldsymbol{\xi}_h,\rho_0\nabla_{\boldsymbol{v_0}}\boldsymbol{v})_K - \langle\rho_0(\boldsymbol{v_0}\cdot\boldsymbol{n})\boldsymbol{\lambda}_h,\boldsymbol{v}\rangle_{\partial K} &= 0,\\ -\omega^2(\rho_0\boldsymbol{\xi}_h,\boldsymbol{r}_h)_K + 2i\omega(\rho_0\nabla_{\boldsymbol{v_0}}\boldsymbol{\xi},\boldsymbol{r}_h)_K + (\rho_0\nabla_{\boldsymbol{v_0}}\boldsymbol{q}_h,\boldsymbol{r}_h)_K + (\nabla p_h,\boldsymbol{r}_h)_K&\\ +i\omega\langle\tau(P_M(\boldsymbol{\xi}_h)\cdot\boldsymbol{n}-\boldsymbol{\lambda}_h\cdot\boldsymbol{n})\boldsymbol{n} + \tau_{\boldsymbol{q}}(\boldsymbol{v_0}\cdot\boldsymbol{n})(P_M\boldsymbol{\xi}_h-\boldsymbol{\lambda}_h),\boldsymbol{r}_h\rangle_{\partial K} &= (\boldsymbol{s},\boldsymbol{r}_h)_K,\end{aligned}$$

for all $(w_h,\boldsymbol{v},\boldsymbol{r}_h)\in W_h(K)\times\boldsymbol{Q}_h(K)\times\boldsymbol{V}_h(K)$. We can also write the *transmission condition* on the interior edges as

$$\langle p_h\boldsymbol{n} + (\boldsymbol{v_0}\cdot\boldsymbol{n})\boldsymbol{q}_h + i\omega\tau(P_M\boldsymbol{\xi}_h\cdot\boldsymbol{n}-\boldsymbol{\lambda}_h\cdot\boldsymbol{n})\boldsymbol{n} + i\omega\tau_{\boldsymbol{q}}(\boldsymbol{v_0}\cdot\boldsymbol{n})(P_M-\boldsymbol{\lambda}_h) + \widehat{\boldsymbol{q}_h},\boldsymbol{\mu}\rangle_{\partial\mathcal{T}_h\setminus\Gamma_D} = 0,$$

for all $\boldsymbol{\mu}\in\boldsymbol{M}_h$.
Here we have chosen to write the transmission condition on the interior edges only, as the choice of boundary conditions for Galbrun's equation is quite difficult.

Due to the lack of theory that we have discussed earlier, we cannot prove that the choices that we presented here to construct a HDG method for Galbrun's equation are actually the best ones, even if they seem reasonable.

Characterization of the global problem Assuming that the local problems are well-posed, we can introduced the *local solvers*

$$\boldsymbol{\xi}_h^K = \boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h) + \boldsymbol{\Xi}^K_{\mathrm{src}}(\boldsymbol{s}), \quad p_h^K = \mathsf{P}^K(\boldsymbol{\lambda}_h) + \mathsf{P}^K_{\mathrm{src}}(\boldsymbol{s}), \quad \boldsymbol{q}_h^K = \mathsf{Q}^K(\boldsymbol{\lambda}_h) + \mathsf{Q}^K_{\mathrm{src}}(\boldsymbol{s}),$$

where $(\mathsf{P}^K(\boldsymbol{\lambda}_h), \mathsf{Q}^K(\boldsymbol{\lambda}_h), \boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h))$ is defined by

$$\left(\frac{\mathsf{P}^K(\boldsymbol{\lambda}_h)}{\rho_0 c_0^2}, w_h\right)_K - \left(\boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h), \nabla w_h\right)_K = -\left\langle \boldsymbol{\lambda}_h \cdot \boldsymbol{n}, w_h \right\rangle_{\partial K}$$

$$\left(\rho_0 \mathsf{Q}^K(\boldsymbol{\lambda}_h), \boldsymbol{v}\right)_K + \left(\boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h), \rho_0 \nabla_{\boldsymbol{v_0}} \boldsymbol{v}\right)_K = \left\langle \rho_0 (\boldsymbol{v_0} \cdot \boldsymbol{n}) \boldsymbol{\lambda}_h, \boldsymbol{v} \right\rangle_{\partial K}$$

$$\begin{aligned} -\omega^2 \left(\rho_0 \boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h), \boldsymbol{r}_h\right)_K + 2i\omega \left(\rho_0 \nabla_{\boldsymbol{v_0}} \boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h), \boldsymbol{r}_h\right)_K \\ + \left(\rho_0 \nabla_{\boldsymbol{v_0}} \mathsf{Q}^K(\boldsymbol{\lambda}_h), \boldsymbol{r}_h\right)_K + \left(\nabla \mathsf{P}^K(\boldsymbol{\lambda}_h), \boldsymbol{r}_h\right)_K \\ + i\omega \left\langle \tau (P_M \boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h) \cdot \boldsymbol{n} - \boldsymbol{\lambda}_h \cdot \boldsymbol{n}) \boldsymbol{n} + \tau_{\boldsymbol{q}} (\boldsymbol{v_0} \cdot \boldsymbol{n})(P_M \boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h) - \boldsymbol{\lambda}_h), \boldsymbol{r}_h \right\rangle_{\partial K} = 0, \end{aligned}$$

for all $(w_h, \boldsymbol{v}, \boldsymbol{r}_h) \in W_h(K) \times Q_h(K) \times W_h(K)$. The local solvers $\mathsf{P}^K_{\mathrm{src}}(\boldsymbol{s})$, $\mathsf{Q}^K_{\mathrm{src}}(\boldsymbol{s})$ and $\boldsymbol{\Xi}^K_{\mathrm{src}}(\boldsymbol{s})$ are defined as in the previous section, and the terms involving those solvers can be moved to the right-hand-side and we will therefore omit them for the sake of conciseness.
Making only minor changes to the proof of Proposition 5.3.2, we can obtain a characterization

$$a_h(\boldsymbol{\lambda}_h, \boldsymbol{\mu}) = \ell(\boldsymbol{\mu}),$$

of the global problem.

Proposition 5.3.3:
The bilinear form a_h associated with the global problem satisfies

$$\begin{aligned} a_h(\boldsymbol{\lambda}_h, \boldsymbol{\mu}) = &\left(\frac{\mathsf{P}^K(\boldsymbol{\lambda}_h)}{\sqrt{\rho_0 c_0^2}}, \frac{\mathsf{P}^K(\boldsymbol{\mu})}{\sqrt{\rho_0 c_0^2}}\right)_{\mathcal{T}_h} - \left(\sqrt{\rho_0}\mathsf{Q}^K(\boldsymbol{\lambda}_h), \sqrt{\rho_0}\mathsf{Q}^K(\boldsymbol{\mu})\right)_{\mathcal{T}_h} - \omega^2 \left(\sqrt{\rho_0}\boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h), \boldsymbol{\Xi}^K(\boldsymbol{\mu})\right)_{\mathcal{T}_h} \\ &+ 2i\omega \left(\rho_0 \nabla_{\boldsymbol{v_0}} \boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h), \boldsymbol{\Xi}^K(\boldsymbol{\mu})\right)_{\mathcal{T}_h} \\ &+ i\omega \left\langle \tau (P_M \boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h) \cdot \boldsymbol{n} - \boldsymbol{\lambda}_h \cdot \boldsymbol{n}) \boldsymbol{n}, \boldsymbol{\Xi}^K(\boldsymbol{\mu}) - \boldsymbol{\mu} \right\rangle_{\partial \mathcal{T}_h} \\ &- i\omega \left\langle \rho_0 (\boldsymbol{v_0} \cdot \boldsymbol{n})(P_M \boldsymbol{\Xi}^K(\boldsymbol{\lambda}_h) - \boldsymbol{\lambda}_h), \boldsymbol{\Xi}^K(\boldsymbol{\mu}) - \boldsymbol{\mu} \right\rangle_{\partial \mathcal{T}_h}, \end{aligned}$$

for all $\boldsymbol{\mu} \in \boldsymbol{M}_h$.

Once again, we would like to point out that we have neglected the boundary terms as the choice of boundary conditions for the Galbrun's equation is a difficult problem.
Using this characterization we can see the major difficulty arising when trying to solve Galbrun's equation. Indeed, the term

$$\left(\frac{\mathsf{P}^K(\boldsymbol{\lambda}_h)}{\sqrt{\rho_0 c_0^2}}, \frac{\mathsf{P}^K(\boldsymbol{\mu})}{\sqrt{\rho_0 c_0^2}}\right)_{\mathcal{T}_h} - \left(\sqrt{\rho_0}\mathsf{Q}^K(\boldsymbol{\lambda}_h), \sqrt{\rho_0}\mathsf{Q}^K(\boldsymbol{\mu})\right)_{\mathcal{T}_h},$$

is equivalent to

$$\left(\sqrt{\rho_0} c_0 \mathrm{div}\left(\boldsymbol{\xi}_h\right), \sqrt{\rho_0} c_0 \mathrm{div}\left(\boldsymbol{r}_h\right)\right)_{\mathcal{T}_h} - \left(\sqrt{\rho_0}\nabla_{\boldsymbol{v_0}} \boldsymbol{\xi}_h, \sqrt{\rho_0}\nabla_{\boldsymbol{v_0}} \boldsymbol{r}_h\right)_{\mathcal{T}_h},$$

in primal form. This last form does not seem correspond to a coercive operator, but the second term is expected to be smaller than the first one if the flow is subsonic, *i.e.* if $c_0^2 > |\boldsymbol{v_0}|^2$. To illustrate this we recall the main step of the analysis of Galbrun's equation in the regularized framework of [Pey13]. In this framework, the solution $\boldsymbol{\xi}$ is looked for in $\boldsymbol{H}^1(\mathcal{O})$ and a regularization term

$$\left(\sqrt{\rho_0}c_0\mathbf{curl}\left(\boldsymbol{\xi}\right), \sqrt{\rho_0}c_0\mathbf{curl}\left(\boldsymbol{r}_h\right)\right)_{\mathcal{T}_h}$$

is added to the variational formulation. The authors can then prove that

$$\left(\sqrt{\rho_0}c_0\mathrm{div}\left(\boldsymbol{\xi}\right), \sqrt{\rho_0}c_0\mathrm{div}\left(\boldsymbol{r}_h\right)\right)_{\mathcal{T}_h} + \left(\sqrt{\rho_0}c_0\mathbf{curl}\left(\boldsymbol{\xi}\right), \sqrt{\rho_0}c_0\mathbf{curl}\left(\boldsymbol{r}_h\right)\right)_{\mathcal{T}_h} = \left(\sqrt{\rho_0}c_0\nabla\boldsymbol{\xi}, \sqrt{\rho_0}c_0\nabla\boldsymbol{r}_h\right)_{\mathcal{T}_h} + \text{compact perturbation},$$

leading to

$$\left(\sqrt{\rho_0}c_0^2\mathrm{div}\left(\boldsymbol{\xi}\right), \sqrt{\rho_0}c_0\mathrm{div}\left(\boldsymbol{\xi}\right)\right)_{\mathcal{T}_h} - \left(\sqrt{\rho_0}\nabla_{\boldsymbol{v_0}}\boldsymbol{\xi}, \sqrt{\rho_0}\nabla_{\boldsymbol{v_0}}\boldsymbol{\xi}\right)_{\mathcal{T}_h} \geqslant \inf_{\mathcal{O}}(c_0^2 - |\boldsymbol{v_0}|^2)\left\|\nabla\boldsymbol{\xi}\right\|^2.$$

However, this identity cannot be used for the low-regularity settings $\boldsymbol{\xi} \in \boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O})$, and the well-posedness of this discrete formulation should be studied using discrete weak T-coercivity.

Conclusion

In this section we have presented how the numerical tools described in the previous chapters could be used or extended to perform realistic simulations of wave propagation inside the Sun.

It seems possible to use the HDG solver of Chapter 3 to perform numerical simulations in solar-like configurations, even if some more precise boundary conditions that work in the presence of convection would be required to obtain better numerical results.

We have also considered the well-posedness of Galbrun's equation obtained in [HH20]. Even if this result seems to fit quite naturally in the framework of Galerkin methods, it is not quite usable yet. As we discussed, well-posedness results based on T-coercivity or its variants can only be used to devise conforming numerical method. However the space $\boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O})$ in which Galbrun's equation is well-posed is too complicated to devise such methods onto and we have to rely on non-conforming numerical methods instead. We have shown that a HDG method can easily be built for the velocity formulation of the Helmholtz equation, which is essentially Galbrun's equation without convection, that has the usual properties of a discretization of a wave equation. Unfortunately, due to the lack of theory linking T-coercivity and non-conforming methods, as well as some more precise regularity estimates for the solution of Galbrun's equation, it seems very difficult to extend this HDG method to the full Galbrun's equation.

Bibliography

[BCD+17] Hélène Barucq, Juliette Chabassier, Marc Duruflé, Laurent Gizon, and Michael Leguèbe. Atmospheric Radiation Boundary Conditions for the Helmholtz Equation. *ESAIM: Mathematical Modelling and Numerical Analysis*, 2017.

[BFP19] Hélène Barucq, Florian Faucher, and Ha Pham. Outgoing solutions to the scalar wave equation in helioseismology. June 2019.

[CC12] Lucas Chesnel and Patrick Ciarlet. T-coercivity and continuous Galerkin methods: Application to transmission problems with sign changing coefficients. April 2012.

[CD16] Juliette Chabassier and Marc Duruflé. High Order Finite Element Method for solving Convected Helmholtz equation in radial and axisymmetric domains. Application to Helioseismology. Research Report RR-8893, Inria Bordeaux Sud-Ouest, March 2016.

[CDA+96] Jorgen Christensen-Dalsgaard, W. Däppen, S. V. Ajukov, E. R. Anderson, H. M. Antia, S. Basu, V. A. Baturin, G. Berthomieu, B. Chaboyer, S. M. Chitre, A. N. Cox, P. Demarque, J. Donatowicz, W. A. Dziembowski, M. Gabriel, D. O. Gough, D. B. Guenther, J. A. Guzik, J. W. Harvey, F. Hill, G. Houdek, C. A. Iglesias, A. G. Kosovichev, J. W. Leibacher, P. Morel, C. R. Proffitt, J. Provost, J. Reiter, E. J. Rhodes, F. J. Rogers, I. W. Roxburgh, M. J. Thompson, and R. K. Ulrich. The Current State of Solar Modeling. *Science*, 272(5266):1286–1292, May 1996.

[Cha19] T. Chaumont-Frelet. Mixed finite element discretizations of acoustic Helmholtz problems with high wavenumbers. *Calcolo*, 56(4):49, December 2019.

[Cia12] Patrick Ciarlet. T-coercivity: Application to the discretization of Helmholtz-like problems. Technical report, March 2012.

[DS19] Shukai Du and Francisco-Javier Sayas. *An Invitation to the Theory of the Hybridizable Discontinuous Galerkin Method: Projections, Estimates, Tools*. SpringerBriefs in Mathematics. Springer International Publishing, Cham, 2019.

[EG04] Alexandre Ern and Jean-Luc Guermond. *Theory and Practice of Finite Elements*. Applied Mathematical Sciences. Springer-Verlag, New York, 2004.

[GBD+17] Laurent Gizon, Hélène Barucq, Marc Durufle, Chris Hanson, Michael Leguèbe, Aaron Birch, Juliette Chabassier, Damien Fournier, Thorsten Hohage, and Emanuele Papini. Computational helioseismology in the frequency domain: Acoustic waves in axisymmetric solar models with flows. *Astronomy and Astrophysics - A&A*, 600:A35, 2017.

[Hal19a] Martin Halla. Electromagnetic Stekloff eigenvalues: Approximation analysis. *arXiv:1909.00689 [cs, math]*, September 2019.

[Hal19b] Martin Halla. Galerkin approximation of holomorphic eigenvalue problems: Weak T-coercivity and T-compatibility. *arXiv:1908.05029 [cs, math]*, August 2019.

[HH20] Martin Halla and Thorsten Hohage. On the well-posedness of the damped time-harmonic Galbrun equation and the equations of stellar oscillations. *arXiv:2006.07658 [math-ph]*, June 2020.

[PE12] Daniele Antonio Di Pietro and Alexandre Ern. *Mathematical Aspects of Discontinuous Galerkin Methods*. Mathématiques et Applications. Springer-Verlag, Berlin Heidelberg, 2012.

[Pey13] Emilie Peynaud. *Rayonnement Sonore Dans Un Écoulement Subsonique Complexe En Régime Harmonique : Analyse et Simulation Numérique Du Couplage Entre Les Phénomènes Acoustiques et Hydrodynamiques*. Thesis, Toulouse, INSA, June 2013.

Chapter 6

A scalar model with gravity

Contents

Introduction

In a previous chapter, we have constructed HDG methods for the convected Helmholtz equation. In Chapter 5, we have used the HDG methods of Chapter 3 to perform some numerical simulations in solar-like settings. To further illustrate how this method can be useful in the context of helioseismology, have chosen to use it on a scalar model involving gravity coming from *asteroseismology* that was introduced in [LG09]. This will therefore show the possibility to use the HDG framework of Chapter 3 to study more complex problems.

6.1 Total field

The scalar model that we will use in this chapter can take gravity and some of the effects of rotation into account. We begin with recalling the Euler's equations in a rotating frame.

6.1.1 Euler's Equations in a rotating frame

We consider an ideal gas in rotating frame that satisfies the following set of equations

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \boldsymbol{v}) = 0, \tag{6.1}$$

$$\rho\left[\frac{\partial \boldsymbol{v}}{\partial t} + \nabla_{\boldsymbol{v}}\boldsymbol{v} + 2\boldsymbol{\Omega}\times\boldsymbol{v}\right] + \nabla p = \rho\left[\boldsymbol{g} - \boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\boldsymbol{x})\right], \tag{6.2}$$

$$\frac{\partial p}{\partial t} + \boldsymbol{v}\cdot\nabla p + \gamma p \operatorname{div}(\boldsymbol{v}) = 0, \tag{6.3}$$

$$p = \mathcal{P}(\rho). \tag{6.4}$$

Equation (6.4) is an equation of state that is required to close the system. One of the most simple example of such equations is the equation of state for a polytropic ideal gas

$$p_0 = K\rho^\gamma,$$

where K is a positive constant and γ is the *adiabatic index*[1].
The centrifugal force is potential, indeed we have

$$\boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\boldsymbol{x}) = -\nabla\left[\frac{1}{2}|\boldsymbol{\Omega}\times\boldsymbol{x}|^2\right].$$

We can therefore define the *effective gravity*

$$\boldsymbol{g_0} = -\nabla\left[\varphi - \frac{1}{2}|\boldsymbol{\Omega}\times\boldsymbol{x}|^2\right]$$

where φ is the gravitational potential solving Poisson's Equation :

$$\Delta\varphi = 4\pi G\rho$$

6.1.2 Can we neglect the Coriolis force ?

As described in [Tas00, Sec 2.2.3], we can study the influence of the Coriolis force by defining the *Rossby number* Ro as

$$Ro := \frac{\text{Acceleration in the rotating frame}}{\text{Acceleration due to the Coriolis force}} = \frac{|\partial_t\boldsymbol{v} + \nabla_{\boldsymbol{v}}\boldsymbol{v}|}{|2\boldsymbol{\Omega}\times\boldsymbol{v}|}.$$

For high Rossby numbers, the contribution of the Coriolis force to the acceleration is small when compared to the acceleration in the rotating frame, and the Coriolis force can therefore be neglected. In the frequency domain, the Rossby number becomes

$$Ro = \frac{|-i\omega\boldsymbol{v} + \nabla_{\boldsymbol{v}}\boldsymbol{v}|}{|2\boldsymbol{\Omega}\times\boldsymbol{v}|}$$

and it is therefore possible to neglect the Coriolis force when the frequency ω is large enough.

In Section 6.2 we will derive a scalar model for acoustic waves in the presence of gravity. However this model is valid only when the Coriolis force $2\boldsymbol{\Omega}\times\boldsymbol{v}$ can be neglected. This can be done by neglecting the effets of rotation altogether, *i.e.* by setting $\boldsymbol{\Omega} = \boldsymbol{0}$, but also when the frequency ω is large enough. We would like to point out that even if the Coriolis force is neglected, the model will still encompass some of the effects of rotation as the centrifugal force is still present.

[1] We have $\gamma = \frac{c_p}{c_v}$ where c_p is the isobaric (constant pressure) specific heat index and c_v is the isochoric (constant volume) specific heat index.

6.2 Reduction to a scalar model when the Coriolis force can be neglected

To obtain the underlying scalar model, we focus on *eulerian perturbations*

$$\varepsilon q'(\boldsymbol{x},t) = q(\boldsymbol{x},t) - q_0(\boldsymbol{x},t) + \mathcal{O}(\varepsilon), \tag{6.5}$$

for a physical quantity $q \in \{\rho, p, \boldsymbol{v}\}$, and ε being a small parameter. Moreover, we will make the *Cowling approximation*: acoustic wave propagation does not create a perturbation of gravity.
The assumptions required to derive the scalar model are stated in Assumption 9.

Assumption 9 (Scalar model with gravity):
To derive the scalar model with gravity, we assume that

- the Coriolis force can be neglected (no rotation or ω large enough),
- the asymptotic expansion (6.5) holds,
- the *Cowling approximation* holds: $\boldsymbol{g}' = \boldsymbol{0}$,
- the background flow is steady-state:

$$\forall q_0 \in \{\rho_0, p_0\}, \quad \frac{\partial q_0}{\partial t} = 0,$$

- the background velocity field is null: $\boldsymbol{v_0} = \boldsymbol{0}$.

6.2.1 Background flow

As it was stated in Assumption 9, we assume that the background velocity field is null and that the background flow is steady-state. Introducing the asymptotic expansion (6.5) into the Euler's equations (6.1), (6.2) et (6.3) and identifying the zeroth order terms in ε, we obtain the equations satisfied by the background flow.
They consist of to the following set of steady-state Euler's equations

$$\nabla p_0 = \rho_0 \boldsymbol{g_0} = \rho_0 \nabla \left[\varphi_0 + \frac{1}{2} |\boldsymbol{\Omega} \times \boldsymbol{x}|^2 \right], \tag{6.6}$$

$$p_0 = \mathcal{P}_0(\rho_0), \tag{6.7}$$

$$\Delta \varphi_0 = 4\pi G \rho_0.$$

It is important to note that if $\boldsymbol{v_0} = \boldsymbol{0}$ and

$$\frac{\partial \rho_0}{\partial t} = 0,$$

then (6.1) degenerates to

$$\operatorname{div}(\boldsymbol{0}) = 0,$$

and there is therefore no further constraints on ρ_0.
We can now define the adiabatic sound speed c_0 using the equation of state (6.7): the square of the adiabatic sound speed is given by

$$c_0^2 = \frac{\partial \mathcal{P}_0}{\partial \rho_0},$$

and we have

$$\rho_0 c_0^2 = \gamma p_0.$$

6.2.2 Linearized Euler's Equations without the Coriolis force

Introducing the asymptotic expansion (6.5) into the Euler's equations (6.1), (6.2) et (6.3) and identifying the first order terms in ε, we obtain the *Linearized Euler's Equations* that describe the acoustic perturbation.

$$\frac{\partial \rho'}{\partial t} + \text{div}\,(\rho_0 \boldsymbol{v}') = 0 \tag{6.8}$$

$$\rho_0 \frac{\partial \boldsymbol{v}'}{\partial t} = -\nabla p' + \rho' \boldsymbol{g_0} \tag{6.9}$$

$$\frac{\partial p'}{\partial t} + \boldsymbol{v}' \cdot \nabla p_0 + \gamma p_0 \text{div}\,(\boldsymbol{v}') = 0 \tag{6.10}$$

Following [LG09], we will now derive a single scalar equation from the system (6.8)–(6.9)–(6.10).
We begin by rewriting (6.10) in a more convenient way

$$\begin{aligned}
\gamma p_0 \text{div}\,(\boldsymbol{v}') &= \rho_0 c_0^2 \text{div}\,(\boldsymbol{v}')\,, \\
&= c_0^2 \left[\text{div}\,(\rho_0 \boldsymbol{v}') - \boldsymbol{v}' \cdot \nabla \rho_0\right], \\
&= c_0^2 \left[-\frac{\partial \rho'}{\partial t} - \boldsymbol{v}' \cdot \nabla \rho_0\right],
\end{aligned}$$

so we finally have

$$\frac{\partial p'}{\partial t} + \boldsymbol{v}' \cdot \nabla p_0 = c_0^2 \left[\frac{\partial \rho'}{\partial t} + \boldsymbol{v}' \cdot \nabla \rho_0\right] \tag{6.11}$$

Élimination of $\boldsymbol{v}'$. We begin by reordering the terms in (6.11)

$$\begin{aligned}
\frac{\partial p'}{\partial t} + \boldsymbol{v}' \cdot \nabla p_0 &= c_0^2 \left[\frac{\partial \rho'}{\partial t} + \boldsymbol{v}' \cdot \nabla \rho_0\right], \\
\frac{\partial p'}{\partial t} &= c_0^2 \frac{\partial \rho'}{\partial t} + \boldsymbol{v}' \cdot \left(c_0^2 \nabla \rho_0 - \nabla p_0\right), \\
\frac{\partial p'}{\partial t} &= c_0^2 \frac{\partial \rho'}{\partial t} + \rho_0 c_0^2 \boldsymbol{v}' \cdot \left(\frac{\nabla \rho_0}{\rho_0} - \frac{\nabla p_0}{\rho_0 c_0^2}\right),
\end{aligned}$$

and we obtain

$$\boxed{\frac{\partial p'}{\partial t} = c_0^2 \frac{\partial \rho'}{\partial t} + \rho_0 c_0^2 \boldsymbol{v}' \cdot \left(\frac{\nabla \rho_0}{\rho_0} - \frac{\nabla p_0}{\gamma p_0}\right)} \tag{6.12}$$

Then, (6.12) is derived with respect to time, leading to

$$\frac{\partial^2 p'}{\partial t^2} = c_0^2 \frac{\partial^2 \rho'}{\partial t^2} + c_0^2 \rho_0 \frac{\partial \boldsymbol{v}'}{\partial t} \cdot \left(\frac{\nabla \rho_0}{\rho_0} - \frac{\nabla p_0}{\gamma p_0}\right).$$

The derivative of $\boldsymbol{v}'$ can now be eliminated thanks to (6.9)

$$\frac{\partial^2 p'}{\partial t^2} + c_0^2 \left(\frac{\nabla \rho_0}{\rho_0} - \frac{\nabla p_0}{\rho_0 c_0^2}\right) \cdot \nabla p' = c_0^2 \left[\frac{\partial^2 \rho'}{\partial t^2} + N_0^2 \rho'\right],$$

where N_0^2 is the *Brunt-Väisälä* or *buyoancy frequency* defined by

$$N_0^2 = \boldsymbol{g_0} \cdot \left(\frac{\nabla \rho_0}{\rho_0} - \frac{\nabla p_0}{\gamma p_0}\right).$$

We would like to point out that even if it is a squared quantity, N_0^2 can be negative. We already discussed this phenomenon when we described the g-modes, or internal gravity waves, in the Introduction as g-modes can only propagate when $N_0^2 > 0$. We recall the explanation for this phenomenon given in [**?**]:

> When a fluid element is displaced upwards in an adiabatic motion, its behavior depends on the density of its new surroundings. If $N_0^2 > 0$, the element is heavier than the fluid and buoyancy forces it back into its original position leading to an oscillatory motion. On the other hand if $N_0^2 < 0$, the element is lighter than the fluid and buoyancy acts to enhance the motion.

Using the background equations (6.6)–(6.7), we notice that

$$\frac{\nabla\rho_0}{\rho_0} - \frac{\nabla p_0}{\gamma p_0} = \boldsymbol{g_0} \cdot \left(\frac{\nabla\rho_0}{\rho_0} - \frac{\nabla p_0}{\gamma p_0}\right) \frac{\boldsymbol{g_0}}{|\boldsymbol{g_0}|^2},$$

and we obtain

$$\boxed{\frac{\partial^2 p'}{\partial t^2} + \frac{c_0^2 N_0^2}{|\boldsymbol{g_0}|^2} \boldsymbol{g_0} \cdot \nabla p' = c_0^2 \left[\frac{\partial^2 \rho'}{\partial t^2} + N_0^2 \rho'\right]}. \tag{6.13}$$

Taking the derivative of (6.8) with respect to time and using (6.9) leads to

$$\boxed{\frac{\partial^2 \rho'}{\partial t^2} + \operatorname{div}\left(\rho' \boldsymbol{g_0}\right) = \Delta p'} \tag{6.14}$$

Time-harmonic solutions. In order to obtain a single scalar equation, we consider time-harmonic solutions

$$q(\boldsymbol{x}, t) = \mathfrak{Re}\left(q(\boldsymbol{x}) e^{-i\omega t}\right).$$

The two final equations of the previous paragraph (6.14) and (6.13) therefore become

$$-\omega^2 \rho' + \operatorname{div}\left(\rho' \boldsymbol{g_0}\right) = \Delta p', \tag{6.15}$$

$$c_0^2 \left[N_0^2 - \omega^2\right] \rho' = -\omega^2 p' + \frac{c_0^2 N_0^2}{|\boldsymbol{g_0}|^2} \boldsymbol{g_0} \cdot \nabla p'. \tag{6.16}$$

From (6.16) we obtain

$$\boxed{\rho' = \frac{-\omega^2}{c_0^2 \left(N_0^2 - \omega^2\right)} p' + \frac{N_0^2}{|\boldsymbol{g_0}|^2 \left(N_0^2 - \omega^2\right)} \boldsymbol{g_0} \cdot \nabla p'}, \tag{6.17}$$

and using (6.17) in (6.15) leads to

$$\left[\frac{\omega^4}{c_0^2 \left(N_0^2 - \omega^2\right)} - \operatorname{div}\left(\frac{\omega^2}{c_0^2 \left(N_0^2 - \omega^2\right)} \boldsymbol{g_0}\right)\right] p' + \frac{N_0^2}{|\boldsymbol{g_0}|^2 \left(N_0^2 - \omega^2\right)} \left(\boldsymbol{g_0} \cdot \nabla\right)^2 p'$$
$$+ \left[\frac{-\omega^2}{\left(N_0^2 - \omega^2\right)} \left(\frac{1}{c_0^2} + \frac{N_0^2}{|\boldsymbol{g_0}|^2}\right) + \operatorname{div}\left(\frac{N_0^2}{|\boldsymbol{g_0}|^2 \left(N_0^2 - \omega^2\right)} \boldsymbol{g_0}\right)\right] \boldsymbol{g_0} \cdot \nabla p' = \Delta p'.$$

We notice that

$$\frac{N_0^2}{|\boldsymbol{g_0}|^2 \left(N_0^2 - \omega^2\right)} \left(\boldsymbol{g_0} \cdot \nabla\right)^2 p' = \operatorname{div}\left(\frac{N_0^2}{|\boldsymbol{g_0}|^2 \left(N_0^2 - \omega^2\right)} \boldsymbol{g_0} \boldsymbol{g_0}^T \nabla p'\right) - \operatorname{div}\left(\frac{N_0^2}{|\boldsymbol{g_0}|^2 \left(N_0^2 - \omega^2\right)} \boldsymbol{g_0}\right) \boldsymbol{g_0} \cdot \nabla p',$$

so the scalar equation becomes

$$\left[\frac{\omega^4}{c_0^2\left(N_0^2-\omega^2\right)}-\operatorname{div}\left(\frac{\omega^2}{c_0^2\left(N_0^2-\omega^2\right)}\boldsymbol{g_0}\right)\right]p'$$
$$-\frac{\omega^2}{N_0^2-\omega^2}\left(\frac{1}{c_0^2}+\frac{N_0^2}{|\boldsymbol{g_0}|^2}\right)\boldsymbol{g_0}\cdot\nabla p'-\operatorname{div}\left(\boldsymbol{K_0}\nabla p'\right)=0,$$

where the anisotropy tensor $\boldsymbol{K_0}$ is defined by

$$\boldsymbol{K_0}:=\mathbf{Id}-\frac{N_0^2}{|\boldsymbol{g_0}|^2\left(N_0^2-\omega^2\right)}\boldsymbol{g_0}\boldsymbol{g_0}^T.$$

As discussed in CHAPTER 3, it is convenient to write the convective term $\boldsymbol{g_0}\cdot\nabla p'$ in divergence form to derive HDG methods. In this regard, we have

$$-\frac{\omega^2}{c_0^2\left(N_0^2-\omega^2\right)}\left(\frac{1}{c_0^2}+\frac{N_0^2}{|\boldsymbol{g_0}|^2}\right)\nabla p'=-\operatorname{div}\left(\frac{\omega^2}{\left(N_0^2-\omega^2\right)}\left(\frac{1}{c_0^2}+\frac{N_0^2}{|\boldsymbol{g_0}|^2}\right)p'\boldsymbol{g_0}\right)$$
$$+p'\operatorname{div}\left(\frac{\omega^2}{\left(N_0^2-\omega^2\right)}\left(\frac{1}{c_0^2}+\frac{N_0^2}{|\boldsymbol{g_0}|^2}\right)\boldsymbol{g_0}\right),$$

so the scalar model becomes

$$\omega^2\left[\frac{\omega^2}{c_0^2\left(N_0^2-\omega^2\right)}+\operatorname{div}\left(\frac{N_0^2}{|\boldsymbol{g_0}|^2\left(N_0^2-\omega^2\right)}\boldsymbol{g_0}\right)\right]p'$$
$$-\operatorname{div}\left(\boldsymbol{K_0}\nabla p'+\frac{\omega^2}{N_0^2-\omega^2}\left(\frac{1}{c_0^2}+\frac{N_0^2}{|\boldsymbol{g_0}|^2}\right)p'\boldsymbol{g_0}\right)=0.$$

Finally we end up with the following model:

Proposition 6.2.1:
In the time-harmonic domain, the *eulerian pressure pertubation* satisfies either

- the equation in *diffusive flux formulation*

$$\left[\frac{\omega^4}{c_0^2\left(N_0^2-\omega^2\right)} - \operatorname{div}\left(\frac{\omega^2}{c_0^2\left(N_0^2-\omega^2\right)}\boldsymbol{g_0}\right)\right]p' - \frac{\omega^2}{N_0^2-\omega^2}\left(\frac{1}{c_0^2}+\frac{N_0^2}{|\boldsymbol{g_0}|^2}\right)\boldsymbol{g_0}\cdot\nabla p' - \operatorname{div}\left(\boldsymbol{K_0}\nabla p'\right) = 0, \tag{6.18a}$$

- or the equation in *total flux formulation*

$$\omega^2\left[\frac{\omega^2}{c_0^2\left(N_0^2-\omega^2\right)} + \operatorname{div}\left(\frac{N_0^2}{|\boldsymbol{g_0}|^2\left(N_0^2-\omega^2\right)}\boldsymbol{g_0}\right)\right]p' - \operatorname{div}\left(\boldsymbol{K_0}\nabla p' + \frac{\omega^2}{N_0^2-\omega^2}\left(\frac{1}{c_0^2}+\frac{N_0^2}{|\boldsymbol{g_0}|^2}\right)p'\boldsymbol{g_0}\right) = 0. \tag{6.18b}$$

The *eulerian density perturbation* satisfies

$$\rho' = \frac{-\omega^2}{c_0^2\left(N_0^2-\omega^2\right)}p' + \frac{N_0^2}{|\boldsymbol{g_0}|^2\left(N_0^2-\omega^2\right)}\boldsymbol{g_0}\cdot\nabla p',$$

and the *eulerian velocity perturbation* satisfies

$$\boldsymbol{v}' = \frac{\nabla p'}{i\omega\rho_0} - \frac{\rho'}{i\omega\rho_0}\boldsymbol{g_0}.$$

The tuple $(\rho', \boldsymbol{v}', p')$ is a solution to the Linearized Euler's Equations (6.8)–(6.9)–(6.10) in time-harmonic domain.

Remark 6.2.1: In [LG09], where this model was introduced, the authors did not directly solve (6.18a) or (6.18b). Instead, they considered the asymptotic behaviour when $\omega^2 \gg N_0^2$ and used the *WKB approximation*

$$p' = \mathfrak{Re}\left(Ae^{i\theta(\boldsymbol{x})-i\omega t}\right),$$

to obtain an *eikonal equation*, that can be solved using *ray methods*.

Relationship with convected Helmholtz equation. We would like to point out that (6.18a) is similar to the convected Helmholtz equation

$$-\rho_0\omega^2 p - 2i\omega\rho_0\boldsymbol{v_0}\cdot\nabla p - \operatorname{div}\left(\boldsymbol{K_0}\nabla p\right) = s.$$

The HDG framework that have constructed in CHAPTER 3 is therefore well-suited to perform some numerical simulations using this model. To emphasize the relationship between these two equations, we can make the following approximations:

- *Approximation 1:* if we assume that $N_0^2 = 0$ (which is relevant in the convective zone, *i.e.* when $r > 0.7R_\odot$ where $R_\odot$ is the solar radius, then (6.18a) becomes

$$\left[-\omega^2 + c_0^2\operatorname{div}\left(\frac{1}{c_0^2}\boldsymbol{g_0}\right)\right]p' + \boldsymbol{g_0}\cdot\nabla p' - c_0^2\Delta p' = 0.$$

- *Approximation 2:* if we consider the asymptotic regime in ω, *i.e.* if we assume that $\omega^2 \gg N_0^2$, then we have

$$\frac{\omega^2}{N_0^2-\omega^2} \simeq -1, \quad \text{and} \quad \frac{N_0^2}{N_0^2-\omega^2} \simeq -\frac{N_0^2}{\omega^2},$$

and (6.18a) therefore becomes

$$\left[-\frac{\omega^2}{c_0^2} + \operatorname{div}\left(\frac{1}{c_0^2}\boldsymbol{g_0}\right)\right] p' + \left(\frac{1}{c_0^2} + \frac{N_0^2}{|\boldsymbol{g_0}|^2}\right) \boldsymbol{g_0}\cdot\nabla p' - \operatorname{div}\left(\left(\mathbf{Id} + \frac{N_0^2}{\omega^2|\boldsymbol{g_0}|^2}\boldsymbol{g_0}\boldsymbol{g_0}^T\right)\nabla p'\right) = 0$$

where the term involving gravity in the second-order operator is small.

However, we would like to point out one important difference in the physics involved within those two models. In CHAPTER 3, the waves were convected by a velocity field $\boldsymbol{v_0}$ that satisfied the local steady-state mass conservation equation

$$\operatorname{div}(\rho_0 \boldsymbol{v_0}) = 0.$$

In (6.18a) on the other hand, the waves are convected by the gravity field $\boldsymbol{g_0}$ for which this mass conservation property does not hold.

Source term. Adding a source term $\boldsymbol{s}$ in the RHS of (6.9) leads to the following source term s in (6.18a) or (6.18b)

$$s := -\frac{\rho_0\omega^2 N_0^2}{|\boldsymbol{g_0}|^2\,(N_0^2-\omega^2)}\boldsymbol{g_0}\cdot\boldsymbol{s} + \operatorname{div}\left(\frac{\rho_0 c_0^2 N_0^2}{|\boldsymbol{g_0}|^2\,(N_0^2-\omega^2)}(\boldsymbol{g_0}\cdot\boldsymbol{s})\boldsymbol{g_0}\right) + \operatorname{div}(\boldsymbol{s}), \tag{6.19}$$

where

- the first term comes from $-\omega^2\rho'$ term in (6.15),
- the second term comes from $\operatorname{div}(\rho'\boldsymbol{g_0})$ term in (6.15),
- the third term comes from $\operatorname{div}(\rho_0\boldsymbol{v'})$ term in (6.8).

We will assume that the source term is compactly supported in $\mathcal{O}$ and regular enough for (6.19) to be well-defined.
If the support of $\boldsymbol{s}$ is entirely included in the convective zone ($r > 0.7$), it may be reasonable to assume that $N_0^2 \simeq 0$ and to use

$$s \simeq \operatorname{div}(\boldsymbol{s})$$

instead of (6.19).

Dimensional analysis. As the model described in PROPOSITION 6.2.1 is not a standard, it seems important to check the homogeneity of (6.18b). We denote the dimension of a physical quantity q by $[q]$, and the dimension of mass, time and length by M, T and L respectively, finally we use $\mathsf{1}$ for physical quantities without dimension. For the background parameters we have

$$[\omega] = \mathsf{T}^{-1}, \quad [\rho_0] = \mathsf{M}\mathsf{L}^{-3}, \quad [c_0] = \mathsf{L}\mathsf{T}^{-1}, \quad [p_0] = \mathsf{M}\mathsf{L}^{-1}\mathsf{T}^{-2}, \quad [\boldsymbol{g_0}] = \mathsf{L}\mathsf{T}^{-2},$$

where $\boldsymbol{g_0}$ is homogenous to the *gravitational acceleration*. We recall that the dimension of the derivatives of a physical quantity q is

$$\left[\frac{\partial q}{\partial x}\right] = \mathsf{L}^{-1}[q].$$

First we notice that

$$\left[N_0^2\right] = [\boldsymbol{g_0}]\frac{[\nabla\rho_0]}{[\rho_0]} = \mathsf{L}\mathsf{T}^{-2}\frac{\mathsf{M}\mathsf{L}^{-4}}{\mathsf{M}\mathsf{L}^{-3}} = \mathsf{T}^{-2},$$

so N_0^2 is homogenous to a squared frequency as expected. To lighten the notations, we rewrite (6.18b) as

$$Ap' - \operatorname{div}\left(\boldsymbol{K_0}\nabla p' - Bp'\boldsymbol{g_0}\right) = s.$$

We can now focus on the reaction term. We have

$$\left[\frac{\omega^4}{c_0^2\left(N_0^2-\omega^2\right)}\right] = \frac{[\omega^2]}{[c_0^2]} = \mathsf{L}^{-2}, \quad \text{as} \quad \left[\frac{\omega^2}{\left(N_0^2-\omega^2\right)}\right] = 1,$$

and we also have

$$\left[\omega^2\operatorname{div}\left(\frac{N_0^2}{|\boldsymbol{g_0}|^2\left(N_0^2-\omega^2\right)}\boldsymbol{g_0}\right)\right] = \mathsf{T}^{-2}\mathsf{L}^{-1}[\boldsymbol{g_0}]^{-1} = \mathsf{L}^{-2},$$

so we have

$$[Ap'] = \mathsf{L}^{-2}[p'].$$

We then work on the Laplace operator. We begin by noticing that

$$[\boldsymbol{K_0}] = 1,$$

so we have

$$[\operatorname{div}\left(\boldsymbol{K_0}\nabla p'\right)] = \mathsf{L}^{-2}[p'].$$

We can now focus on the convection term, and we have

$$\left[\frac{N_0^2}{|\boldsymbol{g_0}|^2}\right] = \frac{1}{\mathsf{T}^2}\frac{\mathsf{T}^4}{\mathsf{L}^2} = \mathsf{T}^2\mathsf{L}^{-2} = \left[\frac{1}{c_0^2}\right] = [B],$$

so we have

$$[B\boldsymbol{g_0}] = \mathsf{T}^2\mathsf{L}^{-2}\mathsf{L}\mathsf{T}^{-2} = \mathsf{L}^{-1},$$

and therefore

$$[\operatorname{div}\left(Bp'\boldsymbol{g_0}\right)] = \mathsf{L}^{-2}[p'].$$

We can finally conclude this dimensional analysis. Indeed, we have

$$[Ap'] = [\operatorname{div}\left(\boldsymbol{K_0}\nabla p'\right)] = [\operatorname{div}\left(Bp'\boldsymbol{g_0}\right)] = \mathsf{L}^{-2}[p'],$$

so (6.18b) is actually homogeneous.

6.2.3 Well-posedness of the scalar model

In this section, we recall some elements of the theory of elliptic partial differential equations to prove the well-posedness of the scalar model of Proposition 6.2.1. This well-posedness result will be proven under the following assumptions.

Assumption 10 (Well-posedness of the scalar model with gravity):
We assume that

- the frequency is large enough: $\omega^2 > N_0^2$,
- the following boundary conditions are enforced

$$\begin{aligned} p' &= \varphi_1, && \text{on } \Gamma_D, \\ \left[\boldsymbol{K_0}\nabla p' + \frac{\omega^2}{\left(N_0^2-\omega^2\right)}\left(\frac{1}{c_0^2}+\frac{N_0^2}{|\boldsymbol{g_0}|^2}\right)p'\boldsymbol{g_0}\right]\cdot\boldsymbol{n} &= \varphi_2, && \text{on } \Gamma_N, \end{aligned}$$

with $\partial\mathcal{O} = \Gamma_D \cup \Gamma_N$ and $\Gamma_D \neq \varnothing$.

A well-posedness result for elliptic PDEs

In this section, we state a well-posedness result for elliptic PDEs. For the sake of simplicity, we consider only homogenous Dirichlet boundary conditions, but the extension to the boundary conditions of Assumption 10 is immediate and can be found *e.g.* in [GT01, Ch. 8 – Notes]

Let $\mathcal{O}$ be an open bounded subset of $\mathbb{R}^n$, a_{ij} be $\mathcal{C}^1(\overline{\mathcal{O}})$ functions and a_i be $\mathcal{C}(\overline{\mathcal{O}})$ functions. We consider the following weak formulation

$$\text{Seek } p' \in H_0^1(\mathcal{O}), \quad \text{s.t} \quad a(p', q) = \ell(q) \quad \forall q \in H_0^1(\mathcal{O}). \tag{6.20}$$

with

$$a(u,v) := \int_{\mathcal{O}} \sum_{i,j} a_{ij} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} \mathrm{d}x + \int_{\mathcal{O}} \sum_i a_i \frac{\partial u}{\partial x_i} v \mathrm{d}x + \int_{\mathcal{O}} a_0 uv \mathrm{d}x,$$
$$\ell(v) := \int_{\mathcal{O}} f v \mathrm{d}x.$$

We recall the following result coming from [Bre87, Th IX.23].

Theorem 11 :

If the variational problem (6.20) satisfies the following *uniform ellipticity condition*

$$\exists \alpha > 0, \ \forall \boldsymbol{x} \in \mathbb{R}^n, \ \sum_{i,j} a_{ij} x_i x_j \geqslant \alpha |\boldsymbol{x}|, \tag{6.21}$$

then the two following propositions are true:

1. If $f = 0$ then there is a d-dimensional linear subspace of $H_0^1(\mathcal{O})$ of solutions to (6.20).
2. If $f \neq 0$, then there exists a d-dimensional linear subset $F \subset L^2(\mathcal{O})$ such that

$$((6.20) \text{ admits a unique solution in } H_0^1(\mathcal{O})) \iff \forall v \in F, \ (f,g)_{\mathcal{O}} = 0.$$

Remark 6.2.2: We would like to point out that Theorem 11 is nothing but a Fredholm alternative and the problem (6.20) has the same behavior as a harmonic wave equation. For the scalar model of Proposition 6.2.1, this means that eigenmodes can exist for some frequencies.

Remark 6.2.3: If the coefficient a_0 in (6.20) is positive, it is possible to prove that the problem is always uniquely solvable using the *maximum principle*. For the scalar model of Proposition 6.2.1, we cannot guarantee the sign of a_0.

A quick way to prove that the ellipticity condition (6.21) holds is to study the spectrum of $\boldsymbol{A} := (a_{ij})_{ij}$. Indeed, we denote by $\lambda_1 < \cdots < \lambda_n$ the eigenvalues of $\boldsymbol{A}$ and $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ be

the associated basis of eigenvectors, we have

$$\begin{aligned}\sum_{i,j} a_{ij}x_ix_j &= \boldsymbol{A}\boldsymbol{x}\cdot\boldsymbol{x},\\ &= \boldsymbol{A}\left(\sum_m \beta_m\boldsymbol{w}_m\right)\cdot\left(\sum_n \beta_n\boldsymbol{w}_n\right),\\ &= \left(\sum_m \beta_m\lambda_m\boldsymbol{w}_m\right)\cdot\left(\sum_n \beta_n\boldsymbol{w}_n\right),\\ &= \sum_m \lambda_m\beta_m^2,\\ &\geqslant \lambda_-|\boldsymbol{x}|^2,\end{aligned}$$

where $\lambda_- := \min_{\boldsymbol{x}\in\mathcal{O}}\lambda_1$. Notice that for the scalar model of this section, $\boldsymbol{A}$ has only three eigenvalues.

Application to the scalar model of Proposition 6.2.1.

We recall the anisotropy tensor of the scalar model is

$$\boldsymbol{K_0} := \mathbf{Id} - \frac{N_0^2}{|\boldsymbol{g_0}|^2\,(N_0^2-\omega^2)}\boldsymbol{g_0}\boldsymbol{g_0}^T,$$

the second-order part of the primal variational formulation is therefore

$$a(p',q) = \int_{\mathcal{O}} \boldsymbol{K_0}\nabla p'\cdot\nabla q^\star \mathrm{d}\boldsymbol{x} + \text{low-order terms},$$

and we now need to study the spectrum of $\boldsymbol{K_0}$.

Lemma 6.2.1:
The spectrum of $\boldsymbol{g_0}\boldsymbol{g_0}^T$ is

$$\mathrm{Sp}(\boldsymbol{g_0}\boldsymbol{g_0}^T) = \left\{0, |\boldsymbol{g_0}|^2\right\},$$

and the spectrum of $\boldsymbol{K_0}$ therefore is

$$\mathrm{Sp}(\boldsymbol{K_0}) = \left\{1, \frac{\omega^2}{\omega^2-N_0^2}\right\}.$$

Proof:
Step 1: Let $\boldsymbol{v}\in\mathbb{R}^n$ be a vector such that $\boldsymbol{g_0}\perp\boldsymbol{v}$, we have

$$\boldsymbol{g_0}\boldsymbol{g_0}^T\boldsymbol{v} = \mathbf{0}.$$

In dimension n and if $\boldsymbol{g_0}\neq\mathbf{0}$, the space of vectors orthogonal to $\boldsymbol{g_0}$ has dimension $n-1$. The eigenspace

$$E_0(\boldsymbol{g_0}\boldsymbol{g_0}^T) := \left\{\boldsymbol{v}\in\mathbb{R}^n \;\middle|\; \boldsymbol{g_0}\boldsymbol{g_0}^T\boldsymbol{v} = 0\right\}$$

therefore has dimension $n-1$. We also have

$$\boldsymbol{g_0}\boldsymbol{g_0}^T\boldsymbol{g_0} = |\boldsymbol{g_0}|^2\boldsymbol{g_0}.$$

The eigenvalues of $\boldsymbol{g_0}\boldsymbol{g_0}^T$ therefore are 0 and $|\boldsymbol{g_0}|^2$.

Step 2: Let $\boldsymbol{v}$ be an eigenvector of $\boldsymbol{K_0}$ and λ be the associated eigenvalue, we have

$$\begin{aligned}\boldsymbol{K_0 v} = \lambda \boldsymbol{v} &\iff \boldsymbol{v} - \frac{N_0^2}{|\boldsymbol{g_0}|^2 (N_0^2 - \omega^2)} \boldsymbol{g_0 g_0^T v} = \lambda \boldsymbol{v}, \\ &\iff \boldsymbol{g_0 g_0^T v} = (1-\lambda) \frac{|\boldsymbol{g_0}|^2 (N_0^2 - \omega^2)}{N_0^2} \boldsymbol{v},\end{aligned}$$

so we have

$$(1-\lambda) \frac{|\boldsymbol{g_0}|^2 (N_0^2 - \omega^2)}{N_0^2} \in \mathrm{Sp}(\boldsymbol{g_0 g_0^T}),$$

and the result is immediate. ∎
This immediately leads to the following result.

Corollary 6.2.1:
Under Assumption 10, the scalar model of Proposition 6.2.1 is well-posed except on a set of resonant frequencies.

Proof: Using the assumption on the frequency, we can see that both the eigenvalues of $\boldsymbol{K_0}$ computed in Lemma 6.2.1 are positive, so we can use Theorem 11. ∎

Remark 6.2.4: The assumption $\omega^2 > N_0^2$ is really important here, indeed if $\omega^2 < N_0^2$ then

$$\frac{\omega^2}{\omega^2 - N_0^2} < 0,$$

and we cannot use Theorem 11.

6.3 Numerical aspects

Axisymmetric domain. To obtain 3D solutions using the 2D solver described in Chapter 3, we consider an *axisymmetric* domain. We denote by (r, θ, z) the 3D-cylindrical coordinates and by $(\boldsymbol{e}_r, \boldsymbol{e}_\theta, \boldsymbol{e}_z)$ the associated basis vector-fields. Numerical results in 3D are obtained by performing 2D numerical simulations in the (r, z)-plane and a *modal decomposition into Fourier series* in the θ-variable.

Assumption 11 (Axisymmetric domain):
We assume that the background quantities only depend on (r, z).

An immediate consequence of Assumption 11 is that

$$\nabla p_0 = \frac{\partial p_0}{\partial r} \boldsymbol{e}_r + \frac{\partial p_0}{\partial z} \boldsymbol{e}_z,$$

together with the background balance of momentum (6.6), we have

$$\boldsymbol{g_0} = g_r \boldsymbol{e}_r + g_z \boldsymbol{e}_z.$$

The anisotropy tensor

$$\boldsymbol{K_0} := \mathbf{Id} - \frac{N_0^2}{|\boldsymbol{g_0}|^2 (N_0^2 - \omega^2)} \boldsymbol{g_0 g_0^T},$$

therefore becomes

$$\boldsymbol{K_0} = \begin{bmatrix} K_{rr} & 0 & K_{rz} \\ 0 & 1 & 0 \\ K_{zr} & 0 & K_{zz} \end{bmatrix},$$

with

$$K_{rr} = 1 - \frac{N_0^2}{|\boldsymbol{g_0}|^2\left(N_0^2-\omega^2\right)} g_r^2 = \frac{|\boldsymbol{g_0}|^2\left(N_0^2-\omega^2\right) - N_0^2 g_r^2}{|\boldsymbol{g_0}|^2\left(N_0^2-\omega^2\right)},$$
$$K_{zz} = 1 - \frac{N_0^2}{|\boldsymbol{g_0}|^2\left(N_0^2-\omega^2\right)} g_z^2 = \frac{|\boldsymbol{g_0}|^2\left(N_0^2-\omega^2\right) - N_0^2 g_z^2}{|\boldsymbol{g_0}|^2\left(N_0^2-\omega^2\right)},$$
$$K_{zr} = K_{rz} = -\frac{N_0^2}{|\boldsymbol{g_0}|^2\left(N_0^2-\omega^2\right)} g_r g_z.$$

Development of the equation in cylindrical coordinates. We recall that the scalar equation (6.18b) reads

$$\omega^2\left[\frac{\omega^2}{c_0^2\left(N_0^2-\omega^2\right)} + \operatorname{div}\left(\frac{N_0^2}{|\boldsymbol{g_0}|^2\left(N_0^2-\omega^2\right)}\boldsymbol{g_0}\right)\right] p' - \operatorname{div}\left(\boldsymbol{K_0}\nabla p' + \frac{\omega^2}{N_0^2-\omega^2}\left(\frac{1}{c_0^2}+\frac{N_0^2}{|\boldsymbol{g_0}|^2}\right) p'\boldsymbol{g_0}\right) = 0,$$

developing the operators in cylindrical coordinates (r,θ,z) leads to

$$Ap' - \frac{1}{r}\frac{\partial}{\partial r}\left[rK_{rr}\frac{\partial p'}{\partial r} + rK_{rz}\frac{\partial p'}{\partial r} - rBg_r p'\right] - \frac{\partial}{\partial z}\left[K_{rz}\frac{\partial p'}{\partial r} + K_{zz}\frac{\partial p'}{\partial z} - Bg_z p'\right] + \frac{1}{r^2}\frac{\partial^2 p'}{\partial \theta^2} = 0, \tag{6.22}$$

where

$$A := \frac{\omega^4}{c_0^2\left(N_0^2-\omega^2\right)} + \omega^2 \operatorname{div}\left(\frac{N_0^2}{|\boldsymbol{g_0}|^2\left(N_0^2-\omega^2\right)}\boldsymbol{g_0}\right),$$
$$B := \frac{\omega^2}{N_0^2-\omega^2}\left(\frac{1}{c_0^2}+\frac{N_0^2}{|\boldsymbol{g_0}|^2}\right).$$

Decomposition into Fourier series. The unknown p' is expanded as

$$p'(r,\theta,z) = \sum_{m\in\mathbb{Z}} p'_m(r,z)e^{im\theta}. \tag{6.23}$$

Introducing (6.23) into (6.22) and identifying each mode leads to the following *modal equation*

$$\left(A+\frac{m^2}{r^2}\right)p'_m - \frac{1}{r}\frac{\partial}{\partial r}\left[rK_{rr}\frac{\partial p'_m}{\partial r} + rK_{rz}\frac{\partial p'_m}{\partial r} - rBg_r p'_m\right] - \frac{\partial}{\partial z}\left[K_{rz}\frac{\partial p'_m}{\partial r} + K_{zz}\frac{\partial p'_m}{\partial z} - Bg_z p'\right] = 0,$$

which can be rewritten as

$$A_m p'_m - \operatorname{div}\left(\boldsymbol{K_0}\nabla p'_m - Bp'_m\boldsymbol{g_0}\right) = 0, \tag{6.24}$$

where $A_m := A + m^2/r^2$ and the differential operators only act on r and z.

First-order formulation. To use a HDG method to solve (6.24), we need to reach a first-order in space formulation. As we have chosen to work with (6.18b), it is natural to use the HDG-$\boldsymbol{\sigma}_h$ method of Chapter 3. We therefore introduce the *total flux*

$$\boldsymbol{\sigma}_m = -\boldsymbol{K_0}\nabla p'_m + Bp'_m\boldsymbol{g_0},$$

leading to the *total flux formulation*

$$\boldsymbol{K_0}^{-1}\boldsymbol{\sigma}_m + \nabla p'_m - Bp'_m\boldsymbol{K_0}^{-1}\boldsymbol{g_0} = 0,$$
$$A_m p'_m + \operatorname{div}\left(\boldsymbol{\sigma}_m\right) = s.$$

Using the *Sherman-Morisson formula* we have

$$\boldsymbol{W_0} := \boldsymbol{K_0}^{-1} = \mathbf{Id} - \frac{N_0^2}{\omega^2|\boldsymbol{g_0}|^2}\boldsymbol{g_0}\boldsymbol{g_0}^T.$$

HDG formulation. We can now write the local problems of the HDG method in polar coordinates. The global problem, which translates the continuity of $\boldsymbol{\sigma}_h$, is the same as the one for the convected Helmholtz equation. On an element $K \in \mathcal{T}_h$, we multiply the system by the test-function $(\boldsymbol{v}, w)$ leading to

$$\begin{aligned}\int_K \left(W_{rr}\sigma_r + W_{rz}\sigma_z - Bp(W_{rr}g_r + W_{rz}g_z)\right) v_r r\mathrm{d}r\mathrm{d}z \\ + \int_K \left(W_{rz}\sigma_r + W_{zz}\sigma_z - Bp(W_{rz}g_r + W_{zz}g_z)\right) v_z \, r\mathrm{d}r\mathrm{d}z - \int_K p\left(\frac{1}{r}\frac{\partial r v_r}{\partial r} + \frac{\partial v_z}{\partial z}\right) r\mathrm{d}r\mathrm{d}z \\ + \int_{\partial K} \hat{p}\boldsymbol{v}\cdot\boldsymbol{n} r\mathrm{d}s = 0,\end{aligned}$$

and

$$\int_K \left(A_m p + \frac{1}{r}\frac{\partial r\sigma_r}{\partial r} + \frac{\partial \sigma_z}{\partial z}\right) w \, r\mathrm{d}r\mathrm{d}z + \int_{\partial K} i\tau(p-\hat{p}) w r\mathrm{d}s = \int_K s w r\mathrm{d}r\mathrm{d}z.$$

Using the matrix notations of Chapter 3 this can be written as

$$\sum_{v\in\{r,z\}} rW_{rv}\mathbb{M}^K\underline{\sigma_v^K} - Br\left(\sum_{v\in\{r,z\}} W_{rv}g_v\right)\mathbb{M}^K\underline{p^K} - r\mathbb{D}_r^K\underline{p^K} - \mathbb{M}^K\underline{p^K} + \sum_{\ell=1}^{3} rn_r^\ell\mathbb{F}_\ell^K\underline{\hat{p}^{g(K,\ell)}} = 0,$$

$$\sum_{v\in\{r,z\}} rW_{zv}\mathbb{M}^K\underline{\sigma_v^K} - rB\left(\sum_{v\in\{r,z\}} W_{zv}g_v\right)\mathbb{M}^K\underline{p^K} - r\mathbb{D}_z^K\underline{p^K} + \sum_{\ell=1}^{3} rn_z^\ell\mathbb{F}_\ell^K\underline{\hat{p}^{g(K,\ell)}} = 0,$$

$$rA_m\mathbb{M}^K\underline{p^K} + r(\mathbb{D}_r^K)^T\underline{\sigma_r^K} + \mathbb{M}^K\underline{\sigma_r^K} + r(\mathbb{D}_z^K)^T\underline{\sigma_z^K} + i\sum_{\ell=1}^{3} r\tau_\ell\left(\mathbb{E}_\ell^K\underline{p^K} - \mathbb{F}_\ell^K\underline{\hat{p}^{g(K,\ell)}}\right) = r\mathcal{S}^K,$$

and we can finally get the *matrix form* of the local problem

$$\mathbb{A}^K\begin{bmatrix}\underline{p^K}\\ \underline{\sigma_r^K}\\ \underline{\sigma_z^K}\end{bmatrix} + \mathbb{C}^K\begin{bmatrix}\underline{\hat{p}^{g(K,1)}}\\ \underline{\hat{p}^{g(K,2)}}\\ \underline{\hat{p}^{g(K,3)}}\end{bmatrix} = \mathcal{S}^K,$$

which allows the elimination of the local unknowns through *static condensation*. The matrices $\mathbb{A}^K$ and $\mathbb{C}^K$ are given by

Matrix $\mathbb{A}^K$: We introduce $\beta_u = Br\sum_v W_{uv}g_v$ to lighten the notations.

$\mathbb{A}^K$	$\underline{p^K}$	$\underline{\sigma_r^K}$	$\underline{\sigma_z^K}$
$\underline{p^K}$	$rA_m\mathbb{M}^K + i\sum_\ell \tau_\ell r\mathbb{E}_\ell^K$	$r(\mathbb{D}_r^K)^T + \mathbb{M}^K$	$r(\mathbb{D}_y^K)^T$
$\underline{\sigma_r^K}$	$-r\mathbb{D}_r^K - \beta_r\mathbb{M}^K - \mathbb{M}^K$	$rW_{0,11}^K\mathbb{M}^K$	$rW_{0,12}^K\mathbb{M}^K$
$\underline{\sigma_z^K}$	$-r\mathbb{D}_z^K - \beta_z\mathbb{M}^K$	$rW_{0,21}^K\mathbb{M}^K$	$rW_{0,22}^K\mathbb{M}^K$

Matrix $\mathbb{C}^K$:

$\mathbb{C}^K$	$\widehat{p}^{g(K,1)}$	$\widehat{p}^{g(K,2)}$	$\widehat{p}^{g(K,3)}$
p^K	$-ir\tau_1\mathbb{F}_1^K$	$-ir\tau_2\mathbb{F}_2^K$	$-ir\tau_3\mathbb{F}_3^K$
σ_r^K	$n_r^{K,1}r\mathbb{F}_1^K$	$n_r^{K,2}r\mathbb{F}_2^K$	$n_r^{K,3}r\mathbb{F}_3^K$
σ_z^K	$n_z^{K,1}r\mathbb{F}_1^K$	$n_z^{K,2}r\mathbb{F}_2^K$	$n_z^{K,3}r\mathbb{F}_3^K$

Evaluation of the physical quantities and nondimensionalization. As we already discussed in CHAPTER 5, we need to perform a nondimensionalization of the problem in order to get stable numerical methods. We denote the solar radius by $R_\odot$, the physical coordinates inside the Sun by $\boldsymbol{x}$ and the normalized coordinates used to perform the numerical simulation by $\widetilde{\boldsymbol{x}} = \boldsymbol{x}/R_\odot$. For a quantity f defined in the Sun, the associated quantity in normalized coordinates is defined by

$$\widetilde{f}(\widetilde{\boldsymbol{x}}) := f(\boldsymbol{x}) = f(R_\odot\widetilde{\boldsymbol{x}}).$$

The values for the physical parameters, ρ_0, p_0 and c_0 come from the Model S of [CDA+96]. We use the same convention as in [FFP20][2], where the scaled pressure and gravitational potential

$$\widetilde{\widetilde{p_0}}(\widetilde{\boldsymbol{x}}) := \frac{p_0(R_\odot\widetilde{\boldsymbol{x}})}{R_\odot^2}, \quad \text{and} \quad \widetilde{\widetilde{\varphi_0}}(\widetilde{\boldsymbol{x}}) := \frac{\varphi_0(R_\odot\widetilde{\boldsymbol{x}})}{R_\odot^2},$$

are used instead of the usual background quantities. As the sound-speed is related to the pressure by

$$c_0^2 = \frac{\gamma p_0}{\rho_0},$$

this naturally leads to a scaled sound-speed

$$\widetilde{\widetilde{c_0}}(\widetilde{\boldsymbol{x}}) := \frac{c_0(R_o dot\widetilde{\boldsymbol{x}})}{R_\odot}.$$

As the unknown p' is homogenous to a pressure, we follow this convention and we define the scaled unknown

$$\widetilde{\widetilde{p'}}(\widetilde{\boldsymbol{x}}) := \frac{p'(R_\odot\widetilde{\boldsymbol{x}})}{R_\odot^2}.$$

If we compute the auxiliary physical parameters using the scaled quantities, *e.g.*

$$\widetilde{\widetilde{\boldsymbol{g_0}}} := -\nabla_{\widetilde{\boldsymbol{x}}}\widetilde{\widetilde{\varphi_0}}, \quad \text{and} \quad \widetilde{\widetilde{N_0^2}} := \widetilde{\widetilde{\boldsymbol{g_0}}} \cdot \left(\frac{\nabla_{\widetilde{\boldsymbol{x}}}\widetilde{\rho_0}}{\widetilde{\rho_0}} - \frac{\nabla_{\widetilde{\boldsymbol{x}}}\widetilde{\widetilde{p_0}}}{\widetilde{\gamma}\widetilde{\widetilde{p_0}}}\right) = \widetilde{N_0^2},$$

then the equation for $\widetilde{\widetilde{p'}}$ is the same as the one for p'. Indeed, if p' satisfies (6.18b), then $\widetilde{\widetilde{p'}}$ satisfies

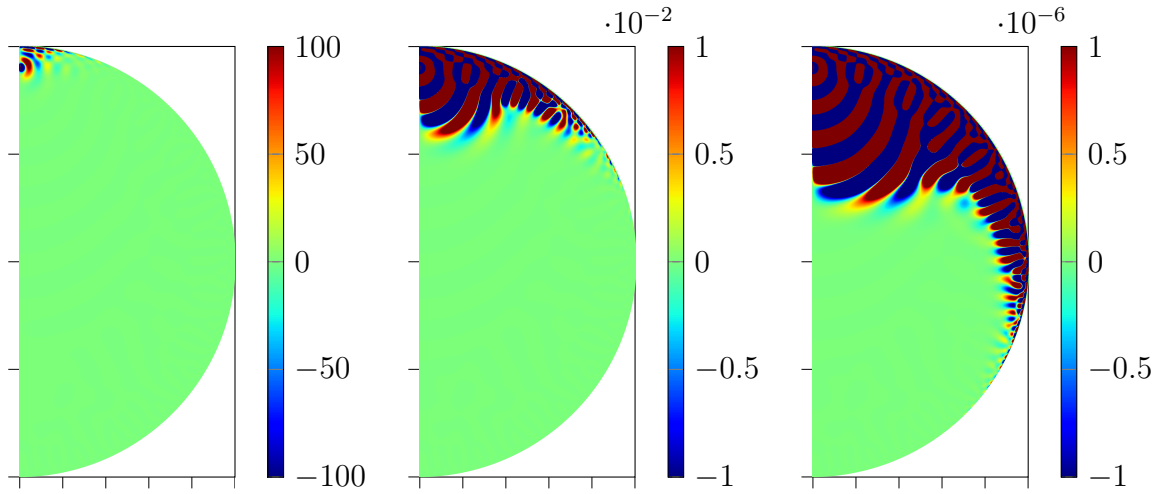
$$\widetilde{\widetilde{A}}\widetilde{\widetilde{p'}} - \mathrm{div}_{\widetilde{\boldsymbol{x}}}\left(\boldsymbol{K_0}\nabla_{\widetilde{\boldsymbol{x}}}\widetilde{\widetilde{p'}} - \widetilde{\widetilde{B}}\widetilde{\widetilde{p'}}\widetilde{\widetilde{\boldsymbol{g_0}}}\right) = \widetilde{s}.$$

[2]The values for those scaled quantities can be obtained here : https://phaidra.univie.ac.at/view/o:1097638.

Numerical results. We end up this chapter with some numerical results. On FIGURE 6.1, we have depicted the mode $m = 0$ in the (r, z)-plane for a frequency of 3mHz. The source term is Dirac mass located close to the axis of symmetry and to the surface at the point $(10^{-3}, 0.9)$. Using three different saturation levels, we can see that there is a wave propagating close to the surface. This is consistent with the fact that the scalar model of this chapter was derived to support surfacic gravity waves, or f-modes. We would also like to point out that the amplitude of the solution decreases as the index m increases. To illustrate this we have depicted the solutions for indexes $m = 2$ and $m = 4$ in FIGURE 6.2. We have also plotted 3D results for modes $m = 0$ in FIGURE 6.3, $m = 1$ in FIGURE 6.4 and $m = 4$ in FIGURE 6.5. As the amplitude of the modes decreases, there is almost no visual difference between the mode $m = 0$ and the truncated sum

$$\sum_{m=-N}^{N} p_m(r, z)e^{im\theta} = p_0(r, z) + 2\sum_{m=1}^{N} p_m(r, z)\cos\theta,$$

that we have depicted in FIGURE 6.6. Notice that the last equality holds as m only appears as m^2 in the modal equations, so we have $p_m(r, z) = p_{-m}(r, z)$.

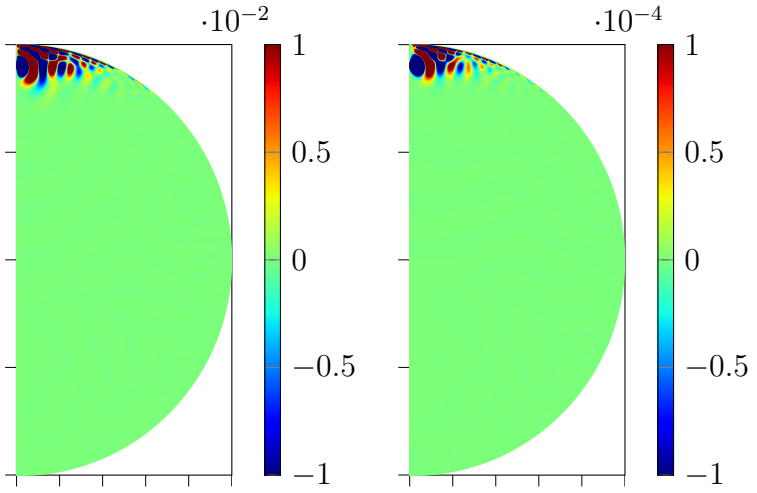


(a) Low saturation (b) Medium saturation (c) High saturation

Figure 6.1: Numerical results for mode $m = 0$ in the (r, z)-plane.

Conclusion

In this chapter we have illustrated the possibility of using the HDG framework introduced in CHAPTER 3 to more complex wave-like equations. To this end, we have considered a scalar model involving gravity for wave propagation inside a star. The coefficients of this model are more complicated than the ones of the convected Helmholtz equation. It is worth noting that this model could be used to validate other numerical solvers. Indeed, when Galbrun's equation is solved for realistic solar-like applications, there is no reference solution to validate the numerical results. This scalar model could be used as a comparison point to validate the results obtained with a numerical solver for Galbrun's equation with gravity. However, we would like to point out that there is still some work to do. In particular, questions regarding boundary conditions and source terms should be addressed carefully before using this model as a reference one.



(a) Mode $m = 2$ (b) Mode $m = 4$

Figure 6.2: Numerical results for modes $m = 2$ and $m = 4$ in the the (r, z)-plane.

Bibliography

[Bre87] Haim Brezis. *Analyse fonctionnelle*. 1987.

[CDA+96] Jorgen Christensen-Dalsgaard, W. Däppen, S. V. Ajukov, E. R. Anderson, H. M. Antia, S. Basu, V. A. Baturin, G. Berthomieu, B. Chaboyer, S. M. Chitre, A. N. Cox, P. Demarque, J. Donatowicz, W. A. Dziembowski, M. Gabriel, D. O. Gough, D. B. Guenther, J. A. Guzik, J. W. Harvey, F. Hill, G. Houdek, C. A. Iglesias, A. G. Kosovichev, J. W. Leibacher, P. Morel, C. R. Proffitt, J. Provost, J. Reiter, E. J. Rhodes, F. J. Rogers, I. W. Roxburgh, M. J. Thompson, and R. K. Ulrich. The Current State of Solar Modeling. *Science*, 272(5266):1286–1292, May 1996.

[FFP20] Florian Faucher, Damien Fournier, and Ha Pham. C2 representations of the solar background coefficients for the model S-AtmoI. *arXiv:2009.01587 [astro-ph]*, September 2020.

[GT01] David Gilbarg and Neil S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Classics in Mathematics. Springer-Verlag, Berlin Heidelberg, second edition, 2001.

[LG09] François Lignières and Bertrand Georgeot. Asymptotic analysis of high-frequency acoustic modes in rapidly rotating stars. *Astronomy & Astrophysics*, 500(3):1173–1192, June 2009.

[Tas00] Jean-Louis Tassoul. *Stellar Rotation*. April 2000.

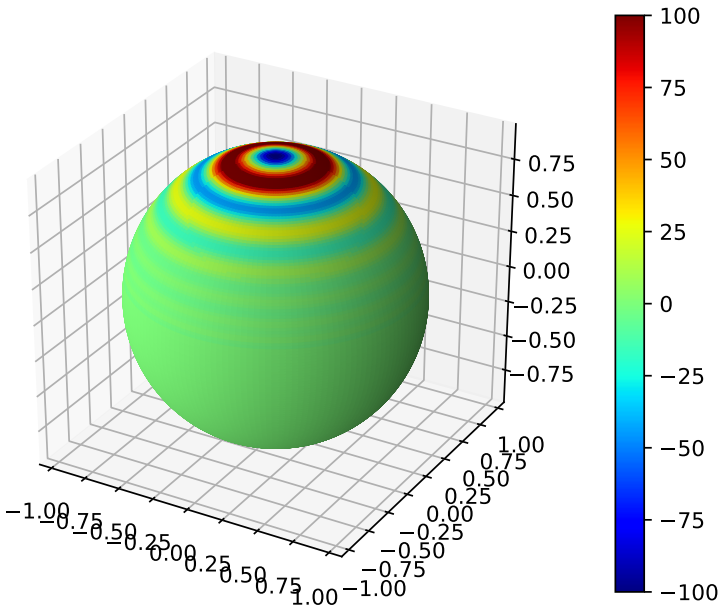


Figure 6.3: 3D numerical result for mode $m = 0$.

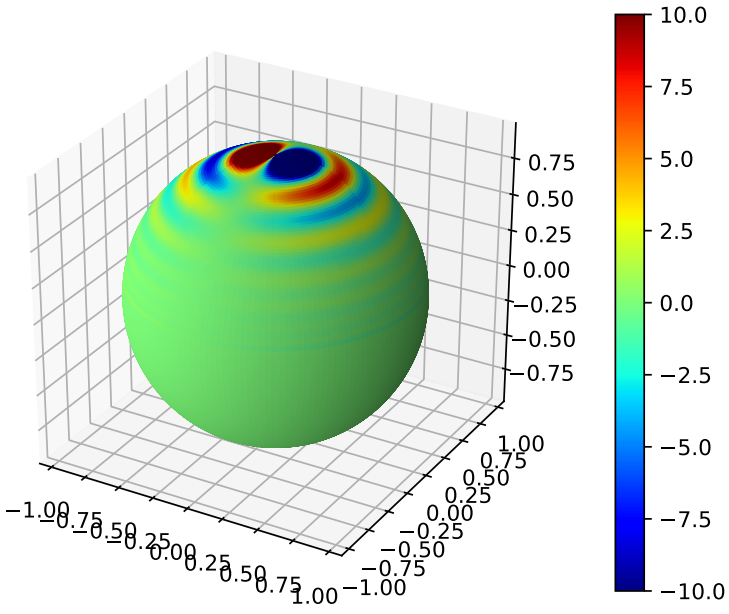


Figure 6.4: 3D numerical result for mode $m = 1$.

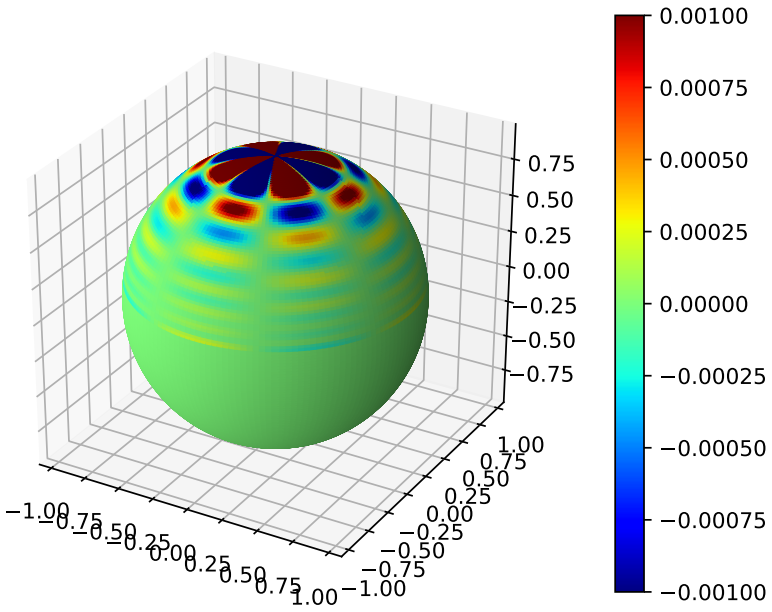


Figure 6.5: 3D numerical results for mode $m = 4$.

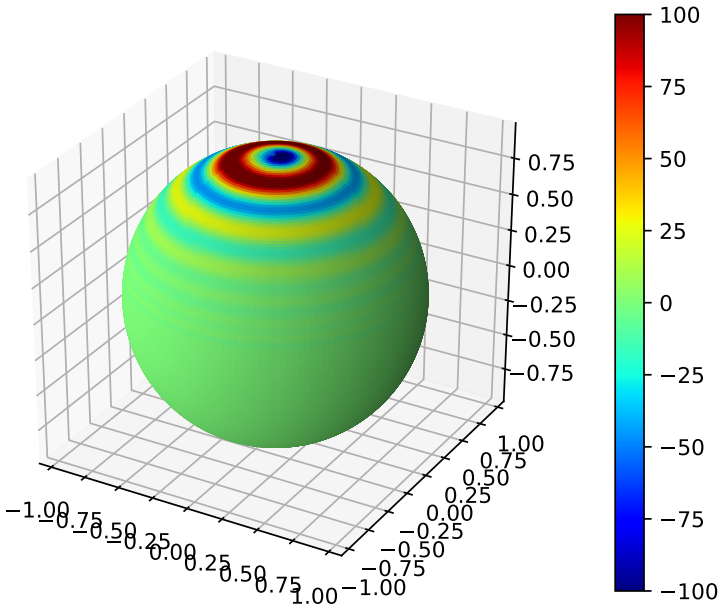


Figure 6.6: 3D numerical result, summing mode for $|m| = 0$ to 6.

Conclusion

The main contribution consists of a computational framework based upon HDG formulations for convected Helmholtz equations. This piece of software has been developed from scratch and is meant to become part of a helioseismic branch of the open-source software `hawen`.

Modelling aspects

Summary of our contributions. During the first part of this study, we have considered the existence of a resonant-like phenomenon for aeroacoustic propagation in a recirculating flow. We are particularly interested in this question because it is crucial for the modelling of the interior of the Sun where flows are recirculating. Previous well-posedness results for time-harmonic aeroacoustic wave models specifically excluded those background flows, while they are common for applications in helioseismology. We have shown that a resonant-like phenomenon can occur if the symbol of the material derivative vanishes inside the computational domain. This phenomenon gives a mathematical interpretation for the spurious modes that have been reported when Lagrange finite-elements are used to solve Galbrun's equation. Finally, when this phenomenon occurs, the equivalence between aeroacoustic models, such as Galbrun's equation and LEE, cannot be guaranteed.

Perspectives: using the low-regularity settings. As we have mentioned, the well-posedness of Galbrun's equation has also been studied in a Hilbert space $\boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O})$ with lower regularity requirements than $\boldsymbol{H}^1(\mathcal{O})$. Functions belonging to $\boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O})$ do not have a trace on the resonant lines that we have computed. The approximation of Galbrun's equation in $\boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O})$ instead of $\boldsymbol{H}^1(\mathcal{O})$ could lead to new numerical methods and better numerical results. At this point, the only results available in this settings is the well-posedness of Galbrun's equation. A next step could be the derivation of sharper regularity results for the solution to allow a rigorous construction of numerical methods.

Perspectives: solving the equations in time-domain. In this thesis, we have considered time-harmonic solutions to aeroacoustic models. The meaning of these solutions is given by the *limiting amplitude principle*: they can be interpreted as the limit of time-domain solutions as time goes to infinity. In other words, we show that time-harmonic solutions correspond to the established regime of time-dependent solutions letting t to infinity So there is a hidden physical assumptions for the time-harmonic solutions to make sense: the physical parameters must be "nice" enough to allow this periodic regime. It turns out that it is not the case for realistic applications in helioseismology. Hence more reliable results could perhaps be obtained by solving these equations in time-domain instead of the time-harmonic domain. The numerical methods constructed in this thesis could be used for the semi-discretization in space in a time-domain aeroacoustic solver.

Perspectives: considering non-linear wave models. The models considered in this thesis are obtained by linearizing the equations of fluid dynamics around an equilibrium state. Some of the difficulties encountered when trying to solve those equations may be linked to the linear nature of those models. Indeed, if an instability occurs in a linear model it can only grow exponentially. On the other hand, some non-linear effects can contain those instabilities as it happens with the *Kelvin-Helmholtz instability*. Taking some of the non-linear effects into account could be a way to obtain better results. In particular, using the second-order asymptotic expansion

$$q = q_0 + \varepsilon q' + \varepsilon^2 q'' + \mathcal{O}(\varepsilon^3),$$

leads to the equations of *weakly non-linear acoustics*. Those equations consist of a coupled system of two Galbrun's equations, one for each perturbation order. It is possible that the non-linear and the coupling terms help controls the resonant-like phenomenon that we have described.

Numerical aspects

Summary of our contributions. In a second part of this work, we have introduced a HDG framework for the convected Helmholtz equation. To the best of our knowledge, this is an original contribution. HDG methods are mixed DG methods that rely on a static condensation process to eliminate the interior degrees of freedom. This allows to keep all the advantages of DG methods for a reduced numerical cost. Among other advantages, DG methods are easily implemented in a parallel way with hp-adaptivity and arbitrary high-order. In particular, their non-conforming nature seems well-suited to work with the low-regularity space $\boldsymbol{H}_{\boldsymbol{v_0}}(\mathcal{O})$. As the theory for the convected Helmholtz equation is well established and quite standard, we were able to derive precise theoretical results on the HDG methods that we have introduced, including well-posedness of both the local and global problems, as well as a detailed convergence analysis. We have also proposed a simple technique to construct Absorbing Boundary Conditions (ABCs) based on the Prandtl-Glauert-Lorentz transformation, that maps the convected Helmholtz equation to the standard one. This provide ABCs that are both efficient for low and intermediate Mach numbers and easy to implement in any finite-element solver. In particular, the resulting ABC is easily included in the HDG formulation which is not obvious in general. Finally we have shown that this HDG framework can be extended to handle more complex problems.

Ongoing work: toward the extension to Galbrun's equation. We have also briefly presented a way to extend this HDG framework to the full Galbrun's equation, which is a vectorial and more realistic aeroacoustic model. As we have introduced a rigorous HDG framework for convected acoustics, it seems possible to extend these methods to more realistic aeroacoustic models. Even if these constructions will be less rigorous, it can lead to efficient numerical methods. One way to avoid the resonant-like phenomenon that we have described is to study this equation in a low regularity settings. The continuity of functions belonging to this space between two mesh elements is interface-dependent, making it difficult to construct conforming approximation spaces. One way to derive suitable numerical methods could be then to use $\boldsymbol{H}_{\mathrm{div}}(\mathcal{O})$-conform finite elements with a penalization term to enforce this interface-dependent continuity between mesh elements. This is an ongoing work in the team of T. Hohage in Göttingen in the framework of project C04, and it turns out that this approach leads to a non-conforming method. If we have to use a non-conforming method, it occurs to us that it is easier and more natural to consider HDG approximations in the

continuity of our work on the convected Helmholtz equation. Whatever the approach is, we will end up with the same difficulty of approximating T-coercive problems with non-conforming methods.

Perspectives: choosing the penalization parameters. One of the main drawbacks of HDG methods is the necessity to choose the penalization parameters. Indeed, if those parameters are badly chosen the quality of the numerical solutions are poor. For the HDG+ variant, the penalization parameters should be very large to ensure optimal convergence, and this method therefore seems less sensitive than standard HDG methods to the choice of penalization parameters. If the solved system is hyperbolic, we have presented a choice penalization parameter with physical meaning by solving a Riemann problem on the interface between two elements. It is not clear how to extend this method to problems whose transient counterpart is not hyperbolic, but it could be very attractive as the method would become parameter-free.

Perspectives: inverse problems. The methods that we have constructed in this thesis were implemented in the `hawen` solver. This solver is dedicated to solving both direct and quantitative inverse problems. Therefore, it offers us the computational framework to perform the inverse problem using the *Full Waveform Inversion* (FWI) technique. In particular, we could reconstruct solar parameters such as the background flow $\boldsymbol{v_0}$, the sound-speed c_0 or the density ρ_0.

ECOLE DOCTORALE :
École Doctorale des Sciences Exactes et de leurs Applications

LABORATOIRE :
Équipe-Projet Commune MAKUTU
Inria Bordeaux – Sud-Ouest
Laboratoire de Mathématiques et de leurs Applications de Pau

Dans cette thèse, nous nous intéressons à des problèmes de propagation d'ondes acoustiques dans un écoulement en vue d'applications en physique solaire. En effet, la structure interne du Soleil peut être étudiée à partir de l'observation de ces ondes sur la surface solaire.

Dans un premier temps, nous rappelons le procédé de linéarisation des équations de la mécanique des fluides permettant de construire des modèles vectoriels de propagation d'ondes en écoulement. Nous étudions alors l'équivalence entre les modèles obtenus à partir des linéarisations eulériennes et lagrangiennes et nous montrons que cette équivalence ne peut pas toujours être garantie en régime harmonique. D'un point de vue pratique, l'exploitation numérique de cette équivalence semble inefficace, notamment pour la reconstruction des perturbations lagrangiennes à partir des perturbations eulériennes. Nous nous concentrons ensuite sur l'étude d'un phénomène de résonance en régime harmonique lorsque l'écoulement porteur a des lignes de courant fermées. Une étude modale montre que les équations considérées dégénèrent sur certaines lignes de courant. Sur ces lignes, il n'est alors pas possible de résoudre les équations. Il semble toutefois possible de résoudre ce problème en étudiant ces équations dans un espace hilbertien dont les propriétés de régularités sont plus faibles que pour les espaces de Hilbert habituellement utilisés pour l'étude des équations d'ondes aéroacoustiques. D'un point de vue numérique, la recherche de la solution dans un tel espace nécessite la construction de nouvelles méthodes numériques.

Dans la deuxième partie de ce travail, nous nous concentrons sur la construction de méthodes numériques pour un modèle aéroacoustique simple : l'équation de Helmholtz convectée. Ce modèle scalaire peut être obtenu à partir des modèles vectoriels lorsque l'écoulement porteur est irrotationnel. Pour cette équation, nous construisons trois variantes de la méthode de Galerkine Discontinue Hybride (HDG). Les méthodes HDG sont des méthodes de Galerkine Discontinues mixtes dont le coût numérique reste raisonnable grâce à un procédé de condensation statique qui permet de réduire le problème à un problème posé uniquement sur le squelette du maillage. Nous avons effectué une analyse détaillée de ces méthodes, en particulier nous avons montré le caractère bien posé des méthodes, ainsi que des estimations de la vitesse de convergence pour des solutions régulières. Enfin, nous avons également discuté le choix du paramètre de pénalisation qui peut exercer une influence importante sur la qualité des résultats numériques. Ces méthodes ont été implémentées dans le code open-source Hawen et les résultats numériques ont permis d'illustrer les conclusions de notre étude théorique. Nous avons également construit des conditions aux limites absorbantes (CLA) d'ordre faible pour l'équation de Helmholtz convectée. Ces CLA sont obtenues par transformation de PrandtL-Glauert-Lorentz de CLA pour l'équation de Helmholtz standard lorsque l'écoulement porteur est uniforme à l'extérieur du domaine. Ces CLA sont performantes pour des écoulements dont le nombre de Mach est faible ou modéré et leur mise en œuvre dans un code éléments finis, notamment HDG, est simple.

Enfin, la troisième partie de ce travail est consacrée à l'extension des méthodes HDG construites pour l'équation de Helmholtz convectée à des modèles plus réalistes. Dans un premier temps, nous décrivons les changements à apporter pour traiter les cas vectoriels. La construction d'une méthode HDG pour le cas vectoriel sans convection semble relativement aisée, mais la prise en compte des phénomènes convectifs pose de nombreuses questions tant théoriques que pratiques. Finalement, nous illustrons les possibilités des méthodes HDG sur un problème scalaire issu de l'astérosismologie qui prend en compte une partie des effets liés à la gravité. Le modèle obtenu est semblable à une équation d'Helmholtz convectée par la gravité.