# THÈSE

**En vue de l'obtention du**

## DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

**Délivré par :** *l'Université Paul Sabatier - Toulouse 3*

**Présentée et soutenue le** 10/12/2020 **par**

**Firas DHAOUADI**

# An augmented Lagrangian approach for Euler-Korteweg type equations

**JURY**

| | | |
|---|---|---|
| Christophe Besse | Université Paul Sabatier - Toulouse 3 | Président |
| Michael Dumbser | Università degli Studi di Trento | Rapporteur |
| Thierry Gallouët | Aix-Marseille Université | Rapporteur |
| Sylvie Benzoni-Gavage | Université Claude Bernard Lyon 1 | Examinatrice |
| Sergey Gavrilyuk | Aix-Marseille Université | Examinateur |
| Jean-Paul Vila | INSA Toulouse | Directeur de Thèse |
| Nicolas Favrie | Aix-Marseille Université | Co-directeur de Thèse |

**École doctorale et spécialité :**
*MITT : Domaine Mathématiques : Mathématiques appliquées*
**Unité de Recherche :**
*Institut de Mathèmatiques de Toulouse (UMR 5219)*
**Directeur(s) de Thèse :**
*Jean-Paul VILA* et *Nicolas FAVRIE*
**Rapporteurs :**
*Michael DUMBSER* et *Thierry GALLOUËT*

# Acknowledgements

This work was supported by the Ecole doctorale Mathématiques, Informatique et Télécommuncations de Toulouse, l'Institut de Mathématiques de Toulouse and financially supported by the Université Toulouse III Paul Sabatier. I am indebted to them for permitting this PhD work to see the light. I am also grateful to the INSA de Toulouse for granting me everything I needed for my research and teaching activities. I am thankful to the Service Hydrographique et Océanographique de la Marine (SHOM) for financing some of my research activities and covering the fees of my conferences and workshop participations. Lastly, I would like to thank Aix-Marseille Université and IUSTI for their collaboration and support.

Besides, this journey would not have any meaning in my life if not for all the wonderful persons I met on the way. These words are but a very small token of gratitude towards them, and I'm certain I'll be forgetting many.

Naturally, I would like to extend my deepest thanks to my PhD advisor Jean-Paul Vila. I have learned a lot by your side, in numerical analysis, mechanics, hydrodynamics, simply in everything that was related to my work. Although, it usually takes me a while to understand your suggestions, they usually proved to be efficient. I'm really grateful for the time you invested into taking care of me, despite all your commitments and extremely busy schedule. Thank you for seeing this work to the end.

Then, I would like to express my gratitude towards my Phd co-advisor Nicolas Favrie. Nicolas, you have always been there, you have been a great help during this experience. I learned a lot, really a lot by your side, be it in mechanics, numerics, coding, writing proposals, making presentations and especially finding solutions and never giving up. Your support was always very important, especially in the hard times and all the stressful periods. Thanks for always treating me as an equal, for always giving me tips and pieces of advice for my career. Thanks for caring about my career than you did for the work itself. This PhD would not have been possible, if not for your perseverance and your faith in me. I hope I did not fail your trust.

Next, I would like to sincerely thank Sergey Gavrilyuk for the great help he provided during all this period of time. Thanks for everything you have taught me, in the most pedagogic way ever. Thanks for answering the insane amount of questions I kept asking, for all the hints and tips you taught me, for all the differential calculus I learned by your side, for the very accurate explanations that got me through hard parts, for the sense of humor that kept us going, for all the inspiring quotes. Thank you for all the time you devoted to my care. Thank you for all the encouragements. I feel truly blessed to have worked by your side. I am deeply grateful for everything.

Now, I would like to express my sincere thanks to the jury members for devoting of their time to examine this work. First, I would like to thank Michael Dumbser and Thierry Gallouët for kindly accepting to review this manuscript. Their insightful comments, questions

and suggestions are surely to improve the quality of this work. Thanks also to Christophe Besse and to Sylvie Benzoni-Gavage for their perspicacious remarks, and for the fruitful discussion, which shed light on interesting aspects to be addressed.

A great part of this PhD work took place at l'INSA de Toulouse, precisely at the GMM department, and I would like to thank all the colleagues, with whom I shared a lot of laughter, interesting discussions and teaching duties. First, I would like to thank my office mate Ilya Peshkov who shared with me every last detail of this PhD experience. Thank you for listening to me complain all the time, for all the helpful advice, and for all the constructive discussions we had. Thanks to Sandrine Bosc for always being there for us and for taking care of any administrative tasks in a record time. Thanks to Victor for the many helpful tips, especially in numerics. Thanks also to you and Clémentine for always listening to my everyday stories. I would have been quite lonely at many occasions if not for you two. Also thanks to Pascal, Fred and Rémy for the interesting discussions and for their continuous interest and support. A big thank you to all my teaching colleagues, especially to Géraldine, Sandrine and Olivier. My first steps in teaching were done under your excellent mentorship. Thank you sincerely for all the advice and tips. I would also like to extend my thanks to my fellow colleagues in numerical analysis, especially to grand chef Robin, to Romain and to Florent. Thanks for sharing the passion and for the great work atmosphere you all provide. A thought goes to all those I meet regularly in the cafeteria, in the corridors, at lunch time or in the afterworks. Thanks for sharing experiences, for sharing a laughter, a cup of coffee, for playing cross words, although I never figured out any of the words. In particular, thanks to Marlène, Mélisande, Maria, Jean-Yves, Béatrice, Stéphane, Simona, Aude, Jérôme, Aldéric, Anthony, Pierre, Cathy and David. Last but not least, I wish all my colleagues currently doing their PhD, the best of luck, although I'm certain they don't need any. In particular, I wish the best of luck to Josué, Armando, Thibault, Morgane, Franck, Mahmoud, Clément and Jessica. May the power be with you!

I would like to acknowledge my colleagues from IUSTI, where I carried on my work for around two years. I would like to thank all the colleagues from the research team there, namely Jacques, Olivier, Eric, Fabien, Aziz, Henri, Joris, Kevin, and Ksenia. Thanks for the constant support, for the many discussions about science, work and life. Then I would like to acknowledge my Bratan Sergey, for sharing most of this experience with me, for the interesting conversations and the philosophical debates, about maths, science and existence. Thanks also to my friends and office mates Nabila and Sanjeev for the very friendly atmosphere, for all the coffee breaks we had and for all the stories we shared. A very big thank you goes to my colleague and dear friend Naima. Thanks for your unconditional support, all the delightful moments we spent together, all the stories we told each other, all the laughter we shared and all the dinners you kindly organized. I'm sure your PhD will end up in a great success, so keep up the hard work! I would now like to express my gratitude towards my friend Pascal Campion for all the comforting talks we had, for all the inspiration in poetry and for the very kind pieces of advice. Special thanks also to my colleague in science and also in music Jeongeun. Thanks for all the nice times when we gathered and played piano together. Last but not least, I would like to thank all the colleagues, be it PhD students, postdoctoral fellows, researchers, professors, administrative and technical staff. In particular I wish the best of luck to all my fellow PhD students. I have faith that every last one of you will be

successful. Best of luck to all of you.

This PhD work would not have been possible without the help and the commitment of many administrative personnel, to whom I am indebted. I would like to especially thank Catherine Stasiulis for a job wonderfully done, for providing excellent advice and for listening to all our problems. My sincere thanks to all agents who worked on my hosting agreement in Marseille as well, in particular, thanks to Patricia, Ariane and Romain and all the agents from Université Paul Sabatier. Also thanks to Agnès Requis and to Mihai Maris from the doctoral school MITT for always replying to my many questions in time, clearly and concisely.

To all my friends whose support and assistance were inexhaustible, I would like to thank you for your concern and sincere friendship. In particular thanks to Ines, Marwen and Adnen for always being there all along and for always caring despite the distance and also for greatly helping out with typos and corrections. I am most grateful to you. I would also like to express my gratitude towards all my friends who kept calling and regularly checking in during the PhD period. Thanks especially to Hamma, Nidhal, Brahim, Amal, Sameh, for keeping me company during both hard and joyful times. Thanks also to all the friends I met during conferences and workshops, conferences and during Cemracs 2019. Thanks in particular to Andrea, Arnaud, Benoit, Roland, Jiao, Khaoula, Dali, Solène (sorry I was really bad at card games) and to all the great ping-pong players (yo Louis and Abraham), thank you for mentoring me.

Lastly, I would like to thank my family : My mother Sarra, my father Sami and my brother Elyes for their unfaltering commitment, their constant support and their unwavering dedication since my existence to make my life better. I would not have made it this far if not for you. For that, I thank you.

# Contents

# List of Figures

# Abstract

An approximate first order quasilinear hyperbolic model for Euler-Korteweg (E-K) equations, describing compressible fluid flows whose energy depend on the gradient of density, is derived. E-K system can be seen as the Euler-Lagrange equations to a Lagrangian submitted to the mass conservation constraint. Due to the presence of the density gradient in the Lagrangian, one recovers high-order derivatives of density in the motion equations. The approach presented here permits us to obtain a system of hyperbolic equations that approximate E-K system. The idea is to introduce a new order parameter which approximates the density via a carefully chosen penalty method. The gradient of this new independent variable will then replace the original gradient of density in the Lagrangian, resulting in the so-called *augmented Lagrangian*. The Euler-Lagrange equations of the augmented Lagrangian result in a first order hyperbolic system with stiff source terms and fast characteristic speeds. Such a system is then analyzed and solved numerically by using IMEX schemes. In particular, this approach was applied to the defocusing nonlinear Schrödinger equation (which can be reduced to the E-K equations via the Madelung transform), for which a comparison with exact and asymptotic solutions, namely gray solitons and dispersive shock waves was performed. Then, the same approach was extended to thin film flows with capillarity, for which comparison of the numerical results with both reference numerical solutions and experimental results was performed. It was shown that the augmented model is also extendable to models with full nonlinear surface tension. In the same setting, a study of stationary droplets on a horizontal solid substrate was conducted in an attempt to classify droplet profiles depending on their energy forms. This also allowed to compare the augmented Lagrangian approach in the case of stationary solutions, and which showed excellent agreement with the reference solutions. Lastly, an independent part of this work is devoted to the study of modified equations associated to numerical schemes for stability purposes. It is shown that for a linear scheme, stability conditions which are obtained from a truncation of the associated modified equation, are only relevant if the corresponding series in Fourier space is convergent for the admissible wavenumbers.

# Résumé

On présente un modèle hyperbolique quasi-linéaire de premier ordre approximant les équations d'Euler-Korteweg (E-K), qui décrivent des écoulements de fluides compressibles dont l'énergie dépend du gradient de la densité. Le système E-K peut être vu comme les équations d'Euler-Lagrange d'un Lagrangien soumis à la conservation de la masse. Vu la présence du gradient de la densité dans le Lagrangien, des dérivées d'ordre élevé de la densité apparaissent dans les équations du mouvement. L'approche présentée ici permet d'obtenir un système d'équations hyperboliques qui approxime le système E-K. L'idée est d'introduire un nouveau paramètre d'ordre qui approxime la densité via une méthode de pénalisation classique. Le gradient de cette nouvelle variable remplace alors le gradient de la densité dans le Lagrangien, ce qui permet de construire le *Lagrangien augmenté*. Les équations d'Euler-Lagrange associées à celui-ci, sont des équations hyperboliques avec des termes sources raides et des vitesses de caractéristiques rapides. Ce système est analysé puis résolu numériquement en utilisant des schémas de type IMEX. En particulier, cette approche a été appliquée à l'équation de Schrödinger nonlinéaire défocalisante (qui peut être réduite au système E-K via la transformée de Madelung), pour laquelle des comparaisons avec des solutions exactes et asymptotiques ont été faites, notamment pour des solitons gris et des ondes de choc dispersives. La même approche a été également appliquée aux équations de filmes minces avec capillarité, pour lesquelles une comparaison avec des résultats numériques de référence et des résultats expérimentaux a été faite. Il a été démontré que le modèle augmenté peut aussi bien s'appliquer pour des modèles dont le terme de capillarité est non-linéaire. Dans ce même cadre, une étude de gouttes stationnaires sur un substrat solide horizontal a été établie afin de classifier les profils possibles de gouttes selon leur énergie. Ceci a permis également de faire des comparaisons du modèle augmenté sur des solutions stationnaires. Enfin, une partie indépendante de ce travail est consacrée à l'étude des équations équivalentes associées aux schémas numériques, où l'on démontre que les conditions de stabilité qui dérivent d'une troncature de l'équation équivalente, n'a du sens que si la série correspondante dans l'espace de Fourier est convergente, sur les longueurs d'onde admissibles dans la pratique.

# Introduction

In 1893, Van der Waals published a paper in the *Verhandelingen* of the Academy, entitled *The thermodynamic theory of capillarity under the hypothesis of a continuous variation of density* [72]. It took sixty years to understand the worth of his work [48]. Regrettably, new ideas are more prone to rejection, no matter how great a contribution they could have been.

Van der Waals assumed in his paper a continuous transition of the density across the layer separating a liquid and its vapor. In such a setting, he stated that equilibrium states where the density distribution is heterogeneous, are compatible with capillarity, only if the free energy at a point, not only depended on its local density, but also on the densities in the neighboring points [72, 48]. Based on the work of Van der Waals, Korteweg formulated a constitutive equation of the Cauchy stress that depended on the density and its spatial gradients [50] in the static case. The dynamical form of the equations, based on Hamilton's principle of stationary action, was obtained by Casal [16, 17] and a year later independently by Eglit [29].

Van der Waals-Korteweg type equations (often called Euler-Korteweg (E-K) equations), appear also in other fields of physics, like capillary fluids, thin films flows and quantum hydrodynamics. It has been the subject of interest in many works in recent decades which include for example [3, 7, 8, 9, 11, 12, 13, 18, 36, 44, 61].

The aim of this thesis is to derive a first order hyperbolic model which approximates the E-K system in an attempt to open new perspectives for its numerical resolution and the understanding of its solutions properties. This manuscript is organized as follows. Chapter one is dedicated to the augmented Lagrangian method for the E-K system in general. Explanations on how the approach works in general, and how it affects the structure of the system are given. Full details for the derivation of the corresponding system of equations are given. The structure of the obtained system is then discussed. This includes the proof of the hyperbolicity of the system and how it approximates the E-K system.

In chapter two, the augmented Lagrangian approach is applied to the defocusing nonlinear Schrödinger equation. Preliminaries on how to cast the equation into a system of conservation laws via the Madelung transform, and how to obtain traveling wave solutions are presented. Then, the augmented formulation of the equations is given and analyzed. In particular, heuristic estimations of the introduced penalty parameters via the dispersion relation are suggested. Finally, the numerical schemes are given in details and numerical results are presented and compared for both exact and asymptotic solutions, namely gray solitons and dispersive shock waves, for which the asymptotic behavior is obtained through Whitham's modulation equations.

In chapter three, the same approach is applied to thin film flows with capillarity. Details on the used model are first reminded, in particular the governing equations, the dispersion relation and the following stability condition. The corresponding augmented model is then

presented and analyzed. The main difference in this case, is that the expressions of the phase velocities could not be obtained explicitly. Instead, an asymptotic expansion of the latter were performed and compared at leading orders with the original system. This permitted to obtain a suitable scaling of the penalty parameters as a function of the long-wave parameter $\varepsilon$, which governs the asymptotics of the original model. Then it is explained, how the augmented Lagrangian approach is extendable to energy forms that have nonlinear surface tension terms. The corresponding system of equations is computed and shown hyperbolic. Finally, numerical results for models with linear and nonlinear surface tension terms are presented and compared with reference numerical solutions obtained in [14] as well as the experimental results of Liu and gollub [57].

Chapter four, deals with stationary droplets on a horizontal solid substrate, under the effects of gravity and surface tension. The governing equilibrium equations are derived through variational principles and cast into a first order differential equation satisfied by the droplet height. Under some assumptions on the energy form, we discuss under which conditions, a droplet may admit a change of convexity. The study is conducted for both smooth and singular profiles. Most of the results are given in generic cases and then supplemented with explicit expressions of the total energy. The study is extended to the augmented Lagrangian model. This permitted to perform comparisons of the augmented and original models for a stationary solutions. This permitted to perform comparisons of the differential equation, as well as the numerical solutions, which showed excellent agreement between both models.

Lastly, chapter five, which is completely independent from the rest, deals with a pure numerical analysis problem, which is stability conditions derived from modified equations. In particular, we discuss some limitations of the modified equations approach as a tool for stability analysis for a class of explicit linear schemes to scalar partial derivative equations. We show that the infinite series obtained by Fourier transform of the modified equation is not always convergent and that in the case of divergence, it becomes unrelated to the scheme. Based on these results, we explain when the stability analysis of a given truncation of a modified equation may yield a reasonable estimation of a stability condition for the associated scheme. We illustrate our analysis by some examples of schemes namely for the heat equation and the transport equation.

The work presented in these chapters, is complemented by four appendixes, which contain most of the computational details. Appendix A includes the variational calculus required to derive the augmented system of equations from the Lagrangian and details about the augmented model's properties. Step-by-step computations to obtain the dispersion relation and obtain an explicit form of the IMEX scheme are also given. Appendix B is a reminder on elliptic functions and integrals. This includes practical examples on their usage in order to obtain periodic solutions of the nonlinear Schrödinger equation, and to describe the structure of a dispersive shock wave. Appendix C concerns mainly the asymptotic analysis performed for thin films equations. Lastly, appendix D contains details on how to compute some radii of convergence of some explicit power series, needed to explain the stability of some considered finite differences schemes.

CHAPTER 1

# A First order Hyperbolic model for Euler Korteweg Systems

In this chapter we explain the core idea behind this PhD work, namely the mechanisms behind the augmented Lagrangian approach for the E-K system. We first explain the concept of the method and how it affects the structure of the system from the variational point of view. Next, we explain how to obtain the corresponding system of augmented equations and then justify how it approximates the original E-K system of equations and finally discuss its hyperbolicity. Before we proceed any further, it seems necessary to remind that E-K systems refer in general to the following system of conservation laws :

$$\begin{cases} \dfrac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{u}) = 0 \\ \dfrac{\partial \rho \mathbf{u}}{\partial t} + \operatorname{div}\left(\rho \mathbf{u} \otimes \mathbf{u} + \Pi\right) = 0 \end{cases} \tag{1.1}$$

where $\Pi$ is given by :

$$\Pi = \left(\rho^2 \epsilon'(\rho) - \frac{1}{2}\left(\rho K'(\rho) + K(\rho)\right)|\nabla \rho|^2 - \rho K(\rho)\Delta \rho\right)\mathbf{Id} + K(\rho)\nabla\rho \otimes \nabla\rho \tag{1.2}$$

Here, $\rho$ is the fluid density, $\mathbf{u}$ is the velocity field, $\epsilon(\rho)$ is the specific hydrodynamic energy and $K(\rho)$ is a given function of $\rho$. This system can be derived as the Euler-Lagrange equations for the Lagrangian :

$$\mathcal{L} = \int_{\Omega_t} \left(\rho \frac{|\mathbf{u}|^2}{2} - \rho\epsilon(\rho) - \frac{1}{2}K(\rho)|\nabla\rho|^2\right) d\Omega \tag{1.3}$$

E-K system admits an additional scalar conservation law for the total energy :

$$\frac{\partial E}{\partial t} + \operatorname{div}\left(E\mathbf{u} + \Pi\mathbf{u} - \dot{\rho}K(\rho)\nabla\rho\right) = 0, \tag{1.4}$$

where the energy $E$ is given by :

$$E = \rho\frac{|\mathbf{u}|^2}{2} + \rho\epsilon(\rho) + \frac{1}{2}K(\rho)|\nabla\rho|^2. \tag{1.5}$$

Compared to classical models, for which the potential energy does not depend explicitly on $\nabla\rho$, we have an additional term in the energy flux $-\dot{\rho}K(\rho)\nabla\rho$ known as *interstitial working* [28]. This system covers the majority of the applications presented in this thesis. Even if the physical meaning of unknowns will be not the same as in E-K system, the mathematical structure of the governing models will remain the same.

## 1.1 Augmented Lagrangian approach

### 1.1.1 The concept

The method we use here follows the same philosophy introduced in [33] for the Serre-Green-Naghdi equations. While many relaxation techniques involve straightforward modifications in the system of equations, we prefer to rather act on the Lagrangian. This has the advantage of preserving the Hamiltonian structure of the obtained equations. Let us proceed step by step from the Lagrangian (1.3) as a starting point. Normally, in order to obtain the E-K equations, one can apply Hamilton's principle in order to minimize the Lagrangian (1.3) under the mass conservation constraint. In terms of variations, mass conservation links the velocity and the density (and consequently the gradient of density) in such a way that their respective Eulerian variations can be expressed in terms of the virtual displacement of the continuum. The idea is to create a new unconstrained variable that will approximate the gradient of density, while being independent from mass conservation. For this purpose, let us consider a new variable, which we call $\eta(\mathbf{x},t)$. The key idea here is to guarantee that $\eta$ converges towards $\rho$ in a certain limit and then substitute $\nabla\rho$ by $\nabla\eta$ in the Lagrangian. Hence, consider the new augmented Lagrangian :

$$\mathcal{L}=\int_{\Omega_t}\left(\rho\frac{|\mathbf{u}|^2}{2}-\rho\epsilon(\rho)-\frac{1}{2}K(\rho)\left|\nabla\eta\right|^2-\frac{\rho}{2\alpha}\left(1-\frac{\eta}{\rho}\right)^2+\frac{\beta}{2}\rho\dot{\eta}^2\right)d\Omega \tag{1.6}$$

Let us explain how this Lagrangian meets the expectations of the above description. First, the so-called augmented Lagrangian is constructed from the original one (1.3) by replacing $\nabla\rho$ by $\nabla\eta$ and adding two terms :

- The term $\dfrac{\rho}{2\alpha}\left(1-\eta/\rho\right)^2$ is a classical penalty term where $\alpha$ is a small parameter. Indeed, when $\alpha\to 0$, the difference $(1-\eta/\rho)$ is expected to vanish, implying that $\eta\to\rho$. The choice of the form of this term follows the discussion in [33] for the Serre-Green-Naghdi equations, where it is shown that some other formulations can result in the loss of hyperbolicity or the cancellation of dispersive effects.

- The term $\dfrac{\beta}{2}\rho\dot{\eta}^2$ ensures a regular time evolution of $\eta$. $\beta$ is a small parameter so that the contribution of this additional term remains negligible. The form of this term is also reminiscent of the dispersion in the Serre-Green-Naghdi equations. So, it can be perceived as a small micro-inertia term.

### 1.1.2 Augmented system of equations

Given the Lagrangian (1.6), let us consider the associated Hamilton's action in an arbitrary time interval $[t_0, t_1]$ :

$$a = \int_{t_0}^{t_1} \mathcal{L} dt \tag{1.7}$$

In a more compact form, we can write the augmented Lagrangian as :

$$\mathcal{L}^e = \int_{\Omega_t} \left( \rho \frac{|\mathbf{u}|^2}{2} - W(\rho, \eta, \dot{\eta}, \nabla\eta) \right) d\Omega. \tag{1.8}$$

where $W$ is the potential given by :

$$W(\rho, \eta, \dot{\eta}, \nabla\eta) = \rho\epsilon(\rho) + \frac{1}{2}K(\rho)\left|\nabla\eta\right|^2 + \frac{\rho}{2\alpha}\left(1 - \frac{\eta}{\rho}\right)^2 - \frac{\beta}{2}\rho\dot{\eta}^2 \tag{1.9}$$

We remind that the variation of Hamilton's action is submitted to the constraint which is the mass conservation law :

$$\frac{\partial\rho}{\partial t} + \operatorname{div}(\rho\mathbf{u}) = 0. \tag{1.10}$$

For the sake of keeping this part light and readable, applying Hamilton's principle and all the related calculus will be carried out in the appendix A.1. Nevertheless, it seems necessary to give a few words on how the differential calculus is done in order to highlight some differences between the augmented model and the original one in terms of variations. In the original Lagrangian, all the variables' variations are linked to the virtual displacement of the continuum. In fact they write :

$$\hat{\delta}\rho = -\operatorname{div}(\rho\delta\mathbf{x}), \quad \hat{\delta}\mathbf{u} = (\dot{\delta\mathbf{x}}) - \frac{\partial\mathbf{u}}{\partial\mathbf{x}}\delta\mathbf{x}, \quad \hat{\delta}\nabla\rho = \nabla\left(\hat{\delta}\rho\right) = -\nabla\left(\operatorname{div}(\rho\delta\mathbf{x})\right) \tag{1.11}$$

Here, the 'hat' means that we use variations at fixed Eulerian variables (see further details about Lagrangian and Eulerian variations in [26]). Note that the variation of the gradient of density is obtained simply by using Schwartz's theorem to interchange the order of derivatives. Thus, in this case applying Hamilton's principle results in one Euler-Lagrange equation that is the momentum balance equations of E-K system .

In the case of the augmented Lagrangian, the gradient $\nabla\eta$ which substitutes the gradient of density is no longer dependent on the density and velocity. Thus, it depends only on the new unconstrained variable $\eta$. Naturally in this case, one needs to separate the variables in two sets depending on their respective variations. The first type corresponds to the usual virtual displacement of the continuum, respecting the mass conservation law and will be denoted by $\delta\mathbf{x}$. The second one corresponds to the variation of $\eta$ as an independent variable and will be denoted by $\delta\eta$. The independent variables upon which the augmented Lagrangian depends classify into these types according to the following schematic :

$$\overbrace{\mathbf{u}, \rho, \underbrace{\dot{\eta}}, \eta, \nabla\eta}^{\delta\mathbf{x}} \qquad \underbrace{\qquad}_{\delta\eta}$$

Note that $\dot{\eta}$ is included in both kinds, as it depends explicitly on derivatives of $\eta$ but also on the velocity field $\mathbf{u}$. Let us now detail how these variations write and express the corresponding Euler-Lagrange equation in each case.

**Variation with respect to $\delta\mathbf{x}$**

The variations of the density, velocity and $\dot{\eta}$ are related to $\delta\mathbf{x}$ by :

$$\hat{\delta}\rho = -\mathrm{div}\left(\rho\delta\mathbf{x}\right), \quad \hat{\delta}\mathbf{u} = (\dot{\delta\mathbf{x}}) - \frac{\partial\mathbf{u}}{\partial\mathbf{x}}\delta\mathbf{x}, \quad \hat{\delta}\dot{\eta} = \hat{\delta}\mathbf{u}\cdot\nabla\eta. \tag{1.12}$$

Applying Hamilton's principle on the action (1.7), under the mass conservation law, yields the following Euler-Lagrange equation (Details in Appendix A.1.1):

$$\boxed{\frac{\partial\rho\mathbf{u}}{\partial t} + \mathrm{div}\left(\rho\mathbf{u}\otimes\mathbf{u} + \left(\rho\frac{\partial W}{\partial\rho} - W\right)\mathbf{Id} + \frac{\partial W}{\partial\nabla\eta}\otimes\nabla\eta\right) = 0} \tag{1.13}$$

**Variation with respect to $\delta\eta$**

By virtue of the Schwartz's theorem, the variations of $\nabla\eta$ and $\dot{\eta}$ are related to $\delta\eta$ by :

$$\delta(\nabla\eta) = \nabla\left(\delta\eta\right) \tag{1.14}$$

$$\delta\dot{\eta} = \delta\left(\frac{\partial\eta}{\partial t}\right) + \frac{\partial\delta\eta}{\partial\mathbf{x}}\mathbf{u} = \frac{\partial\delta\eta}{\partial t} + \frac{\partial\delta\eta}{\partial\mathbf{x}}\mathbf{u} \tag{1.15}$$

Similarly, applying Hamilton's principle on the action (1.7) results in the Euler-Lagrange equation (Details are given in Appendix A.1.2):

$$\boxed{\frac{\partial}{\partial t}\left(\frac{\partial W}{\partial\dot{\eta}}\right) + \mathrm{div}\left(\frac{\partial W}{\partial\dot{\eta}}\mathbf{u} + \frac{\partial W}{\partial\nabla\eta}\right) = \frac{\partial W}{\partial\eta}} \tag{1.16}$$

In order to make the obtained system of equations first order hyperbolic, we perform a last step which is a sort of an order reduction. Indeed, let us consider the variables :

$$\mathbf{p} = \nabla\eta, \quad w = \dot{\eta} \tag{1.17}$$

Under these notations, if we gather the previous equations, one obtains a set of three conservation laws (Mass conservation + 2 Euler-Lagrange equations) that are dependent on five independent variables $(\rho, \mathbf{u}, \eta, \nabla\eta, \dot{\eta})$. Therefore, we need to provide two additional equations in order to close the system. First, by definition of $w$ one gets :

$$\frac{\partial\eta}{\partial t} + \frac{\partial\eta}{\partial\mathbf{x}}\mathbf{u} = w \tag{1.18}$$

This already depicts the evolution of $\eta$ in time. Furthermore, by taking the gradient of this equation and applying Schwartz's theorem, we obtain :

$$\frac{\partial\nabla\eta}{\partial t} + \nabla\left(\frac{\partial\eta}{\partial\mathbf{x}}\mathbf{u} - w\right) = 0 \tag{1.19}$$

which can be put in the form :

$$\frac{\partial \mathbf{p}}{\partial t} + \operatorname{div}\left(\left(\frac{\partial \eta}{\partial \mathbf{x}}\mathbf{u} - w\right)\mathbf{Id}\right) = 0 \tag{1.20}$$

Thus, now the system of equations is completed. Expanding the expression of $W$ in both Euler-Lagrange equations and putting all the equations in their conservative form permits us to obtain the following system:

$$\begin{cases} \dfrac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{u}) = 0, \\ \dfrac{\partial \rho \mathbf{u}}{\partial t} + \operatorname{div}\left(\rho \mathbf{u} \otimes \mathbf{u} + \mathcal{P}\right) = 0, \\ \dfrac{\partial \rho w}{\partial t} + \operatorname{div}\left(\rho w \mathbf{u} - \dfrac{1}{\beta} K(\rho)\mathbf{p}\right) = \dfrac{1}{\alpha\beta}\left(1 - \dfrac{\eta}{\rho}\right), \\ \dfrac{\partial \rho \eta}{\partial t} + \operatorname{div}(\rho \eta \mathbf{u}) = \rho w, \\ \dfrac{\partial \mathbf{p}}{\partial t} + \operatorname{div}\left(\left(\mathbf{p} \cdot \mathbf{u} - w\right)\mathbf{Id}\right) = 0 \end{cases} \tag{1.21}$$

where the tensor $\mathcal{P}$ is given by :

$$\mathcal{P} = \left(\rho^2 \epsilon'(\rho) + \frac{1}{2}(\rho K'(\rho) - K(\rho))|\mathbf{p}|^2 + \frac{\eta}{\alpha}\left(1 - \frac{\eta}{\rho}\right)\right)\mathbf{Id} + K(\rho)\mathbf{p} \otimes \mathbf{p} \tag{1.22}$$

One can derive an additional conservation law for the total energy of the augmented system :

$$\frac{\partial E}{\partial t} + \operatorname{div}\left(E\mathbf{u} + \mathbf{P}\mathbf{u} - K(\rho) w \mathbf{p}\right) = 0,$$

where the total energy expression writes :

$$E = \rho \frac{|\mathbf{u}|^2}{2} + \rho\epsilon(\rho) + \frac{1}{2}K(\rho)|\mathbf{p}|^2 + \frac{1}{2\alpha}\rho\left(\frac{\eta}{\rho} - 1\right)^2 + \frac{\beta}{2}\rho w^2.$$

It is worthy of note, that even for the augmented system, the energy conservation law still contains, as in the case of original system the interstitial working term written here in the form $-K(\rho)w\mathbf{p}$. It remains now to check whether or not the system of equations (1.21) fulfilled all the required conditions, namely, is the obtained system first order hyperbolic and does it approximate Euler-Korteweg equations? Let us first address the latter and compare both original and augmented system.

**Remark 1.1.1.** *In the augmented setting,* $\mathbf{p}$ *is by definition the gradient of* $\eta$*. Thus, it should satisfy the compatibility condition* $\operatorname{curl}(\mathbf{p}) = \mathbf{0}$*. Even though we will focus mainly on one-dimensional applications in the current work, it should be kept in mind that* $\mathbf{p}$ *should be curl-free in the multi-dimensional case.*

**Remark 1.1.2.** *The used hyperbolic reformulation changes the definition of a potential flow. In fact, for Euler-Korteweg fluids, the definition of a potential flow is classical :* $\mathbf{u} = \nabla\phi$ *and hence* $\mathrm{curl}(\mathbf{u}) = 0$*. In the hyperbolic setting, the flow is potential if* $\mathbf{K} = \nabla\phi$ *where* $\mathbf{K}$ *is defined by* [37] *:*

$$\mathbf{K} = \mathbf{u} - \frac{1}{\rho}\frac{\partial W}{\partial w}\mathbf{p} \tag{1.23}$$

*This is a consequence from the fact that the augmented Lagrangian depends on the material derivative of* $\eta$ *(See Appendix* A.4 *for details).*

## 1.2 Comparison with the original equations

All we need to compare are the mass and momentum balance equations. The remaining equations of the augmented system serve mainly as intermediate steps and couplings in the augmentation process. Let us proceed from the third equation in system (1.21). In fact, multiplying it by $\beta$ yields :

$$\beta\frac{\partial \rho w}{\partial t} + \beta \mathrm{div}\left(\rho w \mathbf{u}\right) - \mathrm{div}\left(K(\rho)\mathbf{p}\right) = \frac{1}{\alpha}\left(1 - \frac{\eta}{\rho}\right) \tag{1.24}$$

which means that :

$$\frac{1}{\alpha}\left(1 - \frac{\eta}{\rho}\right) = -\left(K(\rho)\Delta\eta + K'(\rho)\nabla\rho\cdot\nabla\eta\right) + \beta\rho\dot{w} \tag{1.25}$$

Assuming we work on smooth solutions in which no singularities arise so that $\nabla\eta$ and $\Delta\eta$ remain bounded quantities, the last equation gives the relative error between $\rho$ and $\eta$ as :

$$\frac{\rho - \eta}{\rho} = -\alpha\left(K(\rho)\Delta\eta + K'(\rho)\nabla\rho\cdot\nabla\eta\right) + \alpha\beta\rho\dot{w} \tag{1.26}$$

First, as expected, one recovers a vanishing error in the limit $\alpha \to 0$. We can write :

$$\eta = \rho + \mathcal{O}(\alpha) \Rightarrow \begin{cases} \nabla\eta = \nabla\rho + \mathcal{O}(\alpha) \\ \Delta\eta = \Delta\rho + \mathcal{O}(\alpha) \end{cases} \tag{1.27}$$

With this asymptotic approximations taken into account we replace the expression (1.25) into the momentum equation, and we obtain after simplifications (see appendix A.2):

$$\frac{\partial\rho}{\partial t} + \mathrm{div}(\rho\mathbf{u}\otimes\mathbf{u}) + \nabla(p(\rho)) = \rho\nabla\left(K(\rho)\Delta\rho + \frac{1}{2}K'(\rho)\left|\nabla\rho\right|^2\right) + \mathcal{O}(\alpha) + \mathcal{O}(\beta) \tag{1.28}$$

This illustrates the fact that the augmented system approaches the original E-K system in the limits $\alpha \to 0$ and $\beta \to 0$, which answers to one of the requirements of this model. It remains now to justify its hyperbolicity, which we investigate in the following section.

## 1.3 Hyperbolicity of the Augmented E-K equations

While in one dimension of space the question of hyperbolicity of the augmented E-K system is a trivial matter, it requires some additional manipulations in the multidimensional case which we will discuss in the current section. For this purpose, it is more practical to cast the system (1.21) into the form :

$$\frac{\partial \mathbf{U}}{\partial t} + \mathbf{A}\frac{\partial \mathbf{U}}{\partial x} + \mathbf{B}\frac{\partial \mathbf{U}}{\partial y} + \mathbf{C}\frac{\partial \mathbf{U}}{\partial z} = \mathbf{S}(\mathbf{U}) \tag{1.29}$$

In a general framework, let us define hyperbolicity based on such a system of PDEs [21][67] :

**Definition 1.3.1.** *Consider a smooth hypersurface $H(t,x,y,z) = 0$. Then, under the following notations :*

$$\tau = H_t; \quad \xi = H_x; \quad \nu = H_y; \quad \zeta = H_z, \tag{1.30}$$

*the hypersurface is charactersitic if :*

$$\det\left(\tau\mathbf{I} + \xi\mathbf{A} + \nu\mathbf{B} + \zeta\mathbf{C}\right) = 0 \tag{1.31}$$

*The system of equations* (1.29) *is hyperbolic if for every $(\xi,\nu,\zeta)^T$, the matrix $\mathcal{A}$ defined as :*

$$\mathcal{A} = \xi\mathbf{A} + \nu\mathbf{B} + \zeta\mathbf{C} \tag{1.32}$$

*has all its eigenvalues real and admits a basis of eigenvectors.*

When it comes to practice, some considerations may fairly reduce the amount of work needed to show hyperbolicity. In particular, if the system of equations (1.29) is invariant by rotations of the $SO(3)$ Group, it is sufficient to consider the 1D case, that is the conserved variables and their respective fluxes only depend on time and on one dimension of space ($\mathbf{U}(t,x,y,z) = \mathbf{U}(t,x)$)[60]. We can prove that system of equations (1.21) is in fact invariant by rotation (see proof in Appendix A.3) and consequently, we will make use of such simplifications. Lastly and most importantly, there is a subtle change of structure that is required in order to obtain hyperbolicity in the multidimensional case. This last step relies on the fact that by definition, the vector $\mathbf{p}$ is equal to $\nabla\eta$ and is thus irrotational. Hence, adding terms which are proportional to curl($\mathbf{p}$) only alters the form of the system. Following this, we rewrite the momentum conservation equation and $\mathbf{p}$ conservation equation respectively as follows :

$$\frac{\partial \rho\mathbf{u}}{\partial t} + \operatorname{div}\left(\rho\mathbf{u}\otimes\mathbf{u} + \mathcal{P}\right) - K(\rho)\operatorname{curl}(\mathbf{p})\wedge\mathbf{p} = 0 \tag{1.33}$$

$$\frac{\partial \mathbf{p}}{\partial t} + \operatorname{div}\left(\left(\mathbf{p}\cdot\mathbf{u} - w\right)\mathbf{Id}\right) + \operatorname{curl}(\mathbf{p})\wedge\mathbf{u} = 0 \tag{1.34}$$

In this new form, and with all assumptions taken into consideration, the augmented E-K system can be put into the quasilinear form :

$$\frac{\partial \mathbf{U}}{\partial t} + \mathbf{A}(\mathbf{U})\frac{\partial \mathbf{U}}{\partial x} = 0 \tag{1.35}$$

in which the vector of state variables and the quasilinear matrix are respectively given by :

$$\mathbf{U} = \begin{pmatrix} \rho \\ u \\ v \\ w \\ p_1 \\ p_2 \\ \eta \end{pmatrix}, \quad \mathbf{A}(\mathbf{U}) = \begin{pmatrix} u & \rho & 0 & 0 & 0 & 0 & 0 \\ a_{21} & u & 0 & 0 & a_{25} & p_2 K'(\rho) & \frac{\rho - 2\eta}{\alpha\rho^2} \\ \frac{p_1 p_2 K'(\rho)}{\rho} & 0 & u & 0 & \frac{p_2 K(\rho)}{\rho} & 0 & 0 \\ -\frac{p_1 K'(\rho)}{\beta\rho} & 0 & 0 & u & -\frac{K(\rho)}{\beta\rho} & 0 & 0 \\ 0 & p_1 & p_2 & -1 & u & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & u & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & u \end{pmatrix} \tag{1.36}$$

$$a_{21} = \frac{K'(\rho)p_1^2}{\rho} + 2\epsilon'(\rho) + \frac{1}{2}\left(p_1^2 + p_2^2\right)K''(\rho) + \rho\epsilon''(\rho) + \frac{\eta^2}{\alpha\rho^3}, \quad a_{25} = p_1\left(\frac{K(\rho)}{\rho} + K'(\rho)\right)$$

We have chosen here a set of non-conservative equations since it seemed easier for computations. Obviously, this does not affect in any way the hyperbolicity property. Next, a straightforward computation shows that the eigenvalues of $\mathbf{A}$, which correspond to the characteristic speeds of the augmented system are :

$$\xi_{1,2,3} = u, \quad \xi_4^{\pm} = u \pm \sqrt{\frac{b - \sqrt{b^2 - 4c}}{2}}, \quad \xi_5^{\pm} = u \pm \sqrt{\frac{b + \sqrt{b^2 - 4c}}{2}} \tag{1.37}$$

where :

$$b = \rho\left(2\epsilon'(\rho) + \frac{1}{2}\left(p_1^2 + p_2^2\right)K''(\rho) + \rho\epsilon''(\rho) + \frac{\eta^2}{\alpha\rho^3}\right) + \frac{p_1^2 + p_2^2}{\rho}K(\rho) + \frac{K(\rho)}{\beta\rho} + 2p_1^2K'(\rho) \tag{1.38}$$

$$c = \left(p_2^2 + \frac{1}{\beta}\right)\left(K(\rho)\left(2\epsilon'(\rho) + \frac{1}{2}\left(p_1^2 + p_2^2\right)K''(\rho) + \rho\epsilon''(\rho) + \frac{\eta^2}{\alpha\rho^3}\right) - 2p_1^2K'(\rho)^2\right) \tag{1.39}$$

In the general case where $\epsilon(\rho)$ and $K(\rho)$ are abstract functions, the computation of the eigenvectors is quite tedious and results in very heavy expressions. This task will be done separately in the next chapters for explicitly defined systems such as the NLS equation. Nevertheless, we will show here the following proposition:

**Proposition 1.3.2.** *If the total energy associated to the augmented E-K system is convex in the conserved variables, then the system is at least weakly hyperbolic.*

*Proof.* By definition, we need to show that all the eigenvalues (1.37) are real. This only requires us to prove that :

$$\begin{cases} b^2 - 4c \geq 0 \\ b \pm \sqrt{b^2 - 4c} \geq 0 \end{cases} \quad \text{or equivalently} \quad \begin{cases} b^2 - 4c \geq 0 \\ b \geq 0 \\ c \geq 0 \end{cases} \tag{1.40}$$

In order to show this, let us first cast $b$ and $c$ into a more convenient form, in which, the derivatives of the total energy appear explicitly :

$$b = \frac{1}{\rho}\left(\rho^2 E_{\rho\rho} + p_1^2 E_{p_1p_1} + 2p_1\rho E_{p_1\rho} + p_2^2 E_{p_1p_1} + \frac{1}{\beta}E_{p_1p_1}\right) = \frac{1}{\rho}\left(\mathbf{e}^t\mathbf{H}\mathbf{e} + p_2^2 E_{p_2p_2} + \frac{1}{\beta}E_{p_1p_1}\right)$$

$$c = \left(p_2^2 + \frac{1}{\beta}\right)\left(E_{\rho\rho}E_{p_1p_1} - E_{p_1\rho}^2\right)$$

The convexity of the total energy with respect to the conserved variables $(\rho, \rho u, \rho v, \rho w, p_1, p_2, \rho\eta)$ implies its convexity in every sub-base and in particular in $(\rho, p_1)$. The Hessian matrix in this sub-base

$$\mathbf{H} = \begin{pmatrix} E_{\rho\rho} & E_{\rho p_1} \\ E_{\rho p_1} & E_{p_1p_1} \end{pmatrix}$$

is thus semi-definite positive which implies that :

$$b = \frac{1}{\rho}\left(\mathbf{e}^t\mathbf{H}\mathbf{e} + p_2^2 E_{p_2p_2} + \frac{1}{\beta}E_{p_1p_1}\right) \geq 0, \qquad \mathbf{e}^T = (\rho, p_1)$$

$$c = \left(p_2^2 + \frac{1}{\beta}\right)\det(\mathbf{H}) \geq 0$$

It only remains to show that the discriminant $b^2 - 4c \geq 0$. Let us write :

$$\begin{aligned}
b^2 - 4c^2 &= \frac{1}{\rho^2}\left(\rho^2 E_{\rho\rho} + p_1^2 E_{p_1p_1} + 2p_1\rho E_{p_1\rho} + \left(\frac{1}{\beta} + p_2^2\right)E_{p_1p_1}\right)^2 - 4\left(p_2^2 + \frac{1}{\beta}\right)\left(E_{\rho\rho}E_{p_1p_1} - E_{p_1\rho}^2\right) \\
&= \frac{1}{\rho^2}\left(\rho^2 E_{\rho\rho} + p_1^2 E_{p_1p_1} + 2p_1\rho E_{p_1\rho} - \left(\frac{1}{\beta} + p_2^2\right)E_{p_1p_1}\right)^2 - 4\left(p_2^2 + \frac{1}{\beta}\right)\left(E_{\rho\rho}E_{p_1p_1} - E_{p_1\rho}^2\right) \\
&\quad + \frac{4}{\rho^2}\left(p_2^2 + \frac{1}{\beta}\right)\left(\rho^2 E_{\rho\rho}E_{p_1p_1} + p_1^2 E_{p_1p_1}^2 + 2p_1\rho E_{p_1\rho}E_{p_1p_1}\right) \\
&= \frac{1}{\rho^2}\left(\rho^2 E_{\rho\rho} + p_1^2 E_{p_1p_1} + 2p_1\rho E_{p_1\rho} - \left(\frac{1}{\beta} + p_2^2\right)E_{p_1p_1}\right)^2 + \frac{4}{\rho^2}\left(p_2^2 + \frac{1}{\beta}\right)\left(p_1 E_{p_1p_1} + \rho E_{p_1\rho}\right)^2 \geq 0
\end{aligned}$$

Thus all the eigenvalues of $A$ are real which concludes the proof. □

**Remark 1.3.3.** *Obtaining the explicit expression of the whole eigensystem of the augmented system is a nontrivial task in general. The triple eigenvalue $\xi = u$ always admits three independent eigenvectors which are given by :*

$$\begin{aligned}
\mathbf{v_1} &= \left(\frac{(2\eta - \rho)K(\rho)}{\alpha\rho^2\left(a_{21}K(\rho) - a_{25}p_1K'(\rho)\right)}, 0, 0, 0, -\frac{p_1(2\eta - \rho)K'(\rho)}{\alpha\rho^2\left(a_{21}K(\rho) - a_{25}p_1K'(\rho)\right)}, 0, 1\right) \\
\mathbf{v_2} &= \left(-\frac{p_2K(\rho)K'(\rho)}{a_{21}K(\rho) - a_{25}p_1K'(\rho)}, 0, 0, 0, -\frac{p_1p_2B'(\rho)^2}{a_{25}p_1K'(\rho) - a_{21}K(\rho)}, 1, 0\right) \\
\mathbf{v_3} &= (0, 0, 1, p_2, 0, 0, 0)
\end{aligned}$$

*For the remaining eigenvalues, additional analysis is needed, as some of the eigenvalues may coincide with one another. This is actually the case if we have $c = 0$, in which case $\xi_4^{\pm} = u$, or $b^2 = 4c$, in which case $\xi_4^{\pm} = \xi_5^{\pm}$.*

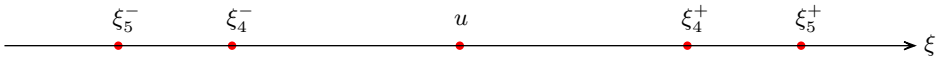


Figure 1.1: Representation of the eigenvalues of the augmented E-K system. $u$ is a triple eigenvalue. $\xi_4^\pm$ can coincide with $u$ and $\xi_4^\pm$ can coincide with $\xi_5^\pm$.

## 1.4 Dispersion Relation

For a dispersionless system, waves propagate at a velocity that is independent from the wavelength. This is not the case in dispersive systems and the fundamental relation that relates the velocity to the wavenumber is called the dispersion relation. Besides providing the expression of the wave velocity, it also permits to quantify the damping or the amplification of small perturbations to the system and consequently check its stability with respect to the perturbation frequency. In the general case, for nonlinear systems such as the one we study here, obtaining the exact dispersion relation is a major challenge, and in most cases it is impossible to obtain. Nevertheless, it is usually possible to derive a linearized version that still provides relevant information for small enough disturbances. Hence, let us consider a monochromatic perturbation of a constant equilibrium state $\mathbf{U}_0 = (\rho_0, 0, \rho_0, 0, 0)^T$ defined by, $\mathbf{U}(x,t) = \mathbf{U_0} + \mathbf{U}'e^{i(kx-\omega t)}$. Plugging this expression in our system of equations and linearizing allows us to obtain an expression of the phase velocity $c_p = \omega/k$ as an eigenvalue of the matrix $\mathbf{A}(\mathbf{U}_0) + i/k\nabla\mathbf{S}(\mathbf{U}_0)$. (See details in Appendix A.5). Its explicit expression is given by :

$$(c_p(k)^\pm)^2 = \frac{1}{2}\left(B(k) \pm \sqrt{B(k)^2 - 4C(k)}\right) \tag{1.41}$$

where the functions $B(k)$ and $C(k)$ are given by :

$$B(k) = \frac{K(\rho)}{\beta\rho} + \frac{1}{\alpha} + \frac{1}{\alpha\beta k^2\rho^2} + \rho^2\epsilon''(\rho) + 2\rho\epsilon'(\rho) \tag{1.42}$$

$$C(k) = (\rho^2\epsilon''(\rho) + 2\rho\epsilon'(\rho))\left(\frac{K(\rho)}{\beta\rho} + \frac{1}{\alpha\beta k^2\rho^2}\right) + \frac{K(\rho)}{\alpha\beta\rho} \tag{1.43}$$

Unlike the original E-K system, there are two possible positive values for the phase velocity which we call $c_p^+$ and $c_p^-$. One can see through a simple limit computation that for $c_p^-$, in the limits $\alpha \to 0$ and $\beta \to 0$, one recovers the original phase velocity for the E-K system which is given by :

$$c_p(k)^2 = \rho^2\epsilon''(\rho) + 2\rho\epsilon'(\rho) + k^2\rho K(\rho). \tag{1.44}$$

The value $c_p^+$ is rather linked to the fact that the system is 'mimicking' a dispersive system by virtue of hyperbolic equations. In fact, while dispersive systems tolerate perturbations that are felt across the whole domain instantaneously, hyperbolic equations only allow for finite propagation velocities and hence bounded phase velocities. Thus, in the limits $\alpha \to 0$ and $\beta \to 0$, the phase velocity $c_p^+$ becomes infinite to exhibit the same instantaneous behavior as a dispersive system. For fixed $\alpha > 0$ and $\beta > 0$, one can say that the approximate augmented

system approaches the instantaneous propagation velocity through very fast characteristics. Lastly, the dispersion relation also proves useful to get an additional estimate for the introduced parameters $\alpha$ and $\beta$. While the estimate remains heuristic, it is based on the fact that both the original system and the augmented one should display close properties in terms of wave velocities, stability regions and amplitude damping, since these aspects are of a significant physical relevancy. Most of these aspects are strongly dependent on the explicit system and are of no concern in the generic case as the E-K system has no inherent dissipation or energy production. Consequently, more explicit analysis will be performed in the next chapters.

CHAPTER 2

# Non-Linear Schrödinger equation

## 2.1 About the NLS equation

### 2.1.1 A brief history

The Non-linear Schrödinger equation (NLSE) is a nonlinear version of the Schrödinger equation that has been a subject of extensive research in the latest 50 years. Since many versions of the equation commonly share the exact same name, it seems necessary, as to avoid ambiguity to point out that we refer precisely to the cubic Non-Linear Schrödinger equation, which is given by :

$$i\psi_t + \frac{1}{2}\Delta\psi + \kappa\left|\psi\right|^2\psi = 0, \tag{2.1}$$

where $\psi(\mathbf{x}, t)$ is a complex-valued field. Although the appellation *Schrödinger equation* may invoke quantum mechanics and therefore lead the reader into associating $\psi$ with a quantum state vector, NLSE is widely used in classical physics and represents a universal model to describe wave propagation in nonlinear media with dispersion. Variants of the equation appeared since 1950, for example in the Ginzburg-Landau theoretical works on superconductivity [53, 52] and in the works of Ginzburg-Pitaevskii on the theory of superfluidity [40]. It was only later that practical applications of NLSE appeared, for example in nonlinear optics with the works of Chiao [19] and Talanov [69] in 1964, describing the propagation of light beams in nonlinear media, allowing for the beam to focus rather than spread, thus creating zones of very high intensity. A few years later in 1968, Zakharov while working on the stability of surface gravity waves in an infinitely deep fluid [76] established that for slowly modulated surface waves, the wave amplitude is governed by the NLS equation. In the same context, Benney and Roskes derived in [6] a variant of NLS, namely the Benney-Roskes equation (also referred to as the Davey-Stewartson system [23]) and obtained similar results for a finite depth fluid.
From the mathematics point of view, NLSE is also a very interesting object of analysis and its popularity among mathematicians increased significantly after the (1+1)-dimensional case was shown integrable by Zakharov and Shabat [68] by virtue of the inverse scattering transform, under which setting, the latter equation becomes a compatibility condition for a system of two linear equations, namely the Zakharov-Shabat system. Another major contribution

to its arising fame in mathematics, is the existence and uniqueness results by Ginibre and Vélo [39], on the associated initial value problem to the NLSE (in a more general case which treats a power-law nonlinearity).
The above list of contributions is but a carefully handpicked selection, and is by no means exhaustive. For more thorough informations, the reader may consult for example [34, 1]. Nevertheless, one can see that the aforementioned examples already involve many fields of physics, of theoretical and practical interest, and thus demonstrate how versatile and worth considering is the NLSE, for both physicists and mathematicians.

### 2.1.2 Focusing versus Defocusing NLSE

At this point it seems necessary to make the distinction between the so-called defocusing and focusing NLSE, as they exhibit fundamentally different behaviors. The term *focusing* seems to find its original meaning in nonlinear optics, in the description of the propagation of beams in nonlinear media. For instance consider a medium whose refractive index $n$ changes accordingly to the amplitude of the applied electric field $E$, namely a Kerr medium [34]:

$$n^2 = n_0^2 + 4n_0n_2|E|^2 \tag{2.2}$$

This means that the medium gets more (*resp.* less) refractive for higher amplitudes of $E$ if $n_2$ is positive (*resp.* negative). In such a medium and for some conditions, the propagation of a continuous wave laser beam is modeled, at leading order by a NLS equation that can be cast into the form [34] :

$$i\psi_z + \Delta_\perp\psi + 4\frac{n_2}{n_0}|\psi|^2\psi = 0, \tag{2.3}$$

where $\psi(x, y, z)$ is the slowly-varying amplitude of the electrical field propagating along the $z$-axis, $\Delta_\perp = \partial_{xx}^2 + \partial_{yy}^2$ denotes the Laplacian in the transverse directions. Equation (2.3) tells that the propagation of the laser beam is only governed by the diffraction term (Laplace operator) and the nonlinear Kerr effects. It is shown by geometric arguments [34] that if both terms have the same sign, that is $n_2 > 0$, the beam will have a tendency to bend towards the axis and increase its intensity. Hence, this is called self-focusing of the light beam as opposed to self-defocusing when $n_2 < 0$. The same terminology holds in the general case so that equation (2.1) is called *focusing* NLSE if $\kappa > 0$ and *defocusing* NLSE if $\kappa < 0$.

### 2.1.3 Madelung Transformation

There exists a reformulation of NLSE which casts it into a system of conservation laws in terms of real variables. This reformulation, introduced by Erwin Madelung [58] and thus bearing his name, is done by decomposing $\psi(x, t)$ into a particular polar form that writes :

$$\psi(\mathbf{x}, t) = \sqrt{\rho(\mathbf{x}, t)}e^{i\theta(\mathbf{x},t)}. \tag{2.4}$$

Here, $\rho(\mathbf{x}, t) \geq 0$ and $\theta(\mathbf{x}, t)$ are real valued functions that represent the square of the modulus and the argument of $\psi(\mathbf{x}, t)$ respectively. Replacing into equation (2.1) gives :

$$\left(\frac{i}{2\rho}\rho_t - \theta_t\right) + \left(\frac{1}{2}\frac{\Delta(\sqrt{\rho})}{\sqrt{\rho}} + \frac{i}{2\rho}\mathrm{div}(\rho\nabla\theta) - \frac{1}{2}\nabla\theta\cdot\nabla\theta\right) + \kappa\rho = 0 \tag{2.5}$$

As a consequence, it is now possible to isolate the real and imaginary parts of the equation. If we further consider a vector field variable $\mathbf{u}(\mathbf{x}, t)$ defined as $\mathbf{u} = \nabla\theta$ , we obtain the following system:

$$\begin{cases} \rho_t + \operatorname{div}(\rho\mathbf{u}) = 0 \\ \mathbf{u}_t + (\mathbf{u}\cdot\nabla)\,\mathbf{u} - \nabla\left(\kappa\rho + \dfrac{\Delta(\sqrt{\rho})}{2\sqrt{\rho}}\right) = 0 \end{cases} \tag{2.6}$$

This formulation of NLSE is referred to as *Madelung equations*, *quantum hydrodynamics equations* or simply as *hydrodynamic form of NLSE*, since it bears similarities in many respects to Euler equations of hydrodynamics. More importantly, it represents a particular case of an Euler-Korteweg system with $K(\rho) = 1/4\rho$. System (2.6) can also be written in conservative form :

$$\begin{cases} \rho_t + \operatorname{div}(\rho\mathbf{u}) = 0, \\ (\rho\mathbf{u})_t + \operatorname{div}\left(\rho\mathbf{u}\otimes\mathbf{u} + \mathbf{\Pi}\right) = 0, \end{cases} \tag{2.7}$$

with the now usual stress tensor :

$$\mathbf{\Pi} = -\left(\kappa\frac{\rho^2}{2} + \frac{1}{4}\Delta\rho\right)\mathbf{Id} + \frac{1}{4\rho}\nabla\rho\otimes\nabla\rho. \tag{2.8}$$

System (2.7) admits an energy conservation law given by :

$$E_t + \operatorname{div}\left(E\mathbf{u} + \mathbf{\Pi}\mathbf{u} - \frac{\dot{\rho}}{4\rho}\nabla\rho\right) = 0, \tag{2.9}$$

with a total energy density expressed as :

$$E = \rho\frac{|\mathbf{u}|^2}{2} - \kappa\frac{\rho^2}{2} + \frac{1}{4\rho}\frac{|\nabla\rho|^2}{2}. \tag{2.10}$$

Consequently, the corresponding Lagrangian is given by :

$$\mathcal{L} = \int_{\Omega_t}\left(\rho\frac{|\mathbf{u}|^2}{2} + \kappa\frac{\rho^2}{2} - \frac{1}{4\rho}\frac{|\nabla\rho|^2}{2}\right)d\Omega. \tag{2.11}$$

It is worth noting that in the dispersionless case, the E-K system reduces to an Euler system, which is hyperbolic in the defocusing case and elliptic in the focusing case. Consequently, even though it may be worthwhile to test our augmented Lagrangian approach on the focusing NLSE, the hyperbolicity of the resulting system is not given. Thus, for the next part and the following numerical tests, we will retain only the defocusing cubic NLSE in (1+1) dimensions, that is :

$$i\psi_t + \frac{1}{2}\psi_{xx} - |\psi|^2\psi = 0, \tag{2.12}$$

to which corresponds the following Euler-Korteweg system :

$$\begin{cases} \rho_t + (\rho u)_x = 0, \\ (\rho u)_t + \left(\rho u^2 + \dfrac{\rho^2}{2} + \dfrac{1}{4\rho}(\rho_x^2 - \rho\rho_{xx})\right)_x = 0 \end{cases} \tag{2.13}$$

## 2.2 Reference solutions

### 2.2.1 General form of periodic solutions

In this part we are looking for periodic solutions to equation (2.12). Precisely, we are looking for traveling wave solutions, meaning they depend on $x$ and $t$ only in the combination :

$$\psi(x,t) = A(\xi)e^{i\theta(\xi)} \quad ; \quad \xi = x - Ut. \tag{2.14}$$

where $A(\xi) = \sqrt{\rho(x,t)}$ and $\phi(\xi) = \theta(x,t)$. Under these assumptions and notations, equation (2.12) written in the real variables $A$ and $\phi$ yields :

$$-iU(A' + i\phi' A) + \frac{1}{2}(A'' - \phi'^2 + i(2A'\phi' + \phi'' A)) - A^3 = 0 \tag{2.15}$$

Separating real and imaginary parts yields:

$$\begin{cases} A'\phi' + \dfrac{1}{2}\phi'' A = UA' \\ UA\phi' + \dfrac{1}{2}(A'' - \phi'^2) - A^3 = 0 \end{cases} \tag{2.16}$$

The first equation of this system can be solved explicitly as an ODE in terms of the variable $\phi'$ which gives :

$$\phi' = U + \frac{q}{A^2} \tag{2.17}$$

where $q$ is a constant of integration. Plugging this expression into the second equation of system (2.16) and integrating yields :

$$A'^2 + U^2A^2 + \frac{q^2}{A^2} - A^4 = d \tag{2.18}$$

where $d$ is another constant of integration. Switching to the hydrodynamic variables $\rho(x,t) = A^2(\xi)$ and $u(x,t) = \phi'(\xi)$, finally gives the system :

$$\begin{cases} \left(\dfrac{d\rho}{d\xi}\right)^2 = 4(\rho^3 - U^2\rho^2 + d\rho - q^2) = P(\rho) \\ u = U + \dfrac{q}{\rho} \end{cases} \tag{2.19}$$

Since we are looking for real periodic solutions, this requires $P(\rho)$ to have two positive roots so that $\rho$ oscillates periodically between them. Thus, we can write :

$$\left(\frac{d\rho}{d\xi}\right)^2 = 4(\rho - b_1)(\rho - b_2)(\rho - b_3). \tag{2.20}$$

Identifying the expressions in (2.19) and (2.20) allows to obtain the relations :

$$q^2 = b_1 b_2 b_3 \quad ; \quad U^2 = b_1 + b_2 + b_3 \tag{2.21}$$

Without loss of generality, we can assume that $b_1 > b_2 > b_3 > 0$. Finally, one can obtain a solution of (2.20) in terms of the Jacobi elliptic function dn (See Appendix B.3 for the details) :

$$\rho(x,t) = b_1 - (b_1 - b_3)\mathrm{dn}^2\left(\sqrt{b_1 - b_3}\,(x - Ut), s\right), \quad u(x,t) = U + \frac{q}{\rho(x,t)}. \tag{2.22}$$

We remind that $s$ is the elliptic modulus satisfying the relation :

$$s^2 = \frac{b_2 - b_3}{b_1 - b_3}, \quad 0 < s < 1. \tag{2.23}$$

Also note that for each fixed value of $0 < s < 1$, solution (2.2.1) is a periodic wave of amplitude $a$ and wavenumber $k$ given by :

$$a = \frac{b_2 - b_3}{2}, \quad k = \frac{\pi}{K(s)}\sqrt{\frac{2a}{s^2}}. \tag{2.24}$$

where $K(s)$ is the complete elliptic integral of the first kind . This shows that the amplitude and the wavelength of the solution are tightly linked. For the limiting values $s \to 0$ and $s \to 1$, it follows from the formulas (2.24) that the solution is no longer periodic. Periodicity is lost differently in each of the limits. In the case $s \to 0$, the amplitude of the oscillations vanishes. The wave number remains at a finite value given by $k \to 2\sqrt{b_1 - b_3}$. In the case $s \to 1$, we get $K(s) \to \infty$ and therefore the solution is no more oscillatory as the wavenumber $k \to 0$. More precisely, Since $\mathrm{dn}(v,s)|_{s\to 1} = 1/\cosh(v)$ the solution behaves like a soliton of amplitude $(b_1 - b_3)$. Below, we will present some exact and asymptotic solutions to the NLS equation based on the representation (2.2.1).

### 2.2.2 Gray solitons

The first case corresponds to the family of solutions obtained from (2.2.1) in the limit $s \to 1$ :

$$\rho(x,t) = b_1 - \frac{b_1 - b_3}{\cosh^2\left(\sqrt{b_1 - b_3}\,(x - Ut)\right)} \qquad u(x,t) = U - \frac{b_1\sqrt{b_3}}{\rho(x,t)} \tag{2.25}$$

This corresponds to a two-parameter family of solitary waves called gray solitons. It consists of a localized density dip, of amplitude $(b_1 - b_3)$ which propagates at a constant velocity $U$ without deforming or collapsing. It maintains its shape due to a perfect balance between the nonlinearity and dispersion. The parameters $b_1$ and $b_3$ define the limit values of the soliton such that :

$$\lim_{|x|\to\infty} \rho(x,t) = b_1 \quad ; \quad \rho_{min} = b_3 \quad ; \lim_{|x|\to\infty} u(x,t) = U - \sqrt{b_3} \quad ; \quad u_{min} = U - \frac{b_1}{\sqrt{b_3}}. \tag{2.26}$$

The shape of the soliton is shown on Figure 2.1.

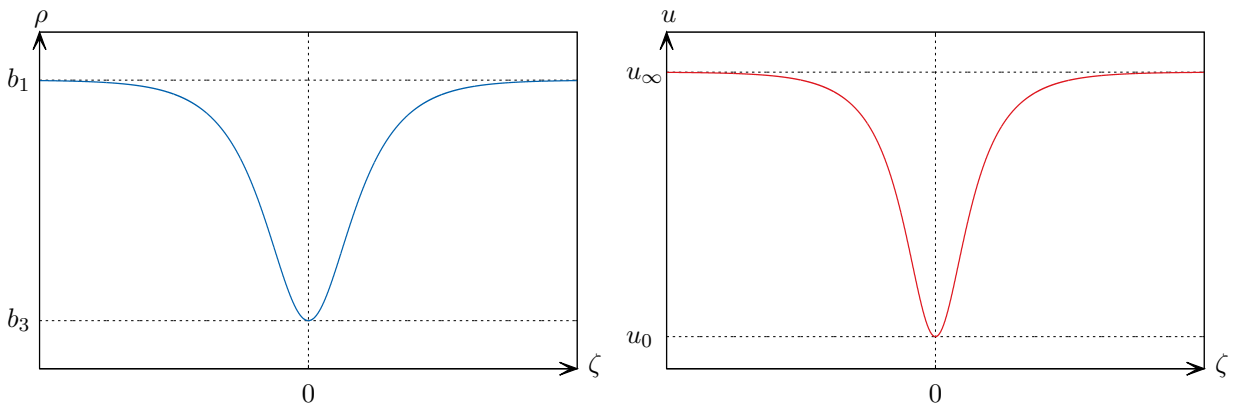


Figure 2.1: Overall shape of the gray soliton solution in terms of the variables $\rho$ and $u$ for arbitrary values of the parameters $b_1$ and $b_3$ at $t = 0$

**Remark 2.2.1.** *The particular case where $b_3 = 0$ is called a dark soliton. Notice that the soliton peak in this case reaches a point where $\rho = 0$ to which corresponds a singularity in the velocity.*

### 2.2.3 Dispersive shockwaves

In classical hydrodynamics described by dispersionless hyperbolic Euler equations, shock waves (strong discontinuities) can appear. In the case of dispersive hydrodynamics, these singularities are resolved by the appearance of an oscillatory wave train in a region of space which expands over time. This is referred to as a dispersive shock wave (or alternatively dissipationless shockwave, collisionless shockwaves). We will denote it by DSW for shortness. Many works in the literature have been devoted to the study and analysis of such solutions, their behaviors and their structures. See for example [43, 42, 32, 30, 31]. Dispersive shockwaves generally display a highly nonlinear behavior, and involve a slowly modulated train of fast oscillations. Whitham's modulation equations [75] are in this case a very powerful tool that provides a precise description of the slow modulation of the wavetrain of oscillations, through averaging techniques. This allows for a proper construction of the asymptotic behavior of the DSW. In particular, it permits us to obtain an asymptotic envelope of the oscillations which will serve later as a reference for the numerical results. In this context, we will focus on DSWs that appear in the case of a Riemann problem for the defocusing NLSE. Thus, consider the initial condition :

$$\begin{cases} \rho(x) = \rho_L & x < 0 \\ \rho(x) = \rho_R & x > 0 \end{cases} \qquad \begin{cases} u(x) = u_L & x < 0 \\ u(x) = u_R & x > 0 \end{cases} \tag{2.27}$$

Such an initial discontinuity gives rise to either a DSW or a rarefaction wave on each side. The evolution of the amplitude of the DSW is described in terms of the Riemann invariants $r_i$ ($r_1 > r_2 > r_3 > r_4$) of the corresponding Whitham's averaged equations [63], [41] :

$$\frac{\partial r_i}{\partial t} + V_i \frac{\partial r_i}{\partial x} = 0, \quad i = 1, 2, 3, 4. \tag{2.28}$$

The characteristic velocities are given by :

$$V_1 = U(r) + \frac{1}{2}(r_1 - r_2)\left(1 - \frac{r_2 - r_4}{r_1 - r_4}\frac{E(s)}{K(s)}\right)^{-1},$$
$$V_2 = U(r) - \frac{1}{2}(r_1 - r_2)\left(1 - \frac{r_1 - r_3}{r_2 - r_3}\frac{E(s)}{K(s)}\right)^{-1},$$
$$V_3 = U(r) + \frac{1}{2}(r_3 - r_4)\left(1 - \frac{r_2 - r_4}{r_2 - r_3}\frac{E(s)}{K(s)}\right)^{-1},$$
$$V_4 = U(r) - \frac{1}{2}(r_3 - r_4)\left(1 - \frac{r_1 - r_3}{r_1 - r_4}\frac{E(s)}{K(s)}\right)^{-1},$$

where the complete elliptic integral of the second kind $E(s)$ is defined as:

$$E(s) = \int_0^{\frac{\pi}{2}} \sqrt{1 - s^2 \sin^2(\theta)}d\theta.$$

The variables $b_i$ and $U$ are linked to the $r_i$ via the relations [63], [41] :

$$b_1 = \frac{1}{16}\left(r_1 + r_2 - r_3 - r_4\right)^2, \tag{2.29}$$
$$b_2 = \frac{1}{16}\left(r_1 + r_3 - r_2 - r_4\right)^2, \tag{2.30}$$
$$b_3 = \frac{1}{16}\left(r_1 + r_4 - r_2 - r_3\right)^2, \tag{2.31}$$
$$U = \frac{1}{4}\left(r_1 + r_2 + r_3 + r_4\right), \tag{2.32}$$
$$s^2 = \frac{(r_1 - r_2)(r_3 - r_4)}{(r_1 - r_3)(r_2 - r_4)}. \tag{2.33}$$

If we consider a self-similar evolution of the Riemann invariants, $r_i$ depend only on $\tau = x/t$ and equations (2.28) reduce to :

$$r_i'(V_i - \tau) = 0, \quad i = 1, 2, 3, 4. \tag{2.34}$$

It means that one of the Riemann invariants, say $r_j$, changes in space and time and its characteristic speed is $V_j = \tau$, while the three other invariants are constants determined by the initial conditions on both sides of the initial discontinuity. Outside of the DSW region, it was shown in [41] that the Whitham equations for the defocusing cubic NLSE degenerate into the Euler equations for shallow water flows. This transition occurs when two of the Riemann invariants $r_i$ merge together, leading to either $s^2 = 0$ or $s^2 = 1$. The two remaining invariants behave like Riemann invariants for the shallow water equations :

$$r_\pm = u \pm 2\sqrt{\rho} \tag{2.35}$$

The solution is then determined via matching of the Riemann invariants at the DSW fronts [43], [41]. Since this procedure depends strongly on the structure of the flow, we will consider

the case in which a rarefaction wave on the left and a DSW on the right are created (see Figure 2.2). At the oscillatory front (corresponding to the leading edge of the DSW, $s^2 = 0$) we get [32], [31]:

$$r_1 = r_2, \quad V_1 = V_2, \quad r_3 = r_+(R), \quad r_4 = r_-(R). \tag{2.36}$$

At the soliton front (corresponding to the trailing edge of the DSW, $s^2 = 1$), we get :

$$r_2 = r_3, \quad V_2 = V_3, \quad r_1 = r_+(0), \quad r_4 = r_-(0). \tag{2.37}$$

Here, $r_\pm(0)$ and $r_\pm(R)$ are the values of the invariants $r_\pm$ at the states '0' (the constant state found solving the Riemann problem for the non-dispersive shallow water equations) and 'R' (the state on the right of the initial discontinuity). Matching these invariants on either side of the constant state allows to obtain the theoretical values of $\rho_0$ and $u_0$ at the central plateau. Indeed we write :

$$r_+(0) = r_1 \;\Rightarrow\; u_0 + 2\sqrt{\rho_0} = u_L + 2\sqrt{\rho_L} \tag{2.38}$$

$$r_-(0) = r_4 \;\Rightarrow\; u_0 - 2\sqrt{\rho_0} = u_R - 2\sqrt{\rho_R} \tag{2.39}$$

Combining both equations permits us to obtain :

$$\rho_0 = \left(\frac{1}{4}\left(u_L - u_R + 2\sqrt{\rho_L} + 2\sqrt{\rho_R}\right)\right)^2 \tag{2.40}$$

$$u_0 = \frac{1}{2}\left(u_L + u_R + 2\sqrt{\rho_L} - 2\sqrt{\rho_R}\right) \tag{2.41}$$

Particularly in the case where $u_R = u_L = 0$ we get :

$$\rho_0 = \left(\frac{1}{2}\left(\sqrt{\rho_L} + \sqrt{\rho_R}\right)\right)^2 \tag{2.42}$$

$$u_0 = \sqrt{\rho_L} - \sqrt{\rho_R} \tag{2.43}$$

These expressions are in agreement with the ones in [46][32]. Next, we shall denote by $\tau_1$ and $\tau_2$ the asymptotic boundaries of the DSW region, and by $\tau_3$ and $\tau_4$ the boundaries of the rarefaction wave when $t \to +\infty$. Their values are given by : (See Appendix B.4 for details)

$$\tau_1 = u_R + \frac{8\rho_0 - 8\sqrt{\rho_0\rho_R} + \rho_R}{2\sqrt{\rho_0} - \sqrt{\rho_R}}, \quad \tau_2 = u_R + \sqrt{\rho_0}, \quad \tau_3 = u_0 - \sqrt{\rho_0}, \quad \tau_4 = u_L - \sqrt{\rho_L}. \tag{2.44}$$

The asymptotic profile of the solution is shown in Figure 2.2. The oscillatory part of the solution is plotted according to the following algorithm :

1. Set the values of $\rho_L,\ \rho_R,\ u_L,\ u_R$.
2. Calculate the values of $\rho_0\ u_0,\ r_1,\ r_3,\ r_4$.
3. Calculate $r_2(s) = (r_1(r_3 - r_4) + s^2(r_1 - r_3)r_4)/(r_3 - r_4 + (r_1 - r_3)s^2)$.
4. Calculate the functions $b_1(s)$, $b_2(s)$, $b_3(s)$, $U(s)$.

5. Calculate $\tau(s) = V_2(s)$.

6. Choose a time instant $t$. The DSW is shown as a parametric plot of (2.2.1) : $\rho(s,t) = \rho(\tau(s),t)$, $0 \leq s \leq 1$.

7. The low and upper boundaries of the oscillatory profile are described by :

$$\rho_{inf}(s) = b_3(s), \quad \rho_{sup}(s) = s^2 b_1(s) + (1-s^2) b_3(s).$$

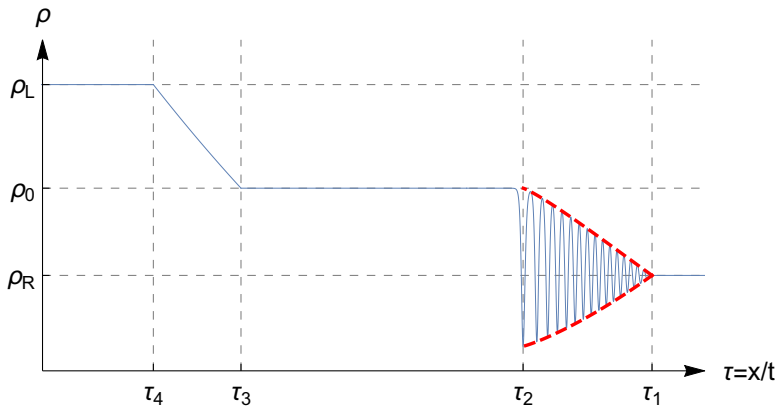


Figure 2.2: Asymptotic profile of the solution to NLS equation (continuous line) for the Riemann problem $\rho_L = 2$, $\rho_R = 1$ , $u_L = u_R = 0$. The boundaries $\tau_i$, $i = 1,2,3,4$ delimit the DSW and the rarefaction wave regions. The modulation of the DSW profile between $\tau_2$ and $\tau_1$ is described by the rarefaction wave solution to the Whitham system (bold dashed line). The oscillatory profile is shown at $t = 70$. The values of $\tau_i$, $i = 1,2,3,4$, are given by $\tau_1 = u_R + \dfrac{8\rho_0 - 8\sqrt{\rho_0\rho_R} + \rho_R}{2\sqrt{\rho_0} - \sqrt{\rho_R}}$, $\tau_2 = u_R + \sqrt{\rho_0}$, $\tau_3 = u_0 - \sqrt{\rho_0}$, $\tau_4 = u_L - \sqrt{\rho_L}$.

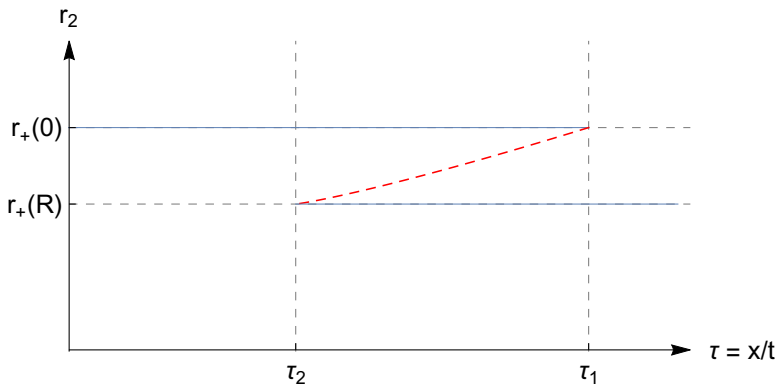


Figure 2.3: Asymptotic profile of invariant $r_2$ for Whitham's system. In the DSW region $\tau_2 < \tau < \tau_1$, $r_2$ (dashed line) varies while the other invariants $r_i$, $i \neq 2$, are constants.

## 2.3 Augmented Lagrangian formulation for NLSE

### 2.3.1 System of Equations

In the same notations as in section 1.1, the augmented Lagrangian for the defocusing NLSE writes :

$$\mathcal{L} = \int_{\Omega_t} \left( \rho \frac{|\mathbf{u}|^2}{2} + \frac{\beta}{2}\rho\dot{\eta}^2 - \frac{\rho^2}{2} - \frac{1}{4\rho}\frac{|\nabla\eta|^2}{2} - \frac{1}{2\alpha}\rho\left(\frac{\eta}{\rho} - 1\right)^2 \right) d\Omega \tag{2.45}$$

and the corresponding system of equations is given by :

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho\mathbf{u}) = 0, \tag{2.46}$$

$$\frac{\partial \rho\mathbf{u}}{\partial t} + \operatorname{div}\left( \rho\mathbf{u}\otimes\mathbf{u} + \left(\frac{\rho^2}{2} - \frac{1}{4\rho}|\mathbf{p}|^2 + \frac{\eta}{\alpha}\left(1 - \frac{\eta}{\rho}\right)\right)\mathbf{Id} + \frac{1}{4\rho}\mathbf{p}\otimes\mathbf{p} \right) = 0, \tag{2.47}$$

$$\frac{\partial \rho\eta}{\partial t} + \operatorname{div}(\rho\eta\mathbf{u}) = \rho w, \tag{2.48}$$

$$\frac{\partial \rho w}{\partial t} + \operatorname{div}\left( \rho w\mathbf{u} - \frac{1}{4\rho\beta}\mathbf{p} \right) = \frac{1}{\alpha\beta}\left(1 - \frac{\eta}{\rho}\right), \tag{2.49}$$

$$\frac{\partial \mathbf{p}}{\partial t} + \operatorname{div}\left( (\mathbf{p}\cdot\mathbf{u} - w)\,\mathbf{Id} \right) = 0; \quad \operatorname{curl}(\mathbf{p}) = 0. \tag{2.50}$$

This system admits the energy conservation law :

$$\frac{\partial E}{\partial t} + \operatorname{div}\left( E\mathbf{u} + \mathcal{P}\mathbf{u} - \frac{1}{4\rho}w\mathbf{p} \right) = 0,$$

where

$$E = \rho\frac{|\mathbf{u}|^2}{2} + \frac{\beta}{2}\rho\dot{\eta}^2 + \frac{\rho^2}{2} + \frac{1}{4\rho}\frac{|\mathbf{p}|^2}{2} + \frac{\rho}{2\alpha}\left(\frac{\eta}{\rho} - 1\right)^2,$$

$$\mathcal{P} = \left(\frac{\rho^2}{2} - \frac{1}{4\rho}|\mathbf{p}|^2 + \frac{\eta}{\alpha}\left(1 - \frac{\eta}{\rho}\right)\right)\mathbf{Id} + \frac{1}{4\rho}\mathbf{p}\otimes\mathbf{p}.$$

We will concentrate on the one-dimensional case where $\mathbf{u} = (u, 0, 0)^T$ and $\mathbf{p} = (p, 0, 0)^T$. In this case, the equations are hyperbolic and the corresponding eigensystem is given by :

$$\begin{array}{lll}
\xi_1 = u & , & \mathbf{v_1} = (\frac{\rho}{\alpha\rho^3+\eta^2}, 0, 0, \frac{p}{\alpha\rho^3+\eta^2}, \frac{1}{2\eta-\rho})^T \\
\xi_2 = u + \frac{1}{2\rho\sqrt{\beta}} & , & \mathbf{v_2} = (0, 0, \sqrt{\beta}, 2, 0)^T \\
\xi_3 = u - \frac{1}{2\rho\sqrt{\beta}} & , & \mathbf{v_3} = (0, 0, -\sqrt{\beta}, 2, 0)^T \\
\xi_4 = u + \sqrt{\rho + \frac{\eta^2}{\alpha\rho^2}} & , & \mathbf{v_4} = (\rho, \sqrt{\rho + \frac{\eta^2}{\alpha\rho^2}}, 0, p, 0)^T \\
\xi_5 = u - \sqrt{\rho + \frac{\eta^2}{\alpha\rho^2}} & , & \mathbf{v_5} = (\rho, -\sqrt{\rho + \frac{\eta^2}{\alpha\rho^2}}, 0, p, 0)^T
\end{array} \tag{2.51}$$

### 2.3.2 Dispersion Relation

Linearizing the governing equations on the constant solution $\rho = \rho_0, u = 0, w = 0, p = 0, \eta = \rho_0$ and looking for the solutions which are proportional to $e^{i(kx-\omega t)}$, where $k$ is the wave number and $\omega$ is the frequency, one can obtain the dispersion relation expressed here in the form $c_p = c_p(k)$, where $c_p = \omega/k$ is the phase velocity :

$$(c_p^{\pm})^2 = \frac{\frac{1}{4\beta\rho_0^2} + \rho_0 + \frac{1}{\alpha} + \frac{1}{\alpha\beta\rho_0^2 k^2} \pm \sqrt{\left(\frac{1}{4\beta\rho_0^2} + \rho_0 + \frac{1}{\alpha} + \frac{1}{\alpha\beta\rho_0^2 k^2}\right)^2 - 4\left(\frac{1}{\alpha\beta\rho_0 k^2} + \frac{\rho_0 + \frac{1}{\alpha}}{4\beta\rho_0^2}\right)}}{2}. \tag{2.52}$$

**Remark 2.3.1.** *One can easily see that in the linear case, at fixed positive values of $\alpha$ and $\beta$, the phase velocity approaches to the characteristic velocity in the limit $k \to \infty$.*

### 2.3.3 Estimation of $\beta$ and $\alpha$

As discussed already, the augmented Lagrangian approach requires a choice of $\beta$ and $\alpha$. This choice can be based, for example, on the fact that the dispersion relations for both augmented and original systems must remain close in a specific range of wave numbers, that is wave numbers of interest. For the equilibrium state defined by $\rho = \rho_0, u = u_0 = 0$, the dispersion relation for the original Euler-Korteweg NLS system (2.13) is :

$$c_p^2 = \rho_0 + k^2/4. \tag{2.53}$$

When $(\beta, \alpha) \to (0, 0)$, the convergence of these dispersion relations is not uniform, that is, the curves almost coincide for low wave numbers but start to stray away from each other beginning from a certain threshold wave number $k_{max}(\beta, \alpha)$. This threshold wave number must be chosen such that the wave numbers that may be present in the solution are contained in $[0, k_{max}]$.

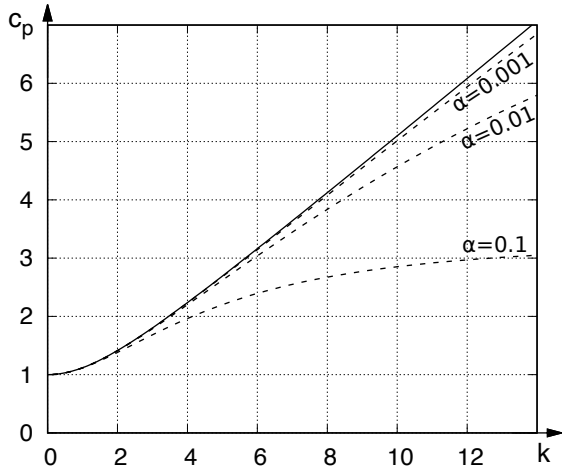


Figure 2.4: The dispersion relation (2.53) (continuous line) and (2.52) for the augmented Lagrangian (dashed lines) for $\beta = 10^{-4}$ and different values of $\alpha$.

## 2.4 Numerical Schemes

In this part, we explain all the necessary steps related to the numerical resolution of the augmented NLS system in one dimension of space. Since the latter is a first order hyperbolic system of equations, any appropriate finite volume scheme should be able to deal with the numerical resolution. In our case we will use two types of schemes, namely a MUSCL-Hancock [70] scheme with a splitting strategy for the source terms, and an Implicit-Explicit type scheme ((2,2,2) Diagonally-Implicit Runge-Kutta scheme) [2]. We will present here details about the algorithms for both schemes are found. Boundary conditions will be discussed later for each test case separately. In a most generic case, the system to solve numerically can be written as :

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{U})}{\partial x} = \mathbf{S}(\mathbf{U}), \tag{2.54}$$

where $\mathbf{U}$, $\mathbf{F}(\mathbf{U})$ and $\mathbf{S}(\mathbf{U})$ are respectively the vector of conservative variables, flux vector and source terms vector. In the case of augmented NLS system, they are given by :

$$\mathbf{U} = \begin{pmatrix} \rho \\ \rho u \\ \rho \eta \\ \rho w \\ p \end{pmatrix}, \quad \mathbf{F}(\mathbf{U}) = \begin{pmatrix} \rho u \\ \rho u^2 + \frac{\rho^2}{2} + \frac{\eta}{\alpha}\left(1 - \frac{\eta}{\rho}\right) \\ \rho \eta u \\ \rho w u - \frac{1}{4\rho\beta} p \\ pu - w \end{pmatrix}, \quad \mathbf{S}(\mathbf{U}) = \begin{pmatrix} 0 \\ 0 \\ \rho w \\ \frac{1}{\alpha\beta}\left(1 - \frac{\eta}{\rho}\right) \\ 0 \end{pmatrix}.$$

### 2.4.1 MUSCL-Hancock method

In a first attempt, we solve the hyperbolic system (2.54) by using the MUSCL-Hancock extension to the Godunov scheme [70]. Since we have a non-zero source term, a splitting strategy is applied [70] in which, at each time step, the numerical resolution is split into a hyperbolic part :

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{U})}{\partial x} = 0, \tag{2.55}$$

and an ordinary differential equation part for the source terms :

$$\frac{\mathrm{d}\mathbf{U}}{\mathrm{dt}} = \mathbf{S}(\mathbf{U}). \tag{2.56}$$

so that at each time step $\Delta t = t^{n+1} - t^n$, the numerical resolution is split as follows [70]:

$$\text{Hyperbolic step} \left\{ \begin{array}{l} \dfrac{\partial \mathbf{U}}{\partial t} + \dfrac{\partial \mathbf{F}}{\partial x} = 0 \\ \text{IC: } \mathbf{U}(x, t^n) = \mathbf{U}^n \end{array} \right\} \Longrightarrow \bar{\mathbf{U}}^{\mathbf{n+1}} \quad \text{ODE step} \left\{ \begin{array}{l} \dfrac{\mathrm{d}\mathbf{U}}{\mathrm{d}t} = \mathbf{S}(\mathbf{U}) \\ \text{IC: } \mathbf{U}(x, t^n) = \bar{\mathbf{U}}^{\mathbf{n+1}} \end{array} \right\} \Longrightarrow \mathbf{U}^{\mathbf{n+1}}$$

Thus the algorithm can be summarized as follows :

1. **Boundary conditions :** Setting the appropriate boundary conditions in the ghost cells $U_0$ and $U_{N+1}$.

2. **Data Reconstruction:** Replaces the constant states $\mathbf{U}_i^n$ by piecewise linear functions $u_i^n(x)$ defined by the extreme points :

$$\mathbf{U}_i^L = u_i(0) = u_i^n - \frac{1}{2}\Delta_i \quad ; \quad \mathbf{U}_i^R = u_i(\Delta x) = u_i^n + \frac{1}{2}\Delta_i \tag{2.57}$$

Where $\Delta_i$ is the slope.

3. **Prediction Step:** Time transition of $\mathbf{U}_i^L$ and $\mathbf{U}_i^R$ by half a time-step $\frac{\Delta t}{2}$ :

$$\begin{cases} \bar{\mathbf{U}}_i^L = \mathbf{U}_i^L + \dfrac{\Delta t}{2\Delta x}(\mathbf{F}(\mathbf{U}_i^L) - \mathbf{F}(\mathbf{U}_i^R)) \\ \bar{\mathbf{U}}_i^R = \mathbf{U}_i^R + \dfrac{\Delta t}{2\Delta x}(\mathbf{F}(\mathbf{U}_i^L) - \mathbf{F}(\mathbf{U}_i^R)) \end{cases}$$

4. **Updating of solution to hyperbolic part:** It is done according to the conservative formula :

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \frac{\Delta t}{\Delta x}\left(\mathbf{F}^*_{i+\frac{1}{2}} - \mathbf{F}^*_{i-\frac{1}{2}}\right), \quad 1 \leq i \leq N. \tag{2.58}$$

The intercell fluxes are obtained by solving the Riemann problem with $\mathbf{U_L} = \bar{\mathbf{U}}_i^R$ and $\mathbf{U_R} = \bar{\mathbf{U}}_i^L$. We use in our case a Rusanov flux given by :

$$\mathbf{F}^*_{i+\frac{1}{2}} = \frac{1}{2}\left(\mathbf{F}(\mathbf{U}_L) + \mathbf{F}(\mathbf{U}_R)\right) - \frac{1}{2}\kappa^n_{i+\frac{1}{2}}\left(\mathbf{U}_R - \mathbf{U}_L\right), \tag{2.59}$$

where $\kappa^n_{i+\frac{1}{2}}$ is obtained by using the Davis approximation [24] :

$$\kappa^n_{i+1/2} = \max_j(|c_j(\mathbf{U}_i^n)|, |c_j(\mathbf{U}_{i+1}^n)|), \tag{2.60}$$

with $c_j$, the eigenvalues of the Jacobian $\dfrac{\partial \mathbf{F}}{\partial \mathbf{U}}$.

5. **Solving the ODE part:** The final stage which corresponds to the taking into account of the source term by solving the differential equation (2.56). In our case, this ordinary differential equation is integrable. Indeed, it writes :

$$\frac{d\rho}{dt} = 0, \quad \frac{d(\rho u)}{dt} = 0, \quad \frac{d(\rho\eta)}{dt} = \rho w, \quad \frac{d(\rho w)}{dt} = \frac{1}{\alpha\beta}\left(1 - \frac{\eta}{\rho}\right), \quad \frac{dp}{dt} = 0 \tag{2.61}$$

This can be cast into :

$$\begin{cases} \dfrac{d\rho}{dt} = 0, \quad \dfrac{d\rho u}{dt} = 0, \quad \dfrac{dp}{dt} = 0 \\ \dfrac{d^2\eta}{dt^2} + \dfrac{1}{\alpha\beta\rho}\eta - \dfrac{1}{\alpha\beta} = 0, \\ w = \dfrac{d\eta}{dt}. \end{cases}$$

Finally, solving the Cauchy problem associated to this step gives the solution :

$$\begin{cases} \rho^{n+1} = \bar{\rho}^n \qquad u^{n+1} = \bar{u}^n \qquad p^{n+1} = \bar{p}^n \\ \eta^{n+1} = \bar{\rho}^n + (\bar{\eta}^n - \bar{\rho}^n)\ \cos(\dfrac{1}{\alpha\beta\rho^2}\Delta t) + \dfrac{\bar{w}^n}{\alpha\beta\rho^2}\ \sin(\dfrac{1}{\alpha\beta\rho^2}\Delta t) \\ w^{n+1} = \dfrac{1}{\alpha\beta\rho^2}(\bar{\rho}^n - \bar{\eta}^n)\ \sin(\dfrac{1}{\alpha\beta\rho^2}\Delta t) + \bar{w}^n \cos\ (\dfrac{1}{\alpha\beta\rho^2}\Delta t) \end{cases}$$

The algorithm loops over the previous steps until an initially defined final time is reached. Since this is a purely first-order hyperbolic system of equations, stability of the scheme is ensured by a usual Courant–Friedrichs–Lewy constraint on the time step $\Delta t$ :

$$\Delta t \leq C_{cfl}\ \frac{\Delta x}{\max\limits_{1\leq i\leq N}(|c(\mathbf{U_i^n})|)} \tag{2.62}$$

The characteristic speeds in the x-direction are given by :

$$c = u,\ (c-u)^2_{\pm} = \frac{\frac{1}{4\beta\rho^2} + \rho + \frac{\eta^2}{\alpha\rho^2} \pm |-\frac{1}{4\beta\rho^2} + \rho + \frac{1\eta^2}{\alpha\rho^2}|}{2}. \tag{2.63}$$

This means that depending on the sign of the quantity $-\frac{1}{4\beta\rho^2} + \rho + \frac{\eta^2}{\alpha\rho^2}$, there are 5 possible values for the characteristic speed given by :

$$c_1 = u,\ c_{2,3} = u \pm \sqrt{\rho + \frac{\eta^2}{\alpha\rho^2}},\ c_{4,5} = u \pm \sqrt{\frac{1}{4\beta\rho^2}}$$

Thus, the characteristic speed is bounded by :

$$\max_{1\leq i\leq N}(|c(\mathbf{U_i})|) = \max_{1\leq i\leq N}\left(\max\left(|u_i|\,, |u_i| + \sqrt{\rho_i + \frac{\eta_i^2}{\alpha\rho_i^2}}, |u_i| + \sqrt{\frac{1}{4\beta\rho_i^2}}\right)\right)$$

In practice, $\beta$ and $\alpha$ are small parameters. Therefore, although the time step scales linearly with the mesh size $\Delta x$, the coefficient which multiplies $\Delta x$ can be very small for small values of $\alpha$ and $\beta$. This remains advantageous for refined meshes.

### 2.4.2 IMEX-(2,2,2) scheme

The name IMEX stands for Explicit-Implicit Runge-Kutta scheme [2][62]. This is a family of multistep schemes which are well adapted to differential equations and partial differential equations with stiff terms. This is particularly the case for hyperbolic equations with relaxation, as the latter often involves small time scales that stiffen source terms. The main concept behind an IMEX Runge-Kutta scheme is to apply an implicit discretization to the stiff part of the system and an explicit one to the remaining part. There are many variants of IMEX schemes, and to help classify them, there is generally an associated triplet of integers $(s, \sigma, p)$ where $s$ characterizes the number of stages of the implicit part, $\sigma$ the number of

stages of the explicit part and $p$ the order of the scheme. Under this notation, the scheme we will use is an IMEX-(2,2,2) scheme. It can be represented by two tables in the usual Butcher notations for RK-schemes :

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \gamma & \gamma & 0 & 0 \\ 1 & \gamma-1 & 2-\gamma & 0 \\ \hline & \gamma-1 & 2-\gamma & 0 \end{array} \qquad \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \gamma & 0 & \gamma & 0 \\ 1 & 0 & 1-\gamma & \gamma \\ \hline & 0 & 1-\gamma & \gamma \end{array} \qquad \gamma = 1 - \frac{1}{\sqrt{2}} \tag{2.64}$$

where the left table corresponds to the explicit part and the right one to the implicit part. When applied to the conservative equation (2.54), this yields the two-stage scheme :

$$\begin{aligned}
&\mathbf{U}^{n,0} = \mathbf{U}^n \\
&\mathbf{U}^{n,1} = \mathbf{U}^\star = \mathbf{U}^n - \gamma\frac{\Delta t}{\Delta x}\left(F^n_{i+\frac{1}{2}} - F^n_{i-\frac{1}{2}}\right) + \gamma\Delta t\mathbf{S}(\mathbf{U}^\star) \\
&\mathbf{U}^{n,2} = \mathbf{U}^{n+1} = \mathbf{U}^n - (\gamma-1)\frac{\Delta t}{\Delta x}\left(F^n_{i+\frac{1}{2}} - F^n_{i-\frac{1}{2}}\right) - (2-\gamma)\frac{\Delta t}{\Delta x}\left(F^\star_{i+\frac{1}{2}} - F^\star_{i-\frac{1}{2}}\right) + (1-\gamma)\Delta t S(\mathbf{U}^\star) + \gamma\Delta t S(\mathbf{U}^{n+1})
\end{aligned}$$

In these formulas, $F^n_{i+\frac{1}{2}}$ are the usual intercell fluxes computed by using a Riemann solver as is the case for a classical Godunov scheme. Note that, for the second step of the scheme, it is required to also compute intercell fluxes $F^\star_{i+\frac{1}{2}}$ for the intermediate state $\mathbf{U}^\star$ (which also requires to apply another MUSCL reconstruction beforehand). The source terms are resolved implicitly, which requires a little bit of additional work when compared to explicit resolution. The summary of the numerical resolution algorithm is given as follows :

1. **Boundary conditions :** Setting the appropriate boundary conditions in the ghost cells $U_0$ and $U_{N+1}$.

2. **MUSCL Reconstruction:** Replaces the constant states $\mathbf{U}^n_i$ by piecewise linear functions $u^n_i(x)$ as explained in the previous section.

3. **Compute the intercell Fluxes $F^n_{i+\frac{1}{2}}$:** This is done using a Rusanov Flux :

$$\mathbf{F}^n_{i+\frac{1}{2}} = \frac{1}{2}\left(\mathbf{F}(\mathbf{U}_L) + \mathbf{F}(\mathbf{U}_R)\right) - \frac{1}{2}\kappa^n_{i+\frac{1}{2}}\left(\mathbf{U}_R - \mathbf{U}_L\right),$$

4. **IMEX intermediate step:** Consists in solving the equation :

$$\mathbf{U}^\star = \underbrace{\mathbf{U}^n - \gamma\frac{\Delta t}{\Delta x}\left(F^n_{i+\frac{1}{2}} - F^n_{i-\frac{1}{2}}\right)}_{=V_1 \ (known)} + \gamma\Delta t\mathbf{S}(\mathbf{U}^\star)$$

   where the only unknown is $\mathbf{U}^\star$. All the variables at the time $t^n$ are known. It can be simply put into the form :

$$\mathbf{U}^\star - \gamma\Delta t\mathbf{S}(\mathbf{U}^\star) - V_1 = 0$$

   This equation can be solved numerically by any suitable root-finding algorithm (e.g Newton-Raphson or a fixed-point method) for an equation $f(\mathbf{U}^\star) = 0$ where $f$ is given by :

$$f(\mathbf{U}^\star) = \mathbf{U}^\star - \gamma\Delta t\mathbf{S}(\mathbf{U}^\star) - V_1$$

5. **MUSCL Reconstruction of the state $\mathbf{U}^\star$:** Applies a MUSCL reconstruction for the piece-wise data $\mathbf{U}^\star$ in order to compute the intercell fluxes $F^\star_{i+\frac{1}{2}}$.

6. **Boundary conditions for the state $\mathbf{U}^\star$:** Applying the boundary conditions for $\mathbf{U}^\star$, to compute the boundary intercell fluxes.

7. **Computation of the intercell Fluxes $F^\star_{i+\frac{1}{2}}$:** The intercell fluxes are obtained by solving again the Riemann problem with $\mathbf{U_L} = \bar{\mathbf{U}}^{\star R}_i$ and $\mathbf{U_R} = \bar{\mathbf{U}}^{\star L}_i$.

8. **IMEX final step:** We now have everything we need to solve the equation :

$$\mathbf{U}^{n+1} = \underbrace{\mathbf{U}^n - (\gamma-1)\frac{\Delta t}{\Delta x}\left(F^n_{i+\frac{1}{2}} - F^n_{i-\frac{1}{2}}\right) - (2-\gamma)\frac{\Delta t}{\Delta x}\left(F^\star_{i+\frac{1}{2}} - F^\star_{i-\frac{1}{2}}\right) + (1-\gamma)\Delta t S(\mathbf{U}^\star)}_{known = V_2}$$
$$+ \gamma \Delta t S(\mathbf{U}^{n+1})$$

   It reduces to exactly the same form of equation to solve as in the intermediate step, that is an equation of the form $g(\mathbf{U}^{n+1}) = 0$ where $g$ is given by :

$$g(\mathbf{U}^{n+1}) = \mathbf{U}^{n+1} - \gamma \Delta t \mathbf{S}(\mathbf{U^{n+1}}) - V_2$$

   This equation is solved identically as above.

Before we move on, a few remarks of practical interest are worth mentioning.

- The time-step is constrained by the usual CFL condition $\Delta t \leq C_{cfl}\ \frac{\Delta x}{\max\limits_{1\leq i\leq N}(\left|c(\mathbf{U^n_i})\right|)}$.
- This IMEX scheme is second order accurate.
- It is possible and more robust to solve the implicit part of the scheme by hand and implement it directly into the code (see Appendix A.6). Although this makes for a less generic code and more manual labor, this also avoids convergence issues of root-finding algorithms and saves computational power.

## 2.5 Numerical Results

### 2.5.1 Gray solitons :

#### Initial condition and boundary conditions

We consider the initial conditions corresponding to the solitary wave solutions of NLS equation :

$$\rho(x,0) = b_1 - \frac{b_1 - b_3}{\cosh^2\left(\sqrt{b_1 - b_3}\ x\right)} \quad u(x,0) = U - \frac{b_1\sqrt{b_3}}{\rho(x,0)}, \tag{2.65}$$
$$\eta(x,0) = \rho(x,0), \quad w(x,0) = -\rho(x,0)u_x(x,0), \quad p(x,0) = \rho_x(x,0).$$

We use periodic boundary conditions in the computational domain. Thus we impose the following relation between numerical fluxes in the first and last numerical cells:

$$\mathbf{F}^*_{\frac{1}{2}} = \mathbf{F}^*_{N+\frac{1}{2}} = \mathbf{F}^* \left( \mathbf{U}^n_N, \mathbf{U}^n_1 \right). \tag{2.66}$$

It means that when the soliton passes through one of the boundaries, it continuously reappears on the other side. This permits us to run simulations for a longer time without having to use large domains. The period of such a configuration is the required time for the soliton to reach back its initial position. Such boundary conditions would normally apply to periodic solutions. The results are shown on Figure 2.5:

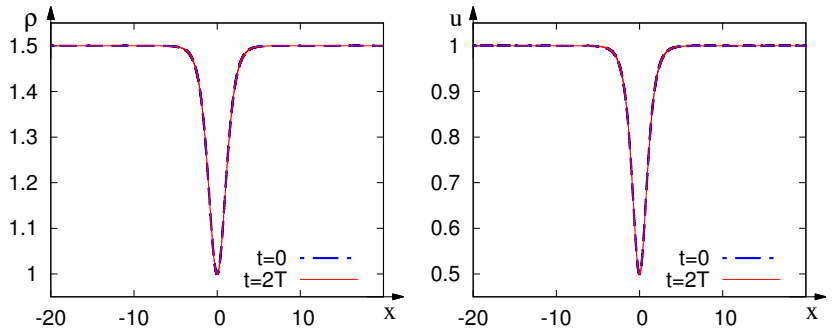


Figure 2.5: Numerical profiles of $\rho$ (left) and $u$ (right) for the grey soliton at $t = 0$ (dot-dashed line) and at $t = 2T$ (continuous line). The used domain is $L = [-20, 20]$ with $\Delta x = 0.0002$, the period is $T = D/U = 20$. Parameters used for the simulation are $b_1 = 1.5,\ b_3 = 1,\ U = 2,\ \epsilon = 1,\ \beta = 10^{-4},\ \alpha = 2.10^{-3}$.

One can see that the shape and the position of the simulated soliton are in perfect agreement with the exact solution. Furthermore, Figure 2.6 shows the solution for different mesh sizes $\Delta x = 0.004$, $\Delta x = 0.002$, $\Delta x = 0.0008$ and $\Delta x = 0.0002$, with a focus on the soliton peak. We can see that for more refined meshes, we get more accurate values for the soliton position and amplitude. This gives a preliminary idea on the convergence of the scheme, which will be studied thoroughly in the following paragraphs.

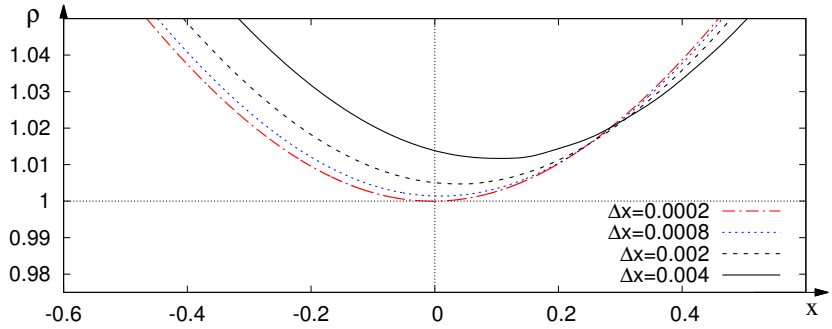


Figure 2.6: Magnified view on the numerical 'grey' soliton peak, for different mesh sizes, at $t = 2T$. The theoretical amplitude and position are $\rho = 1$ and $x = 0$. Parameters used for the simulation are the same as in figure (2.5).

### Convergence study

We measure hereafter the $L_2$ relative error of the soliton's profile across the domain. Since the gray soliton admits non vanishing limits at infinity, we will be considering the quantity $\rho - \rho_\infty$. This quantity is integrable and corresponds to the density of matter being expelled when the soliton propagates. Hence, we consider the following error :

$$\Delta_2(\rho) = \sqrt{\frac{\sum_{i=0}^{N} \left(\rho(i) - \rho_e(i)\right)^2}{\sum_{i=0}^{N} \left(\rho_e(i) - \rho_\infty\right)^2}} \tag{2.67}$$

Where $\rho(i)$ and $\rho_e(i)$ are respectively the simulated value and the exact solution value of $\rho$ taken at the $i^{th}$ cell.

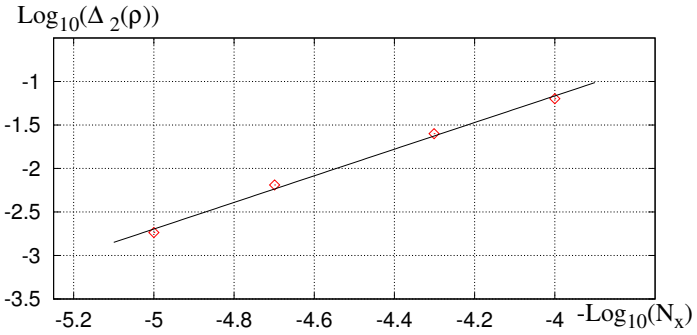


Figure 2.7: The log-log plot of the $L_2$ error. The dots show the measured $L_2$ errors for different mesh sizes $(10000, 20000, 50000, 100000)$ at $t = 40$. The continuous line is a linear interpolation. The measured convergence slope is approximately 1.53.

It is interesting to investigate the phase error as well, that is the difference in the predicted and simulated values of the solitary wave's peak position. This error we measure is given by :

$$\Delta\phi(t) = |x_{cp}(t) - x_{cs}(t)| \tag{2.68}$$

where $x_{cp}(t)$ and $x_{cs}(t)$ are respectively the predicted and simulated position of the soliton's peak at time t. Figure (2.8) shows this error as a function of time for different meshes :

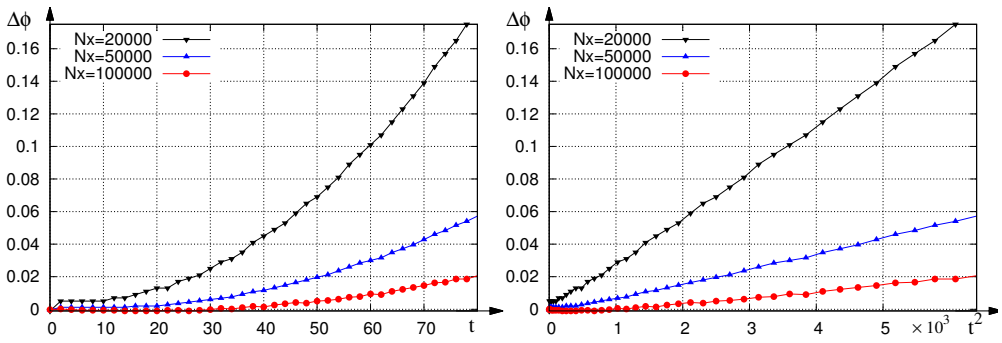


Figure 2.8: The phase error as a function of time (left) and as a function of the time squared (right), for three different mesh sizes : $Nx = 20000, 50000$ and $100000$.

According to these graphs, The phase error seems to be quadratic in time and vanishes with mesh refinement.

### Long time behavior

We carried on calculations until $t = 80$. The same behavior is observed, the soliton still conserves its shape and travels at the same velocity without breaking-down. Figure (2.9) shows the numerical solution for $t = 80$ along previously shown solutions for $t = 0$ and $t = 40$ and Figure (2.10) shows the difference with the initial state for the solutions for $t = 40$ and $t = 80$. The error is growing in time especially at the center of the solitary wave, mainly due to the phase error.

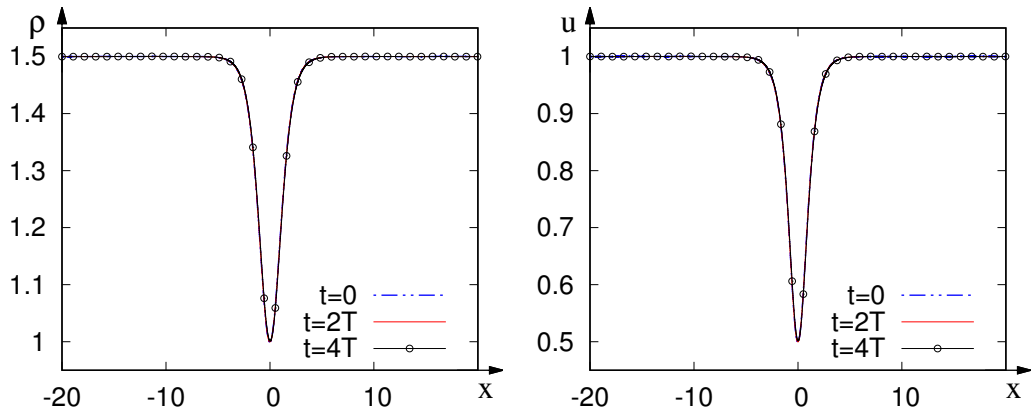


Figure 2.9: Numerical profiles of $\rho$ (left) and $u$ (right) for the gray soliton for $t = 0$, $t = 2T$ and $t = 4T$. The used domain is $L = [-20, 20]$ with $\Delta x = 0.0002$, the period is $T = D/U = 20$. Parameters used for the simulation are $b_1 = 1.5$, $b_3 = 1$, $U = 2$, $\epsilon = 1$, $\beta = 10^{-4}$, $\alpha = 2.10^{-3}$.

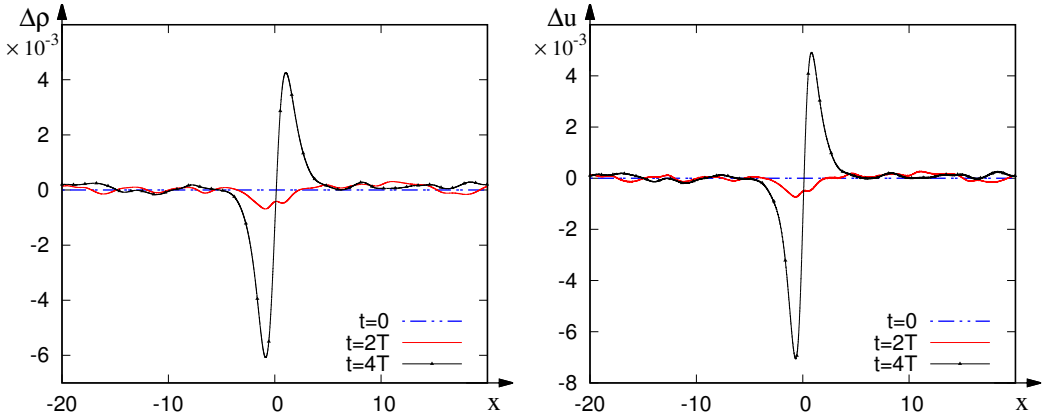


Figure 2.10: Differences $\Delta\rho = \rho(x,t) - \rho(x,0)$ (left) and $\Delta u = u(x,t) - u(x,0)$ (right) for the gray soliton for $t = 0$, $t = 2T$ and $t = 4T$. The parameters are the same ones used above.

### Conserved quantities

The following analysis is carried out in order to see how well the mass, momentum and energy are conserved, we consider a similar normalization to the one introduced in (2.67), meaning we calculate the integral quantities according to :

$$m' = \sum_{i=1}^{N} (\rho_\infty - \rho(i))\,\Delta x;\;\; q' = \sum_{i=1}^{N} (\rho_\infty u_\infty - \rho(i)u(i))\,\Delta x;\;\; E' = \sum_{i=1}^{N} (E_\infty - E(i))\,\Delta x \quad (2.69)$$

where, m', q' and E' correspond respectively to the mass, momentum and total energy, which are expelled by the soliton's propagation. They are estimated with numerical integration over the whole computational domain . We compare these quantities with their initial values, using the following relative errors :

$$\Delta m(t) = \left|\frac{m(t) - m_0}{m_0}\right|; \quad \Delta q(t) = \left|\frac{q(t) - q_0}{q_0}\right|; \quad \Delta E(t) = \frac{E(t) - E_0}{E_0} \quad (2.70)$$

Figure (2.11) shows the evolution of these errors over time.

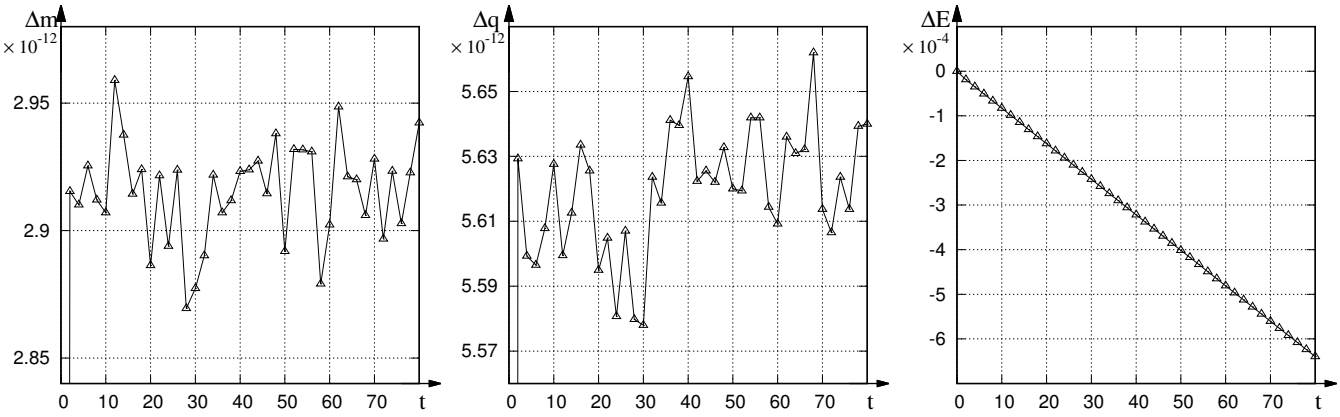


Figure 2.11: Evolution of the errors $\Delta m$, $\Delta q$ and $\Delta E$ over time.

We can see that the mass and the momentum are well conserved. The measured errors remain around the same values even after reaching the long time domain. A different behavior is observed for the energy. Indeed, it steadily decreases from its initial value in a linear manner. This is due to the numerical dissipation of the scheme.

## 2.5.2 Dispersive shockwaves

### MUSCL-Hancock : Smoothed initial step

We present here the results for dispersive Riemann problem. In order to get the case where a DSW propagates to the right and a rarefaction wave to the left, the initial values must satisfy the condition [32] :

$$r_+(L) > r_+(R) > r_-(R) > r_-(L) \quad (2.71)$$

An initial step function trivially induces an infinite gradient at the point of discontinuity. Thus, the gradient of density is proportional to a Dirac Delta function. The latter is not $L^2$ which implies an initially infinite total energy in this case. Besides, since in our model, the gradient is also an independent variable that must be initialized as a Delta function. From the numerical point of view, this is impossible in the sense that machine-wise, real numbers have a finite memory representations and infinite values cannot be stored. Thus, in order to solve the Riemann problem, with properly implemented initial conditions, we will consider a regular step-like function defined by :

$$\rho_0(x) = \frac{\rho_L + \rho_R}{2} + \left(\frac{\rho_L - \rho_R}{2}\right) \tanh\left(\frac{x}{\delta}\right), \tag{2.72}$$

$$u_0(x) = \frac{u_L + u_R}{2} + \left(\frac{u_L - u_R}{2}\right) \tanh\left(\frac{x}{\delta}\right). \tag{2.73}$$

The parameter $\delta$ controls the steepness of the jump, in the sense that the abrupt discontinuity is spread over a region approximately delimited by $[-2\delta, 2\delta]$ as shown in Figure (2.12) :

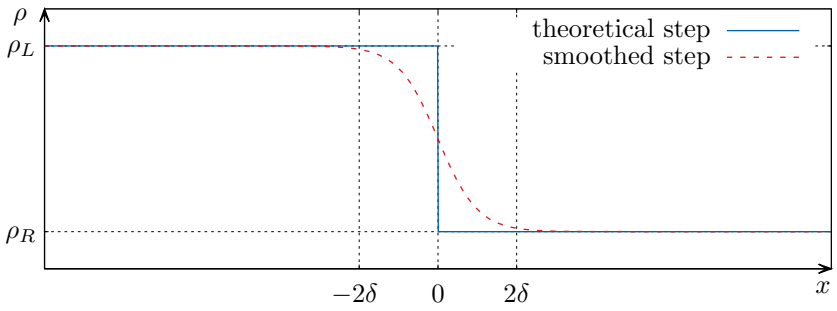


Figure 2.12: Magnified view over the smoothed step (continuous line) for $\delta = 0.1$, $\Delta x = 0.000667$.

Now, given this custom initial condition, we investigate the long time behavior of the solution. We plot the quantities $\rho$ and $u$, obtained numerically, as functions of the self-similar variable $x/t$. The results are shown in figure (2.13). Clearly, the overall structure of the solution complies with the asymptotic one shown on Figure 2.2. The amplitude of the oscillations shows a very good agreement with the asymptotic DSW profile from Whitham's theory of modulations. The oscillations develop in both regions: $\tau > \tau_1$ and $\tau < \tau_1$. These oscillations are part of the solution but they do not appear in the asymptotic (in time) limit because they vanish when $t \to \infty$ as their amplitude $a$ decreases with time as $a \propto t^{-1/2}$ [42]. This behavior is also displayed with the same power law in our results. By measuring the amplitude $a$ of the first oscillation at the vicinity of $\tau_4$, we could plot the function $f(t) = at^{3/2}$ (see Figure (2.14) . The plot shows that $f(t)$ is a linear function of time which implies that $a\sqrt{t}$ is constant.

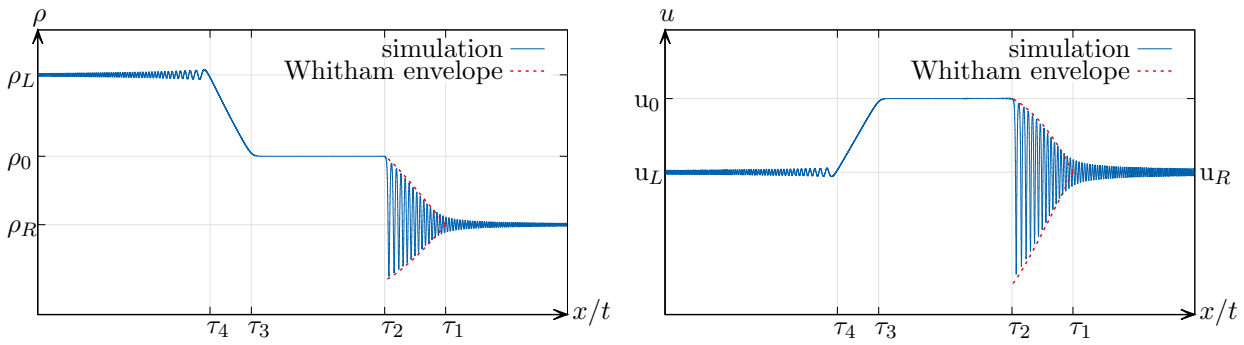


Figure 2.13: Comparison of the numerical results (thin line) with the Whitham modulational profile of the DSW (thick line) at $t = 70$. The left figure shows $\rho = f(x/t)$ and the right figure shows $u = f(x/t)$. The displayed $\tau_i$ are the theoretical boundaries of the DSW and the rarefaction wave. The values of $\rho_0$ and $u_0$ are the theoretical values in the central plateau given by $\rho_0 = \frac{1}{4}(\sqrt{\rho_R} + \sqrt{\rho_L} + \frac{1}{2}(u_L - u_R))^2$ and $u_0 = \frac{1}{2}(u_L + u_R) + \sqrt{\rho_L} - \sqrt{\rho_R}$. The initial values used for this simulation are $\rho_L = 2, \rho_R = 1, u_L = u_R = 0$. The parameters are: $\beta = 2.10^{-5}, \alpha = 3.33.10^{-3}, \Delta x = 0.000667$. The whole computational domain is $[-800, 800]$.

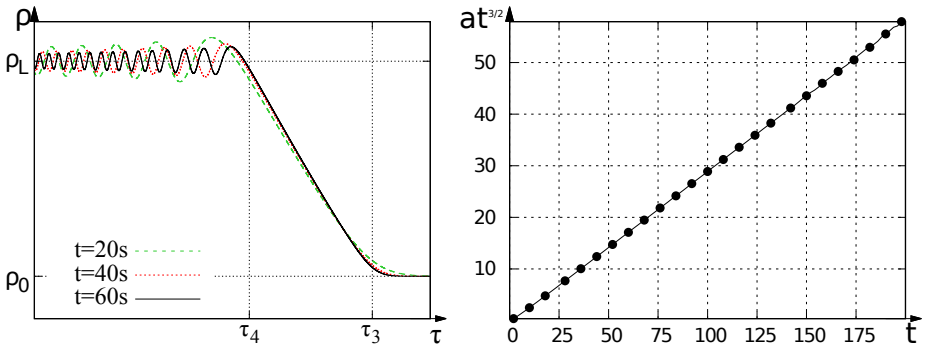


Figure 2.14: Vanishing oscillations at the vicinity of the singular point $\tau = \tau_4$. The figure on the left shows that these oscillations decrease in time. The figure on the right shows that the amplitude of the first oscillation $a$ is such that $at^{3/2}$ is linear. This implies the power law $a \propto t^{-1/2}$.

Another important detail is to check the position of the soliton which arises at the vicinity of $\tau = \tau_2$. Let $\tau_s(t)$ be the soliton position at time $t$. We will compare $\tau_s(t)$ with the asymptotic position $\tau_2$. The relative error is given by :

$$err_{pos}(t) = \left| \frac{\tau_s(t) - \tau_2}{\tau_2} \right|. \tag{2.74}$$

The time evolution of this error is shown on Figure (2.15) :

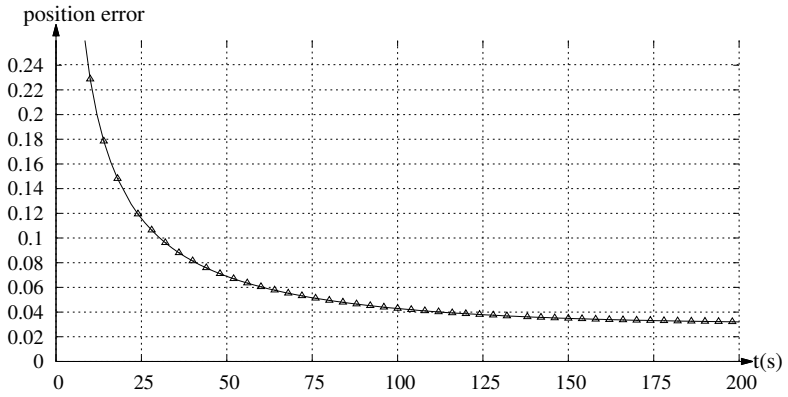


Figure 2.15: Relative error on the position of the first soliton plotted as function of time.

### MUSCL-Hancock : True discontinuity

We come back here to the issue of integrating a true discontinuity. As mentioned previously, one must be careful at the choice of the initial condition, since if we consider a step function for $\rho(x,0)$, we must also choose Dirac's delta function for $p(x,0) = \rho_x(x,0)$. Thus, the safest option would be to choose a $\delta$ that is sufficiently small (smaller than the mesh size) and still use the initial condition :

$$\rho(x,0) = \frac{\rho_L + \rho_R}{2} + \left(\frac{\rho_L - \rho_R}{2}\right) \tanh(\frac{x}{\delta}) \tag{2.75}$$

$$u(x,0) = 0 \tag{2.76}$$

$$\eta(x,0) = \rho(x,0) \tag{2.77}$$

$$w(x,0) = 0 \tag{2.78}$$

$$p(x,0) = \left(\frac{\rho_L - \rho_R}{2\delta}\right)\left(1 - \tanh^2\left(\frac{x}{\delta}\right)\right) \tag{2.79}$$

Under theses assumptions the initial condition is plotted on figure (2.16).

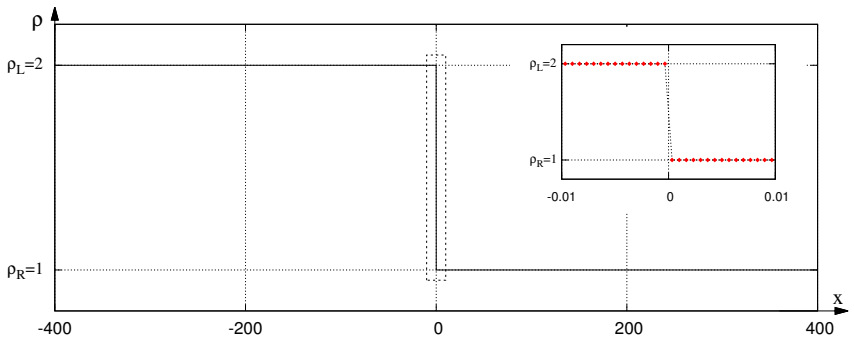


Figure 2.16: The step initial condition considered for the Riemann problem and the parameters $\rho_L = 2$, $\rho_R = 1$,$u_L = u_R = 0$, and $\delta = 10^{-5}$. The whole computation domain is shown here. The inset shows a zoom at the mesh size's level around the initial discontinuity.

We ran simulations using this initial condition. The result is shown on figure (2.18) :

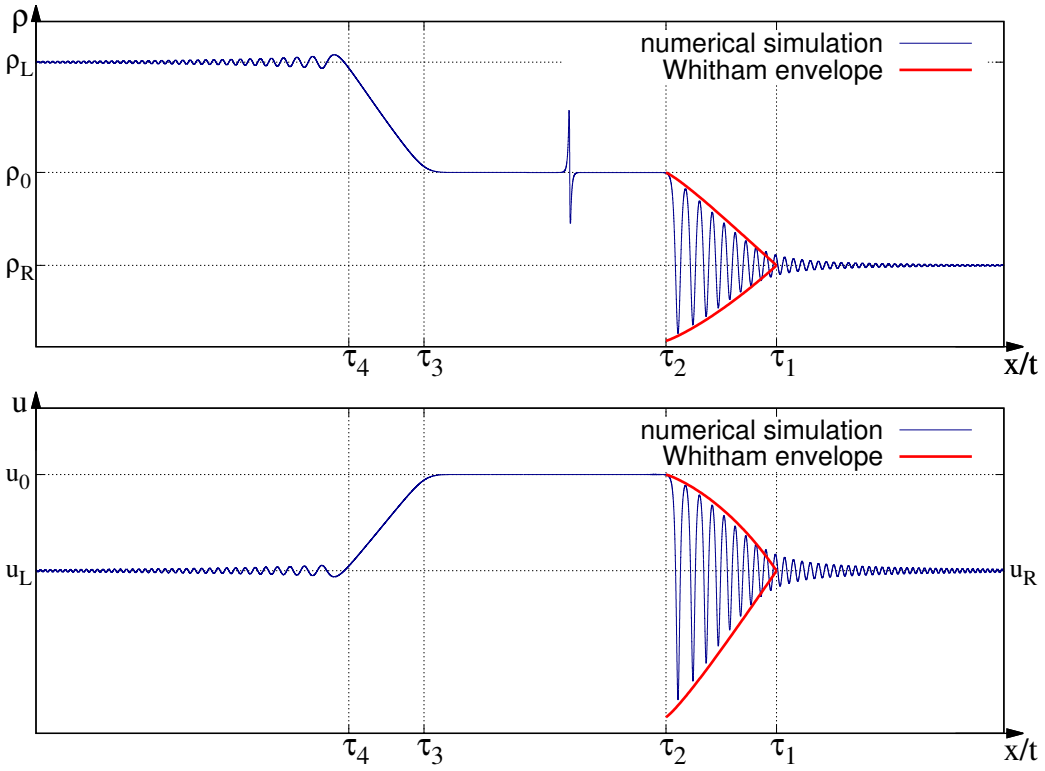


Figure 2.17: Comparison of the numerical results obtained for the previous initial conditions (thin line) with the Whitham modulational profile of the DSW (thick line) at $t = 50$. The upper figure shows $\rho = f(x/t)$ and the lower figure shows $u = f(x/t)$. The parameters are $\beta = 2.10^{-5}, \alpha = 3.33\ 10^{-3}, \Delta x = 0.000667$. The whole computational domain is $[-800, 800]$.

Like the results for the smoothed discontinuity, the numerical results overall comply with the asymptotic solution. However, in the profile of $\rho$, a non-vanishing glitch appears at the central plateau. It seems to be a sort of Gibbs phenomenon which appears at the initial discontinuity. Further refining the mesh does not prevent its appearance. Still, it does not hinder in anyway the propagation of the dispersive shock and rarefaction, nor does it significantly impact the central plateau.

**IMEX result**

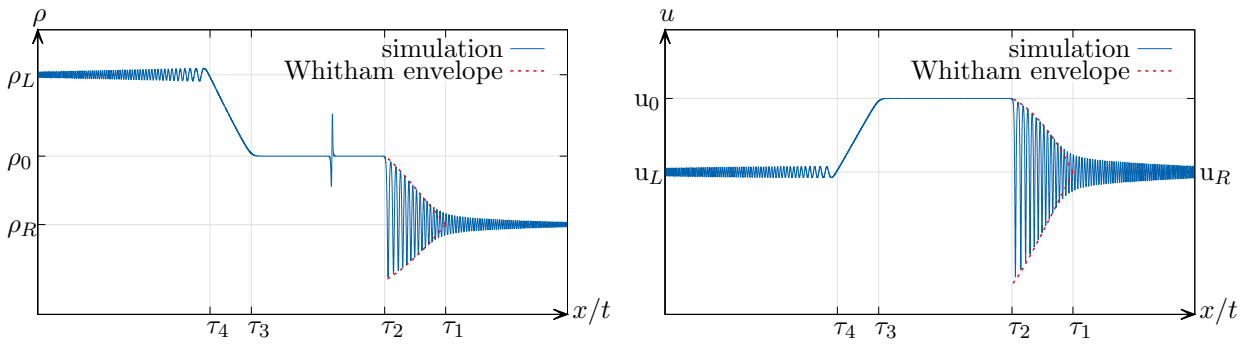


Figure 2.18: Comparison of the numerical results obtained with an IMEX scheme for the previous initial conditions (thin line) with the Whitham modulational profile of the DSW (thick line) at $t = 70$. The left figure shows $\rho = f(x/t)$ and the right figure shows $u = f(x/t)$. The parameters are $\beta = 2.10^{-5}, \alpha = 10^{-3}, \Delta x = 0.01$. The whole computational domain is $[-500, 500]$

## 2.6 Conlusion

We presented in this chapter the application of our augmented Lagrangian approach for the defocusing Nonlinear Schrödinger equation in one dimension of space. The numerical results show a very good agreement for both stationary solutions and non-stationary solutions. There are advantages as well as disadvantages when presenting this approach as a yet another method among others to solve the NLSE in general. First, out of the many methods that exist in the literature, the majority treats directly the NLSE in its complex field form. Most of these methods are spectral methods and are often accompanied by an arsenal of numerical techniques for the resulting ODE in Fourier space. Such techniques include Split-Steps [51][4][5] , Runge-Kutta Sliders [35], Integrating Factors [55] and exponential time differencing [59]. With the effective use and implementations of such methods, one can achieve very high accuracy with minimal computational costs, see for example [49] where 4th order time stepping is considered in combination with Fourier spectral methods in x, to numerically solve the KdV equation and the NLS equation. However, while in terms of performance, the method we present may not have the upper hand, it offers the possibility of solving a system of first-order hyperbolic conservation laws in which the mass and momentum in terms of Madelung variables are conserved. The hyperbolic setting also offers flexibility from the numerical point of view, in the sense that numerical methods for hyperbolic equations are quite universal and do not require too much additional work, especially for boundary conditions.

CHAPTER 3

# Thin film flows with capillarity

## 3.1 Introduction

This chapter is devoted to another interesting application of the Euler-Korteweg system : thin film flows where surface tension effects are not negligible. The subject has many useful industrial applications in relation to windshield defrosting for planes, coatings, painting, cooling of microelectronic equipment, etc. It also offers a variety of challenging problems for the researcher. One of the greatest challenges in this context is to derive an accurate model that is easy to solve. However, very precise models tend to call for complex mathematical structures which are either impossible to solve by hand or require extremely heavy computations. Thus, decreasing the complexity of the problem is a major concern, in order to obtain models which are fairly accurate but also solvable in a reasonable time.

It is in this context, that we would like to test out our augmented Lagrangian approach. In fact, it may prove useful to provide a new family of models which are first order hyperbolic, and which approximate existing models. From the numerical point of view, the hyperbolic setting is relatively comfortable, since the used methods, stability conditions, and scheme implementations are quite universal and do not call for much customization and additional analysis. The approach we offer is also extendable to more complex energy forms. For instance, the surface energy that is commonly used is linearly dependent on $||\nabla h||^2$. While this approximation is generally satisfactory, it becomes less reliable in the case where large gradients of fluid height appear. Since the method allows for it (as will be shown later), it seems beneficial to analyze some cases with non-linearized surface tension term, for which the capillary energy is given by :

$$E_{ca} = \frac{\sigma}{\rho}\sqrt{1 + ||\nabla h||^2} \tag{3.1}$$

We will begin by clarifying the setting, the assumptions and the notations that will be used throughout this part and then proceed to the analysis.

## 3.2 About thin film equations

### 3.2.1 Setting and notations

We consider a thin film that flows over an inclined horizontal plate under the effects of gravity :

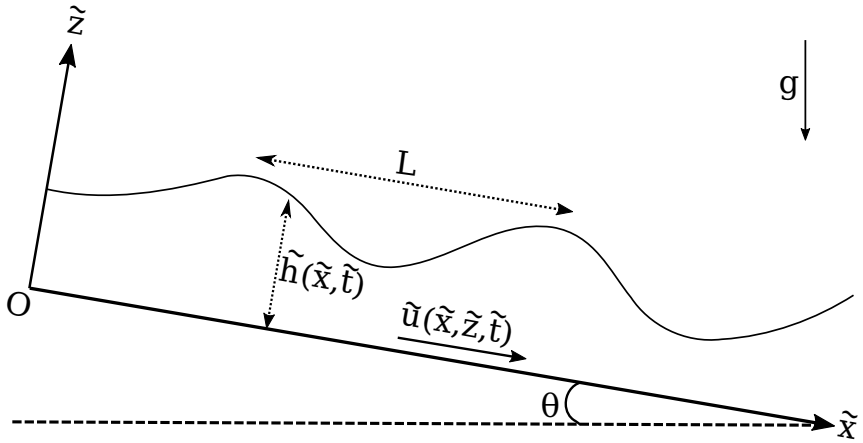


Figure 3.1: Schematic of a flowing thin film over an inclined horizontal plate with the notations used in this section.

Here, the 'tilde' infers that the associated quantity is dimensional. $\tilde{h}(\tilde{x},\tilde{t})$ refers to the fluid height and $\tilde{u}(\tilde{x},\tilde{z},\tilde{t})$ is the local velocity. The $\tilde{x}$ and $\tilde{z}$ respectively denote the direction parallel to the plane and normal to the plane. The transverse direction $\tilde{y}$ is neglected, as the applications considered here exhibit mainly two-dimensional dynamics. The angle $\theta$ is the constant inclination of the plate and $g$ is the acceleration of gravity. The typical distance in the $Ox$ direction will be denoted by $L$. For periodic waves this refers to the wavelength. In the long wave approximation, the ratio $\varepsilon$ between the typical depth and wavelength is considered small.

In order to obtain a non-dimensional form of the equations, one needs to consider some scaling of reference. For that, let us consider the *Nusselt flow* solution. It corresponds to the equilibrium solution, balancing the viscous forces and the gravity-driven acceleration resulting in a semi-parabolic velocity profile :

$$\begin{cases} \tilde{h}(\tilde{x},\tilde{t}) = \tilde{h}_N \\ \tilde{u}(\tilde{x},\tilde{z},\tilde{t}) = \tilde{u}_N(z) = \dfrac{g}{\nu}\sin\theta\tilde{z}\left(\tilde{h}_N - \dfrac{\tilde{z}}{2}\right) \end{cases} \tag{3.2}$$

the velocity is dependent on the fluid height and thus, in order to obtain a relevant characteristic velocity, we need to average it across the depth. For the sake of lightness, the averaged velocity over the depth will be denoted by its uppercase equivalent :

$$\tilde{U}_N = \frac{1}{\tilde{h}_N}\int_0^{\tilde{h}_N} \tilde{u}_N(\tilde{z})d\tilde{z} = \frac{g\tilde{h}_N^2}{3\nu}\sin\theta \tag{3.3}$$

Using this reference solution, we can define the Reynolds number ($Re$), the Froude number ($F$) and the Weber number ($We$) as follows :

$$Re = \frac{\tilde{h}_N \tilde{U}_N}{\nu}, \quad F = \frac{\tilde{U}_N}{\sqrt{g\tilde{h}_N}}, \quad We = \frac{\rho}{\gamma}\tilde{h}_N \tilde{U}_N^2.$$

The long wave parameter is defined as :

$$\varepsilon = \frac{\tilde{h}_N}{L} \tag{3.4}$$

These definitions permit to perform a rescaling of the variables in order to obtain their dimensionless versions:

$$u = \frac{\tilde{u}}{\tilde{U}_N}, \quad U = \frac{\tilde{U}}{\tilde{U}_N}, \quad h = \frac{\tilde{h}}{\tilde{h}_N}, \quad x = \frac{\tilde{x}}{L}, \quad z = \frac{\tilde{z}}{\tilde{h}_N}, \quad t = \frac{\tilde{t}}{L}\tilde{U}_N \tag{3.5}$$

Under these notations, the averaged governing equations write in dimensionless form [54, 65, 64] :

$$\begin{cases} \dfrac{\partial h}{\partial t} + \dfrac{\partial hU}{\partial x} = 0 \\ \dfrac{\partial hU}{\partial t} + \dfrac{\partial}{\partial x}\left(hU^2 + \dfrac{2\lambda^2}{225}h^5 + \dfrac{\cos\theta}{2F^2}h^2\right) = \dfrac{1}{\varepsilon Re}\left(\lambda h - \dfrac{3U}{h}\right) + \dfrac{\kappa}{F^2}hh_{xxx} \end{cases} \tag{3.6}$$

The dimensionless parameters $\lambda$, and $\kappa$ are introduced for convenience, as in [64] and are defined by :

$$\lambda = \frac{Re\sin\theta}{F^2}, \quad \kappa = \frac{\varepsilon^2 F^2}{We} \tag{3.7}$$

**Remark 3.2.1.** *Although the previous definitions imply that $\lambda = 3$ in this case, we will keep its expression as it is. Indeed, some later developments will require the expression instead of the numerical value.*

**Remark 3.2.2.** *When compared to the previously studied NLS equation, this system admits inherent dissipation in the right hand side. It is well-known that, dissipative systems do not derive from a Lagrangian formulation. This problem can be overcome by applying the augmented Lagrangian method to the dissipationless part of the system and then adding back the source terms, once the new augmented system is derived. Although, the term $\lambda h/\varepsilon Re$ may be put into the Lagrangian, as it corresponds to the conservative forces of gravity, it would change nothing to the resulting augmented system, whether it is added before or after the derivation. So we will keep it for later as well.*

**Remark 3.2.3.** *This system of equations is derived through asymptotic expansions in the small parameter $\varepsilon$. As such, it would admit not only the small parameters related to the augmentation method but also the small parameter $\varepsilon$ that governs the asymptotics of the original equations. This problem will be addressed in due time, after the new governing system of equations is established.*

Following these remarks, removing the dispersionless source terms and integrating the dispersive terms into the momentum flux results in the following system :

$$\begin{cases} \dfrac{\partial h}{\partial t} + \dfrac{\partial hU}{\partial x} = 0 \\ \dfrac{\partial hU}{\partial t} + \dfrac{\partial}{\partial x}\left( hU^2 + \dfrac{2\lambda^2}{225}h^5 + \dfrac{\cos\theta}{2F^2}h^2 + \dfrac{\kappa}{2F^2}h_x^2 - \dfrac{\kappa}{F^2}hh_{xx} \right) = 0 \end{cases} \tag{3.8}$$

To these equations, we can associate the following Lagrangian :

$$\mathcal{L} = \int_{\Omega_t} \left( \frac{1}{2}hU^2 - hf(h) - \frac{1}{2}\frac{\kappa}{F^2}h_x^2 \right) d\Omega, \qquad f(h) = \frac{h\cos\theta}{2F^2} + \frac{\lambda^2 h^4}{450} \tag{3.9}$$

and the total energy :

$$E = \frac{1}{2}hU^2 + \frac{h^2\cos\theta}{2F^2} + \frac{\lambda^2 h^5}{450} + \frac{1}{2}\frac{\kappa}{F^2}h_x^2 \tag{3.10}$$

Before proceeding to the augmented Lagrangian approach for this system, let us first analyze the full system of equations (3.6), with all its terms included. Even if the hyperbolization process only concerns the dissipationless part, comparisons between both models shall be done only after reinserting the source terms into the augmented system.

### 3.2.2 Dispersion relation and stability analysis

In this section, we shall derive the dispersion relation for equations (3.6). The latter can be written as :

$$\frac{\partial \mathbf{U}}{\partial t} + A(\mathbf{U})\frac{\partial \mathbf{U}}{\partial x} + B(\mathbf{U})\frac{\partial^3 \mathbf{U}}{\partial x^3} = S(\mathbf{U}) \tag{3.11}$$

where :

$$\mathbf{U} = \begin{pmatrix} h \\ U \end{pmatrix},\ A(\mathbf{U}) = \begin{pmatrix} U & h \\ \dfrac{\cos\theta}{F^2} + \dfrac{\lambda^2 h^3}{45} & U \end{pmatrix}, B(\mathbf{U}) = \begin{pmatrix} 0 & 0 \\ -\dfrac{\kappa}{F^2} & 0 \end{pmatrix}, S(\mathbf{U}) = \begin{pmatrix} 0 \\ \dfrac{\lambda}{\varepsilon Re} - \dfrac{3\lambda U}{\varepsilon h^2 Re} \end{pmatrix}$$

We consider an equilibrium state $\mathbf{U}_0 = (h_0, u_0)$ and look for solutions which are proportional to $e^{i(kx-wt)}$, where $k$ is the wave number and $\omega$ is the frequency. For simplicity, we can take $h_0 = 1$ and $u_0 = 1$ which corresponds to the Nusselt equilibirum solution. We consequently obtain the phase velocities as the eigenvalues of the matrix :

$$M = A(\mathbf{U}_0) - k^2 B(\mathbf{U}_0) + \frac{i}{k}\nabla S(\mathbf{U}_0) \tag{3.12}$$

A straightforward computation shows that the eigenvalues satisfy the equation :

$$\varepsilon(1-c_p)^2 - \frac{3i}{kRe}(1-c_p) - \varepsilon\left( \frac{cos\theta}{F^2} + \frac{2\lambda^2}{45} + \frac{k^2\varepsilon^2}{We} \right) - \frac{6i}{kRe} = 0 \tag{3.13}$$

Equation (3.13) has a complex valued discriminant. Computing the expression of the phase velocities is a trivial task, but separating the real and imaginary parts is not. Thus it brings

no additional information. However, it is possible to recover at least the stability condition for this equation. Since we are considering solutions which are proportional to $e^{i(kx-wt)}$, stability requires the imaginary part of the frequency $\omega$ to be positive, so that the wave amplitude does not grow exponentially in time. This defines the stability criterion :

$$Im(\omega) < 0 \;\Leftrightarrow\; Im(c_p) < 0 \tag{3.14}$$

We can show that this condition comes down to the well-known stability condition (see developments in appendix C.1):

$$\operatorname{cotg}\theta + \frac{\kappa}{\sin\theta}k^2 > \frac{6}{5}Re \tag{3.15}$$

Hence, the neutral stability curve $g(k)$ in the quadrant $(k > 0, Re > 0)$ is defined by :

$$g(k) = \sqrt{\frac{\sin\theta}{\kappa}\left(\frac{6}{5}Re - \operatorname{cotg}\theta\right)} \tag{3.16}$$

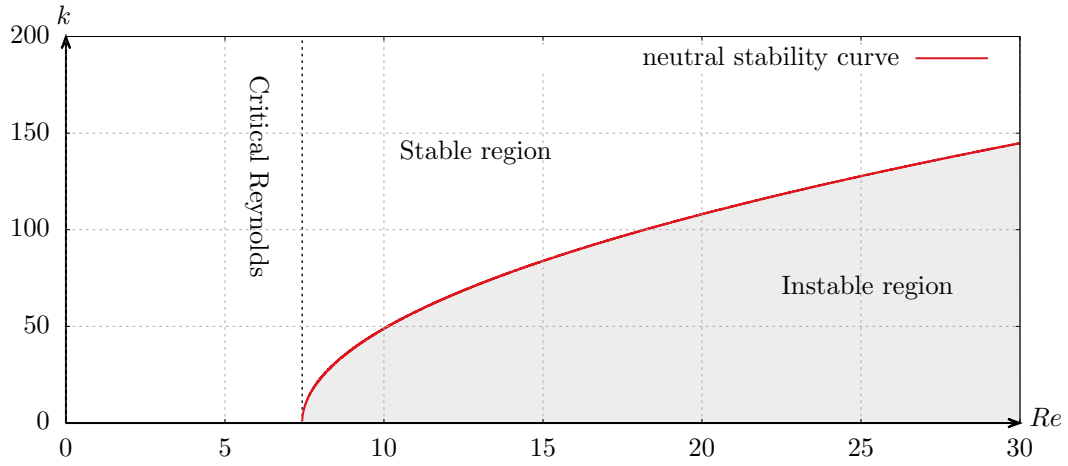


Figure 3.2: Stability regions for equations (3.6) in the $(k, Re)$ plane. The thick red line corresponds to the neutral stability curve. The gray region are instable waves. The white region are stable waves. The parameters used here correspond to the Liu & Gollub's experiments [57] with $\theta = 6.4°$, $We = 0.184$, $F = 0.847$ and $\varepsilon = 0.006$.

### 3.2.3 Asymptotic expansion of phase velocities

Instead of computing the exact expressions of the phase velocities, it is possible and more insightful in this case, to compute their asymptotic expansions in power series of the small parameter $\varepsilon$. In particular, this will prove useful when comparing the augmented model with this one (see section 3.3.2). This gives the leading order behavior of wave propagation. For more simplicity, we consider the variable $X = 1 - c_p$. In a first attempt, assume the following expansion in a regular perturbation series :

$$X = X_0 + \varepsilon X_1 + \varepsilon^2 X_2 + \mathcal{O}(\varepsilon^3) \tag{3.17}$$

Such an expansion is possible if the considered eigenvalue is sufficiently smooth in terms of $\varepsilon$ and bounded when $\varepsilon \to 0$. There will be cases where such assumptions on the regularity are not verified. In these cases, we will use results from the singular perturbation theory to find which form of the expansion is more suitable. Substituting the expansion (3.17) in the characteristic polynomial (3.13) yields :

$$-\left(\frac{3iX_0}{kRe}+\frac{6i}{kRe}\right)-\left(\frac{3iX_1}{kRe}+\frac{cos\theta}{F^2}+\frac{2\lambda^2}{45}-X_0^2\right)\varepsilon+\left(2X_0X_1-\frac{3iX_2}{kRe}\right)\varepsilon^2+\mathcal{O}(\varepsilon^3)=0 \quad (3.18)$$

by successively equating the coefficients multiplying $\varepsilon^n$ to zero for $n \in \{0,1,2\}$ we get :

$$X_0=-2, \quad X_1=\frac{i}{3}kRe\left(\frac{cos\theta}{F^2}+\frac{2\lambda^2}{45}-4\right), \quad X_2=-\frac{4}{9}k^2Re^2\left(\frac{cos\theta}{F^2}+\frac{2\lambda^2}{45}-4\right) \quad (3.19)$$

Hence :

$$X=-2+\frac{i}{3}kRe\left(\frac{cos\theta}{F^2}+\frac{2\lambda^2}{45}-4\right)\varepsilon-\frac{4}{9}k^2Re^2\left(\frac{cos\theta}{F^2}+\frac{2\lambda^2}{45}-4\right)\varepsilon^2+\mathcal{O}(\varepsilon^3)$$

Substituting the expansion in the second degree characteristic polynomial yields only one eigenvalue. This is due to the fact that the considered polynomial clearly degenerates into a first degree equation in the limit $\varepsilon \to 0$. Therefore, one of the eigenvalues was missed because it approaches infinity in the limit $\varepsilon \to 0$. In the general case, in order to find the missing solution, we introduce a rescaled variable $Y(\varepsilon)=X\delta(\varepsilon)$ and substitute it in the characteristic polynomial :

$$\varepsilon\frac{Y^2}{\delta(\varepsilon)^2}-\frac{3i}{kRe}\frac{Y}{\delta(\varepsilon)}-\varepsilon\left(\frac{cos\theta}{F^2}+\frac{2\lambda^2}{45}+\kappa\frac{k^2}{F^2}\right)-\frac{6i}{kRe}=0$$

In order to get a nontrivial solution, we need the leading order terms to have the same order of magnitude, otherwise in the limit $\varepsilon \to 0$, one of them will simply vanish and the equation will degenerate again. Balancing the two first terms yields:

$$\delta(\varepsilon)^2=\varepsilon\delta(\varepsilon) \Rightarrow \delta(\varepsilon)=\varepsilon$$

and so :

$$Y^2-\frac{3i}{kRe}Y-\varepsilon^2\left(\frac{cos\theta}{F^2}+\frac{2\lambda^2}{45}+\frac{k^2\varepsilon^2}{We}\right)-\frac{6i\varepsilon}{kRe}=0$$

Now we take $Y$ in the form :

$$Y=Y_0+\varepsilon Y_1+\varepsilon^2Y_2+Y_3\varepsilon^3+\mathcal{O}(\varepsilon^4) \quad (3.20)$$

In this expansion, we are looking for $Y_0 \neq 0$, otherwise, this will lead again to the previous trivial solution. We truncate at one higher order than previously to remain at a second order approximation when replacing back $X$. Substituting in the characteristic polynomial gives :

$$Y_0^2-\frac{3i}{kRe}Y_0+\left(2Y_0Y_1-\frac{3i}{kRe}Y_1-\frac{6i}{kRe}\right)\varepsilon+\left(Y_1^2+2Y_0Y_2-\frac{3i}{kRe}Y_2-\frac{cos\theta}{F^2}-\frac{2\lambda^2}{45}\right)\varepsilon^2$$
$$+\left(2Y_1Y_2-\frac{3iY_3}{kRe}+2Y_0Y_3\right)+\mathcal{O}(\varepsilon^4)=0$$

once again, by successively equating the coefficients multiplying $\varepsilon^n$ to zero, for $n = \{0, 1, 2, 3\}$ , we get :

$$Y_0 = \frac{3i}{kRe}, \quad Y_1 = 2, \quad Y_2 = i\frac{kRe}{3}\left(4 - \frac{cos\theta}{F^2} - \frac{2\lambda^2}{45}\right), \quad Y_3 = -\frac{4}{9}k^2Re^2\left(4 - \frac{cos\theta}{F^2} - \frac{2\lambda^2}{45}\right)$$

we substitute $X = Y/\varepsilon$ to obtain :

$$X = \frac{3i}{k\varepsilon Re} + 2 + i\frac{kRe}{3}\left(4 - \frac{cos\theta}{F^2} - \frac{2\lambda^2}{45}\right)\varepsilon - \frac{4}{9}k^2Re^2\left(4 - \frac{cos\theta}{F^2} - \frac{2\lambda^2}{45}\right)\varepsilon^2 + \mathcal{O}(\varepsilon^3) \quad (3.21)$$

To conclude, the expansions of two phase velocities $c_p = 1 - X$ are given by:

$$c_{p_1} = 3 - \frac{i}{3}kRe\left(\frac{cos\theta}{F^2} + \frac{2\lambda^2}{45} - 4\right)\varepsilon + \frac{4}{9}k^2Re^2\left(\frac{cos\theta}{F^2} + \frac{2\lambda^2}{45} - 4\right)\varepsilon^2 + \mathcal{O}(\varepsilon^3)$$

$$c_{p_2} = -\frac{3i}{k\varepsilon Re} - 1 + \frac{i}{3}kRe\left(\frac{cos\theta}{F^2} + \frac{2\lambda^2}{45} - 4\right)\varepsilon - \frac{4}{9}k^2Re^2\left(\frac{cos\theta}{F^2} + \frac{2\lambda^2}{45} - 4\right)\varepsilon^2 + \mathcal{O}(\varepsilon^3)$$

## 3.3 Augmented Lagrangian formulation

### 3.3.1 Governing equations

In this section, we will apply the augmented Lagrangian approach to thin film equations. We remind that the augmentation process concerns the sourceless system of equations (3.8). The source terms shall be added by hand after the derivation of the equations. The augmented Lagrangian based on the Lagrangian (3.22) writes :

$$\mathcal{L} = \int_{\Omega_t}\left(\frac{1}{2}hU^2 + \frac{\beta}{2}h\dot{\eta}^2 - hf(h) - \frac{1}{2}\frac{\kappa}{F^2}p^2 - \frac{h}{2\alpha}\left(\frac{\eta}{h} - 1\right)^2\right)d\Omega \quad (3.22)$$

To this Lagrangian corresponds the following system of augmented equations :

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(hU) = 0, \quad (3.23)$$

$$\frac{\partial hU}{\partial t} + \frac{\partial}{\partial x}\left(hU^2 + \frac{h^2}{2F^2}cos\theta + \frac{2\lambda^2h^5}{225} + \frac{\kappa}{2F^2}p^2 + \frac{\eta}{\alpha}\left(1 - \frac{\eta}{h}\right)\right) = 0, \quad (3.24)$$

$$\frac{\partial h\eta}{\partial t} + \frac{\partial}{\partial x}(h\eta U) = hw, \quad (3.25)$$

$$\frac{\partial hw}{\partial t} + \frac{\partial}{\partial x}\left(hwU - \frac{\kappa}{\beta F^2}p\right) = \frac{1}{\alpha\beta}\left(1 - \frac{\eta}{h}\right), \quad (3.26)$$

$$\frac{\partial p}{\partial t} + \frac{\partial}{\partial x}\left(pU - w\right) = 0 \quad (3.27)$$

Now, all that remains is to add back, by hand, the previously removed source terms to the momentum equation, so that it becomes :

$$\frac{\partial hU}{\partial t} + \frac{\partial}{\partial x}\left(hU^2 + \frac{h^2}{2F^2}cos\theta + \frac{2\lambda^2h^5}{225} + \frac{\kappa}{2F^2}p^2 + \frac{\eta}{\alpha}\left(1 - \frac{\eta}{h}\right)\right) = \frac{1}{\varepsilon Re}\left(\lambda h - \frac{3U}{h}\right) \quad (3.28)$$

**Remark 3.3.1.** *As to relieve ambiguity, the augmented system will now refer to equations* (3.23)–(3.27)*, where we substitute the momentum equation* (3.43) *by* (3.28)*.*

In this case, the characteristic speeds $\xi_i$ are given by :

$$\begin{cases} \xi_1 = u \\ \xi_{2,3,4,5} = U \pm \sqrt{(1+\beta p^2)a + b \pm \sqrt{((1+\beta p^2)a+b)^2 - 4ab}} \end{cases} \tag{3.29}$$

where $a$ et $b$ are positive quantities given by :

$$a = \frac{\kappa}{2\beta h F^2}, \quad b = \frac{\lambda^2 h^4}{45} + \frac{\eta^2}{2\alpha h^2} + \frac{h\cos\theta}{2F^2} \tag{3.30}$$

The five eigenvalues are real and distinct, unless we have :

$$p = 0, \quad \text{and} \quad \frac{2\lambda^2 h^5}{45} + \frac{\eta^2}{\alpha h} + \frac{h^2\cos\theta}{F^2} - \frac{\kappa}{\beta F^2} = 0 \tag{3.31}$$

in which case the eigenvalues become :

$$\xi_1 = U, \quad \xi_{2,3} = U + \sqrt{b}, \quad \xi_{4,5} = U - \sqrt{b} \tag{3.32}$$

Even in this case, it is easy to check that we still obtain a full set of linearly independent eigenvectors.

**Remark 3.3.2.** *It would have been straightforward to obtain the hyperbolicity of the system if its energy was convex with respect to the conservative variables. This is not the case here mainly due to the penalty term.*

### 3.3.2 Dispersion Relation: $\alpha$ and $\beta$ scaling

Like for the NLS equation, we must assign values to the relaxation parameters $\alpha$ and $\beta$. Unlike the NLS equation, it proved impossible in the current case, to obtain an explicit expression of the phase velocities. The task requires an unreasonable number of assumptions on the parameters. Instead, since we performed an asymptotic expansion of the phase velocities for the original equations, we will do equivalently for the augmented system and try to conserve the asymptotic behavior up to a given order. First, let us cast the augmented system into a quasilinear form :

$$\frac{\partial \mathbf{U}}{\partial t} + A(\mathbf{U})\frac{\partial \mathbf{U}}{\partial x} = S(\mathbf{U}) \tag{3.33}$$

where :

$$\mathbf{U} = \begin{pmatrix} h \\ U \\ \eta \\ w \\ p \end{pmatrix}, \; A(\mathbf{U}) = \begin{pmatrix} U & h & 0 & 0 & 0 \\ C + \frac{1}{\alpha h} & U & \frac{1}{\alpha}\left(1 - \frac{2\eta}{h}\right) & 0 & \frac{\kappa p}{F^2 h} \\ 0 & 0 & U & 0 & 0 \\ 0 & 0 & 0 & U & -\frac{\kappa}{\beta F^2 h} \\ 0 & p & 0 & -1 & U \end{pmatrix}, S(\mathbf{U}) = \begin{pmatrix} 0 \\ \frac{\lambda}{\varepsilon Re} - \frac{3\lambda U}{\varepsilon h^2 Re} \\ w \\ \frac{1}{\alpha\beta h}\left(1 - \frac{\eta}{h}\right) \\ 0 \end{pmatrix}$$

where $C = \frac{\cos\theta}{F^2} + \frac{2\lambda^2 h^3}{45}$. We consider an equilibrium state $\mathbf{U}_0$ such that $h = h_0, U = U_0, \eta = h_0, w = 0, p = 0$ and look for solutions which are proportional to $e^{i(kx-wt)}$. Similarly to the previous section, we will perform the analysis in terms of the variable $X = u_0 - c_p$ which makes computations slightly lighter. We will also assume the Nusselt flow as reference equilibrium solution and thus take $h_0 = 1$ and $U_0 = 1$. Under these assumptions, the phase velocities are given as the eigenvalues of the matrix $A(\mathbf{U}_0) + \frac{i}{k}\nabla S(\mathbf{U}_0)$ which writes :

$$A(\mathbf{U}_0) + \frac{i}{k}\nabla S(\mathbf{U}_0) = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ C + \frac{1}{\alpha} + \frac{6i}{\varepsilon k Re} & 1 - \frac{3i}{k\varepsilon Re} & -\frac{1}{\alpha} & 0 & 0 \\ 0 & 0 & 1 & \frac{i}{k} & 0 \\ \frac{i}{\alpha\beta k} & 0 & \frac{-i}{\alpha\beta k} & 1 & -\frac{\kappa}{\beta F^2} \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix} \tag{3.34}$$

to which corresponds the characteristic polynomial, written in terms of the variable $X$:

$$\frac{1}{\alpha\beta\epsilon k^3} X \left( A_0 + A_1 X + A_2 X^2 + A_3 X^3 + A_4 X^4 \right) = 0 \tag{3.35}$$

where the coefficients are given by :

$$A_0 = \frac{6i}{Re} + C\varepsilon k + \frac{6i\alpha}{ReWe} k^2\varepsilon^2 + \frac{1+\alpha C}{We} k^3\varepsilon^3, \quad A_1 = \frac{3i}{Re} + \frac{3i\alpha}{ReWe} k^2\varepsilon^2$$
$$A_2 = -\left( \frac{6i\alpha}{Re}\beta k^2 + \left(1 + \beta k^2 + \alpha C \beta k^2\right) k\varepsilon + \frac{\alpha}{We} k^3\varepsilon^3 \right), \quad A_3 = -\frac{3i\alpha\beta k^2}{Re}, \quad A_4 = \alpha\beta\varepsilon k^3$$

The main purpose of this analysis, is to obtain the asymptotic behavior of the phase velocities in the small parameter $\varepsilon$, in order to compare with original system. The characteristic polynomial (3.35) has five complex roots. One root is trivial which is $X = 0$. Out of the four remaining roots, two should be approximately the same as in the original system. We denote the corresponding phase velocities by $c'_{p_1}$ and $c_{p'_2}$. The two remaining roots are rather linked to the fast characteristics of the augmented system. Before we proceed, it would be safest to set all small parameters as a function of $\varepsilon$. Having more than one small parameter may result in an erroneous asymptotic analysis. Thus a choice of $\alpha$ and $\beta$ as functions of $\varepsilon$ is necessary. A reasonable choice would be to make sure that the phase velocities of the augmented system match the original one, at least up to first order in $\varepsilon$. It is shown in appendix C.2 that the latter statement is valid if $\alpha$ and $\beta$ given by :

$$\alpha = \mathcal{O}(\varepsilon), \quad \beta = \mathcal{O}(\varepsilon^3) \tag{3.36}$$

Under such a scaling, the phase velocities, which correspond to the roots of (3.35), can be expanded into power series of $\varepsilon$ as follows :

$$\begin{cases} c_{p_0} = 1 \\ c'_{p_1} = 3 - \frac{i}{3} kRe \left( \frac{cos\theta}{F^2} + \frac{2\lambda^2}{45} - 4 \right) \varepsilon + \frac{4}{9} k^2 Re^2 \left( \frac{cos\theta}{F^2} + \frac{2\lambda^2}{45} - 4 \right) \varepsilon^2 + \mathcal{O}(\varepsilon^3) \\ c'_{p_2} = -\frac{3i}{k\varepsilon Re} - 1 + \frac{i}{3} kRe \left( \frac{cos\theta}{F^2} + \frac{2\lambda^2}{45} - 4 \right) \varepsilon - \frac{4}{9} k^2 Re^2 \left( \frac{cos\theta}{F^2} + \frac{2\lambda^2}{45} - 4 \right) \varepsilon^2 + \mathcal{O}(\varepsilon^3) \\ c_{p_{3,4}} = \pm \frac{1}{k\varepsilon^2} + o\left(\frac{1}{\varepsilon^2}\right) \end{cases} \tag{3.37}$$

### 3.3.3 Neutral stability analysis

Another important property of the model to be checked is its stability curve. If the neutral stability curves of both systems are significantly different, then discrepancies in their respective behaviors may be observed for the same frequencies. Unstable waves in the original system could become stable for the augmented model or vice-verse. Avoiding this inconsistency requires us to check the stability condition $Im(c_p) < 0$ (equivalently $Im(X) > 0$). This task requires tedious calculations but can be greatly simplified if we instead look for the critical curve $Im(c_p) = 0$. This literally means that we search for real roots of the characteristic polynomial (3.35). We discard the root $X = 0$ and we consider a nontrivial root $X \in \mathbb{R}^*$ of (3.35). Equating the real and imaginary parts of (3.35) respectively to zero yields the equations (after normalization) :

$$\begin{cases} Q_{re}(X) = X^4 - \left(\frac{1}{\alpha} + C + \frac{\kappa}{\beta F^2} + \frac{1}{\alpha\beta k^2}\right) X^2 + \frac{C}{\beta}\left(\frac{\kappa}{F^2} + \frac{1}{\alpha k^2}\right) + \frac{\kappa}{\alpha\beta F^2} = 0 \\ Q_{im}(X) = X^3 + 2X^2 - \left(\frac{\kappa}{\beta F^2} + \frac{1}{\alpha\beta k^2}\right) X - 2\left(\frac{\kappa}{\beta F^2} + \frac{1}{\alpha\beta k^2}\right) = 0 \end{cases} \tag{3.38}$$

where $X$ must be a common root to the two polynomials. $Q_{im}$ admits three real roots:

$$X_1 = -2 \quad ; \quad X_{2,3} = \pm\sqrt{\frac{\kappa}{\beta F^2} + \frac{1}{\alpha\beta k^2}} \tag{3.39}$$

Let us plug the obtained roots in $Q_{Re}$. The latter is biquadratic implying that $X_2$ and $X_3$ are equivalent. We obtain :

$$Q_{re}(X_1) = \frac{1}{k^2}\left(\frac{\kappa k^2}{F^2} - 4 + C\right) + \alpha\left(4\beta - \frac{\kappa}{F^2}\right)(4 - C) - 4\beta$$

$$Q_{re}(X_{2,3}) = -\frac{1}{\alpha^2\beta}$$

Clearly, $Q_{re}(X_{2,3})$ is always a non-zero value .This leaves $X_1$ as the unique possible common root. It remains to replace $C$ by its value :

$$C = \frac{\cos\theta}{F^2} + \frac{2\lambda^2}{45} = \frac{\cos\theta}{F^2} + \frac{6Re\sin\theta}{45F^2}$$

Simplifying gives the neutral stability curve:

$$\left(\mathrm{cotg}\theta + \frac{\kappa k^2}{\sin\theta} - \frac{6Re}{5}\right) + k^2(\mathcal{O}(\alpha) + \mathcal{O}(\beta)) = 0 \tag{3.40}$$

This equation is consistent with the stability condition (3.15) in the limits $\alpha \to 0, \beta \to 0$. The error terms are proportional to the square of the wavenumber implying that convergence is not uniform with respect to $k$. This is confirmed in figure 3.3 where neutral stability curves for the original system and the augmented one are compared for several values of $\alpha$ and $\beta$. The values which were obtained through the asymptotic analysis show a very good agreement

in terms of stability regions. Less restrictive values of either $\alpha$ or $\beta$ are shown to exhibit a different behavior in the sense that the instability region in these cases is significantly wider. In terms of convergence, it is clear that for low frequencies the curves match very well up to a certain cut-off frequency, starting from which there is a notable increase in the error of approximation. This is similar to what we had observed for the NLS equation for the phase velocity.

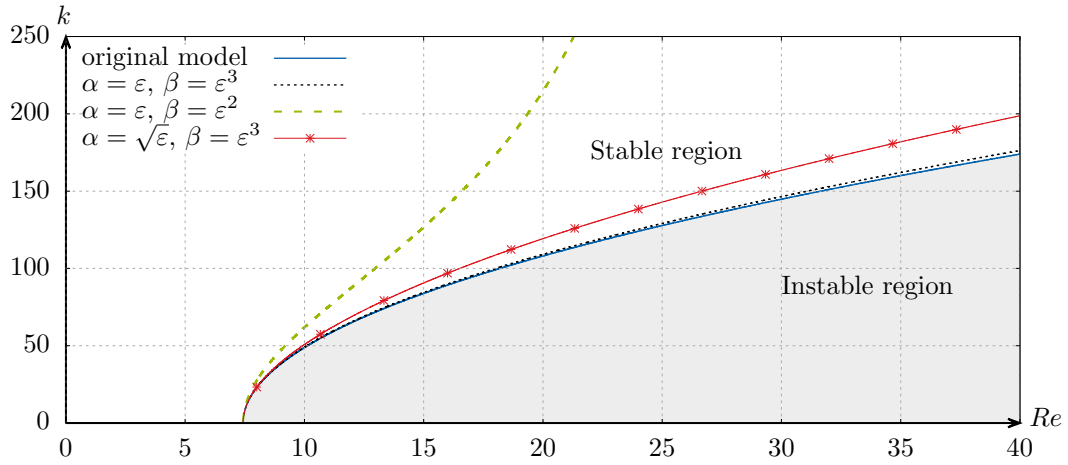


Figure 3.3: Neutral stability curves in the $(k, Re)$ plane for the original model (continuous blue line) and the augmented model for various scalings of $\alpha$ and $\beta$ with respect to $\varepsilon$. The parameters used here are the same as in figure 3.2.

## 3.4 Nonlinear surface tension

We will address now the problem of nonlinear surface energy. We will keep to all the previous setting but the capillary terms. We revert back to the total energy of a thin film without surface tension linearization :

$$E_{nl} = \frac{1}{2}hU^2 + \frac{h^2\cos\theta}{2F^2} + \frac{\lambda^2 h^5}{450} + \frac{\kappa}{\varepsilon^2 F^2}\sqrt{1+\varepsilon^2 h_x^2} \tag{3.41}$$

This form is more suitable to larger gradients of fluid height, for which the usually used Taylor expansion, simplifying the surface energy term, becomes less relevant. We show that this adds almost no difficulty when using the augmented Lagrangian approach. This form of energy is still compatible with the generic Lagrangian (A.1), for which we derived the Euler-Lagrange equations. Thus, applying Hamilton's principle to the associated augmented Lagrangian :

$$\mathcal{L} = \int_{\Omega_t}\left(\frac{1}{2}hU^2 + \frac{\beta}{2}h\dot{\eta}^2 - \frac{h^2\cos\theta}{2F^2} - \frac{\lambda^2 h^5}{450} - \frac{\kappa}{\varepsilon^2 F^2}\sqrt{1+\varepsilon^2 p^2} - \frac{1}{2\alpha}h\left(\frac{\eta}{h}-1\right)^2\right)d\Omega \tag{3.42}$$

submitted to the mass conservation law, and adding the usual closure equations results in the augmented system :

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(hU) = 0, \tag{3.43}$$

$$\frac{\partial hU}{\partial t} + \frac{\partial}{\partial x}\left(hU^2 + \frac{h^2}{2F^2}cos\theta + \frac{2\lambda^2 h^5}{225} - \frac{\kappa}{F^2\varepsilon^2\sqrt{1+\varepsilon^2 p^2}} + \frac{\eta}{\alpha}\left(1-\frac{\eta}{h}\right)\right) = 0, \tag{3.44}$$

$$\frac{\partial h\eta}{\partial t} + \frac{\partial}{\partial x}(h\eta u) = hw, \tag{3.45}$$

$$\frac{\partial hw}{\partial t} + \frac{\partial}{\partial x}\left(hwU - \frac{\kappa p}{\beta F^2\sqrt{1+\varepsilon^2 p^2}}\right) = \frac{1}{\alpha\beta}\left(1-\frac{\eta}{h}\right), \tag{3.46}$$

$$\frac{\partial p}{\partial t} + \frac{\partial}{\partial x}\left(pU - w\right) = 0 \tag{3.47}$$

The characteristic speeds $\xi_i$ of this system are given by :

$$\begin{cases} \xi_1 = U \\ \xi_{2,3,4,5} = U \pm \sqrt{a'(1+\beta p^2) + b \pm \sqrt{(a'(1+\beta p^2)+b)^2 - 4a'b}} \end{cases} \tag{3.48}$$

where $b$ is defined as in equation (3.30) and $a'$ is given by :

$$a' = \frac{\kappa}{2\beta h F^2(1+p^2)^{3/2}} = a(1+p^2)^{-3/2} > 0, \tag{3.49}$$

The hyperbolicity of this system is obtained exactly as in the linear surface tension case.

## 3.5 Numerical simulations

### 3.5.1 Test for a Gaussian initial data

In a first attempt, we compare with the results obtained in [14] for a Gaussian initial data. It offers the advantage of being one of the few tests that was compared for both linear and nonlinear surface tensions in the non-stationary case. Thus, for the setting of this test, we consider the following Lagrangian, in dimensioned variables :

$$\mathcal{L} = \int_{\Omega_t}\left(\frac{1}{2}\tilde{h}\tilde{U}^2 - \frac{1}{2}g\tilde{h}^2 - \frac{\sigma}{\rho}\sqrt{1+\tilde{h}_{\tilde{x}}^2}\right)\, d\Omega \tag{3.50}$$

to which corresponds, in the augmented setting, the Lagrangian :

$$\mathcal{L} = \int_{\Omega_t}\left(\frac{1}{2}\tilde{h}\tilde{U}^2 + \frac{1}{2}\tilde{\beta}\tilde{h}\tilde{w}^2 - \frac{1}{2}g\tilde{h}^2 - \frac{\sigma}{\rho}\sqrt{1+\tilde{p}^2} + \frac{\tilde{h}}{2\tilde{\alpha}}\left(1-\frac{\tilde{\eta}}{\tilde{h}}\right)^2\right)\, d\Omega \tag{3.51}$$

We consider the initial data :

$$\begin{cases} \tilde{h}(\tilde{x},\tilde{t}=0) = h_0 + h_1 e^{-\frac{x^2}{2(b/b_0)^2}} \\ \tilde{\eta}(\tilde{x},\tilde{t}=0) = \tilde{h}(\tilde{x},\tilde{t}=0) \\ \tilde{p}(\tilde{x},\tilde{t}=0) = -\frac{h_1}{(b/b_0)^2}\tilde{x}e^{-\frac{x^2}{2(b/b_0)^2}} \\ \tilde{U}(\tilde{x},\tilde{t}=0) = 0 \\ \tilde{w}(\tilde{x},\tilde{t}=0) = 0 \end{cases} \tag{3.52}$$

Here, $h_0$ is the water elevation at rest. $b$, $b_0$ and $h_1$ are parameters that define the shape of the deformation as can be seen in figure 3.4.

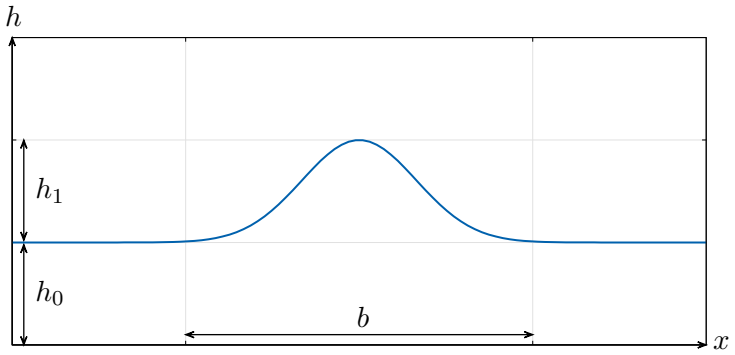


Figure 3.4: Shape of the Gaussian initial data.

Starting from the initial condition (3.52), we compute the time evolution of the deformation, which generates both gravity and capillary waves. For the simulation, we take a domain of $L = [-50mm, 50mm]$. The numerical results, at $t = 5ms$, is displayed in figure 3.5, with both linear and nonlinear capillarity terms :

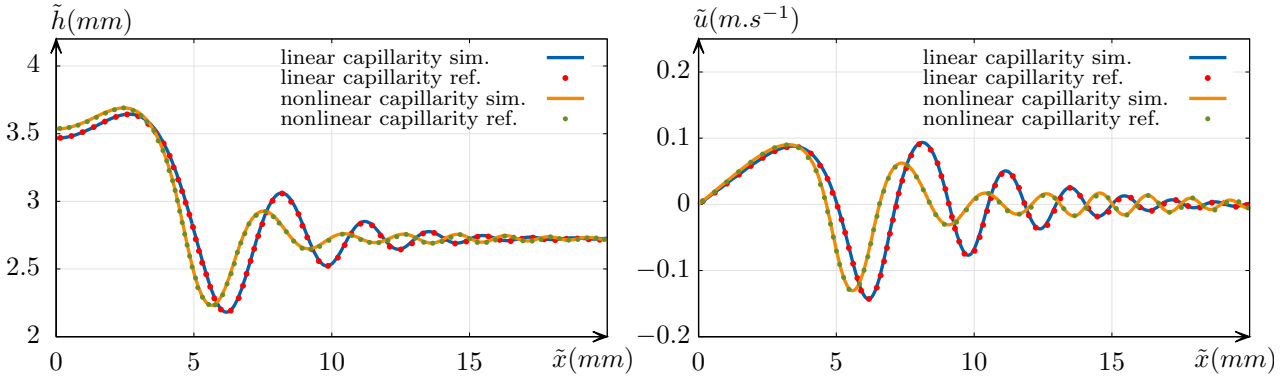


Figure 3.5: Comparison of the obtained numerical results (solid lines) with the converged numerical solutions proposed in [14] (dots), for the Gaussian initial data (3.52) at $t = 5ms$. Parameters used here are $g = 9.81m.s^{-2}$, $\sigma = 0.0728Kg.s^{-2}$, $\rho = 1000Kg.m^{-3}$, $h_0 = 2.725mm$, $h_0 = h_1$, $b = 1.5h_1$ and $b_0 = 4.29193$. $\tilde{\alpha} = 10^{-3}m^{-2}s^2$ and $\tilde{\beta} = 10^{-5}$. Results are shown with a mesh resolution of $n = 5000$.

The shown results are obtained by using an IMEX-2,2,2 scheme with MUSCL reconstruction. The $cfl$ is set to 0.8. Unlike what was used in [14], there is no need to adjust the $cfl$ to be able to capture the capillary waves, as the characteristic velocity of the model, naturally includes the fast oscillations scale. The comparison shows a perfect agreement, in terms of wave amplitude, wave frequency and wave speed. The comparison of $\tilde{p}$ with $\tilde{h}_{\tilde{x}}$, computed through centered finite differences is given in the next figure for both capillarity forms:

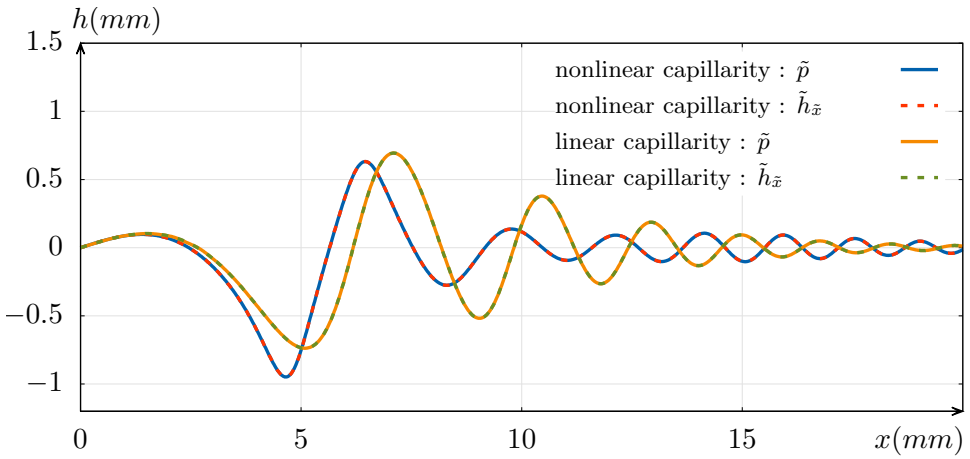


Figure 3.6: Comparison of $\tilde{h}_{\tilde{x}}$ (dashed lines) and $\tilde{p}$ (continuous lines) for the same values as above for both linear and nonlinear surface tension models. The curves coincide perfectly.

This test shows that the numerical results obtained for both linear and nonlinear surface tension models match perfectly with the converged numerical solutions of [14]. Although the reference we compare with, is not an exact solution, this test shows that the augmented model in this setting, approximates the original model with excellent accuracy.

### 3.5.2 Liu & Gollub's experiment

We study here the experiments done by Liu and Gollub [57]. It consists in a two-meters long canal, inclined with a constant angle $\theta$, containing a thin film of water (around one millimeter). The fluid is initially at rest. At $t = 0$, a periodic perturbation, with an arbitrary amplitude, is imposed at one of its boundaries. Depending on the frequency of the perturbation, the thus generated waves have amplitudes that may decay or grow in time. The frequencies of interest in this experiment are the unstable frequencies, for which it was observed that the wave amplitude grows until it reaches a stationary state whose behavior is strongly dependent on the imposed frequency. An example is given in figure 3.7 for a frequency of $1.5Hz$. The perturbation is imposed on the left boundary. One can see that the initial oscillations grow as they propagate until a wave-train of small capillary oscillations appears in the vicinity of the leading front which stabilizes the time evolution.

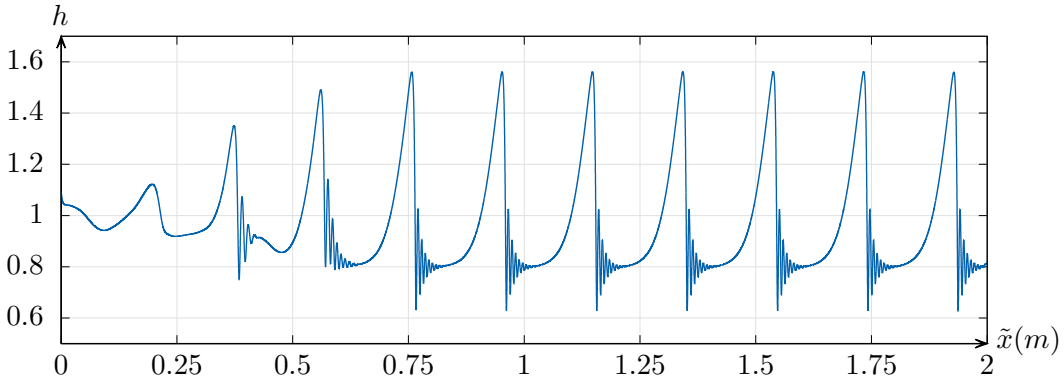


Figure 3.7: Dimensionless water height as a function of space (dimensioned), in the setting of the Liu & Gollub experiment, for an imposed frequency of $1.5Hz$. (Obtained through numerical simulation). Parameters used here are : $Re = 19.33$, $\kappa = 1.440.10^{-4}$, $Fr = 0.8476$,$\theta = 6.4°$

An important feature of this experiment is that, for different imposed frequencies, the generated wave displays significantly different features (presented in figure 3.8). For higher frequencies, the front and tail of the waves become closer which hinders the development of the wave precursors and may lead to strong interaction between the successive pulses.

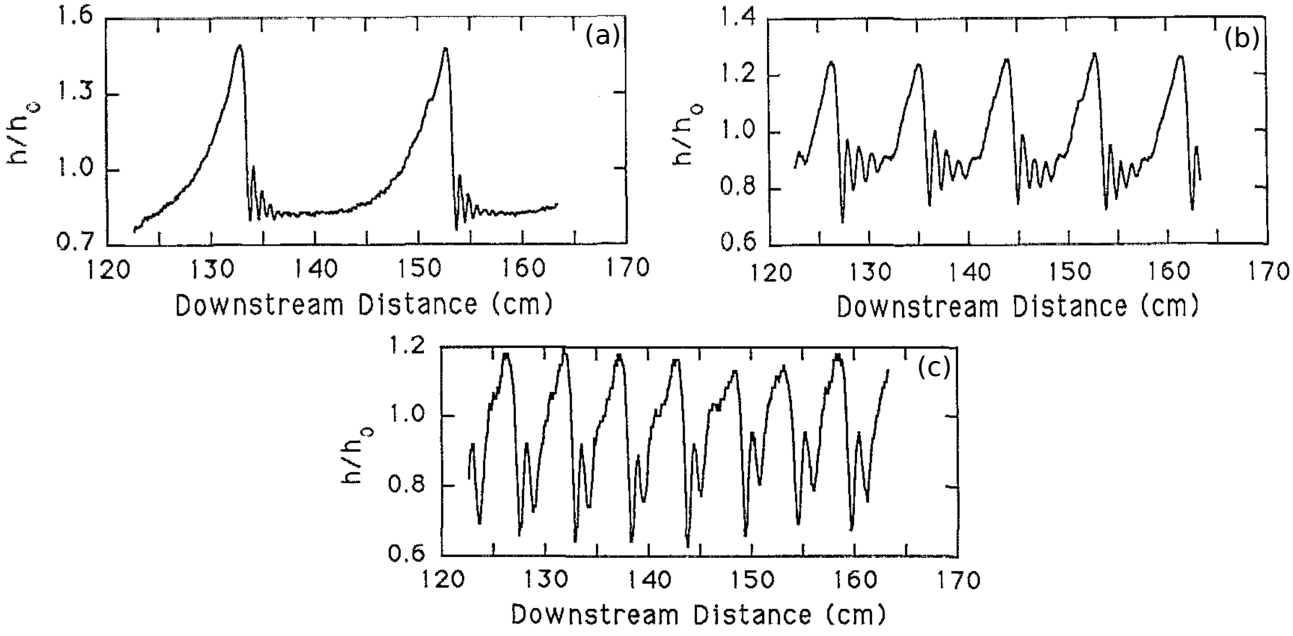


Figure 3.8: Dimensionless water depth as a function of the downstream distance $\tilde{x}$ for different frequencies : (a) $\tilde{f} = 1.5Hz$, (b) $\tilde{f} = 3.0Hz$ and (c) $\tilde{f} = 4.5Hz$ in the experiments of Liu & gollub [57]. Figures reprinted from J. Liu and J. P. Gollub. "Solitary wave dynamics of film flows". In: *Physics of Fluids* 6.5 (1994), pp. 1702–1712, with the permission of AIP Publishing.

Before proceeding to the numerical resolution of this experiment by means of the augmented system, it is obviously unreasonable to attempt a comparison without taking viscosity into account in the system of equations. The modification only concerns the momentum equation which writes instead :

$$\frac{\partial hU}{\partial t}+\frac{\partial}{\partial x}\left(hU^2+\frac{h^2cos\theta}{2F^2}+\frac{2h^5}{25}+\frac{\kappa p^2}{2F^2}+\frac{\eta}{\alpha}\left(1-\frac{\eta}{h}\right)\right)=\frac{1}{\varepsilon Re}\left(\lambda h-\frac{3U}{h}\right)+\frac{9\varepsilon}{2Re}\frac{\partial}{\partial x}\left(h\frac{\partial U}{\partial x}\right) \tag{3.53}$$

From the numerical point of view, this additional diffusive term is solved by an explicit finite differences scheme given by :

$$\frac{\partial}{\partial x}\left(h\frac{\partial U}{\partial x}\right)\Big|_{x=x_i}^{t=t^n}\simeq\frac{1}{4\Delta x^2}((h_i^n+h_{i+1}^n)(U_{i+1}^n-U_i^n)-(h_i^n+h_{i-1}^n)(U_i^n-U_{i-1}^n)) \tag{3.54}$$

To summarize, the system to solve numerically writes fully :

$$\begin{aligned}
&\frac{\partial h}{\partial t}+\frac{\partial}{\partial x}(hU)=0,\\
&\frac{\partial hU}{\partial t}+\frac{\partial}{\partial x}\left(hU^2+\frac{h^2cos\theta}{2F^2}+\frac{2h^5}{25}+\frac{\kappa p^2}{2F^2}+\frac{\eta}{\alpha}\left(1-\frac{\eta}{h}\right)\right)=\frac{1}{\varepsilon Re}\left(\lambda h-\frac{3U}{h}\right)+\frac{9\varepsilon}{2Re}\frac{\partial}{\partial x}\left(h\frac{\partial U}{\partial x}\right),\\
&\frac{\partial h\eta}{\partial t}+\frac{\partial}{\partial x}(h\eta U)=hw,\\
&\frac{\partial hw}{\partial t}+\frac{\partial}{\partial x}\left(hwU-\frac{\kappa}{\beta F^2}p\right)=\frac{1}{\alpha\beta}\left(1-\frac{\eta}{h}\right),\\
&\frac{\partial p}{\partial t}+\frac{\partial}{\partial x}\left(pU-w\right)=0
\end{aligned}$$

Note that solving this system with the IMEX-222 scheme along the finite differences (3.54) for the additional viscous term, requires a stability condition of the form :

$$\max\left(\xi_m\frac{\Delta t}{\Delta x},\frac{9\varepsilon}{2Re}\frac{\Delta t}{\Delta x^2}\right)<1 \tag{3.55}$$

where $\xi_m$ is the maximum characteristic velocity. In order to simulate the experiments of Liu & Gollub, we consider the following initial condition :

$$h(x,0)=\eta(x,0)=U(x,0)=1,\quad w(x,0)=p(x,0)=0 \tag{3.56}$$

which corresponds to the Nusselt flow solution. The boundary conditions are given in this case by :

$$x=0:\begin{cases}h_0^n=1+0.1\sin(2\pi ft^n)\\u_0^n=1\\\eta_0^n=h_0^n\\w_0^n=0.2\pi f\cos(2\pi ft^n)\\p_0^n=(\eta_1^n-\eta_0^n)/\Delta x\end{cases}\qquad x=L:\begin{cases}h_{N+1}^n=h_N^n\\u_{N+1}^n=u_N^n\\\eta_{N+1}^n=\eta_N^n\\w_{N+1}^n=w_N^n\\p_{N+1}^n=p_N^n\end{cases} \tag{3.57}$$

The dimensionless frequency is given by $f = \tilde{f}L/\tilde{U}_N$, where $\tilde{f}$ is the imposed dimensioned frequency. In all the below presented numerical simulations, we take the following values of the relevant physical parameters:

| Fluid properties | | Reference quantities | | Dimensionless numbers | |
|---|---|---|---|---|---|
| $g$ | $9.81m/s^2$ | $\tilde{h}_N$ | $1.279mm$ | $Re$ | 19.33 |
| $\sigma$ | $0.067Kg/s^2$ | $\tilde{U}_N$ | $94.94mm/s$ | $F$ | 0.8476 |
| $\nu$ | $6.28.10^{-6}m^2/s$ | $L$ | $210.5mm$ | $\kappa$ | 0.000144 |
| $\rho$ | $1080Kg/m^3$ | $\theta$ | $6.4°$ | $\varepsilon$ | 0.006076 |

The augmented model parameters are taken equal to $\alpha = \varepsilon$ and $\beta = \varepsilon^3$. The numerical results are given below :

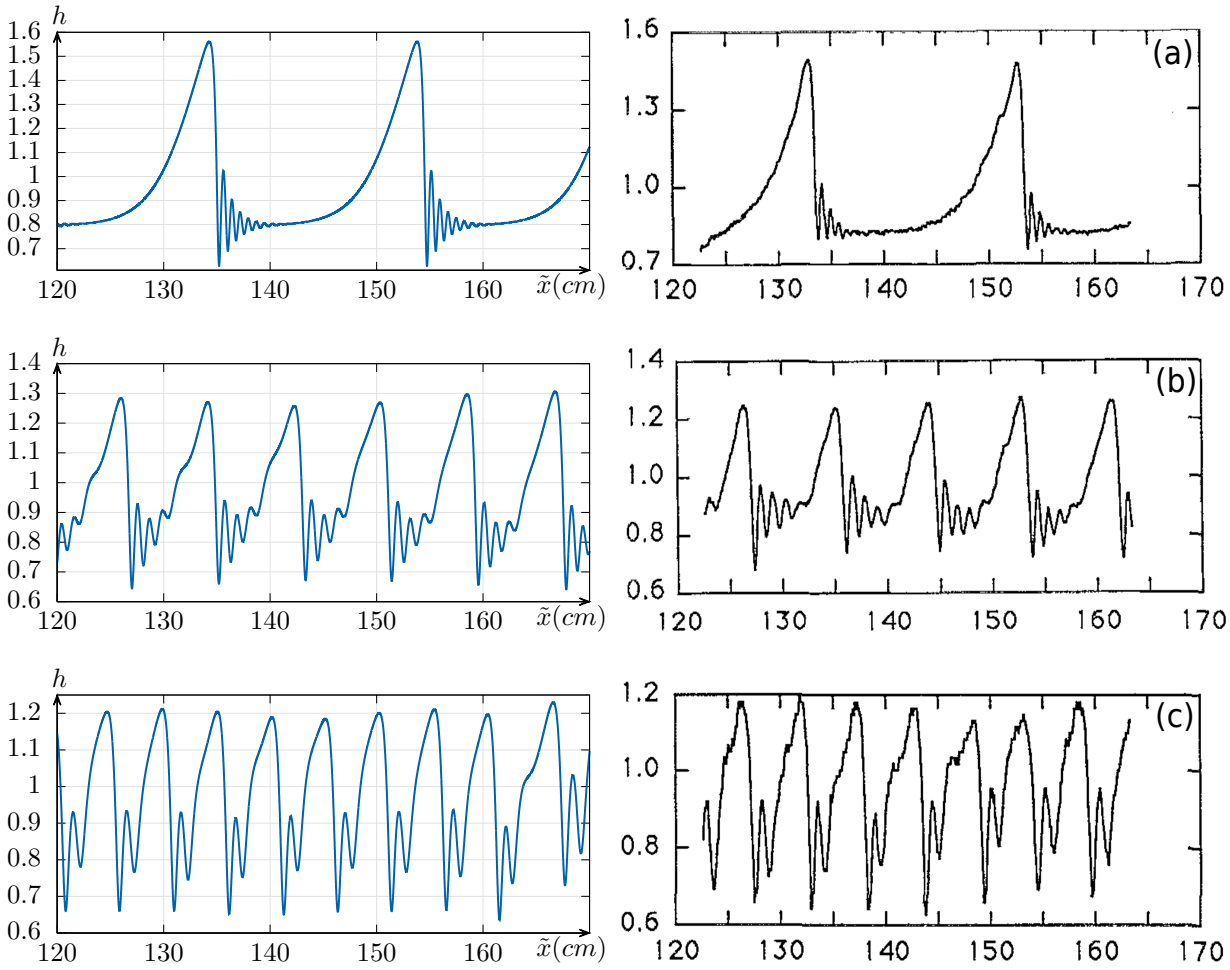


Figure 3.9: Experimental results of Liu & gollub [57] (right) along the obtained numerical results (left), for different forcing frequencies, respectively equal to $(a) : \tilde{f}_1 = 1.5Hz$, $(b) : \tilde{f}_2 = 3.0Hz$ and $(c) : \tilde{f} = 4.5Hz$. The number of mesh points is $n = 40000$. Experimental results figures are reprinted from J. Liu and J. P. Gollub. "Solitary wave dynamics of film flows". In: *Physics of Fluids* 6.5 (1994), pp. 1702–1712, with the permission of AIP Publishing.

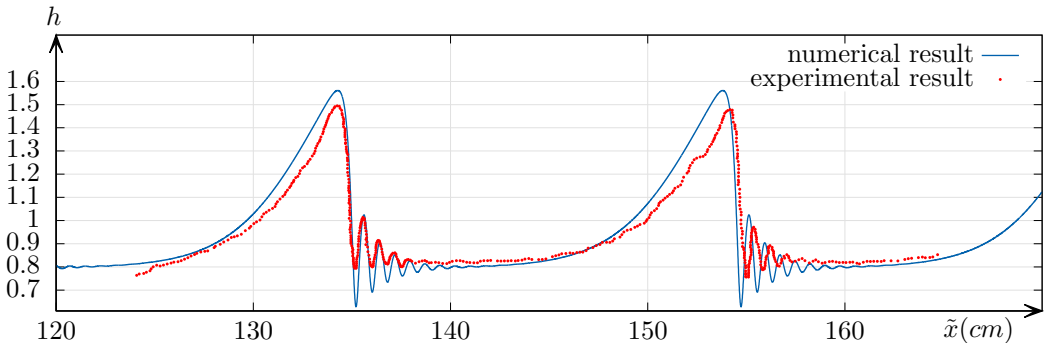


Figure 3.10: Superimposed numerical simulation with the experimental result for $\tilde{f} = 1.5Hz$.

**Remark 3.5.1.** *It may seem that the value* $Re = 19.33$ *used here is in discrepancy with the one used in [57] which is* $Re' = 29$*. The fact that* $Re' = 3Re/2$ *is due to the fact that they are defined with different reference velocities. In fact, we consider here as reference velocity the averaged equilibrium velocity over the depth* $\tilde{U}_N$*, while the reference value used in [57] is the value of the velocity at the water surface* $\tilde{u}_N(h_N) = 3\tilde{U}_N/2$*. (See equations* (3.3) *and* (3.2) *)*

For the different forcing frequencies, the numerically simulated waves display similar features to the experimental measures for the chosen values of $\alpha = \varepsilon$ and $\beta = \varepsilon^3$. The agreement is qualitatively very good. One can see on figure 3.10 that the amplitude and the wavelength of both the wave peak and the capillary ripples, are approximately well recovered through the augmented model for the chosen values of $\alpha$ and $\beta$.

## 3.6 Conclusion

The numerical results obtained through the augmented Lagrangian approach permitted to obtain good numerical results for examples involving thin film flows. The approach was shown to be easily adaptable to both models with linear and nonlinear surface tensions, as the structure of the obtained hyperbolic equations remains almost the same in both cases. This approach, in its current version, is unable to deal with flows involving dry or nearly dry regions. This is due to the fact that the maximum characteristic velocity tends to infinity when the fluid depth $h$ tends to 0. The treatment of dry zones will need significant efforts, on both the modeling and numerical aspects.

CHAPTER 4

# Stationary droplets on a solid substrate

## 4.1 Setting and assumptions

In this chapter, we consider a droplet of liquid, resting at rest on a plane horizontal solid substrate (see figure 4.1). We are interested in the stationary shape of the droplet submitted to its own weight and surface energy. In such a setting, it is reasonable to consider that the droplet is isotropic in the plane of the substrate so that we can restrict the analysis to one dimension of space. This assumption will make calculations more straightforward but not any less rigorous.

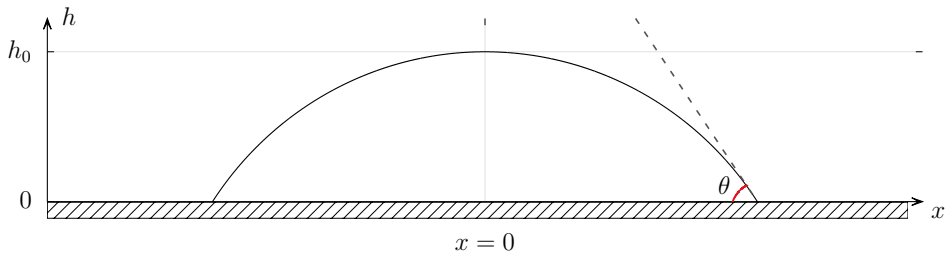


Figure 4.1: One-dimensional profile of the droplet.

In this setting, we are looking for an admissible droplet profile $h(x)$ where $h$ designates the local height of the droplet. If we assume that its center is located at $x = 0$, we can further narrow down the analysis to the domain $x \geq 0$, assuming that $h(x) = h(-x)$. In what follows, we denote by $h(x = 0) = h_0$, the central height and we assume that the edge of the droplet connects to the solid substrate ($h = 0$) forming a contact angle $\theta$ known a priori. We assume that the droplet spreads over a finite length equal to $2L$, which is unknown. Under these assumptions and notations, we have the following constraints at the boundaries

of the droplet :

$$x = 0 \begin{cases} h = h_0 \\ h_x = 0 \end{cases} \qquad x = L \begin{cases} h = 0 \\ h_x = -\tan(\theta) \end{cases} \tag{4.1}$$

where $h_x$ denotes the derivative of $h$ with respect to $x$. Finally, we consider that the total energy of our system writes :

$$E(h, h_x) = \frac{1}{2} g h^2 + \frac{\sigma}{\rho} \sqrt{1 + h_x^2} + P(h) \tag{4.2}$$

where $g$ is the gravity acceleration, $\sigma$ is the surface tension of the liquid, $\rho$ is the density of the liquid. $P(h)$, as introduced in [25] in general and following Derjaguin [27] is often called disjoining energy and is a sort of correction to the surface tension that should be such that $P(\infty) = 0$.

## 4.2 Equilibrium equations

In this part, we establish the differential equation satisfied by the profile $h(x)$ in the most generic case. In fact, for any energy that is dependent on $h$ and $h_x$, we can write Hamilton's action as :

$$a = \int_{-\infty}^{+\infty} E(h, h_x)\ dx. \tag{4.3}$$

The latter is submitted to the mass conservation constraint, which writes:

$$\delta h = -(h \delta x)_x \tag{4.4}$$

We now apply Hamilton's principle of stationary action :

$$\delta a = 0 \Longleftrightarrow \int_{-\infty}^{+\infty} (\delta E(h, h_x))\ \ dx = 0 \tag{4.5}$$

$$\Longleftrightarrow \int_{-\infty}^{+\infty} \left( \frac{\partial E}{\partial h} \delta h + \frac{\partial E}{\partial h_x} \delta h_x \right)\ \ dx = 0 \tag{4.6}$$

We plug in the mass conservation constraint and use Shwartz's theorem $\delta h_x = (\delta h)_x$ to obtain :

$$\int_{-\infty}^{+\infty} \left( -\frac{\partial E}{\partial h} (h \delta x)_x - \frac{\partial E}{\partial h_x} (h \delta x)_{xx} \right)\ \ dx = 0 \tag{4.7}$$

Integrating by parts gives:

$$\int_{-\infty}^{+\infty} \left( \frac{d}{dx} \left( \frac{\partial E}{\partial h} \right) h \delta x + \frac{d}{dx} \left( \frac{\partial E}{\partial h_x} \right) (h \delta x)_x \right)\ \ dx = 0 \tag{4.8}$$

We integrate again the second term by parts to obtain :

$$\int_{-\infty}^{+\infty} \left( \frac{d}{dx} \left( \frac{\partial E}{\partial h} - \frac{d}{dx} \left( \frac{\partial E}{\partial h_x} \right) \right) h \delta x \right)\ \ dx = 0\ \forall \delta x \tag{4.9}$$

Therefore the Euler-Lagrange equation writes:

$$h\frac{d}{dx}\left(\frac{\partial E}{\partial h}-\frac{d}{dx}\left(\frac{\partial E}{\partial h_x}\right)\right)=0 \tag{4.10}$$

Since we are looking for nontrivial solutions, we would like to investigate the case where $h(x)$ is not zero over a continuous region. Thus:

$$\frac{d}{dx}\left(\frac{\partial E}{\partial h}-\frac{d}{dx}\left(\frac{\partial E}{\partial h_x}\right)\right)=0 \tag{4.11}$$

Integrating this equation gives :

$$\frac{\partial E}{\partial h}-\frac{d}{dx}\left(\frac{\partial E}{\partial h_x}\right)=C \tag{4.12}$$

It is possible to integrate again the equation. In fact multiplying (4.12) by $h_x$ gives:

$$h_x\frac{\partial E}{\partial h}-h_x\frac{d}{dx}\left(\frac{\partial E}{\partial h_x}\right)=Ch_x \tag{4.13}$$

$$\Rightarrow \quad \frac{d}{dx}\left(E-h_x\frac{\partial E}{\partial h_x}\right)=Ch_x \tag{4.14}$$

So that finally we have :

$$E-h_x\frac{\partial E}{\partial h_x}=Ch+D \tag{4.15}$$

**Remark 4.2.1.** *Both $C$ and $D$ are constants of integration that are independent of $x$.*

**Remark 4.2.2.** *Equation (4.15) is compatible with the equilibrium state of the momentum equation in the Euler-Korteweg system, associated with the potential energy (4.2). Indeed, at equilibrium the momentum balance writes :*

$$\frac{\partial}{\partial x}(h_xE_{h_x}+hE_h-E-(hE_{h_x})_x)=0 \tag{4.16}$$

*Integrating once yields :*

$$h\frac{\partial E}{\partial h}-h\frac{d}{dx}\left(\frac{\partial E}{\partial h_x}\right)-E+h_x\frac{\partial E}{\partial h_x}=cst \tag{4.17}$$

*Which is exactly obtained by replacing $C$ in equation (4.15) by its expression given in equation (4.12).*

In our case of interest, that is when the energy is expressed as in equation (4.2), we obtain :

$$gh^2/2+P(h)+\frac{\sigma}{\rho\sqrt{1+h_x^2}}=Ch+D \tag{4.18}$$

which implies that $h(x)$ satisfies the differential equation :

$$h_x^2 = \left( \frac{\sigma/\rho}{Ch + D - gh^2/2 - P(h)} \right)^2 - 1. \tag{4.19}$$

In order to render the formulas less cumbersome, we denote by $Q(h)$ the function :

$$Q(h) = 1 + \frac{\rho}{\sigma} \left( \frac{gh^2}{2} + P(h) - (Ch + D) \right) \tag{4.20}$$

So that the differential equation (4.19) becomes :

$$h_x^2 = \left( \frac{1}{1 - Q(h)} \right)^2 - 1 \tag{4.21}$$

Naturally, since we are looking for real-valued solutions, it is necessary to have :

$$Q(h) \geq 0 \quad \forall h \in [0, h_0] \tag{4.22}$$

Depending on the setting which defines the properties of the function $Q(h)$, we can have different admissible droplet shapes. The regularity of the solution strongly depends on the allowed values of the contact angle $\theta$. If we were to consider *hydrophobic* substrates, which comes down to imposing $\theta > \pi/2$, then the droplet must bend at some point in order to form an obtuse angle in the vicinity of the substrate. In this case, proper description of the water height requires two profiles $h_1(x)$ and $h_2(x)$ separated by a singularity in which the derivative of the droplet height reaches infinity. This singular case will be considered later and we will focus first on smooth droplets.

## 4.3 Smooth droplet profiles

In a first attempt, we will restrict to contact angles $\theta < \pi/2$. Since we look for smooth solutions, we impose :

$$Q(h) < 1 \quad \forall h \in [0, h_0] \tag{4.23}$$

Taking into account the boundary conditions (4.1), and plugging them in equation (4.18) yields:

$$\begin{cases} P(0) + \dfrac{\sigma}{\rho\sqrt{1 + \tan(\theta)^2}} = D \\ Ch_0 + D = gh_0^2/2 + \dfrac{\sigma}{\rho} + P(h_0) \end{cases} \tag{4.24}$$

which allows us to obtain the expressions of $C$, $D$ and consequently $Q(h)$ as follows :

$$\begin{cases} D = P(0) + \dfrac{\sigma}{\rho}\cos(\theta) \\ C = gh_0/2 + \dfrac{\sigma}{\rho h_0}(1 - cos(\theta)) + \dfrac{P(h_0) - P(0)}{h_0} \\ Q(h) = \left( \dfrac{1 - cos(\theta)}{h_0} - \dfrac{\rho}{\sigma} \left( \dfrac{gh}{2} + \dfrac{P(h_0) - P(h)}{(h_0 - h)} - \dfrac{P(h_0) - P(0)}{h_0} \right) \right) (h_0 - h) \end{cases} \tag{4.25}$$

It is useful to note that the constraints (4.22) and (4.23) write explicitly in this case :

$$gh/2 + \frac{P(h_0) - P(h)}{h_0 - h} \leq \frac{\sigma}{\rho h_0}(1 - cos(\theta)) + \frac{P(h_0) - P(0)}{h_0} \quad \forall h \in [0, h_0] \tag{4.26}$$

$$gh/2 + \frac{P(h_0) - P(h)}{h_0 - h} > \frac{\sigma}{\rho h_0}(1 - cos(\theta)) - \frac{\sigma}{\rho(h_0 - h)} + \frac{P(h_0) - P(0)}{h_0} \quad \forall h \in [0, h_0] \tag{4.27}$$

These constraints are not local and do not have a simpler formulation in the general case as there are no restraints on the behavior of $P(h)$. However, in the few cases we will consider, it is possible to extract more practical necessary or sufficient conditions from them.

Now that all is set, we are interested in classifying the possible profile structures that are obtainable through solving equation (4.19) under the assumptions (4.1). Depending on the setting, *i.e.* the explicit form of $P(h)$ and the values of the physical parameters, there can be different smooth structures of the droplet we are seeking. Two examples are shown in figure (4.2) below :

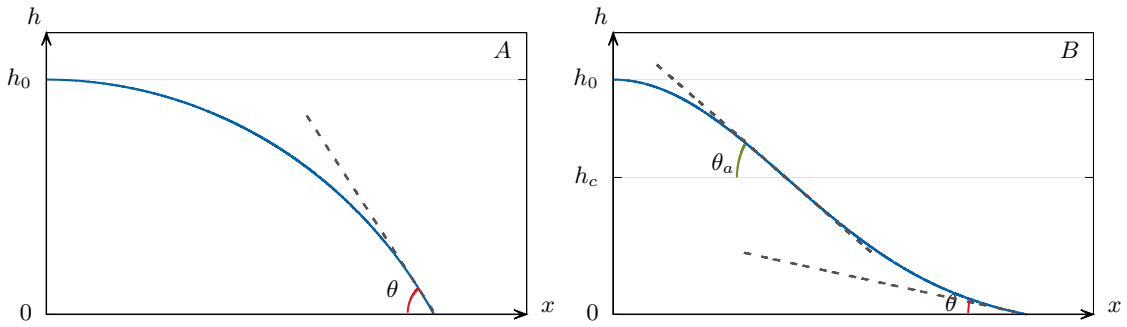


Figure 4.2: Sketches of the overall shape of some droplets. To the left : the droplet is completely concave. Its edge forms a contact angle with the solid substrate underneath. To the right, the droplet changes convexity for some value $h = h_c$ and flattens in the vicinity of the contact angle, forming what we may call a precursor film.

The main difference in both shown profiles is the change in convexity that appears in the rightmost graphic of figure 4.2. In this case, although the edge of the droplet still forms an angle $\theta$ with the solid substrate, and which is in agreement with the imposed constraints (4.1), there is meaning into considering a secondary angle, $\theta_a$, which appears at the inflection point location. Generally, the convex part of such a droplet lies in the close vicinity of the contact line ($h_c << h_0$). This means that, from the macroscopic point of view, it would be more reasonable to consider $\theta_a$ as the "apparent" contact angle. Following these remarks, in order to investigate whether any changes in convexity are likely to happen, it is important to have a look at the sign of $h_{xx}$ inside the interval $[0, h_0]$. The expression of the latter is obtained by deriving both sides of the differential equation (4.21) with respect to $x$ and simplifying by $h_x$. This allows us to write :

$$h_{xx} = -\frac{Q'(h)}{(1 - Q(h))^3} \tag{4.28}$$

Accordingly, the profile $h(x)$ admits as many inflection points as the number of local extrema of the function $Q(h)$ in the interval $[0, h_0]$. Therefore let us further analyze this function and its derivatives. First, let us recall its expression:

$$Q(h) = 1 + \frac{\rho}{\sigma}\left(\frac{gh^2}{2} + P(h) - (Ch + D)\right) \tag{4.29}$$

its first and second derivative are given by :

$$Q'(h) = \frac{\rho}{\sigma}\left(gh + P'(h) - C\right), \quad Q''(h) = \frac{\rho}{\sigma}(g + P''(h)) \tag{4.30}$$

Since conditions (4.1) impose that :

$$\begin{cases} h_x = 0 \text{ at } h = h_0 \\ h_x = -\tan(\theta) \leq 0 \text{ at } h = 0 \end{cases} \tag{4.31}$$

this results in:

$$Q(h_0) = 0 \quad \text{and} \quad Q(0) = 1 - \cos(\theta) \tag{4.32}$$

Since $Q(h) \geq 0$ and $Q(h_0) = 0$ then $Q(h_0)$ is a global minimum in $[0, h_0]$ and:

$$Q'(h_0) \leq 0 \tag{4.33}$$

As a consequence, for a strictly concave droplet, since $h_{xx}$ does not change curvature, $Q'(h)$ admits no roots and thus does not change its sign. Therefore, a necessary condition for a concave droplet is $Q'(0) < 0$. Otherwise, there exists at least one root of the equation $Q'(h) = 0$ inside the interval $]0, h_0[$, meaning there is at least an inflection point.

## First case : $P''(h) > 0$

In this case, the conclusions are rather straightforward. In fact, it implies that :

$$Q''(h) > 0 \quad \forall h \in [0, h_0] \tag{4.34}$$

Therefore, $Q'(h)$ is an increasing function over $[0, h_0]$. Since $Q'(h_0) < 0$ then $Q'(h) < 0 \; \forall h \in [0, h_0]$. Therefore, $h(x)$ admits no inflection points and only profile $A$ of figure 4.2 is admissible in this case.

## Second case : $P''(h) < 0$

This includes some explicit forms of $P(h)$ that are commonly used in practice. This situation allows for many possible scenarios, as there are no bounds on the number of roots of $Q'(h)$ as long as $P(h)$ is not explicitly given. A case of interest is when the equation $Q''(h) = 0$ admits at most one root in all the domain $[0, +\infty[$. Such is the case for instance, if $P''(h)$ is monotonic. We show in this particular setting that only the cases $A$ and $B$ displayed in figure 4.2 are admissible.
Indeed, if the equation $Q''(h) = 0$ admits no more than one root, then $Q'(h)$ admits at most two roots in $[0, h_0]$. Let us investigate the possible situations :

1. If $Q'(h)$ has no roots in $]0, h_0[$, then $h(x)$ is strictly concave (case A).
2. If $Q'(h)$ has one root $h_c \in [0, h_0]$, then $h(x)$ changes convexity once (case B).
3. Assume $Q'(h)$ has two roots in $]0, h_0[$ then $Q''(h)$ admits a root $h_s$ in $]0, h_0[$ (Rolle's theorem). Since there is one change of sign of $Q'(h)$ then $Q'(0) < 0$. Since we assumed that $Q''(h)$ does not admit any more than one root and that $Q''(\infty) = g > 0$, then this implies that $Q'(h)$ is decreasing in $[0, h_s]$ and increasing in $[h_s, h_0]$ and therefore $Q'(h) < 0 \ \forall h \in [0, h_0]$ which contradicts the assumption that $Q'(h)$ admits roots in $[0, h_0]$.

Therefore, if $P''(h)$ is monotonic and negative, then only admissible profiles of $h(x)$ are case $A$ and case $B$. In particular :

- If $Q'(0) \leq 0$ then the droplet is strictly concave.
- If $Q'(0) < 0$ then the droplet admits an inflection point.

**Remark 4.3.1.** *Since $P(\infty) = 0$ and $P(h)$ is concave in this case, it follows trivially that:*

$$P(h) < 0 \quad \forall h \in [0, h_0]; \quad P'(h) > 0 \quad \forall h \in [0, h_0] \tag{4.35}$$

*and that :*

$$\frac{P(h_0) - P(h)}{h_0 - h} \leq \frac{P(h_0) - P(0)}{h_0} \quad \forall h \in [0, h_0] \tag{4.36}$$

This permits to obtain more practical information from the constraints (4.26) and (4.27). Indeed this implies that :

1. A sufficient condition for the constraint (4.26) to be respected is :

$$gh/2 \leq \frac{\sigma}{\rho h_0}(1 - \cos(\theta)) \quad \forall h \in [0, h_0] \quad \Longleftrightarrow \quad h_0^2/2 \leq \frac{\sigma}{\rho g}(1 - \cos(\theta)) \tag{4.37}$$

   which can be cast into the form:

$$h_0 \leq 2 l_c \sin(\theta/2) \tag{4.38}$$

   where $l_c$ is the capillary length given by $l_c = \sqrt{\frac{\sigma}{\rho g}}$

2. A necessary condition for the constraint (4.27) to be respected is :

$$\rho g h/2 + \frac{\sigma}{h_0 - h} > \frac{\sigma}{h_0}(1 - \cos(\theta)) \quad \forall h \in [0, h_0] \tag{4.39}$$

   Since the left-hand side of the inequality is an increasing function of $h$, then it is enough to check the inequality for $h = 0$ which gives :

$$\cos(\theta) > 0 \tag{4.40}$$

   Which is a natural consequence of a smooth droplet assumption, since it does not allow for $\theta = \pi/2$.

## 4.4 Explicit example: $P(h) = -A/(h + h_\star)$

This form of $P(h)$ satisfies the assumptions $P''(h) > 0$ and $P(\infty) = 0$. Its behavior is also discussed in [25] (case(b) on p.92). Out of concave functions reaching to 0 in infinity with a finite value for $h = 0$, the one we choose here may be the simplest in terms of calculations. It is also possible to consider the exponential function, which qualitatively delivers the same behavior. Without further ado, let us detail some of this potential's properties.

$$P(h) = -\frac{A}{h + h_\star} \tag{4.41}$$

where $A$ and $h_\star < h_0$ are positive constants. It follows from it that :

$$P'(h) = \frac{A}{(h + h_\star)^2}; \quad P''(h) = -\frac{2A}{(h + h_\star)^3} < 0 \tag{4.42}$$

$$P(0) = -\frac{A}{h_\star}; \quad P'(0) = \frac{A}{h_\star^2} > 0 \tag{4.43}$$

The graph of $P(h)$ and its derivative are given in the figure 4.3 below:

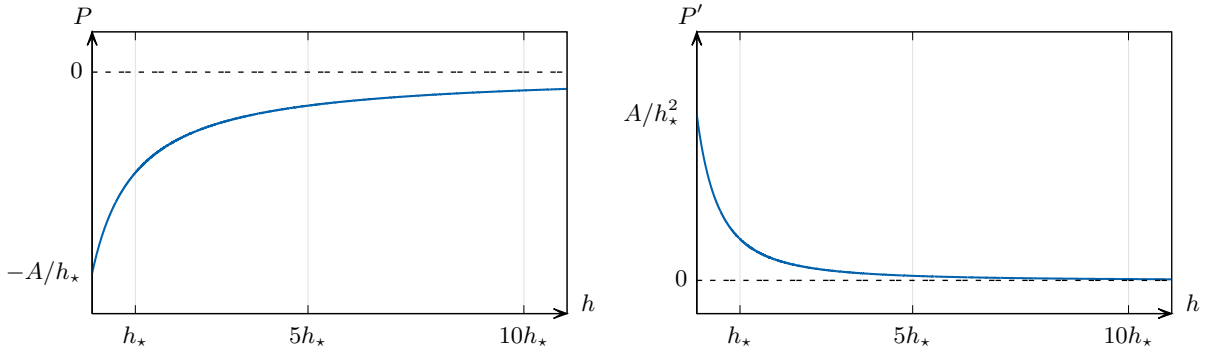


Figure 4.3: sketch of the profiles of $P(h)$ (left) and $P'(h)$ (right) as a function of h.

Replacing $P(h)$ in the expressions of $Q'(h)$ and $Q''(h)$ yields :

$$Q'(h) = \frac{\rho}{\sigma}\left(gh + \frac{A}{(h + h_\star)^2} - C\right), \tag{4.44}$$

$$Q''(h) = \frac{\rho}{\sigma}(g - \frac{2A}{(h + h_\star)^3}) \tag{4.45}$$

In this case, we can clearly see that $Q''(h)$ is a strictly increasing function. As a matter of fact, it follows from the analysis done in the previous section that the profile of the droplet only depends on the sign of $Q'(0)$. Let us consider the dimensionless constant $K = 2A/gh_\star^3$. The equality $Q'(0) < 0$ can be cast into the form :

$$K > K_c, \quad \text{where} \quad K_c = \left(1 + 4\frac{l_c^2}{h_0^2}\sin^2(\theta/2)\right)(1 + h_0/h_\star) \tag{4.46}$$

If this equality holds, then $h(x)$ admits an inflection point. Otherwise it is strictly concave. By solving numerically the differential equation (4.21), for different values of the parameter $K$ (varying $A$ at fixed $h^\star$) we can plot the following figure :

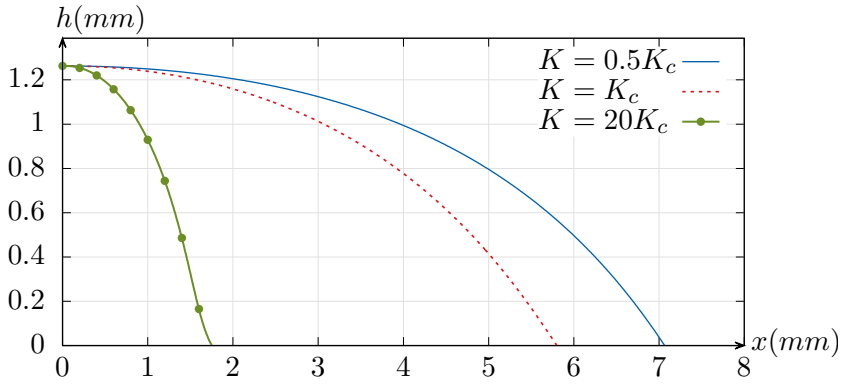


Figure 4.4: Different drop shapes for different values of $K$, obtained through numerically solving equation (4.21). Parameters used here are : $g = 9.8ms^{-2}$, $\sigma = 0.072Kg.s^{-2}$, $h_0 = 1.262mm$, $h_\star = 0.252mm$ and $\theta = 30°$.

While, theoretically, a change of convexity appears for any $K > K_c$, it becomes only discernible at sufficiently high values of $K$. For values of $K$ slightly higher than $K_c$, the inflection point remains in the extremity of the droplet edge. It is also possible to fix $A$ and then vary the contact angle instead. We obtain the following figure :

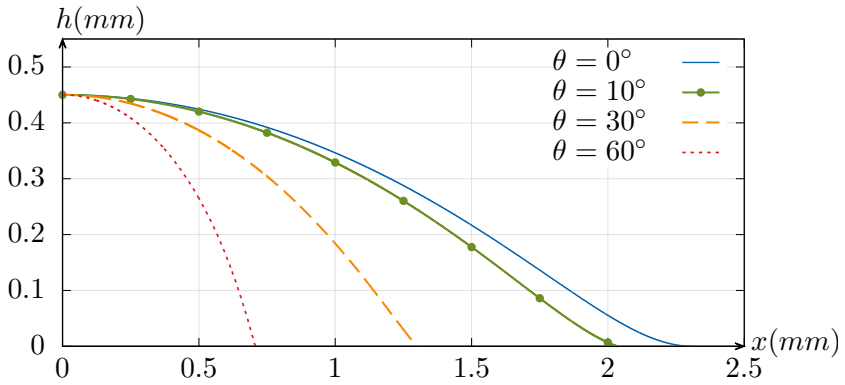


Figure 4.5: Different drop shapes for different values of $\theta$, obtained through numerically solving equation (4.21). Parameters used here are : $g = 9.8ms^{-2}$, $\sigma = 0.072Kg.s^{-2}$, $h_0 = 0.45mm$, $h_\star = 0.09mm$ and $A = 10^{-9}m^4s^{-2}$.

## 4.5 Remarks on droplets with singularities

Before moving on to testing these results under the augmented formulation, we would like to give a few brief remarks on some cases with singularities, with points where $h_x$ reaches

an infinite value. Clearly, by looking at the differential equation (4.21), such a singularity occurs if there exists $h_s \in ]0, h_0[$ such that $Q(h_s) = 1$. This equality writes :

$$\frac{gh_s^2}{2} + P(h_s) = Ch_s + D \tag{4.47}$$

We will disregard the case where multiple roots of this equation exist and focus on the case where only one root $h_s$ exists. The curve $h_x^2 = f(h)$ in this case is plotted in the following figure :

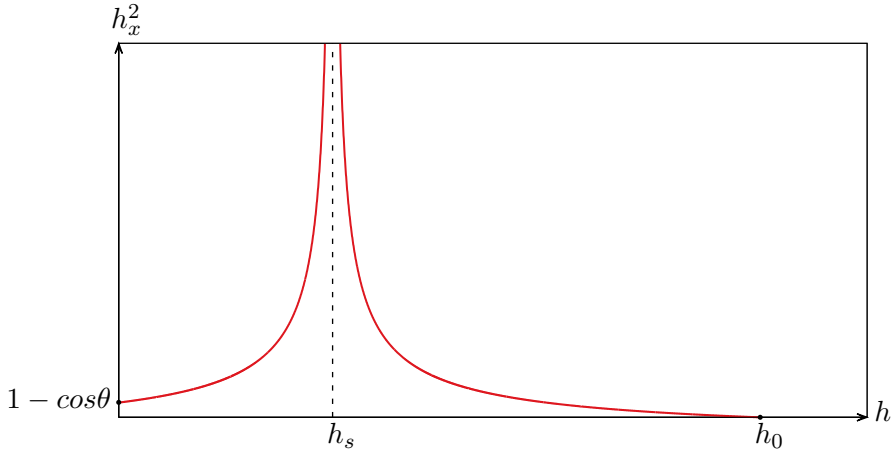


Figure 4.6: graphic of $h_x^2$ as a function of $h$ in the case of a singularity

Such a setting suggests two possible phase portraits which are displayed below :

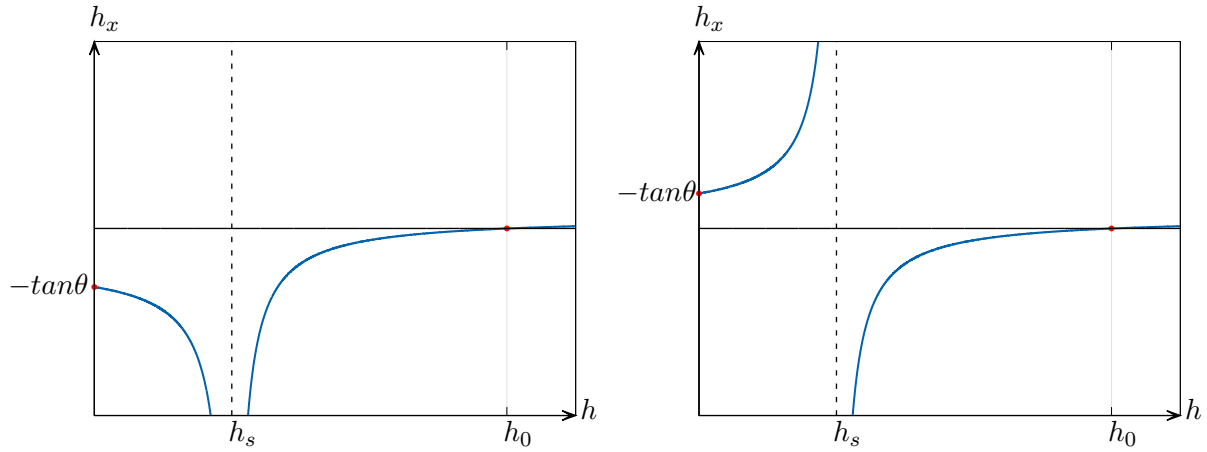


Figure 4.7: Admissible phase portraits in the case of a single root $h_s$. The left phase portrait corresponds to a contact angle $\theta < \pi/2$ with a singularity in the middle. The right phase portrait corresponds to a contact angle $\theta > \pi/2$.

As shown in figure 4.7, there are two admissible phase portraits. The left graphic corresponds to a droplet on a hydrophilic substrate, given the sign of $\tan\theta$ and the sign of $h_x$ in

the region $[0, h_0]$. This means that after the singularity, there is a change of convexity and the droplet edge connects with solid substrate, similarly as for a regular droplet with an inflection point, forming a contact angle $\theta < \pi/2$ (see figure 4.9). The second possibility, which is in the right part of figure 4.7, corresponds to a droplet on a hydrophobic solid substrate. The singularity that occurs at $h = h_s$ corresponds to the turnaround point so that the droplet can bend over to form an angle $\theta > \pi/2$ (see figure 4.8). In these circumstances, the sign of $\cos\theta$ is undetermined. The initially chosen contact angle imposes which branch to take in this case. The constants $C$ and $D$ write more generally :

$$\begin{cases} D = P(0) + \dfrac{\sigma}{\rho}\,|\cos(\theta)| \\ C = gh_0/2 + \dfrac{\sigma}{\rho h_0}(1 - |\cos(\theta)|) + \dfrac{P(h_0) - P(0)}{h_0} \end{cases} \tag{4.48}$$

so that equation (4.47) explicitly writes :

$$gh_s/2 + \frac{\sigma}{\rho h_0}\left(|\cos\theta| + \frac{h_s}{h_0 - h_s}\right) = \frac{P(h_0) - P(0)}{h_0} - \frac{P(h_0) - P(h_s)}{h_0 - h_s} \tag{4.49}$$

The left-hand side of this equality is positive and so must be the right-hand side. This case is clearly incompatible with convex forms of $P(h)$ since the considered slope difference is always negative. For a concave $P(h)$ we may have roots $h_s$. For example, for the explicit case of $P(h)$ we have used previously, this equation writes explicitly :

$$gh_s/2 + \frac{\sigma}{\rho h_0}\left(|\cos\theta| + \frac{h_s}{h_0 - h_s}\right) = \frac{Ah_s}{h_\star(h_0 + h_\star)(h_s + h_\star)} \tag{4.50}$$

Or in dimensionless form :

$$\frac{1}{2} + \frac{l_c^2}{h_0 h_s}\left(|\cos\theta| + \frac{h_s}{h_0 - h_s}\right) = \frac{K}{(1 + h_0/h_\star)(1 + h_s/h_\star)} \tag{4.51}$$

In the case where this equation admits exactly one root, as discussed earlier, there are two possible outcomes. While the governing differential equation is independent of the sign of $\cos\theta$, what makes the difference is the boundary condition $h_x = -\tan\theta$ at $x = 0$. If $\theta > \pi/2$, and if we note by $x_s$ the space coordinate of the singularity, we need to solve :

$$\begin{cases} h_x = -\sqrt{\left(\frac{1}{1-Q(h)}\right)^2 - 1}\ \forall x \in [0, x_s[ \\ h(0) = h_0 \end{cases} \qquad \begin{cases} h_x = \sqrt{\left(\frac{1}{1-Q(h)}\right)^2 - 1}\ \forall x \in [L, x_s[ \\ h(x_s) = h_s \end{cases} \tag{4.52}$$

Solving successively these equations yields the following profile :

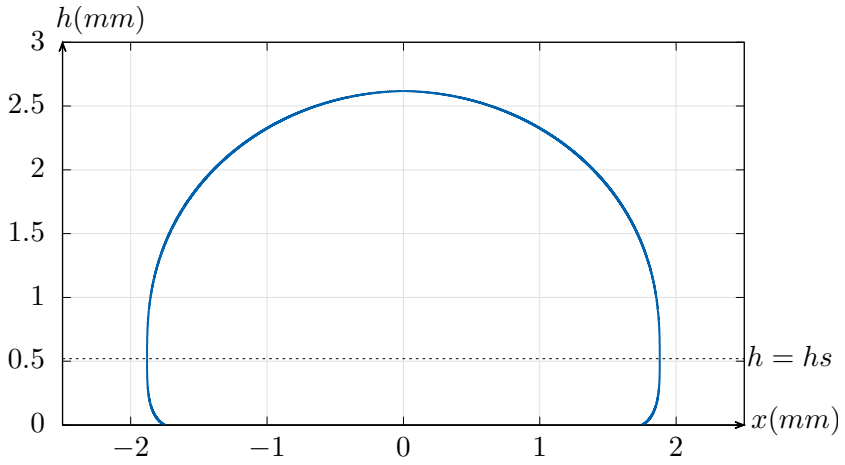


Figure 4.8: Full shape of a droplet for $\theta = 150°$, obtained through the above numerical algorithm. Parameters used here are : $g = 9.8ms^{-2}$, $\sigma = 0.072Kg.s^{-2}$, $h_0 = 2.61mm$, $h_\star = 0.261mm$, $A = 3.761.10^{-8}m^4s^{-2}$ and $\Delta x = 80nm$.

The measured contact angle complies with the imposed value. For the same configuration but the contact angle, which we take as $\theta = \pi/6$, in order to obtain the corresponding profile we need to solve :

$$\begin{cases} h_x = -\sqrt{\left(\frac{1}{1-Q(h)}\right)^2 - 1} \ \forall x \in [0, x_s[ \\ h(0) = h_0 \end{cases} \qquad \begin{cases} h_x = -\sqrt{\left(\frac{1}{1-Q(h)}\right)^2 - 1} \ \forall x \in ]x_s, L] \\ h(x_s) = h_s \end{cases} \tag{4.53}$$

We obtain the following profile :

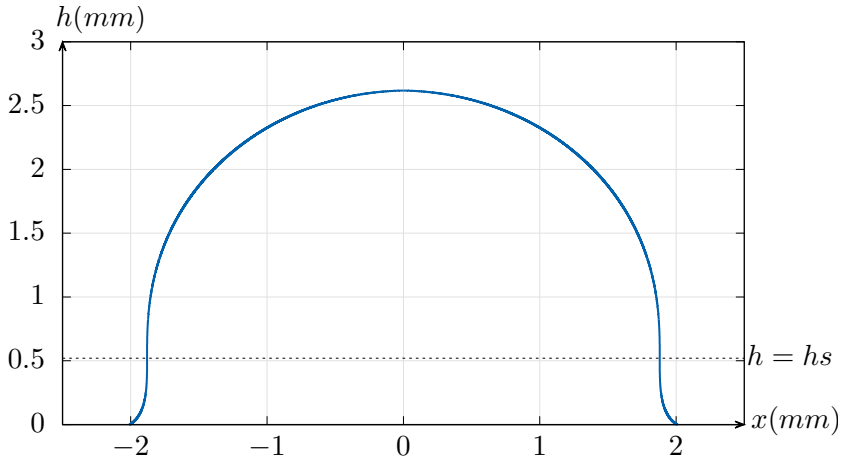


Figure 4.9: Full shape of a droplet for $\theta = 30°$, obtained through numerical simulation. Parameters used here are : $g = 9.8ms^{-2}$, $\sigma = 0.072Kg.s^{-2}$, $h_0 = 2.61mm$, $h_\star = 0.261mm$, $A = 3.761.10^{-8}m^4s^{-2}$ and $\Delta x = 80nm$.

**Remark 4.5.1.** *From the numerical point of view, $h_s$ is obtained through solving equation* (4.51) *numerically. The first differential is solved until $h$ reaches $h_s$. The second equation is solved until $h$ reaches* 0. *So there is no need to compute $x_s$ and $L$ beforehand.*

## 4.6 Augmented model analysis

In this section, we would like to extend the results of the previous section to the augmented model. This requires us to properly define the equations governing the stationary state in this case. The main difference here is that the energy of the augmented system does not depend on $h_x$ anymore, but depends on the unconstrained variable $\eta$ and its derivative $\eta_x$ instead. It writes :

$$E(h,\eta,\eta_x) = gh^2/2 + \frac{h}{2\alpha}\left(1-\frac{\eta}{h}\right)^2 + \frac{\sigma}{\rho}\sqrt{1+\eta_x^2} + P(h) \tag{4.54}$$

Given that only $h$ is constrained through the mass conservation law, it is straightforward to prove that the governing equations are consequent from :

$$\begin{cases} \dfrac{d}{dx}\left(\dfrac{\partial E}{\partial h}\right) = 0 \Longleftrightarrow \dfrac{\partial E}{\partial h} = C_\alpha \\ \dfrac{\partial E}{\partial \eta} - \dfrac{d}{dx}\left(\dfrac{\partial E}{\partial \eta_x}\right) = 0 \end{cases} \tag{4.55}$$

Since we are looking for non-trivial solutions, that is $\eta(x)$ different from the zero function, we multiply the second equation by $\eta_x$ :

$$\eta_x \frac{\partial E}{\partial \eta} - \eta_x \frac{d}{dx}\left(\frac{\partial E}{\partial \eta_x}\right) = 0$$

Then using the identity :

$$\frac{\partial E}{\partial \eta}\eta_x = \frac{dE}{dx} - \frac{\partial E}{\partial h}h_x - \frac{\partial E}{\partial \eta_x}\eta_{xx} \tag{4.56}$$

which implies :

$$\frac{dE}{dx} - \frac{\partial E}{\partial h}h_x - \frac{\partial E}{\partial \eta_x}\eta_{xx} - \eta_x \frac{d}{dx}\left(\frac{\partial E}{\partial \eta_x}\right) = 0 \tag{4.57}$$

Thus we can write :

$$\frac{dE}{dx} - C_\alpha h_x - \frac{d}{dx}\left(\eta_x \frac{\partial E}{\partial \eta_x}\right) = 0 \tag{4.58}$$

Integrating the equation with respect to $x$ finally gives the system :

$$\begin{cases} \dfrac{\partial E}{\partial h} = C_\alpha \\ E - \eta_x \dfrac{\partial E}{\partial \eta_x} = C_\alpha h + D_\alpha \end{cases} \tag{4.59}$$

**Remark 4.6.1.** *This system of equation is not specific to the energy form* (4.54) *but to any energy that depends explicitly on* $(h,\eta,\eta_x)$ *where* $h$ *is constrained by the mass conservation law and* $\eta$ *is a free variable. In particular, it can be used regardless of whether the surface tension term is linearized or not.*

**Remark 4.6.2.** *This final system is consistent with the Euler-Lagrange equations* (4.55) *as well as with the stationary state of the augmented system conservation laws. The main advantage of the provided form is that it only requires to solve a first order differential equation instead of a second order one which is more convenient in practice.*

For the energy (4.54) we obtain the system of equations:

$$\begin{cases} gh + P'(h) + \dfrac{1}{2\alpha}\left(1 - \dfrac{\eta}{h}\right)\left(1 - \dfrac{2\eta}{h}\right) = C_\alpha & (4.60) \\ \dfrac{\sigma}{\rho}\dfrac{1}{\sqrt{1+\eta_x^2}} = C_\alpha h + D_\alpha - gh^2/2 - \dfrac{h}{2\alpha}\left(1 - \dfrac{\eta}{h}\right)^2 - P(h) & (4.61) \end{cases}$$

Now, let us establish the differential equation satisfied by $h$ in this case. Multiplying equation (4.60) by $2\alpha h^2$ and simplifying yields :

$$2\eta^2 - 3h\eta + h^2(1 + 2\alpha(gh + P'(h) - C_\alpha)) = 0 \tag{4.62}$$

which, solved in terms of $\eta$ yields the roots :

$$\eta_\pm = \frac{h}{4}\left(3 \pm \sqrt{1 - 8\alpha(gh + P'(h) - C_\alpha)}\right) \tag{4.63}$$

Clearly, the root of interest is the one that permits to recover $\eta \to h$ when $\alpha \to 0$. Thus we take :

$$\eta = \frac{h}{4}\left(3 + \sqrt{1 - 8\alpha(gh + P'(h) - C_\alpha)}\right) \tag{4.64}$$

We derive the latter with respect to $x$ to obtain :

$$\eta_x = f_\alpha(h)h_x \tag{4.65}$$

where $f_\alpha(h)$ is given by :

$$f_\alpha(h) = \frac{1 + 4\alpha(2C_\alpha - 3gh - 2P'(h) - hP''(h)) + 3\sqrt{1 - 8\alpha(gh + P'(h) - C_\alpha)}}{4\sqrt{1 - 8\alpha(gh + P'(h) - C_\alpha)}} \tag{4.66}$$

Lastly, equation (4.64) also permits to write :

$$1 - \frac{\eta}{h} = 2\alpha\varphi_\alpha(h); \quad \text{where} \quad \varphi_\alpha(h) = \frac{(gh + P'(h) - C_\alpha)}{\left(1 + \sqrt{1 - 8\alpha(gh + P'(h) - C_\alpha)}\right)} \tag{4.67}$$

Thus, putting equation (4.61) into a more convenient form and substituting $\eta$ and $\eta_x$ via the above formulas yields the differential equation :

$$h_x^2 = \frac{1}{(f_\alpha(h))^2}\left(\left(\frac{\sigma/\rho}{C_\alpha h + D_\alpha - gh^2/2 - 2\alpha h(\varphi_\alpha(h))^2 - P(h)}\right)^2 - 1\right) \tag{4.68}$$

Before solving this system of equations, it remains to define the values of $C_\alpha$ and $D_\alpha$ through imposing suitable boundary values. We use the same boundary conditions as in the original case, that is:

$$x = 0 \begin{cases} h = h_0 \\ h_x = 0 \end{cases} \qquad x = L \begin{cases} h = 0 \\ h_x = -\tan(\theta) \end{cases} \tag{4.69}$$

We recall that the value $x = L$ is unknown, that is we only impose $h_x = -\tan(\theta)$ when the droplet height reaches 0. Under these assumptions, the constants $C_\alpha$ and $D_\alpha$ can be obtained as follows. Equation (4.61) can be cast into the form :

$$\frac{\sigma}{\rho}\frac{1}{\sqrt{1+(f_\alpha(h)h_x)^2}} = C_\alpha h + D_\alpha - gh^2/2 - 2\alpha h(\varphi_\alpha(h))^2 - P(h) \tag{4.70}$$

which allows us to compute the limit $h \to 0$:

$$D_\alpha = P(0) + \frac{\sigma}{\rho\sqrt{1+(f_\alpha(0)\tan(\theta))^2}} \tag{4.71}$$

As $f_\alpha(0)$ is dependent on $C_\alpha$, the above expression links $D_\alpha$ and $C_\alpha$. It only remains to calculate $C_\alpha$, by considering equation (4.68) in the point $x = 0$. It follows that :

$$C_\alpha + \frac{\sigma/(\rho h_0)}{\sqrt{1+(f_\alpha(0)\tan(\theta))^2}} - \frac{2\alpha(gh_0 + P'(h_0) - C_\alpha)^2}{\left(1+\sqrt{1-8\alpha(gh_0+P'(h_0)-C_\alpha)}\right)^2} = \frac{\sigma}{\rho h_0} + gh_0/2 + \frac{P(h_0)-P(0)}{h_0} \tag{4.72}$$

This equation is solved numerically to obtain the value of $C_\alpha$ and consequently $D_\alpha$. One can see that in the limit $\alpha \to 0$ we recover the same values as for the original system :

$$\begin{cases} \lim\limits_{\alpha\to 0} C_\alpha = \dfrac{\sigma}{\rho h_0}(1-\cos\theta) + gh_0/2 + \dfrac{P(h_0)-P(0)}{h_0} = C & (4.73) \\ \lim\limits_{\alpha\to 0} D_\alpha = P(0) + \dfrac{\sigma}{\rho}\cos\theta = D & (4.74) \end{cases}$$

We can show that the convergence rate is linear in $\alpha$ as shown in figure 4.10:

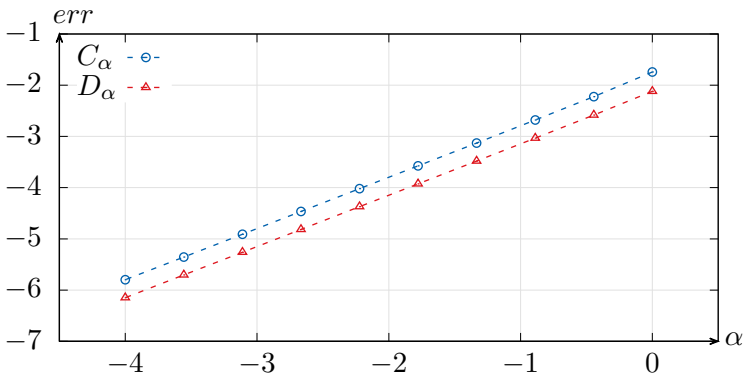


Figure 4.10: Convergence rates of $C_\alpha$ and $D_\alpha$ towards $C$ and $D$ respectively. The plots represent the relative errors $\left|\frac{C_\alpha - C}{C}\right|$ (blue circular points) and $\left|\frac{D_\alpha - D}{D}\right|$ (red triangular points), as a function of $\alpha$ in a log-log representation. the measured slopes are respectively $r_1 = 1.009$ and $r_2 = 1.005$ for $C_\alpha$ and $D_\alpha$.

Finally, by taking the Cauchy problem

$$\begin{cases} h_x^2 = \dfrac{1}{(f_\alpha(h))^2}\left(\left(\dfrac{\sigma/\rho}{C_\alpha h + D_\alpha - gh^2/2 - 2\alpha h(\varphi_\alpha(h))^2 - P(h)}\right)^2 - 1\right) \\ h(x=0) = h_0 \end{cases} \tag{4.75}$$

we can see that in the limit $\alpha \to 0$, we recover exactly its equivalent counterpart for the original system of equations. The numerical resolution of this differential equation, for the same form of $P(h)$ introduced in the previous section, yields the following results :

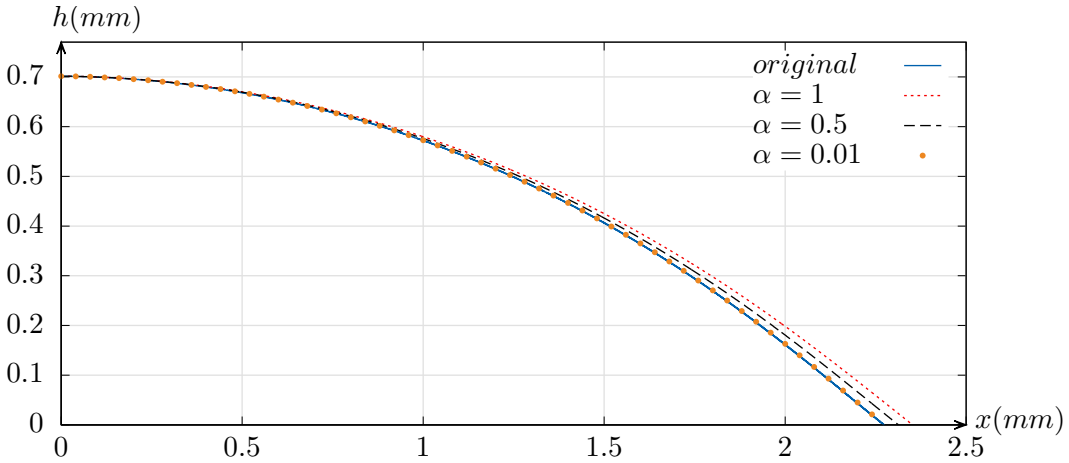


Figure 4.11: Comparison of the overall shape of the droplet for the augmented system (dashed/dotted lines) and the reference model (blue continuous line) for several values of $\alpha$. Parameters used here are $g = 9.8ms^{-2}$, $\sigma = 0.072Kg.s^{-2}$, $\theta = 30°$, $h_0 = 0.7mm$, $h_\star = 0.14mm$ and $A = 10^{-9}m^4s^{-2}$. The mesh size is $\Delta x = 0.1\mu m$.

The mesh size is taken small enough as not to pollute the measures of the contact angle through a numerical derivation of $h$. Clearly, for $\alpha = 0.01$, the curves coincide already. When comparing for higher values of $\alpha$ ,the most significant dissimilarity lies in the vicinity of the contact angle. The angle itself seems to be conserved independently of $\alpha$. The length $L$ is also different and seems to be an increasing function of $\alpha$, that is the higher $\alpha$ is, the wider the drop spreads.

## 4.7 Concluding remarks

We have addressed in this part droplet profiles which were derived through Hamilton's principle of stationary action. We restricted ourselves to some specific assumptions on $P(h)$ which were fully addressed and some criteria that may help classify the droplets' shape were established in a some cases. Although the droplet structures that were discussed here are mathematically admissible and compatible with Hamilton's principle, this does infer in any way that they are mechanically stable configurations, in which case their existence in nature would be also likely. This question shall be addressed in due time. This investigation of

droplets also provided a good setting of comparison for the augmented model in a stationary case, in which it is possible to obtain an explicit expression of $\eta$ as a function of $h$. The comparison of the resulting differential equations, allowed for a better understanding of the penalty method and how close it approximates the original system.

Besides, we have chosen here to impose the central height of the droplet along the contact angle. The resulting profiles, correspond to droplets with different masses. Thus, it would be also interesting to conduct a similar study in which the mass and the contact angle are fixed beforehand for instance, which gives different values of $h_0$ as a consequence. This could give another perspective to the problem.

CHAPTER 5

# On the stability of modified equations for linear schemes

Modified equations have been the subject of investigations and debates in numerical analysis. And yet, despite being relatively easy to establish, their use has been regarded with a lot of skepticism due to the lack of theoretical justification and the trickiness of their analysis. The first use of this technique for stability purposes is due to Hirt [45]. He provided the first practical examples for which modified equations give relevant information, in a heuristic manner. The pionneering work of Warming and Hyett [73] introduced how to obtain these equations in the general case and mostly clarified the link between the stability of the scheme and the modified equations. They provided a simple and efficient way to obtain the exact Von Neumann stability conditions from a finite number of the modified equation coefficients, for linear scalar schemes. An alternate way to obtain modified equations was introduced in [15]. The method permits to obtain the same modified equation through a series expansion of a an explicitly known function rather than the elimination technique of [73]. Other works that tried to tackle stability analysis through modified equations include [56, 71].

In this independent chapter, we investigate the modified equations for some finite difference schemes, in an attempt to make it clear why this technique often fails to provide relevant information on stability. First, we present a reminder on how to obtain modified equations for a given linear scalar scheme. We explain the basics of the heuristic stability theory and its limitations through some examples. In the second part, we present the technical framework and tools needed for stability analysis. We then clarify the mathematical reasons behind the frequent failures of the heuristic stability theory by investigating the convergence of the Fourier transform of the modified equation. We show that the latter is conditionally convergent and only then does it give significant results. Then, we compare the stability conditions of truncated modified equations with the corresponding scheme in the region of convergence. Finally, we present some examples to justify our analysis.

## 5.1 On modified equations and heuristic stability

### 5.1.1 Obtaining the equations

Consider as an example, the linear scalar transport equation given by :

$$\frac{\partial \tilde{u}}{\partial t} + c\frac{\partial \tilde{u}}{\partial x} = 0 \tag{5.1}$$

with positive velocity $c$. In order to solve this equation with a finite difference scheme, we first introduce a uniform grid of points defined as usual by $(x_j = j\Delta x, t^n = n\Delta t)$ and we denote by $u_j^n = u(x_j, t^n)$ the value of the numerical solution in the corresponding grid point. Under these notations, take for instance the upwind Euler scheme for equation (5.1) :

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + c\frac{u_j^n - u_{j-1}^n}{\Delta x} = 0 \tag{5.2}$$

Generally, in order to get relevant information on the consistency of the scheme [66] or its convergence rate, we assume the existence of a smooth, infinitely differentiable numerical solution $u(x,t)$ that satisfies $u(x_j, t^n) = u_j^n$ in each grid point. Provided this solution, we expand each term of the scheme in Taylor series in the vicinity of $(x_j, t^n)$, for instance :

$$u_j^{n+1} = u(x_i, t^{n+1}) = u(x_j, t^n + \Delta t) = u_j^n + \Delta t\frac{\partial u_j^n}{\partial t} + \frac{\Delta t^2}{2}\frac{\partial^2 u_j^n}{\partial t^2} + \dots \tag{5.3}$$

$$u_{j-1}^n = u(x_{i-1}, t^n) = u(x_j - \Delta x, t^n) = u_j^n - \Delta x\frac{\partial u_j^n}{\partial x} + \frac{\Delta x^2}{2}\frac{\partial^2 u_j^n}{\partial x^2} + \dots \tag{5.4}$$

which leads to an equation satisfied by $u$ containing an infinite number of partial derivatives :

$$\frac{\partial u}{\partial t} + \frac{\Delta t}{2}\frac{\partial^2 u}{\partial t^2} + \frac{\Delta t^2}{6}\frac{\partial^3 u}{\partial t^3} + \dots = -c\frac{\partial u}{\partial x} + \frac{c\Delta x}{2}\frac{\partial^2 u}{\partial x^2} + \frac{c\Delta x^2}{6}\frac{\partial^3 u}{\partial x^3} + \dots \tag{5.5}$$

This equation as is, is sufficient to prove consistency. In fact, we can clearly see that the in the limit $\Delta t \to 0$ and $\Delta x \to 0$ we recover the original transport equation. However, for further analysis of the numerical effects induced by the scheme, it would be more intuitive to consider an evolution equation with only space derivatives. This would deliver an evolution equations that is more amenable to physical interpretation. To do that, we use the elimination procedure introduced by Warming and Hyett in [73], that is we repeatedly use linear combinations of the equation (5.5) and its derivatives in order to eliminate time derivatives of order higher than one. Let us demonstrate how this works. First we derive (5.5) with respect to $t$ and multiply it by $\Delta t/2$ to obtain :

$$\frac{\Delta t}{2}\frac{\partial^2 u}{\partial t^2} = -\frac{\Delta t^2}{4}\frac{\partial^3 u}{\partial t^3} - \frac{\Delta t^3}{12}\frac{\partial^4 u}{\partial t^4} - c\frac{\Delta t}{2}\frac{\partial u^2}{\partial x \partial t} + \frac{c^2\Delta x\Delta t}{4}\frac{\partial^3 u}{\partial x^2 \partial t} + \frac{c^2\Delta x^2\Delta t}{12}\frac{\partial^4 u}{\partial x^3 \partial t} + \dots \tag{5.6}$$

replacing this expression in the (5.5) and truncating to the same order of consistency only does half of its expected job, as it eliminates the second derivative in time but there appeared

instead a mixed derivative with respect to $x$ and $t$ :

$$\frac{\partial u}{\partial t} - \frac{\Delta t^2}{4}\frac{\partial^3 u}{\partial t^3} - c\frac{\Delta t}{2}\frac{\partial u^2}{\partial x \partial t} + \frac{c^2 \Delta x \Delta t}{4}\frac{\partial^2 u}{\partial x^2} + \frac{\Delta t^2}{6}\frac{\partial^3 u}{\partial t^3} + \ldots = -c\frac{\partial u}{\partial x} + \frac{c\Delta x}{2}\frac{\partial^2 u}{\partial x^2} + \frac{c\Delta x^2}{6}\frac{\partial^3 u}{\partial x^3} + \ldots \tag{5.7}$$

This term can be eliminated by using the combination (5.7) $+ c\frac{\Delta t}{2}\frac{\partial}{\partial x}$(5.5) and so on. The same procedure is applied successively until we finally obtain the equation :

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = c\frac{\Delta x}{2}\left(1 - c\frac{\Delta t}{\Delta x}\right)\frac{\partial^2 u}{\partial x^2} - \frac{\Delta x^2}{6}\left(1 - c\frac{\Delta t}{\Delta x}\right)\left(1 - 2c\frac{\Delta t}{\Delta x}\right)\frac{\partial^3 u}{\partial x^3}\ldots \tag{5.8}$$

This is called the modified equation associated with the upwind Euler scheme for the transport equation (5.1). It is worth noting that this is not a partial derivative equation in the conventional sense as it does not have an order or a finite amount of partial derivatives. For a solution to exist, one also needs a proper definition of infinitely many boundary conditions. Therefore, we restrict our analysis to solutions that are periodic [73].

### 5.1.2 Heuristic stability and limitations

Besides proving consistency, the modified equation quantifies explicitly the additional numerical effects. For example, it tells that up to first order in $\Delta x$ and $\Delta t$, the numerical solution rather satisfies a convection diffusion equation with a numerical dissipation coefficient given by $c\frac{\Delta x}{2}\left(1 - c\frac{\Delta t}{\Delta x}\right)$. Note however that the sign of the latter can be negative in which case this equation becomes unstable and admits solutions that grow exponentially in time rather than decay. It follows from this that a necessary and sufficient condition of stability for the solutions of this convection-diffusion equivalent is :

$$c\frac{\Delta t}{\Delta x} \leq 1 \tag{5.9}$$

which turns out to be exactly the CFL condition for the scheme (5.2). This gives a lot of potential for the modified equations to be a practical tool for stability analysis. Although the analysis is completely heuristic and has no rigorous foundation, its results make sense for a large class of schemes. However, there are also many examples for which this analysis fails to provide any practical stability condition. Consider for example the one dimensional heat equation :

$$\frac{\partial u}{\partial t} - \alpha\frac{\partial^2 u}{\partial x^2} = 0 \tag{5.10}$$

where $\alpha > 0$ is the diffusion coefficient. We discretize this equation using centered finite differences :

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} - \alpha\frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{\Delta x^2} = 0 \tag{5.11}$$

Using the same elimination procedure we obtain the following modified equation up to $4th$ order :

$$\frac{\partial u}{\partial t} - \alpha\frac{\partial^2 u}{\partial x^2} = \frac{\alpha\Delta x^2}{12}\left(1 - 6\frac{\Delta t}{\Delta x^2}\right)\frac{\partial^4 u}{\partial x^4} + \ldots \tag{5.12}$$

If we choose to stop at the first non-zero truncation term as previously, the conclusions are already less straightforward as we have a non-vanishing second order term that comes from the heat equation itself and a $4th$ order term that comes from the numerical effects. Therefore in order to look into stability of this equation, let us shift to Fourier space. Let $v(k,t)$ be the Fourier transform of $u(x,t)$ then under these notations, equation (5.12) yields :

$$\frac{\partial v}{\partial t} = -\alpha k^2 \left(1 - \frac{k^2\Delta x^2}{12}\left(1 - 6\frac{\Delta t}{\Delta x^2}\right)\right) v \tag{5.13}$$

It is reasonable to only consider the wavenumbers that are bound by $|k\Delta x| \leq \pi$, since these are the only wavenumbers admissible by the discrete mesh. In this setting, it is easy to verify that :

$$-\alpha k^2 \left(1 - \frac{k^2\Delta x^2}{12}\left(1 - 6\frac{\Delta t}{\Delta x^2}\right)\right) \leq 0 \quad \forall k \in \left[-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}\right] \tag{5.14}$$

which implies that the considered truncation is unconditionally stable for all admissible $\Delta t$ and $\Delta x$. This result is obviously erroneous as it is well known that the Von Neumann stability analysis proves that a necessary and sufficient stability condition for the scheme (5.11) is :

$$\alpha\frac{\Delta t}{\Delta x^2} \leq \frac{1}{2} \tag{5.15}$$

Thus, the heuristic analysis of this truncation of the modified equation failed to provide any practical stability condition for the scheme and truncating the equation at higher orders does not seem to do any better. This is one of the examples that demonstrates limitations of the method. In what follows, we first introduce all the necessary notations and setting for stability analysis before attempting to explain why the stability of the truncation is sometimes incoherent with the stability of the scheme.

## 5.2 Theory of stability through modified equations

### 5.2.1 Notations and assumptions

We will consider for our analysis partial derivative equations that are linear, first order in time and of arbitrary order in space :

$$\frac{\partial \tilde{u}}{\partial t} + \sum_{p=1}^{P} A_p \frac{\partial^p \tilde{u}}{\partial x^p} = 0 \tag{5.16}$$

where $A_p$ are constants. We consider explicit in time schemes, that are consistent with equation (5.16) and can be written as :

$$u_j^{n+1} = u_j^n + \Delta t \sum_{p=-n_l}^{n_r} b_p\left(\Delta x\right) u_{j+p}^n \tag{5.17}$$

where $n_l$ and $n_r$ are the number of mesh points to the left and to the right of $x_j$ respectively, used in every iteration. The coefficients $b_p$ verify :

$$\sum_{p=-n_l}^{n_r} b_p(\Delta x) = 0 \tag{5.18}$$

for consistency purposes, so that constant solutions, which are solution of the PDE (5.16) remain as such for the scheme. Sometimes it is preferable to cast the scheme (5.17) into an equivalent form :

$$u_j^{n+1} = u_j^n + \frac{\Delta t}{\Delta x^q} \sum_{p=-n_l}^{n_r} B_p\left(\Delta x\right) u_{j+p}^n \tag{5.19}$$

where $q$ is the highest power of $1/\Delta x$ present in the summation. Very often, when using standard finite differences, the coefficients $b_p$ are polynomials in $1/\Delta x$ and $q$ is the highest degree among these polynomials which is frequently equal to the order $P$ of the PDE (5.16). This formulation can be practical for stability analysis since most stability conditions for explicit schemes are given by bounds on the quantity $\lambda_q = \Delta t/\Delta x^q$ in the limit $\Delta x \to 0$ and $\Delta t \to 0$. Therefore, setting $\lambda_q$ as a non-vanishing parameter is a reasonable assumption that permits to reduce the number of free small parameters, that is we can take $\Delta t = g_q(\Delta x) = \lambda_q \Delta x^q$ and consider instead $\lambda_q$ and $\Delta x$ as free independent parameters. In what follows, in order to perform stability analysis in Fourier space, we consider a space continuous counterpart of the scheme (5.19) :

$$u^{n+1}(x) = u^n(x) + \lambda_q \sum_{p=-n_l}^{n_r} B_p\left(\Delta x\right) u^n(x + p\Delta x) \tag{5.20}$$

Let $v(k,t)$ be the Fourier transform in space of $u(x,t)$. If we take $\theta = k\Delta x$, then the Fourier transform of equation (5.20) yields :

$$v^{n+1}(k) = \left(1 + \lambda_q \sum_{p=-n_l}^{n_r} B_p(\Delta x) \mathrm{e}^{\mathrm{i}p\theta}\right) v^n(k) = S(\theta, \lambda_q, \Delta x) v^n(k) \tag{5.21}$$

Since we will be operating most of the time in Fourier space, it seems necessary to recall consistency of the scheme in the same setting. Therefore if we assume that the exact solution to the PDE (5.16) satisfies in Fourier space:

$$\tilde{v}(k, t + \Delta t) = \exp\left(\Delta t \sum_{p=1}^{P} (\mathrm{i}k)^p A_p\right) \tilde{v}(k,t) = \tilde{G}(\theta, \Delta t) \tilde{v}(k,t) \tag{5.22}$$

then the scheme (5.17) is consistent with the PDE (5.16) to order $s$ if and only if [22] :

$$\left| \frac{\tilde{G}(\theta, \Delta t) - S(\theta, \lambda_q, \Delta x)}{\Delta t} \right| = \mathcal{O}(\Delta t^s) \tag{5.23}$$

Lastly, we denote the modified equation associated to the scheme (5.17) by :

$$u_t = \sum_{p=1}^{\infty} \mu_p(\lambda_q, \Delta x) \frac{\partial^p u}{\partial x^p} \tag{5.24}$$

where $\mu_p$ are constants depending on $\Delta x$ and $\lambda_q$.

### 5.2.2 Fourier stability analysis

In this part, we focus on the link between the modified equation and the scheme [73]. The amplification factor of the scheme is none other than the modulus of $S(\theta, \lambda_q, \Delta x)$. Now, in order to recover an equivalent expression in the continuous time for the modified equation, we apply the Fourier transform to (5.24). This implies that $v(k,t)$ satisfies the differential equation:

$$\frac{dv}{dt} = \left( \sum_{p=1}^{\infty} (ik)^p \mu_p(\lambda_q, \Delta x) \right) v(k,t) = \left( \sum_{p=1}^{\infty} \alpha_p(\lambda_q, \Delta x) \theta^p \right) v(k,t) = G(\theta, \lambda_q, \Delta x) v(k,t) \tag{5.25}$$

where $\alpha_p(\lambda_q, \Delta x) = \mu_p(\lambda_q, \Delta x) i^p / \Delta x^p$. For convenience, we will say that the modified equation is stable if, for an initial condition $v(k, t=0) = v_0(k)$, the solution to the Cauchy problem :

$$\begin{cases} \dfrac{dv}{dt} = G(\theta, \lambda_q, \Delta x) v(k,t) \\ v(k,0) = v_0(k) \end{cases} \tag{5.26}$$

that is given by:

$$v(k,t) = \mathrm{e}^{tG(\theta, \lambda_q, \Delta x)} v_0(k) \tag{5.27}$$

remains bounded $\forall t < T$, where $T > 0$. Given this solution, one can obviously write :

$$v(k, t + \Delta t) = \mathrm{e}^{\Delta t G(\theta, \lambda_q, \Delta x)} v(k,t) \tag{5.28}$$

Now, since the solution to the scheme (5.17) with the same initial condition is also an exact solution to the modified equation (5.24) [73], uniqueness of this solution gives :

$$\mathrm{e}^{\Delta t G(\theta, \lambda_q, \Delta x)} = \mathrm{e}^{\lambda_q \Delta x^q G(\theta, \lambda_q, \Delta x)} = S(\theta, \lambda_q, \Delta x) \tag{5.29}$$

It follows from that the proposition :

**Proposition 5.2.1.** *For any scheme that writes as (5.17) and that is consistent with (5.16), there exists a positive number $\theta_m$ that depends only on $\lambda_q$ and $\Delta x$ such that, $\forall \theta \in ]-\theta_m, \theta_m[$ the expansion*

$$G(\theta, \lambda_q, \Delta x) = -\sum_{p=1}^{\infty} \frac{(1 - S(\theta, \lambda_q, \Delta x))^p}{p \lambda_q \Delta x^q}$$

*holds and the series $\sum_{p=1}^{\infty} \alpha_p(\lambda_q, \Delta x) \theta^p$ converges to $\frac{1}{\lambda_q \Delta x^q} \ln(S(\theta, \lambda_q, \Delta x))$.*

*Proof.* The scheme (5.17) is consistent with the PDE (5.16) and so $|S(0, \lambda_q, \Delta x) - 1| = 0 < 1$. Since $S$ is a continuous function with respect to $\theta$, $\exists \theta_m(\lambda_q, \Delta x) > 0$ such that $|S(\theta, \lambda_q, \Delta x) - 1| < 1 \; \forall \theta \in ]-\theta_m, \theta_m[$, and consequently, the principal logarithm $\ln(S(\theta, \lambda_q, \Delta x))$ defined by the series expansion :

$$\ln(S(\theta, \lambda_q, \Delta x)) = \ln(1 - (1 - S(\theta, \lambda_q, \Delta x))) = \sum_{p=1}^{\infty} -\frac{(1 - S(\theta, \lambda_q, \Delta x))^p}{p} \tag{5.30}$$

is convergent and $e^{\ln(S(\theta,\lambda_q,\Delta x))} = S(\theta,\lambda_q,\Delta x)$. This proves that the function:

$$G(\theta,\lambda_q,\Delta x) = \frac{1}{\lambda_q \Delta x^q}\ln(S(\theta,\lambda_q,\Delta x)) \tag{5.31}$$

is a solution of equation (5.29) for $\theta \in ]-\theta_m,\theta_m[$. The exponential function is locally invertible in the vicinity of 0 and so this expansion is unique in this vicinity [15] and therefore $\forall\theta \in ]-\theta_m,\theta_m[$ . Moreover, since $S(\theta,\lambda_q,\Delta x)$ is entire (as a finite sum of exponential functions) then further expanding $S$ into power series of $\theta$ for $\theta \in ]-\theta_m,\theta_m[$ finally yields:

$$G(\theta,\lambda_q,\Delta x) = \sum_{p=1}^{\infty}\alpha_p\theta^p = \frac{1}{\lambda_q \Delta x^q}\ln(S(\theta,\lambda_q,\Delta x)) \tag{5.32}$$

and hence, $\sum_{p=1}^{\infty}\alpha_p(\lambda_q,\Delta x)\theta^p$ is the series expansion of $\frac{1}{\lambda_q\Delta x^q}\ln(S(\theta,\lambda_q,\Delta x))$. This concludes our proof. □

**Remark 5.2.2.** *Equality (5.32) also implies that the coefficients $\alpha_p(\lambda_q,\Delta x)$ are given by :*

$$\alpha_p(\lambda_q,\Delta x) = \frac{\partial^p}{\partial\theta^p}\left(\frac{\ln(S(\theta,\lambda_q,\Delta x))}{p!\lambda_q\Delta x^q}\right)\bigg|_{\theta=0} \tag{5.33}$$

**Remark 5.2.3.** *$\theta_m$ is not the radius of convergence of the series $G(\theta,\lambda_q,\Delta x)$. If we denote by $R$ the radius of convergence, then we have $R \geq \theta_m$. The exact radius is not trivial to find in practice in the general case, since $G(\theta,\lambda_q,\Delta x)$ is a series expansion of a composite function. However it is sometimes possible to give estimates or exact values of the radius for some examples as will be shown later.*

**Remark 5.2.4.** *In practice, when looking for convergence, we are mainly searching for constraints in $\lambda_q$ and $\Delta x$, for which the series $G(\theta,\lambda_q,\Delta x)$ is convergent $\forall\theta \in [-\pi,\pi]$, that is we want $R > \pi$. A sufficient condition is $\theta_m > \pi$.*

So far, we have shown that for fixed parameters $\Delta x$ and $\lambda_q$, the series $G(\theta,\lambda_q,\Delta x)$ converges if $\theta \in ]-\theta_m,\theta_m[$. This being a sufficient condition of convergence, we do not know what happens for $|\theta| \geq \theta_m$. On the other hand it is worth noting that, for $|\theta| > R$, the series is divergent and equality (5.29) does not hold anymore. Therefore, the modified equation stability is not linked to that of the scheme for $|\theta| > R$.

### 5.2.3 Scheme stability domain and series convergence domain

Generally, in the Von Neumann setting, stability conditions of the scheme are given by constraints linking $\lambda_q$ and $\Delta x$ so that the inequality

$$|S(\theta,\lambda_q,\Delta x)| \leq 1 \quad \forall\theta \in [-\pi,\pi] \tag{5.34}$$

is verified. These constraints define a region of stability $\mathcal{R}_s$ in the $(\lambda_q,\Delta x)$ plane, that is :

$$\mathcal{R}_s = \left\{(\lambda_q,\Delta x) \in \mathbb{R}_+^2 : \forall\theta \in [-\pi,\pi] : |S(\theta,\lambda_q,\Delta x)| \leq 1\right\} \tag{5.35}$$

In the same manner, we can define a region $\mathcal{R}_c$ of the same plane, in which the series $G(\theta, \lambda_q, \Delta x)$ converges :

$$\mathcal{R}_c = \Big\{(\lambda_q, \Delta x) \in \mathbb{R}_+^2 : \forall \theta \in [-\pi, \pi] : |G(\theta, \lambda_q, \Delta x)| < \infty\Big\} \tag{5.36}$$

which is also equivalent to :

$$\mathcal{R}_c = \Big\{(\lambda_q, \Delta x) \in \mathbb{R}_+^2 : R(\lambda_q, \Delta x) \geq \pi\Big\} \tag{5.37}$$

In practice, it is not always possible to exactly determine $\mathcal{R}_c$. It is however easier to explicitly find a subset of this region, that is :

$$\Omega_c = \Big\{(\lambda_q, \Delta x) \in \mathbb{R}_+^2 : \forall \theta \in [-\pi, \pi] : |1 - S(\theta, \lambda_q, \Delta x)| < 1\Big\} \subset \mathcal{R}_c \tag{5.38}$$

It is worth mentioning that $\Omega_c$ is always a non-empty set. Indeed, for $\lambda_q = 0$, we have $S(\theta, 0, \Delta x) = 1$ and consequently $\forall \Delta x \in \mathbb{R}_+$ there exists in $\mathbb{R}^2$ a neighborhood of $(\Delta x, \lambda_q)$ in which we have $|1 - S(\theta, \lambda_q, \Delta x)| < 1$. This means that we always have convergence for sufficiently small values of $\lambda_q$. Lastly, we denote by $\mathcal{R}_m$ the region of stability of the modified equation :

$$\mathcal{R}_m = \Big\{(\lambda_q, \Delta x) \in \mathbb{R}_+^2 : \Big|\mathrm{e}^{\lambda_q \Delta x^q G(\theta, \lambda_q, \Delta x)}\Big| \leq 1\Big\} = \Big\{(\lambda_q, \Delta x) \in \mathbb{R}_+^2 : \mathrm{Re}(G(\theta, \lambda_q, \Delta x) \leq 0\Big\} \tag{5.39}$$

Provided these definitions, we can distinguish two cases :

1. If $\mathcal{R}_s \subset \mathcal{R}_c$ then the stability of the modified equation provides reliable and complete information regarding the scheme stability.

2. If any subset of $\mathcal{R}_s$ lies outside of the convergence domain $\mathcal{R}_c$, this means that there is information on the stability limit of the scheme that is missed by the modified equations since its Fourier transform is non existing outside of $\mathcal{R}_c$.

The following proposition is a direct consequence :

**Proposition 5.2.5.** *For any scheme that writes as (5.17) and that is consistent with (5.16) we have $(\mathcal{R}_m \cap \mathcal{R}_c) \subset \mathcal{R}_s$, that is if the modified equation is stable and its Fourier series is convergent then the scheme is also stable.*

This proves that inside the convergence domain $\mathcal{R}_c$, the stability of the modified equation is a sufficient stability condition for the scheme. Furthermore, we shall add that if $\mathcal{R}_s \subset \mathcal{R}_c$ then this condition is also necessary. This result, literally, is not of practical interest as the stability of the full series $S(\theta, \lambda_q, \Delta x)$ is either nontrivial or impossible to obtain. But we will show that the above classification permits to justify whether a truncated version of the modified equation yields significant information on stability.

### 5.2.4 Link between the stability of the scheme and the stability of a truncation

Instead of the full series expansion, let us consider a truncated modified equation to an arbitrary order $N > P$:

$$u_t = \sum_{p=1}^{N} \mu_p(\lambda_q, \Delta x) \frac{\partial^p u}{\partial x^p}, \quad \mu_N \neq 0 \tag{5.40}$$

We recall that our main purpose is to know in which case does the stability of this truncation provide relevant information on the stability of the corresponding scheme. This differs from the approach of Warming and Hyett [73] in the sense that they showed how to reconstruct the exact Von Neumann amplification factor using a finite amount of coefficients $\mu_p$ without actually analyzing the stability of the truncated version. Under the Fourier transform, the previous equation writes :

$$\frac{dv}{dt} = \left( \sum_{p=1}^{N} \alpha_p(\lambda_q, \Delta x) \theta^p \right) v(k,t) = P_N(\theta, \lambda_q, \Delta x) v(k,t) \tag{5.41}$$

Here, $P_N(\theta, \lambda_q, \Delta x)$ is a polynomial of $\theta$ of degree $N$ which is trivially a truncation of the series $G(\theta, \lambda_q, \Delta x)$. In the same manner as previously, the ordinary differential equation (5.41) yields :

$$v(k, t + \Delta t) = \mathrm{e}^{\Delta t P_N(\theta, \lambda_q, \Delta_x)} v(k,t) = S_N(\theta, \lambda_q, \Delta x) v(k,t) \tag{5.42}$$

In contrast to the full series, the stability conditions of the truncation are obtainable in most cases through the analysis of the polynomial function $P_N(\theta, \lambda_q, \Delta x)$. Let $R_N(\theta, \lambda_q, \Delta x)$ be the rest of the series defined by :

$$R_N(\theta, \lambda_q, \Delta_x) = G(\theta, \lambda_q, \Delta_x) - P_N(\theta, \lambda_q, \Delta_x) = \sum_{p=N+1}^{\infty} \alpha_p(\lambda_q, \Delta x) \theta^p \tag{5.43}$$

In the convergence domain $\mathcal{R}_c$, the rest $R_N(\theta, \lambda_q, \Delta x)$ is bounded and we have :

$$\lim_{N \to +\infty} R_N(\theta, \lambda_q, \Delta x) \equiv 0 \quad \text{and} \quad \lim_{N \to +\infty} S_N(\theta, \lambda_q, \Delta x) \equiv S(\theta, \lambda_q, \Delta x) \tag{5.44}$$

In this setting we can state the following result :

**Proposition 5.2.6.** *Assume an initial condition satisfying supp$(v_0) \in [-M, M]$ and an arbitrary truncation order $N > P$, then for any $(\Delta x, \lambda_q) \in \mathcal{R}_c$, if the truncated modified equation is stable in the sense that there exists $C > 0$ such that:*

$$|S_N(\theta, \lambda_q, \Delta x)| \leq 1 + C\Delta t \tag{5.45}$$

*then the scheme is also stable in the same sense.*

*Proof.* For $(\Delta x, \lambda_q) \in \mathcal{R}_c$ we have :

$$S(\theta, \lambda_q, \Delta x) = e^{\Delta t G(\theta, \lambda_q, \Delta x)} = S_N(\theta, \lambda_q, \Delta x) e^{\Delta t R_N(\theta, \lambda_q, \Delta x)}$$

Thus, since $v^n = S^n v_0$, we can write :

$$\begin{aligned}
|v^n| = |S^n v_0| &\leq |S_N|^n \left| e^{n \sum_{N+1}^{\infty} \alpha_p \theta^p} v_0 \right| \\
&\leq (1 + C\Delta t)^n \left| e^{nA\theta^{N+1}} v_0 \right| \\
&\leq (1 + C\Delta t)^n \left| e^{nA(k\Delta x)^{N+1}} v_0 \right| \\
&\leq (1 + C\Delta t)^n \left| e^{nA\Delta x^{N+1-q} \Delta t k^{N+1}/\lambda_q} v_0 \right| \\
&\leq e^{CT} \left| e^{AT\Delta x^{N+1-q} M^{N+1}/\lambda_q} v_0 \right| \\
&\leq e^{CT} e^{AT\Delta x^{N+1-q} M^{N+1}/\lambda_q} \left| v_0 \right|
\end{aligned}$$

That is, $v^n$ is $L^2$-stable for initial conditions that are of compact support in the frequency domain, that is $k \in [-M, M]$. □

## 5.3 Examples

### 5.3.1 Heat equation - centered finite differences

Consider the centered finite differences scheme for the heat equation. It can be cast into the form :

$$u_i^{n+1} = u_i^n + \alpha\lambda_2 \left( u_{i-1}^n - 2u_i^n + u_{i-1}^n \right) \tag{5.46}$$

We take for simplicity $\alpha = 1$. The modified equations associated to this scheme up to $8th$ order for example is given by :

$$\begin{aligned}
\frac{\partial u}{\partial t} =& \frac{\partial^2 u}{\partial x^2} + \frac{\Delta x^2}{12} (1 - 6\lambda_2) \frac{\partial^4 u}{\partial x^4} + \frac{\Delta x^4}{360} \left( 1 - 30\lambda_2 \left( 1 - 4\lambda_2 \right) \right) \frac{\partial^6 u}{\partial x^6} \\
&+ \frac{\Delta x^6}{20160} \left( 1 - 42\lambda_2 \left( 3 - 40\lambda_2 \left( 1 - 3\lambda_2 \right) \right) \right) \frac{\partial^8 u}{\partial x^8} + \dots
\end{aligned} \tag{5.47}$$

Straightforward computations yield :

$$S(\theta, \lambda_2, \Delta x) = 1 + 4\lambda_2 \sin^2(\theta/2) \tag{5.48}$$

$$\mathcal{R}_s = \left\{ (\lambda_q, \Delta x) \in \mathbb{R}^2 : \lambda_2 \leq \frac{1}{2} \right\} \quad ; \quad \Omega_c = \left\{ (\lambda_q, \Delta x) \in \mathbb{R}^2 : \lambda_2 \leq \frac{1}{4} \right\} \tag{5.49}$$

$$\begin{aligned}
\Delta t P_8(\theta, \lambda_2, \Delta x) =& -\lambda_2 \left( \theta^2 - \frac{1 - 6\lambda_2}{12} \theta^4 + \frac{1 - 30\lambda_2 \left( 1 - 4\lambda_2 \right)}{360} \theta^6 \right) \\
&+ \lambda_2 \frac{1 - 42\lambda_2 (3 - 40\lambda_2 (1 - 3\lambda_2))}{20160} \theta^8
\end{aligned} \tag{5.50}$$

The domains of stability and convergence only depend on the parameter $\lambda_2$ independently of $\Delta x$. This permits to take an arbitrary value of $\Delta x = 1$ and carry on the analysis, based on only $\lambda_2$. Furthermore since $|S(\theta, \lambda_2, \Delta x)|$ is an even function with respect to $\theta$ it suffices to look at $\theta \in [0, \pi]$. Figure 5.1 shows a comparison between $|S(\theta, \lambda_2, \Delta x)|$ and $|S_N(\theta, \lambda_2, \Delta x)|$ for $N = 2$ and $N = 8$ in two cases. In the limit of stability $\lambda_2 = 1/2$, the scheme is stable but the series $G(\theta, \lambda_2, \Delta x)$ is not convergent for $\theta > R = \pi/2$ (See Appendix D.1). As displayed in the left-hand side of figure 5.1, the curves begin aligned in the low frequencies, and then the truncation curves begin to diverge completely from the function $|S(\theta, \lambda_2, \Delta x)|$ once $\theta$ surpasses the threshold $R$. This is not the case for $\lambda_2 = 1/4$. For this value we can calculate the radius of convergence $R = \pi$ (See Appendix D.2). In fact, we can see on the right-hand side of the figure that for $\lambda_2$, the curves remain very close $\forall \theta \in [0, \pi]$. For $N = 8$, the two curves almost overlap.

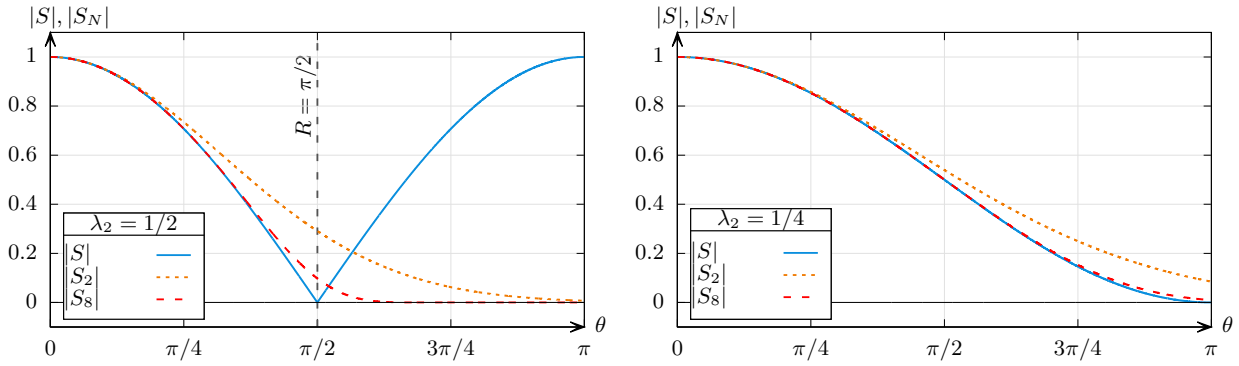


Figure 5.1: Plot of the function $|S(\theta, \lambda_2, \Delta x)|$ along $|S_2(\theta, \lambda_2, \Delta x)|$ and $|S_8(\theta, \lambda_2, \Delta x)|$ for the values of $\lambda_2 = 1/2$ (left) and $\lambda_2 = 1/4$ (right). We can see that for $\lambda_2 = 1/2$, which lies outside of the convergence domain, the truncation curves stray away from the curve of $S$ starting from $\theta = R$. For $\lambda_2 = 1/4$, the truncations match well with $S$.

### 5.3.2 Transport equation - Upwind Euler

The scheme writes :

$$u_i^{n+1} = u_i^n - c\lambda_1(u_i^n - u_{i-1}^n) \tag{5.51}$$

We take $c = 1$. In this case, the modified equation up to $4th$ order for example is given by :

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = (1-\lambda_1)\left(\frac{\Delta x}{2}\frac{\partial^2 u}{\partial x^2} - \frac{\Delta x^2}{6}(1-2\lambda_1)\frac{\partial^3 u}{\partial x^3} + \frac{\Delta x^3}{24}(1-6\lambda_1(1-\lambda_1))\frac{\partial^4 u}{\partial x^4}\right) + \dots \tag{5.52}$$

Straightforward computations yield :

$$S(\theta, \lambda_1, \Delta x) = 1 - \lambda_1\left(1 - \mathrm{e}^{-\mathrm{i}\theta}\right) \tag{5.53}$$

$$\mathcal{R}_s = \left\{(\lambda_q, \Delta x) \in \mathbb{R}^2 : \lambda_1 \leq 1\right\} \quad ; \quad \Omega_c = \left\{(\lambda_q, \Delta x) \in \mathbb{R}^2 : \lambda_1 \leq \frac{1}{2}\right\} \tag{5.54}$$

$$\Delta t P_4(\theta, \lambda_1, \Delta x) = -\mathrm{i}\lambda_1\theta + \lambda_1(1-\lambda_1)\left(-\frac{1}{2}\theta^2 + \frac{1}{6}(1-2\lambda_1)\mathrm{i}\theta^3 + \frac{1}{24}(1-6\lambda_1(1-\lambda_1))\theta^4\right) + \dots \tag{5.55}$$

In contrast to the previous example for the heat equation, we could not compute explicitly the radius of convergence for values of interest of $\lambda_1$. Nevertheless, it is possible to extend $\Omega_c$ to cover all the values of $\lambda_1$ that are in $\mathcal{R}_s$ (See appendix D.3). Hence, as shown in figure 5.2, for values of $\lambda_1 \leq 1$, the amplification factor of the truncation to $6th$ order $|S_6(\theta, \lambda_1, \Delta x)|$ seems to be a very good approximation of the scheme amplification factor $|S(\theta, \lambda_1, \Delta x)|$, even for values of $\lambda_1$ that are in the vicinity of the stability threshold.

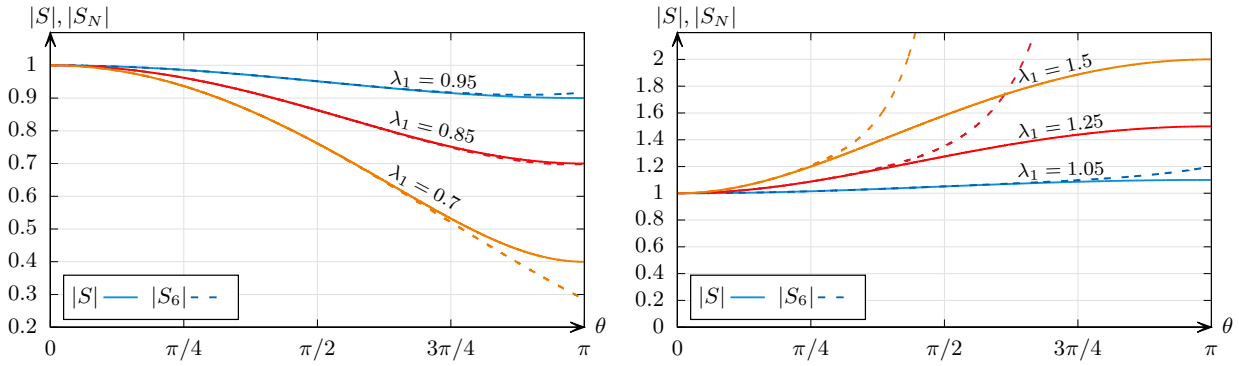


Figure 5.2: Plot of the function $|S(\theta, \lambda_1, \Delta x)|$ along $|S_6(\theta, \lambda_1, \Delta x)|$ for different values of $\lambda_1$ in the stable region (left) and in the unstable region (right). The continuous lines stand for the amplification factor $|S(\theta, \lambda_1, \Delta x)|$ and the dashed lines represent the truncated modified equation amplification factor $|S_6(\theta, \lambda_1, \Delta x)|$. We can see that for $\lambda_1 \leq 1$ the truncation curves match well with the exact amplification factor even in the boundaries of stability. Inversely, it seems according to the right-hand graphic that $\lambda_1 = 1$ marks the threshold of convergence.

# Conclusion and perspectives

We could explain throughout this work that one of the main reasons behind the failures of the modified equations technique is non other than series divergence. Although the analysis only provides sufficient conditions in general, it lifts some well-known ambiguities and provides some assumptions required by the technique to be of a more justified practical use. We show that in these settings, the stability of a truncation gives already reasonable approximations to stability conditions of the scheme. An extension of the obtained results to the case of systems is underway. It would be also interesting to extend this approach to cover a larger class of explicit and also implicit schemes.

# Conclusion and perspectives

The main contribution of this work is the development of a new first order hyperbolic system of equations that approximates Euler-Korteweg type systems. The approach was shown efficient for two cases, namely for the defocusing cubic nonlinear Schrödinger equation and for thin films flows with capillarity. The obtained numerical results showed a very good agreement for a variety of solutions, including exact and asymptotic solutions, reference numerical solutions and experimental results in the one-dimensional case.

A natural continuation of this work would be to extend the results presented here to the multi-dimensional case. This requires the use of schemes with structure preserving properties in order to obtain results which are compatible with the curl-free constraint on $\mathbf{p}$. Such a problem has been addressed for example in a recent work on a hyperbolic reformulation of compressible flows with surface tension [20], which exhibits similar properties as the augmented Euler-Korteweg model.

Another interesting point to investigate is to generalize this approach to systems that are not only dispersive but also inherently dissipative and in particular, systems of the Navier-Stokes-Korteweg type. This can provide a more efficient way to numerically address the thin film equations, for which we employed classical finite differences for the diffusive terms. This would also generalize the approach to a wider class of equations.

It would be interesting to study the extension of the augmented Lagrangian approach to systems with a non-convex free energy such as the focusing NLSE. Some recent advance in that direction has been done for the study of bistable tapespring in [10] where the non-convexity zone is bounded.

Lastly, it would be helpful to properly implement transparent boundary conditions for this approach. This point was not addressed in the current work and was avoided by considering sufficiently large computational domains when necessary, to avoid the spurious oscillations which would reflect on the boundaries.

# Bibliography


[1] M. Ablowitz and B. Prinari. "Nonlinear Schrodinger systems: continuous and discrete". In: *Scholarpedia* 3.8 (2008). revision #137230, p. 5561. DOI: 10.4249/scholarpedia.5561.

[2] U. M. Ascher, S. J. Ruuth, and R. J. Spiteri. "Implicit-explicit Runge-Kutta methods for time-dependent partial differential equations". In: *Applied Numerical Mathematics* 25.2-3 (1997), pp. 151–167. DOI: 10.1016/S0168-9274(97)00056-1.

[3] C. Audiard and B. Haspot. "Global well-posedness of the Euler–Korteweg system for small irrotational data". In: *Communications in Mathematical Physics* 351.1 (2017), pp. 201–247.

[4] W. Bao, S. Jin, and P. A. Markowich. "On time-splitting spectral approximations for the Schrödinger equation in the semiclassical regime". In: *Journal of Computational Physics* 175.2 (2002), pp. 487–524.

[5] W. Bao, S. Jin, and P. A. Markowich. "Numerical study of time-splitting spectral discretizations of nonlinear Schrödinger equations in the semiclassical regimes". In: *SIAM Journal on Scientific Computing* 25.1 (2003), pp. 27–64.

[6] D. J. Benney and G. J. Roskes. "Wave Instabilities". In: *Studies in Applied Mathematics* 48.4 (1969), pp. 377–385. DOI: 10.1002/sapm1969484377.

[7] S. Benzoni-Gavage. "Planar traveling waves in capillary fluids". In: *Differential and Integral Equations* 26.3-4 (2013), pp. 439–485.

[8] S. Benzoni-Gavage, R. Danchin, and S. Descombes. "On the well-posedness for the Euler-Korteweg model in several space dimensions". In: *Indiana University Mathematics Journal* 56.4 (2007), pp. 1499–1580. DOI: 10.1512/iumj.2007.56.2974.

[9] S. Benzoni-Gavage, S. Descombes, D Jamet, and L Mazet. "Structure of Korteweg models and stability of diffuse interfaces". In: *Interfaces and free boundaries* 7.4 (2005), pp. 371–414.

[10] S. Bourgeois, N. Favrie, and B. Lombard. "Dynamics of a regularized and bistable Ericksen bar using an extended Lagrangian approach". In: *International Journal of Solids and Structures* (2020). DOI: https://doi.org/10.1016/j.ijsolstr.2020.09.031.

[11] D. Bresch, M. Gisclon, and I. Lacroix-Violet. "On Navier–Stokes–Korteweg and Euler–Korteweg systems: application to quantum fluids models". In: *Archive for Rational Mechanics and Analysis* 233.3 (2019), pp. 975–1025.

[12] D. Bresch, F. Couderc, P. Noble, and J.-P. Vila. "New extended formulations of euler-korteweg equations based on a generalization of the quantum bohm identity". In: *arXiv preprint arXiv:1503.08678* (2015).

[13] D. Bresch, F. Couderc, P. Noble, and J.-P. Vila. "A generalization of the quantum Bohm identity: Hyperbolic CFL condition for Euler–Korteweg equations". In: *Comptes Rendus Mathematique* 354.1 (2016), pp. 39–43.

[14] D. Bresch, N. Cellier, F. Couderc, M. Gisclon, P. Noble, G.-L. Richard, C. Ruyer-Quil, and J.-P. Vila. "Augmented skew-symmetric system for shallow-water system with surface tension allowing large gradient of density". In: *Journal of Computational Physics* 419 (2020), p. 109670.

[15] R. Carpentier, A. de La Bourdonnaye, and B. Larrouturou. "On the derivation of the modified equation for the analysis of linear numerical methods". In: *ESAIM: Mathematical Modelling and Numerical Analysis* 31.4 (1997), pp. 459–470.

[16] P Casal. "Capillarité interne en mécanique des millieux continus". In: *Compt. Rend* 256.3 (1961).

[17] P Casal. "La capillarité interne". In: *Cahier du groupe français de rhéologie, CNRS VI* 3 (1961), pp. 31–37.

[18] P Casal and H Gouin. "A representation of liquid-vapor interfaces by using fluids of second grade". In: *Annales de Physique*. Vol. 13. 2. 1988, pp. 3–12.

[19] R. Y. Chiao, E. Garmire, and C. H. Townes. "Self-trapping of optical beams". In: *Physical Review Letters* 13.15 (1964), pp. 479–482. DOI: 10.1103/PhysRevLett.13.479.

[20] S. Chiocchetti, I. Peshkov, S. Gavrilyuk, and M. Dumbser. "High order ADER schemes and GLM curl cleaning for a first order hyperbolic formulation of compressible flow with surface tension". In: *arXiv preprint arXiv:2002.08818* (2020).

[21] C. M. Dafermos. *Hyperbolic conservation laws in continuum physics, volume 325 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. 2010.

[22] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology: Volume 6 Evolution Problems II*. Springer Science & Business Media, 2012.

[23] A Davey and K. Stewartson. "On three-dimensional packets of surface waves". In: *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* 338.1613 (1974), pp. 101–110. DOI: 10.1098/rspa.1974.0076.

[24] S. F. Davis. "Simplified second-order Godunov-type methods". In: *SIAM J. Scientific and Statistical Computing* 9.3 (1988), pp. 445–473.

[25] P.-G. De Gennes, F. Brochard-Wyart, and D. Quéré. *Capillarity and wetting phenomena: drops, bubbles, pearls, waves*. Springer Science & Business Media, 2013.

[26] F Dell'Isola and S. L. Gavrilyuk. *Variational Models and Methods in Solid and Fluid Mechanics*. Vol. 535. Springer Science & Business Media, 2011.

[27] B. V. Derjaguin, N. V. Churaev, and V. M. Muller. *Surface Forces*. 1987. DOI: 10.1007/978-1-4757-6639-4.

[28] J. E. Dunn and J Serrin. "On the thermomechanics of interstitial working". In: *Archive Rat. Mech. Anal.* 88.2 (1985), pp. 95–133.

[29] M. E. Eglit. "A generalization of the model of an ideal compressible fluid". In: *J. Appl. Math. Mech.* 29.2 (1965), pp. 395–399.

[30] G. A. El, R. H. J. Grimshaw, and N. F. Smyth. "Unsteady undular bores in fully nonlinear shallow-water theory". In: *Physics of Fluids* 18 (2006), p. 27104.

[31] G. A. El and M. A. Hoefer. "Dispersive shock waves and modulation theory". In: *Physica D: Nonlinear Phenomena* 333 (2016), pp. 11–65.

[32] G. A. El, V. V. Geogjaev, A. V. Gurevich, and A. L. Krylov. "Decay of an initial discontinuity in the defocusing NLS hydrodynamics". In: *Physica D: Nonlinear Phenomena* 87.1-4 (1995), pp. 186–192.

[33] N Favrie and S Gavrilyuk. "A rapid numerical method for solving Serre–Green–Naghdi equations describing long free surface gravity waves". In: *Nonlinearity* 30.7 (2017), p. 2718.

[34] G. Fibich. *The Nonlinear Schrödinger Equation*. Vol. 192. Applied Mathematical Sciences. Cham: Springer International Publishing, 2015. DOI: 10.1007/978-3-319-12748-4.

[35] B. Fornberg and T. A. Driscoll. "A fast spectral algorithm for nonlinear wave equations with linear dispersion". In: *Journal of Computational Physics* 155.2 (1999), pp. 456–467.

[36] S. L. Gavrilyuk and S. M. Shugrin. "Media With Equations of State That Depend on derivatives". In: 37.2 (1996), pp. 177–189.

[37] S. L. Gavrilyuk and V. M. Teshukov. "Generalized vorticity for bubbly liquid and dispersive shallow water equations". In: *Continuum Mechanics and Thermodynamics* 13.6 (2001), pp. 365–382.

[38] S. Gavrilyuk and H. Gouin. "A new form of governing equations of fluids arising from Hamilton's principle". In: *International Journal of Engineering Science* 37.12 (1999), pp. 1495–1520.

[39] J. Ginibre and G. Velo. "On a class of nonlinear Schrödinger equations. I. The Cauchy problem, general case". In: *Journal of Functional Analysis* 32.1 (1979), pp. 1–32. DOI: 10.1016/0022-1236(79)90076-4.

[40] V. L. Ginzburg and L. P. Pitaevskii. "On the theory of superfluidity". In: *Sov. Phys. JETP* 7.5 (1958), pp. 858–861.

[41] A. V. Gurevich and A. L. Krylov. "Dissipationless shock waves in media with positive dispersion". In: *Zh. Eksp. Teor. Fiz.* 92 (1987), p. 1684.

[42] A. V. Gurevich and A. P. Meshcherkin. "Expanding self-similar discontinuities and shock waves in dispersive hydrodynamics". In: *Zh. Eksp. Teor. Fiz.* 87 (1984), pp. 1277–1292.

[43] A. V. Gurevich and L. P. Pitaevskii. "Nonstationary structure of a collisionless shock wave". In: *Zh. Eksp. Teor. Fiz.* 65 (1973), pp. 590–604.

[44] B. Haspot. "Existence of strong solutions for nonisothermal Korteweg system". In: *Annales Mathématiques Blaise Pascal*. Vol. 16. 2. 2009, pp. 431–481.

[45] C. W. Hirt. "Heuristic stability theory for finite-difference equations". In: *Journal of Computational Physics* 2.4 (1968), pp. 339–355.

[46] M. A. Hoefer. "Shock waves in dispersive Eulerian fluids". In: *Journal of Nonlinear Science* 24.3 (2014), pp. 525–577.

[47] C. G. J. Jacobi. *Fundamenta nova theoriae functionum ellipticarum. Auctore D. Carolo Gustavo Iacobo Iacobi...* sumtibus fratrum Borntræger, 1829.

[48] A. Y. Kipnis, A. I. Kipnis, B. E. IAvelov, and J. S. Rowlinson. *Van der Waals and molecular science*. Oxford University Press, 1996.

[49] C. Klein. "Fourth order time-stepping for low dispersion Korteweg-de Vries and nonlinear Schrödinger equation". In: *Electron. Trans. Numer. Anal* 29 (2008), pp. 116–135.

[50] D. J. Korteweg. “Sur la forme que prennent les équations du mouvements des fluides si l’on tient compte des forces capillaires causées par des variations de densité considérables mais connues et sur la théorie de la capillarité dans l’hypothèse d’une variation continue de la densité”. In: *Archives Néerlandaises des Sciences exactes et naturelles* 6 (1901), pp. 1–24.

[51] S. K.∼A.∼Bagrinovskii. “Difference schemes for multidimensional problems”. In: *Dokl. Akad. Nauk SSSR* 115.3 (1957), pp. 431–433.

[52] L. D. Landau. “On the theory of superconductivity”. In: *Collected Papers of L.D. Landau*. Elsevier, 1965, pp. 546–568. DOI: 10.1016/B978-0-08-010586-4.50078-X.

[53] L. D. Landau and V. L. Ginzburg. “On the theory of superconductivity”. In: *Zh. Eksp. Teor. Fiz.* 20 (1950), p. 1064.

[54] G. Lavalle. “Integral modeling of liquid films sheared by a gas flow”. PhD thesis. 2014.

[55] J. D. Lawson. “Generalized Runge-Kutta processes for stable systems with large Lipschitz constants”. In: *SIAM Journal on Numerical Analysis* 4.3 (1967), pp. 372–380.

[56] J. Li and Z. Yang. “The von Neumann analysis and modified equation approach for finite difference schemes”. In: *Applied Mathematics and Computation* 225 (2013), pp. 610–621.

[57] J. Liu and J. P. Gollub. “Solitary wave dynamics of film flows”. In: *Physics of Fluids* 6.5 (1994), pp. 1702–1712.

[58] E Madelung. “Quantentheorie in Hydrodynamischer Form”. In: *Zeitschrift für Physik* 40.3 (1927), pp. 322–326.

[59] B. V. Minchev and W. Wright. *A review of exponential integrators for first order semi-linear problems*. Tech. rep. 2005.

[60] S Ndanou, N Favrie, and S Gavrilyuk. “Criterion of hyperbolicity in hyperelasticity in the case of the stored energy in separable form”. In: *J. Elast.* 115.1 (2014), pp. 1–25.

[61] P. Noble and J.-P. Vila. “Stability theory for difference approximations of some dispersive shallow water equations and application to thin film flows”. In: *arXiv preprint arXiv:1304.3805* (2013).

[62] L. Pareschi and G. Russo. “Implicit-explicit Runge-Kutta schemes and applications to hyperbolic systems with relaxation”. In: *Journal of Scientific Computing* 25.1 (2005), pp. 129–155. DOI: 10.1007/s10915-004-4636-4.

[63] M. V. Pavlov. “Nonlinear Schrödinger equation and the Bogolyubov-Whitham method of averaging”. In: *Teoreticheskaya i Matematicheskaya Fizika* 71.3 (1987), pp. 351–356.

[64] G. L. Richard, C. Ruyer-Quil, and J.-P. Vila. “A three-equation model for thin films down an inclined plane”. In: *Journal of Fluid Mechanics* 804 (2016), pp. 162–200. DOI: 10.1017/jfm.2016.530.

[65] G. L. Richard, M. Gisclon, C. Ruyer-Quil, and J.-P. Vila. “Optimization of consistent two-equation models for thin film flows”. In: *European Journal of Mechanics, B/Fluids* 76.1967 (2019), pp. 7–25. DOI: 10.1016/j.euromechflu.2019.01.004.

[66] R. D. Richtmyer and K. W. Morton. “Difference methods for initial-value problems”. In: *Malabar, Fla.: Krieger Publishing Co.,| c1994, 2nd ed.* (1994).

[67] D. Serre. *Systems of Conservation Laws 1: Hyperbolicity, entropies, shock waves*. Cambridge University Press, 1999.

[68] A Shabat and V. E. Zakharov. “Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media”. In: *Soviet physics JETP* 34.1 (1972), p. 62.

[69] V. I. Talanov. “Self-focusing of wave beams in nonlinear media”. In: *JETP Lett* 2.5 (1965), pp. 138–141.

[70] E. F. Toro. *Riemann solvers and numerical methods for fluid dynamics: a practical introduction*. Springer Science & Business Media, 2013.

[71] L. Tyler. “Heuristic analysis of convective finite difference techniques”. In: *Proceedings of the Second International Conference on Numerical Methods in Fluid Dynamics*. Springer. 1971, pp. 314–319.

[72] J. D. van der Waals. “The thermodynamic theory of capillarity under the hypothesis of a continuous variation of density”. In: *Journal of Statistical Physics* 20.2 (1979), pp. 200–244.

[73] R. F. Warming and B. J. Hyett. “The Modified Equation Approach to the Stability and Accuracy Analysis of Finite-Difference Methods”. In: *Journal of Computational Physics* 179 (1974), pp. 159–179.

[74] E. W. Weisstein. *"Euler Polynomial" From MathWorld–A Wolfram Web Resource*. URL: `http://mathworld.wolfram.com/EulerPolynomial.html` (visited on 01/02/2020).

[75] G. B. Whitham. *Linear and nonlinear waves*. John Wiley & Sons, 2011.

[76] V. E. Zakharov. “Stability of periodic waves of finite amplitude on the surface of a deep fluid”. In: *Journal of Applied Mechanics and Technical Physics* 9.2 (1968), pp. 190–194.

APPENDIX A

# Calculus and developments for the augmented E-K system

In this part of the appendix, we will detail all the necessary calculus for deriving motion equations from a given Lagrangian. This will be done in particular for both the original Lagrangian of the Euler-Korteweg system and for the associated augmented Lagrangian as well. In each case, we will first remind how the variations write and then proceed to a step-by-step computation. Several details which would have been too tedious to put into the main text are also exposed later, for example on how to compare the augmented model with the original one and how to compute the dispersion relation.

## A.1 Augmented Lagrangian Calculus

Here, we will proceed in a generic manner, meaning that will we consider a very general form of the Lagrangian. This choice makes for much lighter and less cumbersome computations which serve the same purpose, while also providing more flexibility. Let us consider a Lagrangian whose expression can be put into the form :

$$\mathcal{L} = \int_{\Omega_t} \left( \rho \frac{|\mathbf{u}|^2}{2} - W(\rho, \eta, \dot{\eta}, \nabla\eta) \right) d\Omega. \tag{A.1}$$

Assume that this Lagrangian is submitted to a differential constraint that is the mass conservation law :

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{u}) = 0.$$

We associate to the given Lagrangian, Hamilton's action for an arbitrary time interval $[t_0, t_1]$:

$$a = \int_{t_0}^{t_1} \mathcal{L} dt \tag{A.2}$$

We recall that the variables in this case are separated into two types depending on their variations in the sense that the variation $\delta\eta$ involves $\eta$, $\nabla\eta$ and $\dot{\eta}$, while the variation $\delta\mathbf{x}$ involves $\rho$, $\mathbf{u}$ and $\dot{\eta}$. In both cases, it is supposed that the variations are vanishing at the boundary of $[t_0, t_1] \times \Omega_t$.

### A.1.1 Euler-Lagrange equation associated with $\delta\eta$

We recall that the variations of $\nabla\eta$ and $\dot{\eta}$ write :

$$\delta(\nabla\eta) = \nabla(\delta\eta), \quad \delta\dot{\eta} = \frac{\partial\delta\eta}{\partial t} + \frac{\partial\delta\eta}{\partial \mathbf{x}}\mathbf{u} \tag{A.3}$$

By applying Hamilton's principle of stationary action to (A.2), we can write :

$$\delta a = 0 \;\Rightarrow\; \delta\left(\int_{t_0}^{t_1}\int_{\Omega_t}\left(\rho\frac{|\mathbf{u}|^2}{2} - W(\rho,\eta,\dot{\eta},\nabla\eta)\right)d\Omega\; dt\right) = 0 \tag{A.4}$$

consequently:

$$\begin{aligned}
\delta a &= \int_{t_0}^{t_1}\int_{\Omega_t}\delta\left(\rho\frac{|\mathbf{u}|^2}{2} - W(\rho,\eta,\dot{\eta},\nabla\eta)\right)\; d\Omega dt \\
&= \int_{t_0}^{t_1}\int_{\Omega_t} -\left(\frac{\partial W}{\partial\eta}\delta\eta + \frac{\partial W}{\partial\dot{\eta}}\delta\dot{\eta} + \frac{\partial W}{\partial\nabla\eta}\delta\nabla\eta\right)d\Omega dt \\
&= \int_{t_0}^{t_1}\int_{\Omega_t} -\left(\frac{\partial W}{\partial\eta}\delta\eta + \frac{\partial W}{\partial\dot{\eta}}\frac{\partial\delta\eta}{\partial t} + \frac{\partial W}{\partial\dot{\eta}}\frac{\partial\delta\eta}{\partial \mathbf{x}}\mathbf{u} + \frac{\partial W}{\partial\nabla\eta}\delta\nabla\eta\right)d\Omega dt.
\end{aligned}$$

Using Gauss-Ostrogradski's theorem and taking into account that the variation $\delta\eta$ is vanishing on the boundary of $[t_0,t_1]\times\Omega_t$, we can eliminate derivatives of variations. Let us demonstrate how this process works on one of the terms in the last integral. We write for example:

$$\begin{aligned}
\int_{t_0}^{t_1}\int_{\Omega_t}\left(\frac{\partial W}{\partial\dot{\eta}}\frac{\partial\delta\eta}{\partial t}\right)d\Omega_t dt &= \int_{t_0}^{t_1}\int_{\Omega_t}\left(\frac{\partial}{\partial t}\left(\frac{\partial W}{\partial\dot{\eta}}\delta\eta\right) - \frac{\partial}{\partial t}\left(\frac{\partial W}{\partial\dot{\eta}}\right)\delta\eta\right)d\Omega_t dt \\
&= \underbrace{\int_{\Omega_t}\left[\frac{\partial W}{\partial\dot{\eta}}\delta\eta\right]_{t0}^{t1} d\Omega_t}_{=0} - \int_{t_0}^{t_1}\int_{\Omega_t}\frac{\partial}{\partial t}\left(\frac{\partial W}{\partial\dot{\eta}}\right)\delta\eta \quad d\Omega_t dt \\
&= -\int_{t_0}^{t_1}\int_{\Omega_t}\frac{\partial}{\partial t}\left(\frac{\partial W}{\partial\dot{\eta}}\right)\delta\eta \quad d\Omega_t dt
\end{aligned}$$

Applying the same technique for the remaining terms, yields:

$$\delta a = \int_{t_0}^{t_1}\int_{\Omega_t}\left(-\frac{\partial W}{\partial\eta} + \frac{\partial}{\partial t}\left(\frac{\partial W}{\partial\dot{\eta}}\right) + \mathrm{div}\left(\frac{\partial W}{\partial\dot{\eta}}\mathbf{u}\right) + \mathrm{div}\left(\frac{\partial W}{\partial\nabla\eta}\right)\right)\delta\eta\; d\Omega dt = 0, \quad \text{for any } \delta\eta.$$

Thus, we obtain Euler-Lagrange's equation for $\eta$:

$$\frac{\partial}{\partial t}\left(\frac{\partial W}{\partial\dot{\eta}}\right) + \mathrm{div}\left(\frac{\partial W}{\partial\dot{\eta}}\mathbf{u} + \frac{\partial W}{\partial\nabla\eta}\right) = \frac{\partial W}{\partial\eta}. \tag{A.5}$$

### A.1.2 Euler-Lagrange equation associated with $\delta\mathbf{x}$

In this case, the variations write :

$$\hat{\delta}\rho = -\mathrm{div}\left(\rho\delta\mathbf{x}\right), \quad \hat{\delta}\mathbf{u} = \left(\dot{\delta\mathbf{x}}\right) - \frac{\partial\mathbf{u}}{\partial\mathbf{x}}\delta\mathbf{x}, \quad \hat{\delta}\dot{\eta} = \hat{\delta}\mathbf{u}\cdot\nabla\eta. \tag{A.6}$$

Hamilton's action writes :

$$\begin{aligned}
\delta a &= \int_{t0}^{t_1}\int_{\Omega_t}\delta\left(\rho\frac{|\mathbf{u}|^2}{2} - W(\rho,\eta,\dot{\eta},\nabla\eta)\right)d\Omega dt \\
&= \int_{t_0}^{t_1}\int_{\Omega_t}\left(\frac{|\mathbf{u}|^2}{2}\hat{\delta}\rho + \rho\mathbf{u}\cdot\hat{\delta}\mathbf{u} - \frac{\partial W}{\partial\rho}\hat{\delta}\rho - \frac{\partial W}{\partial\dot{\eta}}\hat{\delta}\dot{\eta}\right)d\Omega dt \\
&= \int_{t_0}^{t_1}\int_{\Omega_t}\left(\frac{|\mathbf{u}|^2}{2}\hat{\delta}\rho + \rho\mathbf{u}\cdot\hat{\delta}\mathbf{u} + \frac{\partial W}{\partial\rho}\mathrm{div}(\rho\delta\mathbf{x}) - \frac{\partial W}{\partial\dot{\eta}}(\delta\mathbf{u}\cdot\nabla\eta)\right)d\Omega dt \\
&= \int_{t_0}^{t_1}\int_{\Omega_t}\left(\frac{|\mathbf{u}|^2}{2}\hat{\delta}\rho + \rho\mathbf{u}\cdot\hat{\delta}\mathbf{u} - \rho\nabla\left(\frac{\partial W}{\partial\rho}\right)\cdot\delta\mathbf{x}\right)d\Omega dt - \int_{t_0}^{t_1}\int_{\Omega_t}\left(\frac{\partial W}{\partial\dot{\eta}}\left(\frac{\partial\delta\mathbf{x}}{\partial t} + \left(\mathbf{u}\cdot\nabla\right)\delta\mathbf{x} - \nabla\mathbf{u}\delta\mathbf{x}\right)\cdot\nabla\eta\right)d\Omega dt
\end{aligned}$$

For now let us split Hamilton's action into two separate integrals, as each of them still requires some work and term rearrangements. Thus, we write :

$$\delta a = \delta a_1 - \delta a_2$$

where :

$$\delta a_1 = \int_{t_0}^{t_1}\int_{\Omega_t}\left(\frac{|\mathbf{u}|^2}{2}\hat{\delta}\rho + \rho\mathbf{u}\cdot\hat{\delta}\mathbf{u} - \rho\nabla\left(\frac{\partial W}{\partial\rho}\right)\cdot\delta\mathbf{x}\right)d\Omega dt \tag{A.7}$$

$$\delta a_2 = \int_{t_0}^{t_1}\int_{\Omega_t}\left(\frac{\partial W}{\partial\dot{\eta}}\left(\frac{\partial\delta\mathbf{x}}{\partial t} + \left(\mathbf{u}\cdot\nabla\right)\delta\mathbf{x} - \nabla\mathbf{u}\delta\mathbf{x}\right)\cdot\nabla\eta\right)d\Omega dt \tag{A.8}$$

We develop :

$$\begin{aligned}
\delta a_1 &= \int_{t_0}^{t_1}\int_{\Omega_t}\left(-\frac{|\mathbf{u}|^2}{2}\,\mathrm{div}(\rho\delta\mathbf{x}) + \rho\mathbf{u}\cdot\left(\left(\dot{\delta\mathbf{x}}\right) - \frac{\partial\mathbf{u}}{\partial\mathbf{x}}\delta\mathbf{x}\right) - \rho\nabla\left(\frac{\partial W}{\partial\rho}\right)\cdot\delta\mathbf{x}\right)d\Omega dt \\
&= \int_{t_0}^{t_1}\int_{\Omega_t}\left(\rho\nabla\left(\frac{|\mathbf{u}|^2}{2}\right)\cdot\delta\mathbf{x} + \rho\mathbf{u}\cdot\left(\frac{\partial\delta\mathbf{x}}{\partial t} + \frac{\partial\delta\mathbf{x}}{\partial\mathbf{x}}\mathbf{u} - \frac{\partial\mathbf{u}}{\partial\mathbf{x}}\delta\mathbf{x}\right) - \rho\nabla\left(\frac{\partial W}{\partial\rho}\right)\cdot\delta\mathbf{x}\right)d\Omega dt \\
&= \int_{t_0}^{t_1}\int_{\Omega_t}\left(\rho\nabla\left(\frac{|\mathbf{u}|^2}{2}\right)\cdot\delta\mathbf{x} - \frac{\partial\rho\mathbf{u}}{\partial t}\cdot\delta\mathbf{x} - \mathrm{div}(\rho\mathbf{u}\otimes\mathbf{u})\cdot\delta\mathbf{x} - \rho\mathbf{u}\cdot\frac{\partial\mathbf{u}}{\partial\mathbf{x}}\delta\mathbf{x} - \rho\nabla\left(\frac{\partial W}{\partial\rho}\right)\cdot\delta\mathbf{x}\right)d\Omega dt \\
&= \int_{t_0}^{t_1}\int_{\Omega_t}\left(\rho\nabla\left(\frac{|\mathbf{u}|^2}{2}\right)\cdot\delta\mathbf{x} - \rho\mathbf{u}\cdot\frac{\partial\mathbf{u}}{\partial\mathbf{x}}\delta\mathbf{x} - \left(\frac{\partial\rho\mathbf{u}}{\partial t} + \mathrm{div}(\rho\mathbf{u}\otimes\mathbf{u}) + \rho\nabla\left(\frac{\partial W}{\partial\rho}\right)\right)\cdot\delta\mathbf{x}\right)d\Omega dt.
\end{aligned}$$

For lightness, let us denote by $A$ the quantity:

$$A = -\left(\frac{\partial\rho\mathbf{u}}{\partial t} + \mathrm{div}(\rho\mathbf{u}\otimes\mathbf{u}) + \rho\nabla\left(\frac{\partial W}{\partial\rho}\right)\right)$$

and let us pursue the development of the remaining terms :

$$\begin{aligned}
\delta a_1 &= \int_{t_0}^{t_1}\int_{\Omega_t}\left(\rho\nabla\left(\frac{|\mathbf{u}|^2}{2}\right)\cdot\delta\mathbf{x} - \rho\mathbf{u}\cdot(\mathrm{div}(\delta\mathbf{x}\otimes\mathbf{u}) - \mathrm{div}(\delta\mathbf{x})\mathbf{u})) + A\cdot\delta\mathbf{x}\right)d\Omega dt \\
&= \int_{t_0}^{t_1}\int_{\Omega_t}\left(\left(\rho\frac{\partial\mathbf{u}}{\partial\mathbf{x}}\mathbf{u}\right)\cdot\delta\mathbf{x} + tr\left((\delta\mathbf{x}\otimes\mathbf{u})\frac{\partial\rho\mathbf{u}}{\partial\mathbf{x}}\right) + \rho|\mathbf{u}|^2\,\mathrm{div}(\delta\mathbf{x}) + A\cdot\delta\mathbf{x}\right)d\Omega dt \\
&= \int_{t_0}^{t_1}\int_{\Omega_t}\left(\left(\rho\frac{\partial\mathbf{u}}{\partial\mathbf{x}}\mathbf{u}\right)\cdot\delta\mathbf{x} + \mathbf{u}\cdot\left(\frac{\partial\rho\mathbf{u}}{\partial\mathbf{x}}\delta\mathbf{x}\right) + \rho|\mathbf{u}|^2\,\mathrm{div}(\delta\mathbf{x}) + A\cdot\delta\mathbf{x}\right)d\Omega dt \\
&= \int_{t_0}^{t_1}\int_{\Omega_t}\left(\left(\rho\frac{\partial\mathbf{u}}{\partial\mathbf{x}}\mathbf{u}\right)\cdot\delta\mathbf{x} + \mathbf{u}\cdot(\mathrm{div}(\delta\mathbf{x}\otimes\rho\mathbf{u}) - \rho\mathbf{u}\,\mathrm{div}(\delta\mathbf{x})) - \nabla\left(\rho|\mathbf{u}|^2\right)\cdot\delta\mathbf{x} + A\cdot\delta\mathbf{x}\right)d\Omega dt \\
&= \int_{t_0}^{t_1}\int_{\Omega_t}\left(\left(\rho\frac{\partial\mathbf{u}}{\partial\mathbf{x}}\mathbf{u}\right)\cdot\delta\mathbf{x} + \mathbf{u}\cdot\mathrm{div}(\delta\mathbf{x}\otimes\rho\mathbf{u}) + A\cdot\delta\mathbf{x}\right)d\Omega dt \\
&= \int_{t_0}^{t_1}\int_{\Omega_t}((\mathrm{div}(\mathbf{u}\otimes\mathbf{u}) - \mathrm{div}(\mathbf{u})\mathbf{u})\cdot(\rho\delta\mathbf{x}) + \mathbf{u}\cdot\mathrm{div}(\delta\mathbf{x}\otimes\rho\mathbf{u}) + A\cdot\delta\mathbf{x})\,d\Omega dt \\
&= \int_{t_0}^{t_1}\int_{\Omega_t}(\mathrm{div}((\mathbf{u}\otimes\mathbf{u})\rho\delta\mathbf{x}) + A\cdot\delta\mathbf{x})\,d\Omega dt = \int_{t_0}^{t_1}\int_{\Omega_t}\left(-\left(\frac{\partial\rho\mathbf{u}}{\partial t} + \mathrm{div}(\rho\mathbf{u}\otimes\mathbf{u}) + \rho\nabla\left(\frac{\partial W}{\partial\rho}\right)\right)\cdot\delta\mathbf{x}\right)d\Omega dt
\end{aligned}$$

This concludes the computation of $\delta a_1$. Using the same techniques and hints, let us carry on by developing $\delta a_2$. We write :

$$\begin{aligned}
\delta a_2 &= \int_{t_0}^{t_1}\int_{\Omega_t}\left(\frac{\partial W}{\partial\dot\eta}\left(\frac{\partial\delta\mathbf{x}}{\partial t} + \left(\frac{\partial\delta\mathbf{x}}{\partial\mathbf{x}}\mathbf{u} - \nabla\mathbf{u}\delta\mathbf{x}\right)\cdot\nabla\eta\right)\right)d\Omega dt \\
&= \int_{t_0}^{t_1}\int_{\Omega_t}\left(\frac{\partial W}{\partial\dot\eta}\left(\frac{\partial\delta\mathbf{x}}{\partial t}\cdot\nabla\eta + \left(\frac{\partial\delta\mathbf{x}}{\partial\mathbf{x}}\mathbf{u}\right)\cdot\nabla\eta - (\nabla\mathbf{u}^T\nabla\eta)\cdot\delta\mathbf{x}\right)\right)d\Omega dt \\
&= \int_{t_0}^{t_1}\int_{\Omega_t}\left(\frac{\partial W}{\partial\dot\eta}\left(\frac{\partial\delta\mathbf{x}}{\partial t}\cdot\nabla\eta + (\mathrm{div}(\mathbf{u}\otimes\delta\mathbf{x}) - \delta\mathbf{x}\mathrm{div}(\mathbf{u}))\cdot\nabla\eta - (\nabla\mathbf{u}^T\nabla\eta)\cdot\delta\mathbf{x}\right)\right)d\Omega dt \\
&= \int_{t_0}^{t_1}\int_{\Omega_t}\left(\frac{\partial W}{\partial\dot\eta}\left(\frac{\partial\delta\mathbf{x}}{\partial t}\cdot\nabla\eta + \left(\mathrm{div}(\mathbf{u}\otimes\delta\mathbf{x}\nabla\eta) - tr\left(\mathbf{u}\otimes\delta\mathbf{x}\frac{\partial\nabla\eta}{\partial\mathbf{x}}\right) - \mathrm{div}(\mathbf{u})\nabla\eta\cdot\delta\mathbf{x}\right) - (\nabla\mathbf{u}^T\nabla\eta)\cdot\delta\mathbf{x}\right)\right)d\Omega dt \\
&= \int_{t_0}^{t_1}\int_{\Omega_t}\left(\frac{\partial W}{\partial\dot\eta}\left(\frac{\partial\delta\mathbf{x}}{\partial t}\cdot\nabla\eta + \left(\mathrm{div}((\nabla\eta\cdot\delta\mathbf{x})\mathbf{u}) - \left(\frac{\partial\nabla\eta}{\partial\mathbf{x}}\mathbf{u}\right)\cdot\delta\mathbf{x} - \mathrm{div}(\mathbf{u})\nabla\eta\cdot\delta\mathbf{x}\right) - (\nabla\mathbf{u}^T\nabla\eta)\cdot\delta\mathbf{x}\right)\right)d\Omega dt.
\end{aligned}$$

Now we use the Green-Ostrogradski's theorem, taking into account that $\delta\mathbf{x} = 0$ on the boundary $\partial\Omega_t$, to obtain :

$$\delta a_2 = -\int_{t_0}^{t_1}\int_{\Omega_t}\left(\frac{\partial}{\partial t}\left(\frac{\partial W}{\partial\dot\eta}\nabla\eta\right) + \mathbf{u}\cdot\nabla\left(\frac{\partial W}{\partial\dot\eta}\right)\nabla\eta + \frac{\partial W}{\partial\dot\eta}\left(\frac{\partial\nabla\eta}{\partial\mathbf{x}}\mathbf{u}\right) + \frac{\partial W}{\partial\dot\eta}\mathrm{div}(\mathbf{u})\nabla\eta + \left(\frac{\partial W}{\partial\dot\eta}\nabla\mathbf{u}^T\nabla\eta\right)\right)\cdot\delta\mathbf{x}d\Omega dt$$

Reuniting both parts into Hamilton's action $\delta a$ we finally get :

$$\int_{t_0}^{t_1}\int_{\Omega_t}\left(\frac{\partial}{\partial t}\left(\frac{\partial W}{\partial\dot\eta}\nabla\eta\right) + \mathbf{u}\cdot\nabla\left(\frac{\partial W}{\partial\dot\eta}\right)\nabla\eta + \frac{\partial W}{\partial\dot\eta}\left(\frac{\partial\nabla\eta}{\partial\mathbf{x}}\mathbf{u}\right) + \frac{\partial W}{\partial\dot\eta}\mathrm{div}(\mathbf{u})\nabla\eta + \left(\frac{\partial W}{\partial\dot\eta}\nabla\mathbf{u}^T\nabla\eta\right) - A\right)\cdot\delta\mathbf{x}d\Omega dt = 0\ \forall\delta\mathbf{x}$$

Hence,

$$
\begin{aligned}
&-\frac{\partial \rho \mathbf{u}}{\partial t}-\operatorname{div}(\rho \mathbf{u}\otimes\mathbf{u})-\rho\nabla\left(\frac{\partial W}{\partial \rho}\right)+\left(\frac{\partial}{\partial t}\left(\frac{\partial W}{\partial \dot\eta}\right)+\operatorname{div}\left(\frac{\partial W}{\partial \dot\eta}\mathbf{u}\right)\right)\nabla\eta+\frac{\partial W}{\partial \dot\eta}\left(\nabla\left(\frac{\partial}{\partial t}\eta+\mathbf{u}\nabla\eta\right)\right)=0\\
&-\frac{\partial \rho \mathbf{u}}{\partial t}-\operatorname{div}(\rho \mathbf{u}\otimes\mathbf{u})-\rho\nabla\left(\frac{\partial W}{\partial \rho}\right)+\left(\frac{\partial W}{\partial \eta}-\operatorname{div}\left(\frac{\partial W}{\partial \nabla\eta}\right)\right)\nabla\eta+\frac{\partial W}{\partial \dot\eta}\left(\nabla\left(\frac{\partial}{\partial t}\eta+\mathbf{u}\nabla\eta\right)\right)=0\\
&-\frac{\partial \rho \mathbf{u}}{\partial t}-\operatorname{div}(\rho \mathbf{u}\otimes\mathbf{u})-\rho\nabla\left(\frac{\partial W}{\partial \rho}\right)+\left(\frac{\partial W}{\partial \eta}-\operatorname{div}\left(\frac{\partial W}{\partial \nabla\eta}\right)\right)\nabla\eta+\frac{\partial W}{\partial \dot\eta}\nabla\dot\eta=0\\
&-\frac{\partial \rho \mathbf{u}}{\partial t}-\operatorname{div}(\rho \mathbf{u}\otimes\mathbf{u})-\rho\nabla\left(\frac{\partial W}{\partial \rho}\right)-\operatorname{div}\left(\frac{\partial W}{\partial \nabla\eta}\right)\nabla\eta+\left(\frac{\partial W}{\partial \eta}\nabla\eta+\frac{\partial W}{\partial \dot\eta}\nabla\dot\eta\right)=0\\
&-\frac{\partial \rho \mathbf{u}}{\partial t}-\operatorname{div}(\rho \mathbf{u}\otimes\mathbf{u})-\rho\nabla\left(\frac{\partial W}{\partial \rho}\right)-\operatorname{div}\left(\frac{\partial W}{\partial \nabla\eta}\right)\nabla\eta+\left(\nabla W-\frac{\partial W}{\partial \rho}\nabla\rho-\left(\frac{\partial \nabla\eta}{\partial \mathbf{x}}\right)\frac{\partial W}{\partial \nabla\eta}\right)=0\\
&-\frac{\partial \rho \mathbf{u}}{\partial t}-\operatorname{div}(\rho \mathbf{u}\otimes\mathbf{u})-\left(\rho\nabla\left(\frac{\partial W}{\partial \rho}\right)+\frac{\partial W}{\partial \rho}\nabla\rho-\nabla W\right)-\left(\operatorname{div}\left(\frac{\partial W}{\partial \nabla\eta}\right)\nabla\eta+\left(\frac{\partial \nabla\eta}{\partial \mathbf{x}}\right)\frac{\partial W}{\partial \nabla\eta}\right)=0\\
&\frac{\partial \rho \mathbf{u}}{\partial t}+\operatorname{div}(\rho \mathbf{u}\otimes\mathbf{u})+\operatorname{div}\left(\left(\rho\frac{\partial W}{\partial \rho}-W\right)Id\right)+\operatorname{div}\left(\frac{\partial W}{\partial \nabla\eta}\otimes\nabla\eta\right)=0
\end{aligned}
$$

Finally , we obtain the Euler-Lagrange equation:

$$
\boxed{\frac{\partial \rho \mathbf{u}}{\partial t}+\operatorname{div}\left(\rho\mathbf{u}\otimes\mathbf{u}+\left(\rho\frac{\partial W}{\partial \rho}-W\right)\mathbf{Id}+\frac{\partial W}{\partial \nabla\eta}\otimes\nabla\eta\right)=0} \tag{A.9}
$$

## A.2 Asymptotics of the Augmented Momentum equation

In this section we show that the momentum balance equation for the augmented system and which we remind here :

$$
\frac{\partial \rho \mathbf{u}}{\partial t}+\operatorname{div}\left(\rho\mathbf{u}\otimes\mathbf{u}+\left(\rho^2\epsilon'(\rho)+\frac{1}{2}(\rho K'(\rho)-K(\rho))|\mathbf{p}|^2+\frac{\eta}{\alpha}\left(1-\frac{\eta}{\rho}\right)\right)\mathbf{Id}+K(\rho)\mathbf{p}\otimes\mathbf{p}\right)=0
$$

approaches the original one :

$$
\frac{\partial \rho \mathbf{u}}{\partial t}+\operatorname{div}(\rho \mathbf{u}\otimes\mathbf{u})+\nabla(\rho^2\epsilon'(\rho))=\rho\nabla\left(K(\rho)\Delta\rho+\frac{1}{2}K'(\rho)\left|\nabla\rho\right|^2\right)
$$

in the limits of the small parameters $\alpha\to 0$ and $\beta\to 0$. Let us proceed by replacing in the momentum equation, accordingly to equation (1.25), the term :

$$
\frac{1}{\alpha}\left(1-\frac{\eta}{\rho}\right)=-\left(K(\rho)\Delta\eta+K'(\rho)\nabla\rho\cdot\nabla\eta\right)+\beta\rho\dot w
$$

we obtain :

$$
\frac{\partial \rho \mathbf{u}}{\partial t}+\operatorname{div}\left(\rho\mathbf{u}\otimes\mathbf{u}+\left(\rho^2\epsilon'(\rho)+\frac{1}{2}(\rho K'(\rho)-K(\rho))|\mathbf{p}|^2-\rho\left(K(\rho)\Delta\eta+K'(\rho)\nabla\rho\cdot\nabla\eta\right)+\beta\rho\dot w\right)\mathbf{Id}+K(\rho)\nabla\eta\otimes\nabla\eta\right)=0
$$

By making use of the asymptotic formulas :

$$\eta = \rho + \mathcal{O}(\alpha), \quad \nabla\eta = \nabla\rho + \mathcal{O}(\alpha), \quad \Delta\eta = \Delta\rho + \mathcal{O}(\alpha)$$

we get :

$$\frac{\partial \rho \mathbf{u}}{\partial t} + \mathrm{div}\,(\rho \mathbf{u} \otimes \mathbf{u}) + \nabla\left(\rho^2 \epsilon'(\rho)\right) = \nabla\left(\frac{1}{2}(\rho K'(\rho) + K(\rho))|\nabla\rho|^2 + \rho K(\rho)\Delta\rho\right) - \mathrm{div}\,(K(\rho)\nabla\rho \otimes \nabla\rho) + \mathcal{O}(\alpha) + \mathcal{O}(\beta)$$

Thus it only remains to show that :

$$\nabla\left(\frac{1}{2}(\rho K'(\rho) + K(\rho))|\nabla\rho|^2 + \rho K(\rho)\Delta\rho\right) - \mathrm{div}\,(K(\rho)\nabla\rho \otimes \nabla\rho) = \rho\nabla\left(K(\rho)\Delta\rho + \frac{1}{2}K'(\rho)\,|\nabla\rho|^2\right)$$

For this purpose, let us expand the left hand side :

$$\begin{aligned}
&\nabla\left(\frac{1}{2}(\rho K'(\rho) + K(\rho))|\nabla\rho|^2 + \rho K(\rho)\Delta\rho\right) - \mathrm{div}\,(K(\rho)\nabla\rho \otimes \nabla\rho) \\
&= \nabla\left(\rho\left(\frac{1}{2}K'(\rho)|\nabla\rho|^2 + K(\rho)\Delta\rho\right) + \frac{1}{2}K(\rho)|\nabla\rho|^2\right) - \mathrm{div}\,(K(\rho)\nabla\rho)\,\nabla\rho - K(\rho)\frac{\partial\nabla\rho}{\partial\mathbf{x}}\nabla\rho \\
&= \rho\nabla\left(\frac{1}{2}K'(\rho)|\nabla\rho|^2 + K(\rho)\Delta\rho\right) + \left(\frac{1}{2}K'(\rho)|\nabla\rho|^2 + K(\rho)\Delta\rho\right)\nabla\rho + \frac{1}{2}\nabla\left(K(\rho)|\nabla\rho|^2\right) - \mathrm{div}\,(K(\rho)\nabla\rho)\,\nabla\rho - K(\rho)\frac{\partial\nabla\rho}{\partial\mathbf{x}}\nabla\rho \\
&= \rho\nabla\left(\frac{1}{2}K'(\rho)|\nabla\rho|^2 + K(\rho)\Delta\rho\right) + \left(\frac{1}{2}K'(\rho)|\nabla\rho|^2 + K(\rho)\Delta\rho\right)\nabla\rho + \frac{1}{2}K'(\rho)|\nabla\rho|^2\nabla\rho + K(\rho)\frac{\partial\nabla\rho}{\partial\mathbf{x}}\nabla\rho \\
&\quad - K'(\rho)|\nabla\rho|^2\nabla\rho - K(\rho)\Delta\rho\nabla\rho - K(\rho)\frac{\partial\nabla\rho}{\partial\mathbf{x}}\nabla\rho \\
&= \rho\nabla\left(\frac{1}{2}K'(\rho)|\nabla\rho|^2 + K(\rho)\Delta\rho\right)
\end{aligned}$$

This ends the computation and shows that the augmented E-K momentum equation writes as:

$$\frac{\partial \rho \mathbf{u}}{\partial t} + \mathrm{div}\,(\rho \mathbf{u} \otimes \mathbf{u}) + \nabla(\rho^2 \epsilon'(\rho)) = \rho\nabla\left(K(\rho)\Delta\rho + \frac{1}{2}K'(\rho)\,|\nabla\rho|^2\right) + \mathcal{O}(\alpha) + \mathcal{O}(\beta)$$

## A.3 Invariance by rotations of the group $\mathcal{SO}(3)$

In order to simplify the hyperbolicity check, we will show that our system is invariant by rotation. Let $\mathrm{O} \in \mathcal{SO}(3)$ be a constant orthogonal transformation ($\mathrm{OO}^T = Id$). The aim is to show that the set of equations (1.21) remains unchanged after applying the orthogonal transformation O. Let us consider the variables in the new transformed base, denoted with primes :

$$\mathbf{x}' = \mathrm{O}\mathbf{x} \quad \mathbf{u}' = \mathrm{O}\mathbf{u} \quad \mathbf{p}' = \mathrm{O}\mathbf{p} \quad , \quad t' = t \quad \rho' = t \quad w' = w \quad \eta' = \eta \tag{A.10}$$

For the mass conservation equation for example we have :

$$\begin{aligned}
&\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{u}) = 0 \\
&\frac{\partial \rho'}{\partial t'} + \operatorname{div}(\rho' \mathrm{O}^T \mathbf{u}') = 0 \\
&\frac{\partial \rho'}{\partial t'} + tr\left(\mathrm{O}^T \frac{\partial}{\partial \mathbf{x}} (\rho' \mathbf{u}')\right) = 0 \\
&\frac{\partial \rho'}{\partial t'} + tr\left(\mathrm{O}^T \frac{\partial}{\partial \mathbf{x}'} (\rho' \mathbf{u}') \mathrm{O}\right) = 0 \\
&\frac{\partial \rho'}{\partial t'} + \operatorname{div}'(\rho' \mathbf{u}') = 0
\end{aligned}$$

The denotation $\operatorname{div}'$ means that we derive with respect to the transformed base coordinates given by $\mathbf{x}' = \mathrm{O}\mathbf{x}$. More generally for any vector $\mathbf{v}$ and scalar $\lambda$ we give the following relations :

$$\operatorname{div}(\mathbf{v}) = \operatorname{div}\left(\mathrm{O}^T \mathbf{v}'\right) = tr\left(\mathrm{O}^T \frac{\partial \mathbf{v}'}{\partial \mathbf{x}}\right) = tr\left(\mathrm{O}^T \frac{\partial \mathbf{v}'}{\partial \mathbf{x}'} \mathrm{O}\right) = tr\left(\frac{\partial \mathbf{v}'}{\partial \mathbf{x}'}\right) = \operatorname{div}'(\mathbf{v}')) \tag{A.11}$$

$$\frac{\partial \lambda}{\partial \mathbf{x}} = \frac{\partial \lambda'}{\partial \mathbf{x}'} \mathrm{O} \quad \Leftrightarrow \quad \nabla' \lambda' = \mathrm{O}^T \nabla \lambda \tag{A.12}$$

This permits to easily prove that the system of equations (1.21) are invariant by rotation.

## A.4 Analog of Helmholtz's equation for vorticity

Given that the potential $W$ depends explicitly on the material derivative of the macroscopic variable $\eta$, one can derive an analog of Helmholtz's equation for vorticity. We proceed as in [37], [38] and define the quantity :

$$\mathbf{K} = \mathbf{u} - \frac{1}{\rho} \frac{\partial W}{\partial w} \mathbf{p}. \tag{A.13}$$

We would like to show that $\mathbf{K}$ satisfies the PDE :

$$\frac{\partial \mathbf{K}}{\partial t} + \operatorname{curl}(\mathbf{K}) \wedge \mathbf{u} + \nabla\left(\frac{1}{2} |\mathbf{u}|^2 - \frac{1}{\rho} \frac{\partial W}{\partial w} \mathbf{p} \cdot \mathbf{u} + \frac{\partial W}{\partial \rho}\right) = 0$$

Let us start from the augmented system, written here in its general form :

$$
\begin{aligned}
&\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{u}) = 0 \\
&\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{u}}{\partial \mathbf{x}}\mathbf{u} + \frac{1}{\rho}\operatorname{div}\left(\frac{\partial W}{\partial \mathbf{p}}\right)\mathbf{p} + \frac{1}{\rho}\frac{\partial \mathbf{p}}{\partial \mathbf{x}}\frac{\partial W}{\partial \mathbf{p}} + \frac{1}{\rho}\nabla\left(\rho\frac{\partial W}{\partial \rho} - W\right) = 0 \\
&\frac{\partial \mathbf{p}}{\partial t} + \frac{\partial \mathbf{p}}{\partial \mathbf{x}}\mathbf{u} + \left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\right)^T \mathbf{p} - \nabla w = 0 \\
&\frac{\partial}{\partial t}\left(\frac{\partial W}{\partial w}\right) + \operatorname{div}\left(\frac{\partial W}{\partial w}\mathbf{u}\right) + \operatorname{div}\left(\frac{\partial W}{\partial \mathbf{p}}\right) - \frac{\partial W}{\partial \eta} = 0
\end{aligned}
$$

Hence :

$$
\begin{aligned}
\frac{\partial \mathbf{K}}{\partial t} =& \frac{\partial \mathbf{u}}{\partial t} + \frac{1}{\rho^2}\frac{\partial W}{\partial w}\frac{\partial \rho}{\partial t}\mathbf{p} - \frac{1}{\rho}\frac{\partial}{\partial t}\left(\frac{\partial W}{\partial w}\right)\mathbf{p} - \frac{1}{\rho}\frac{\partial W}{\partial w}\frac{\partial \mathbf{p}}{\partial t} \\
=& -\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\mathbf{u} - \frac{1}{\rho}\operatorname{div}\left(\frac{\partial W}{\partial \mathbf{p}}\right) - \frac{1}{\rho}\frac{\partial \mathbf{p}}{\partial \mathbf{x}}\frac{\partial W}{\partial \mathbf{p}} - \frac{1}{\rho}\nabla\left(\rho\frac{\partial W}{\partial \rho} - W\right) \\
&- \frac{1}{\rho^2}\frac{\partial W}{\partial w}\operatorname{div}(\rho\mathbf{u})\,\mathbf{p} + \frac{1}{\rho}\operatorname{div}\left(\frac{\partial W}{\partial w}\mathbf{u}\right)\mathbf{p} + \frac{1}{\rho}\operatorname{div}\left(\frac{\partial W}{\partial \mathbf{p}}\right)\mathbf{p} \\
&- \frac{1}{\rho}\frac{\partial W}{\partial \eta}\mathbf{p} + \frac{1}{\rho}\frac{\partial W}{\partial w}\frac{\partial \mathbf{p}}{\partial \mathbf{x}}\mathbf{u} + \frac{1}{\rho}\frac{\partial W}{\partial w}\left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\right)^T \mathbf{p} - \frac{1}{\rho}\frac{\partial W}{\partial w}\nabla w \\
=& -\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\mathbf{u} - \frac{1}{\rho}\frac{\partial \mathbf{p}}{\partial \mathbf{x}}\frac{\partial W}{\partial \mathbf{p}} - \frac{1}{\rho}\frac{\partial W}{\partial \rho}\nabla\rho - \nabla\left(\frac{\partial W}{\partial \rho}\right) + \frac{1}{\rho}\nabla W - \frac{1}{\rho^2}\frac{\partial W}{\partial w}\operatorname{div}(\rho\mathbf{u})\,\mathbf{p} \\
&+ \frac{1}{\rho}\operatorname{div}\left(\frac{\partial W}{\partial w}\mathbf{u}\right)\mathbf{p} - \frac{1}{\rho}\frac{\partial W}{\partial \eta}\mathbf{p} + \frac{1}{\rho}\frac{\partial W}{\partial w}\frac{\partial \mathbf{p}}{\partial \mathbf{x}}\mathbf{u} + \frac{1}{\rho}\frac{\partial W}{\partial w}\left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\right)^T \mathbf{p} - \frac{1}{\rho}\frac{\partial W}{\partial w}\nabla w \\
=& -\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\mathbf{u} - \nabla\left(\frac{\partial W}{\partial \rho}\right) - \frac{1}{\rho^2}\frac{\partial W}{\partial w}\operatorname{div}(\rho\mathbf{u})\,\mathbf{p} + \frac{1}{\rho}\operatorname{div}\left(\frac{\partial W}{\partial w}\mathbf{u}\right)\mathbf{p} \\
&+ \frac{1}{\rho}\frac{\partial W}{\partial w}\frac{\partial \mathbf{p}}{\partial \mathbf{x}}\mathbf{u} + \frac{1}{\rho}\frac{\partial W}{\partial w}\left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\right)^T \mathbf{p} \\
=& -\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\mathbf{u} - \nabla\left(\frac{\partial W}{\partial \rho}\right) + \frac{\partial W}{\partial w}\left(\nabla\left(\frac{1}{\rho}\right)\cdot\mathbf{u}\right)\mathbf{p} - \frac{1}{\rho}\frac{\partial W}{\partial w}\operatorname{div}(\mathbf{u})\,\mathbf{p} \\
&+ \frac{1}{\rho}\operatorname{div}\left(\frac{\partial W}{\partial w}\mathbf{u}\right)\mathbf{p} + \frac{1}{\rho}\frac{\partial W}{\partial w}\nabla(\mathbf{p}\cdot\mathbf{u}) \\
=& -\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\mathbf{u} - \nabla\left(\frac{\partial W}{\partial \rho}\right) + \left(\frac{\partial W}{\partial w}\nabla\left(\frac{1}{\rho}\right)\cdot\mathbf{u}\right)\mathbf{p} + \left(\frac{1}{\rho}\nabla\left(\frac{\partial W}{\partial w}\right)\cdot\mathbf{u}\right)\mathbf{p} + \frac{1}{\rho}\frac{\partial W}{\partial w}\nabla(\mathbf{p}\cdot\mathbf{u}) \\
=& -\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\mathbf{u} - \nabla\left(\frac{\partial W}{\partial \rho}\right) + \left(\nabla\left(\frac{1}{\rho}\frac{\partial W}{\partial w}\right)\cdot\mathbf{u}\right)\mathbf{p} + \frac{1}{\rho}\frac{\partial W}{\partial w}\nabla(\mathbf{p}\cdot\mathbf{u})
\end{aligned}
$$

We calculate now $\mathrm{curl}\,(\mathbf{K}) \wedge \mathbf{u}$ as follows

$$\mathrm{curl}\,(\mathbf{K}) \wedge \mathbf{u} = \left( \frac{\partial \mathbf{K}}{\partial \mathbf{x}} - \left( \frac{\partial \mathbf{K}}{\partial \mathbf{x}} \right)^T \right) \mathbf{u}$$

Which gives :

$$\mathrm{curl}\,(\mathbf{K}) \wedge \mathbf{u} = \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \mathbf{u} - \left( \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)^T \mathbf{u} - \left( \nabla \left( \frac{1}{\rho} \frac{\partial W}{\partial w} \right) \cdot \mathbf{u} \right) \mathbf{p} + (\mathbf{p} \cdot \mathbf{u}) \nabla \left( \frac{1}{\rho} \frac{\partial W}{\partial w} \right)$$

Hence, summing both equations yields :

$$\frac{\partial \mathbf{K}}{\partial t} + \mathrm{curl}\,(\mathbf{K}) \wedge \mathbf{u} = - \left( \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)^T \mathbf{u} - \nabla \left( \frac{\partial W}{\partial \rho} \right) + \frac{1}{\rho} \frac{\partial W}{\partial w} \nabla (\mathbf{p} \cdot \mathbf{u}) + (\mathbf{p} \cdot \mathbf{u}) \nabla \left( \frac{1}{\rho} \frac{\partial W}{\partial w} \right)$$

Finally, we obtain Lamb's form of the momentum equation :

$$\frac{\partial \mathbf{K}}{\partial t} + \mathrm{curl}\,(\mathbf{K}) \wedge \mathbf{u} + \nabla \left( \frac{1}{2} |\mathbf{u}|^2 - \frac{1}{\rho} \frac{\partial W}{\partial w} \mathbf{p} \cdot \mathbf{u} + \frac{\partial W}{\partial \rho} \right) = 0$$

Applying the curl operator to this equation we get :

$$\boldsymbol{\Omega}_t + \mathrm{curl}\,(\boldsymbol{\Omega} \wedge \mathbf{u}) = 0, \tag{A.14}$$

where $\boldsymbol{\Omega} = \mathrm{curl}\,(\mathbf{K})$ is called a *generalized vorticity* [37]. The previous equation implies that if $\boldsymbol{\Omega} = 0$ initially then it remains 0 for all times.

## A.5 Derivation of the dispersion relation

We consider the pde :

$$\frac{\partial \mathbf{U}}{\partial t} + \mathbf{A}(\mathbf{U}) \frac{\partial \mathbf{U}}{\partial x} = \mathbf{S}(\mathbf{U}) \tag{A.15}$$

Here, the vector $\mathbf{U} = (\rho, u, w, p, \eta)$ is the vector of state variables. The order is chosen as such only because it seemed easier to compute eigenvalues in this basis. Next we assume a constant equilibrium state defined by $\mathbf{U}_0 = (\rho_0, 0, 0, 0, \rho_0)$ where $\rho_0$ is arbitrary and we consider a small monochromatic perturbation of this state in the $x$ direction so that :

$$\mathbf{U} = \mathbf{U}_0 + \varepsilon \mathbf{U}_1 \exp(ikx - i\omega t) \tag{A.16}$$

We plug this expression of $\mathbf{U}$ into the previous pde and we linearize around the equilibrium state $\mathbf{U}_0$, that is we neglect powers of $\varepsilon$ that are quadratic or higher. We obtain the following equation :

$$-i\omega \mathbf{U}_1 + ik\mathbf{A}(\mathbf{U}_0)\mathbf{U}_1 = (\nabla \mathbf{S}(\mathbf{U}_0))\,\mathbf{U}_1 \tag{A.17}$$

where :

$$\mathbf{A}(\mathbf{U}_0) = \begin{pmatrix} 0 & \rho & 0 & 0 & 0 \\ 2\epsilon'(\rho) + \rho\epsilon''(\rho) + \frac{1}{\alpha\rho} & 0 & 0 & 0 & -\frac{1}{\alpha\rho} \\ 0 & 0 & 0 & -\frac{K(\rho)}{\beta\rho} & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \nabla(\mathbf{S}(\mathbf{U}_0) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \frac{1}{\alpha\beta\rho^2} & 0 & 0 & 0 & \frac{-1}{\alpha\beta\rho^2} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

Thus we obtain :

$$\left(-c_p\mathbf{Id} + \mathbf{A}(\mathbf{U}_0) - \frac{i}{k}\left(\nabla\mathbf{S}(\mathbf{U}_0)\right)\right)\mathbf{U}_1 = 0, \quad \text{where } c_p = \omega/k \tag{A.18}$$

which implies :

$$\det\left(-c_p\mathbf{Id} + \mathbf{A}(\mathbf{U}_0) - \frac{i}{k}\left(\nabla\mathbf{S}(\mathbf{U}_0)\right)\right) = 0 \tag{A.19}$$

The last equation means that $c_p$ is an eigenvalue of the matrix $\mathbf{A}(\mathbf{U}_0) - \frac{i}{k}\nabla\mathbf{S}(\mathbf{U}_0)$ :

$$\mathbf{A}(\mathbf{U}_0) - \frac{i}{k}\nabla\mathbf{S}(\mathbf{U}_0) = \begin{pmatrix} 0 & \rho & 0 & 0 & 0 \\ 2\epsilon'(\rho) + \rho\epsilon''(\rho) + \frac{1}{\alpha\rho} & 0 & 0 & 0 & -\frac{1}{\alpha\rho} \\ -\frac{i}{k\alpha\beta\rho^2} & 0 & 0 & -\frac{K(\rho)}{\beta\rho} & \frac{i}{k\alpha\beta\rho^2} \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & -i/k & 0 & 0 \end{pmatrix} \tag{A.20}$$

Thus one obtains the characteristic polynomial :

$$-c_p\left(c_p^4 - B(k)c_p^2 + C(k)\right) = 0$$

where $B(k)$ and $C(k)$ are the same as in section 1.4, which we remind here :

$$B(k) = \left(\frac{1}{\alpha} + \frac{1}{k^2\alpha\beta\rho^2} + \frac{K(\rho)}{\beta\rho} + 2\rho e'(\rho) + \rho^2 e''(\rho)\right)$$

$$C(k) = \left(\frac{K(\rho)}{\alpha\beta\rho} + (2\rho e'(\rho) + \rho^2 e''(\rho))\left(\frac{K(\rho)}{\beta\rho} + \frac{1}{k^2\alpha\beta\rho^2}\right)\right)$$

The eigenvalues are therefore given by :

$$c_p = 0, \quad (c_p(k)^{\pm})^2 = \frac{1}{2}\left(B(k) \pm \sqrt{B(k)^2 - 4C(k)}\right) \tag{A.21}$$

The phase velocities are all real since the augmented E-K system admits no inherent dissipation. In fact, a trivial computation shows that :

$$B(k)^2 - 4C(k) = \left(\frac{1}{\alpha} + \frac{1}{k^2\alpha\beta\rho^2} + \frac{K(\rho)}{\beta\rho} + 2\rho e'(\rho) + \rho^2 e''(\rho)\right)^2 + \frac{4}{\alpha^2\beta\rho^2k^2} > 0$$

## A.6 IMEX scheme : General explicit form

We use the same IMEX scheme presented in section 2.4.2. Nevertheless, we detail its explicit formulation that does not require the risky use of root-finding algorithms. We use here the generic case, in which we put no additional source terms. In this setting, the conserved variables, fluxes and sources terms are respectively given by :

$$\mathbf{U} = \begin{pmatrix} \rho \\ \rho u \\ \rho \eta \\ \rho w \\ p \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} \rho u \\ \rho u^2 + \rho W_\rho - W + p W_p \\ \rho \eta u \\ \rho w u - W_p \\ p u - w \end{pmatrix}, \quad \mathbf{S} = \begin{pmatrix} 0 \\ 0 \\ \rho w \\ (1 - \eta/\rho)/\alpha\beta \\ 0 \end{pmatrix} \tag{A.22}$$

We recall that at each time step $t^n$, the two stages of the scheme write as follows :

$$\mathbf{U}^\star = \mathbf{U}^n - \gamma \frac{\Delta t}{\Delta x} \left( F^n_{i+\frac{1}{2}} - F^n_{i-\frac{1}{2}} \right) + \gamma \Delta t \mathbf{S}(\mathbf{U}^\star)$$
$$\mathbf{U}^{n+1} = \mathbf{U}^n - (\gamma - 1) \frac{\Delta t}{\Delta x} \left( F^n_{i+\frac{1}{2}} - F^n_{i-\frac{1}{2}} \right) - (2 - \gamma) \frac{\Delta t}{\Delta x} \left( F^\star_{i+\frac{1}{2}} - F^\star_{i-\frac{1}{2}} \right) + (1 - \gamma) \Delta t S(\mathbf{U}^\star) + \gamma \Delta t S(\mathbf{U}^{n+1})$$

with $\gamma = \frac{\sqrt{2}}{2} - 1$ a constant parameter. The vector of intercell fluxes is computed as usual using the appropriate Riemann solver of choice. For all what follows, we shall denote the intercell fluxes by $\mathbf{F^n_{i+\frac{1}{2}}} = (f^n_{\rho_{i+\frac{1}{2}}}, f^n_{\rho u_{i+\frac{1}{2}}}, f^n_{\rho \eta_{i+\frac{1}{2}}}, f^n_{\rho w_{i+\frac{1}{2}}}, f^n_{p_{i+\frac{1}{2}}})^T$. With these notations, the let us detail how the scheme writes. The first stage of the scheme writes :

$$\begin{cases} \rho^\star = \rho^n - \gamma \frac{\Delta t}{\Delta x} \left( f^n_{\rho_{i+\frac{1}{2}}} - f^n_{\rho_{i-\frac{1}{2}}} \right) \\ u^\star = \left( \rho^n u^n - \gamma \frac{\Delta t}{\Delta x} \left( f^n_{\rho u_{i+\frac{1}{2}}} - f^n_{\rho u_{i-\frac{1}{2}}} \right) \right) / \rho^\star \\ \eta^\star = \left( \rho^n \eta^n - \gamma \frac{\Delta t}{\Delta x} \left( f^n_{\rho \eta_{i+\frac{1}{2}}} - f^n_{\rho \eta_{i-\frac{1}{2}}} \right) + \gamma \Delta t \rho^\star w^\star \right) / \rho^\star \\ w^\star = \left( \rho^n w^n - \gamma \frac{\Delta t}{\Delta x} \left( f^n_{\rho w_{i+\frac{1}{2}}} - f^n_{\rho w_{i-\frac{1}{2}}} \right) + \gamma \frac{\Delta t}{\alpha \beta} (1 - \eta^\star / \rho^\star) \right) / \rho^\star \\ p^\star = \rho^n u^n - \gamma \frac{\Delta t}{\Delta x} \left( f^n_{\rho u_{i+\frac{1}{2}}} - f^n_{\rho u_{i-\frac{1}{2}}} \right) \end{cases} \tag{A.23}$$

Thus, we first compute $\rho^\star$ and $p^\star$ since their expressions are explicit. Then we use the value of $\rho^\star$ to compute $u^\star$. It remains to solve the implicit part for $\eta^\star$ and $w^\star$. we denote by $Y^n$ and $Z^n$ the quantities :

$$\begin{cases} Y^n = \left( \rho^n \eta^n - \gamma \frac{\Delta t}{\Delta x} \left( f^n_{\rho \eta_{i+\frac{1}{2}}} - f^n_{\rho \eta_{i-\frac{1}{2}}} \right) \right) / \rho^\star \\ Z^n = \left( \rho^n w^n - \gamma \frac{\Delta t}{\Delta x} \left( f^n_{\rho w_{i+\frac{1}{2}}} - f^n_{\rho w_{i-\frac{1}{2}}} \right) \right) / \rho^\star \end{cases} \tag{A.24}$$

so that the expressions of $\eta^\star$ and $w^\star$ reduce to :

$$\begin{cases} \eta^\star = Y^n + \gamma \Delta t \omega^\star \\ w^\star = Z^n + \frac{\gamma \Delta t}{\alpha \beta \rho^\star} \left( 1 - \frac{\eta^\star}{\rho^\star} \right) \end{cases} \tag{A.25}$$

Solving this linear system yields :

$$\begin{cases} \eta^\star = \frac{1}{1+\frac{\gamma^2\Delta t^2}{\alpha\beta\rho^{\star 2}}}\left(Y^n + \gamma\Delta t Z^n + \frac{\gamma^2\Delta t^2}{\alpha\beta\rho^\star}\right) \\ w^\star = \frac{1}{1+\frac{\gamma^2\Delta t^2}{\alpha\beta\rho^{\star 2}}}\left(Z^n - \frac{\gamma\Delta t}{\alpha\beta\rho^{\star 2}}Y^n + \frac{\gamma\Delta t}{\alpha\beta\rho^\star}\right) \end{cases} \tag{A.26}$$

which concludes the first step of the scheme. After using the obtained values of the vector $\mathbf{U}^\star$ to calculate new intercell fluxes $\mathbf{F}^\star_{\mathbf{i+\frac{1}{2}}} = (f^\star_{\rho_{i+\frac{1}{2}}}, f^\star_{\rho u_{i+\frac{1}{2}}}, f^\star_{\rho\eta_{i+\frac{1}{2}}}, f^\star_{\rho w_{i+\frac{1}{2}}}, f^\star_{p_{i+\frac{1}{2}}})^T$, we proceed to the second step which writes :

$$\begin{cases} \rho^{n+1} = \rho^n - (\gamma-1)\frac{\Delta t}{\Delta x}\left(f^n_{\rho_{i+\frac{1}{2}}} - f^n_{\rho_{i-\frac{1}{2}}}\right) + (\gamma-2)\frac{\Delta t}{\Delta x}\left(f^\star_{\rho_{i+\frac{1}{2}}} - f^\star_{\rho_{i-\frac{1}{2}}}\right) \\ u^{n+1} = \left(\rho^n u^n - (\gamma-1)\frac{\Delta t}{\Delta x}\left(f^n_{\rho u_{i+\frac{1}{2}}} - f^n_{\rho u_{i-\frac{1}{2}}}\right) + (\gamma-2)\frac{\Delta t}{\Delta x}\left(f^\star_{\rho u_{i+\frac{1}{2}}} - f^\star_{\rho u_{i-\frac{1}{2}}}\right)\right)/\rho^{n+1} \\ \eta^{n+1} = \left(\rho^n \eta^n - (\gamma-1)\frac{\Delta t}{\Delta x}\left(f^n_{\rho\eta_{i+\frac{1}{2}}} - f^n_{\rho\eta_{i-\frac{1}{2}}}\right) + (\gamma-2)\frac{\Delta t}{\Delta x}\left(f^\star_{\rho\eta_{i+\frac{1}{2}}} - f^\star_{\rho\eta_{i-\frac{1}{2}}}\right)\right)/\rho^{n+1} \\ \qquad + \left((1-\gamma)\Delta t\rho^\star w^\star + \gamma\Delta t\rho^{n+1}w^{n+1}\right)/\rho^{n+1} \\ w^{n+1} = \left(\rho^n w^n - (\gamma-1)\frac{\Delta t}{\Delta x}\left(f^n_{\rho w_{i+\frac{1}{2}}} - f^n_{\rho w_{i-\frac{1}{2}}}\right) + (\gamma-2)\frac{\Delta t}{\Delta x}\left(f^\star_{\rho w_{i+\frac{1}{2}}} - f^\star_{\rho w_{i-\frac{1}{2}}}\right)\right)/\rho^{n+1} \\ \qquad + \left((1-\gamma)\frac{\Delta t}{\alpha\beta}\left(1-\eta^\star/\rho^\star\right) + \gamma\frac{\Delta t}{\alpha\beta}\left(1-\eta^{n+1}/\rho^{n+1}\right)\right)/\rho^{n+1} \\ p^\star = \rho^n u^n - \gamma\frac{\Delta t}{\Delta x}\left(f^n_{\rho u_{i+\frac{1}{2}}} - f^n_{\rho u_{i-\frac{1}{2}}}\right) \end{cases} \tag{A.27}$$

Solving the implicit part goes exactly as in the first steps, since we can write :

$$\begin{cases} \eta^{n+1} = Y^\star + \gamma\Delta t\omega^{n+1} \\ w^{n+1} = Z^\star + \frac{\gamma\Delta t}{\alpha\beta\rho^\star}\left(1 - \frac{\eta^{n+1}}{\rho^{n+1}}\right) \end{cases} \tag{A.28}$$

with $Y^\star$ and $Z^\star$ defined by :

$$\begin{cases} Y^\star = \left(\rho^n\eta^n - (\gamma-1)\frac{\Delta t}{\Delta x}\left(f^n_{\rho\eta_{i+\frac{1}{2}}} - f^n_{\rho\eta_{i-\frac{1}{2}}}\right) + (\gamma-2)\frac{\Delta t}{\Delta x}\left(f^\star_{\rho\eta_{i+\frac{1}{2}}} - f^\star_{\rho\eta_{i-\frac{1}{2}}}\right) + (1-\gamma)\Delta t\rho^\star w^\star\right)/\rho^{n+1} \\ Z^\star = \left(\rho^n w^n - (\gamma-1)\frac{\Delta t}{\Delta x}\left(f^n_{\rho w_{i+\frac{1}{2}}} - f^n_{\rho w_{i-\frac{1}{2}}}\right) + (\gamma-2)\frac{\Delta t}{\Delta x}\left(f^\star_{\rho w_{i+\frac{1}{2}}} - f^\star_{\rho w_{i-\frac{1}{2}}}\right) + (1-\gamma)\frac{\Delta t}{\alpha\beta}\left(1-\eta^\star/\rho^\star\right)\right)/\rho^{n+1} \end{cases}$$

This reduces the explicit expressions of $\eta^{n+1}$ and $w^{n+1}$ to :

$$\begin{cases} \eta^{n+1} = \frac{1}{1+\frac{\gamma^2\Delta t^2}{\alpha\beta\rho^{\star 2}}}\left(Y^\star + \gamma\Delta t Z^\star + \frac{\gamma^2\Delta t^2}{\alpha\beta\rho^\star}\right) \\ w^{n+1} = \frac{1}{1+\frac{\gamma^2\Delta t^2}{\alpha\beta\rho^{\star 2}}}\left(Z^\star - \frac{\gamma\Delta t}{\alpha\beta\rho^{\star 2}}Y^\star + \frac{\gamma\Delta t}{\alpha\beta\rho^\star}\right) \end{cases} \tag{A.29}$$

APPENDIX B

# About Elliptic integrals and functions

While usual trigonometric functions can be defined as coordinates with respect to the unit circle, elliptic functions can be seen as the generalization of the latter for other conics, particularly ellipses in our case. Historically, they owe their existence to Carl Gustav Jakob Jacobi [47], who introduced them in 1829, as inverse functions to elliptic integrals. This appendix is developed in an attempt to summarize the main properties of the canonical elliptic functions sn, cn and dn. This task naturally requires notions from the theory of elliptic integrals which we first recall. A complete example providing details on how to compute exact periodic solution to NLSE is presented afterward as a reference to demonstrate the practical usage of such functions.

## B.1 Elliptic integrals

We consider the Lagrange forms of elliptic integrals. The incomplete elliptic integral of the first kind is defined by :

$$F(\varphi, s) = \int_0^{\varphi} \frac{d\theta}{\sqrt{1 - s^2 \sin^2(\theta)}} \tag{B.1}$$

where $0 < s < 1$ is a parameter usually referred to as the *elliptic modulus*. Similarly, the incomplete elliptic integral of the second kind is defined by :

$$E(\varphi, s) = \int_0^{\varphi} \sqrt{1 - s^2 \sin^2(\theta)}\, d\theta \tag{B.2}$$

For a fixed value of $\varphi = \pi/2$, these integrals become functions of only $s$, called the complete integrals of first and second kind, respectively :

$$K(s) = \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - s^2 \sin^2(\theta)}} \tag{B.3}$$

$$E(s) = \int_0^{\pi/2} \sqrt{1 - s^2 \sin^2(\theta)}\, d\theta \tag{B.4}$$

The asymptotic expansions of both complete integrals is given below in the vicinity of the limiting values. For $s \to 0$ we have :

$$K(s) = \frac{\pi}{2}\left(1 + \frac{s^2}{4} + \frac{9s^2}{64}\right) + \cdots \tag{B.5}$$

$$E(s) = \frac{\pi}{2}\left(1 - \frac{s^2}{4} + \frac{3s^2}{64}\right) + \cdots \tag{B.6}$$

and for $s \to 1$ :

$$K(s) = \ln\left(\frac{4}{\sqrt{1-s^2}}\right) + \frac{1}{4}\left(\ln\frac{4}{\sqrt{1-s^2}} - 1\right)(1-s^2) + \cdots \tag{B.7}$$

$$E(s) = 1 + \frac{1}{2}\left(\ln\left(\frac{4}{\sqrt{1-s^2}}\right) - \frac{1}{2}\right)(1-s^2) + \cdots \tag{B.8}$$

For more insight, we plot the graphs of both functions in the figure below :

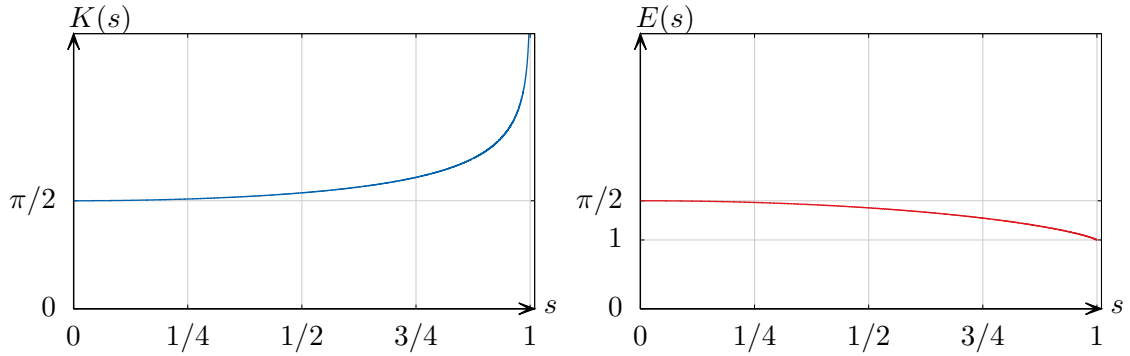


Figure B.1: Plot of the overall behavior of the complete elliptic integrals of first (left) and second (right) kind.

## B.2 Jacobi's elliptic functions

Let us first introduce the *amplitude* function. For a fixed value of the *elliptic modulus* $s$, we can consider the inverse function to $F(\varphi, s)$ denoted as:

$$\mathrm{am}(u, s) = \varphi \tag{B.9}$$

This allows us to define the elliptic functions in terms of $\mathrm{am}(u, s)$, by :

$$\mathrm{sn}(u, s) = \sin(\mathrm{am}(u, s)) \tag{B.10}$$

$$\mathrm{cn}(u, s) = \cos(\mathrm{am}(u, s)) \tag{B.11}$$

$$\mathrm{dn}(u, s) = \sqrt{1 - s^2 \sin^2(\mathrm{am}(u, s))} \tag{B.12}$$

These functions are called respectively *elliptic sine*, *elliptic cosine* and *delta amplitude*. The definition we use does not provide an explicit expression of any of the three functions since the amplitude function am is defined implicitly. Therefore, we will try to summarize the main properties.

1. **Periodicity** : For any fixed value of $0 < s < 1$, *sn* and *cn* are periodic in terms of the real variable $u$, of period $4K(s)$. *dn* is also periodic in the same setting, of period $2K(s)$.

2. **Limiting behaviors** :

$$\begin{aligned} \mathrm{cn}(u,s)|_{s\to 0} &= \cos(u) & \mathrm{cn}(u,s)|_{s\to 1} &= 1/\cosh(u) \\ \mathrm{sn}(u,s)|_{s\to 0} &= \sin(u) & \mathrm{sn}(u,s)|_{s\to 1} &= \tanh(u) \\ \mathrm{dn}(u,s)|_{s\to 0} &= 1 & \mathrm{dn}(u,s)|_{s\to 1} &= 1/\cosh(u) \end{aligned}$$

3. **Useful Relations**

$$\mathrm{cn}^2(u,s) + \mathrm{sn}^2(u,s) = 1 \tag{B.13}$$
$$\mathrm{dn}^2(u,s) + s^2\mathrm{sn}^2(u,s) = 1 \tag{B.14}$$

4. **Differential equations**

$$\mathrm{cn}: \quad y'^2 = (1-y^2)(1-s^2+s^2y^2) \tag{B.15}$$
$$\mathrm{sn}: \quad y'^2 = (1-y^2)(1-s^2y^2) \tag{B.16}$$
$$\mathrm{dn}: \quad y'^2 = (1-y^2)(y^2-1+s^2) \tag{B.17}$$

5. **Graphs**

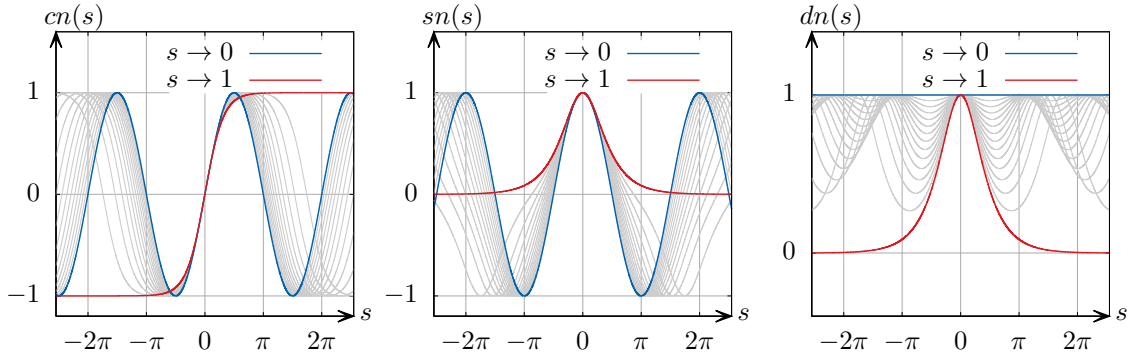


Figure B.2: Overall behavior of the Jacobi elliptic functions for different values of the elliptic modulus $s$. Represented from left to right are sn, cn and dn. For each plot, the blue and red line represent the limiting behaviors for $s \to 0$ and $s \to 1$ respectively. the gray lines correspond to arbitrary values of $0 < s < 1$.

## B.3 Practical example: Periodic solution of NLS equation

We established in section 2.2 that a periodic travelling wave solution to NLSE satisfies the differential equation :

$$\left(\frac{d\rho}{d\xi}\right)^2 = 4(\rho - b_1)(\rho - b_2)(\rho - b_3) = P(\rho) \tag{B.18}$$

with $b_3 < b_2 < b_1$. As long as these roots are simple (*i.e.* $P'(b_i) \neq 0$), equation (B.18) can have a real-valued periodic solution corresponding to the oscillations of $\rho$ between $b_2$ and $b_3$, given that $P(u) \geq 0$ in this interval, as shown in the figure below:

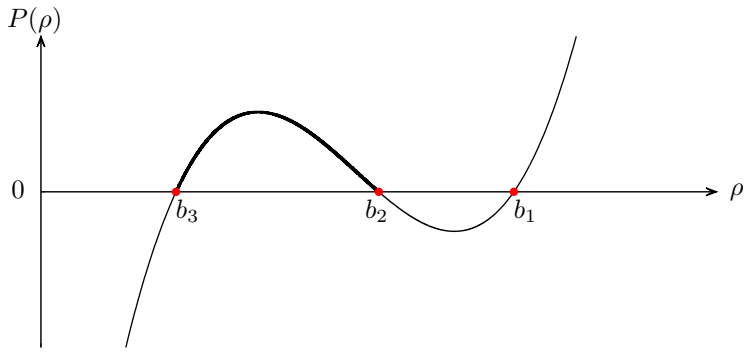


Figure B.3: Representative graphic of $P(\rho)$ for arbitrary parameters $b_1 > b_2 > b_3$.

In this setting, the wavelength of $\rho(\xi)$ corresponds to the distance it takes $\rho$ to move back and forth from $b_3$ to $b_2$. Let us denote by $\xi_3$ and $\xi_2$ the values of $\xi$ such that :

$$\rho(\xi_3) = b_3; \quad \rho(\xi_2) = b_2. \tag{B.19}$$

While $\xi_3$ is arbitrary, we take $\xi_2$ as the smallest value of $\xi$ verifying $\xi_2 > \xi_3$. Let $\xi \in [\xi_3, \xi_2]$. Then, it follows from equation (B.18) that:

$$\xi - \xi_3 = \int_{\xi_3}^{\xi} d\xi = \int_{b_3}^{\rho} \frac{d\rho}{\rho_\xi} = \int_{b_3}^{\rho} \frac{d\rho}{2\sqrt{(\rho - b_1)(\rho - b_2)(\rho - b_3)}} \tag{B.20}$$

Now, we consider the classical change of variables :

$$\rho = b_3 - (b_3 - b_2)\sin^2(\varphi); \quad \varphi \in [0, \pi/2] \tag{B.21}$$

$$d\rho = 2(b_2 - b_3)\sin(\varphi)\cos(\varphi)d\varphi \tag{B.22}$$

the factors inside the square root become :

$$\rho - b_1 = -(b_1 - b_3)\left(1 - s^2\sin^2(\varphi)\right) \tag{B.23}$$

$$\rho - b_2 = -(b_2 - b_3)(1 - \sin^2(\varphi)) = -(b_2 - b_3)\cos^2(\varphi) \tag{B.24}$$

$$\rho - b_3 = (b_2 - b_3)\sin^2(\varphi) \tag{B.25}$$

where $s = \sqrt{\dfrac{b_2 - b_3}{b_1 - b_3}}$. Under these variables, we can write :

$$\int_{b_3}^{\rho} \frac{d\rho}{2\sqrt{(\rho - b_1)(\rho - b_2)(\rho - b_3)}} = \int_0^{\varphi} \frac{(b_2 - b_3)\sin(\varphi)\cos(\varphi)d\varphi}{\sqrt{(\rho - b_1)(\rho - b_2)(\rho - b_3)}} \tag{B.26}$$

$$= \int_0^{\varphi} \frac{d\varphi}{\sqrt{(b_1 - b_3)\left(1 - s^2\sin^2(\varphi)\right)}} \tag{B.27}$$

$$= \frac{1}{\sqrt{b_1 - b_3}} F(\varphi, s) \tag{B.28}$$

So that :

$$\sqrt{b_1 - b_3}(\xi - \xi_3) = F(\varphi, s) \implies \mathrm{am}\left(\sqrt{b_1 - b_3}(\xi - \xi_3), s\right) = \varphi \tag{B.29}$$

Thus :

$$\rho = b_3 - (b_3 - b_2)\sin^2(\varphi) = b_3 - (b_3 - b_2)\mathrm{sn}^2\left(\sqrt{b_1 - b_3}(\xi - \xi_3), s\right) \tag{B.30}$$

Or equivalently :

$$\rho(\xi) = b_1 - (b_1 - b_3)\mathrm{dn}^2\left(\sqrt{b_1 - b_3}(\xi - \xi_3), s\right) \tag{B.31}$$

Since this is a travelling wave solution, we can take $\xi_3 = 0$ to obtain the solution :

$$\rho(\xi) = b_1 - (b_1 - b_3)\mathrm{dn}^2\left(\xi\sqrt{b_1 - b_3}, s\right) \tag{B.32}$$

## B.4 Asymptotic structure of a DSW : Values of $\tau_i$

We recall that we consider the case of an initial discontinuity that leads to a DSW propagating to the right and a rarefaction wave propagating to the left. At the leading edge of the DSW ($s^2 = 0$) we have :

$$r_1 = r_2 = u_0 + 2\sqrt{\rho_0}; \quad r_3 = u_R + 2\sqrt{\rho_R}; \quad r_4 = u_R - 2\sqrt{\rho_R} \tag{B.33}$$

In this setting, and inside of the DSW region, the merging of the Riemann invariants $r_1$ and $r_2$ trivially implies the equality of their corresponding characteristic velocities :

$$\lim_{s^2 \to 0} V_1(s) = \lim_{s^2 \to 0} V_2(s) = \lim_{s^2 \to 0} \tau(s) = \tau_1. \tag{B.34}$$

Referring back to the expressions of the characteristic velocities, let us calculate $\lim\limits_{s^2 \to 0} V_1(s)$ :

$$\lim_{s^2 \to 0} V_1(s) = \lim_{s^2 \to 0} \left( U + \frac{1}{2}(r_1 - r_2)\frac{1}{1 - \frac{r_2 - r_4}{r_1 - r_4}\frac{E(s)}{K(s)}} \right)$$

Besides we have :

$$U = \frac{1}{4}(r_1 + r_2 + r_3 + r_4) = \frac{1}{4}(2r_1 + u_R + 2\sqrt{\rho_R} + u_R - 2\sqrt{\rho_R}) = \frac{1}{2}(r_1 + u_R)$$

$$r_1 - r_2 = \frac{(r_1 - r_3)(r_2 - r_4)}{r_3 - r_4} s^2$$

Thus :

$$\begin{aligned}
\lim_{s^2\to 0} V_1(s) &= \lim_{s^2\to 0}\left(U + \frac{1}{2}(r_1 - r_2)\frac{(r_1 - r_4)K(s)}{(r_1 - r_4)K(s) - (r_2 - r_4)E(s)}\right) \\
&= \lim_{s^2\to 0}\left(U + \frac{1}{2}(r_1 - r_2)\frac{(r_1 - r_4)(1 + s^2/4)}{(r_1 - r_4)(1 + s^2/4) - (r_2 - r_4)(1 - s^2/4)}\right) \\
&= \lim_{s^2\to 0}\left(U + \frac{1}{2}(r_1 - r_2)\frac{(r_1 - r_4)(1 + s^2/4)}{r_1 - r_2 + \frac{s^2}{4}(r_1 - 2r_4 + r_2)}\right) \\
&= \lim_{s^2\to 0}\left(U + \frac{1}{2}\frac{(r_1 - r_3)(r_2 - r_4)}{r_3 - r_4} s^2 \frac{(r_1 - r_4)(1 + s^2/4)}{\frac{(r_1 - r_3)(r_2 - r_4)}{r_3 - r_4} s^2 + \frac{s^2}{4}(r_1 - 2r_4 + r_2)}\right) \\
&= U + \frac{(r_1 - r_4)(r_1 - r_3)(r_2 - r_4)}{2(r_1 - r_3)(r_2 - r_4) + (r_2 - r_4)(r_3 - r_4)} = U + \frac{(r_1 - r_4)(r_1 - r_3)}{2(r_1 - r_3) + (r_3 - r_4)}
\end{aligned}$$

Substituting the values of the Riemann invariants and using the equality $u_R - 2\sqrt{\rho_R} = u_0 - 2\sqrt{\rho_0}$, we finally obtain :

$$\tau_1 = \lim_{s^2\to 0} V_1(s) = u_R + \frac{8\rho_0 - 8\sqrt{\rho_0\rho_R} + \rho_R}{2\sqrt{\rho_0} - \sqrt{\rho_R}} \tag{B.35}$$

In a similar fashion, we can calculate the asymptotic value $\tau_2$ which corresponding to the trailing edge ($s^2 = 1$). In this setting we have :

$$r_2 = r_3 = u_R + 2\sqrt{\rho_R}, \quad V_2 = V_3, \quad r_1 = u_0 + 2\sqrt{\rho_0}, \quad r_4 = u_0 - 2\sqrt{\rho_0}. \tag{B.36}$$

In this case, the Riemann invariants who merge are $r_2$ and $r_3$ and thus :

$$\lim_{s^2\to 1} V_3(s) = \lim_{s^2\to 1} V_2(s) = \lim_{s^2\to 1} \tau(s) = \tau_2. \tag{B.37}$$

and we have :

$$r_2 - r_3 = \frac{(r_1 - r_3)(r_2 - r_4)}{(r_1 - r_4)}(1 - s^2) \tag{B.38}$$

$$\begin{aligned}
\lim_{s^2\to 1} V_3(s) &= \lim_{s^2\to 1}\left(U + \frac{1}{2}(r_3 - r_4)\frac{1}{1 - \frac{r_2 - r_4}{r_2 - r_3}\frac{E(s)}{K(s)}}\right) \\
&= \lim_{s^2\to 1}\left(U + \frac{1}{2}(r_3 - r_4)\frac{1}{1 - \frac{r_2 - r_4}{r_2 - r_3}\frac{1}{\ln(4/\sqrt{1-s^2})}}\right) \\
&= \lim_{s^2\to 1}\left(U + \frac{1}{2}(r_3 - r_4)\frac{1}{1 - \frac{r_1 - r_4}{(r_1 - r_3)(1-s^2)\ln(4/\sqrt{1-s^2})}}\right)
\end{aligned}$$

It might be more intuitive to consider the variable $s'^2 = 1 - s^2$, in which case, the previous equality yields :

$$\lim_{s'^2 \to 0} V_3(s') = \lim_{s'^2 \to 0} \left( U + \frac{1}{2}(r_3 - r_4) \frac{s'^2 \left( \ln(4) - \frac{1}{2}\ln(s'^2) \right)}{s'^2 \left( \ln(4) - \frac{1}{2}\ln(s'^2) \right) - \frac{r_1 - r_4}{r_1 - r_3}} \right)$$

since $\lim_{s'^2 \to 0} \left( s'^2 \left( \ln(4) - \frac{1}{2}\ln(s'^2) \right) \right) = 0$ then we get :

$$\tau_2 = \lim_{s'^2 \to 0} V_3(s') = U = \frac{1}{4}(r_1 + r_2 + r_3 + r_4) = u_R + \sqrt{\rho_0} \tag{B.39}$$

The values of $\tau_3$ and $\tau_4$ which delimit the rarefaction wave, are obtained exactly as in the case of Euler equations with a pressure $p(\rho) = \rho^2/2$ :

$$\tau_3 = u_R - 2\sqrt{\rho_R}, \quad \tau_4 = u_L - 2\sqrt{\rho_L} \tag{B.40}$$

APPENDIX C

# Asymptotics for the augmented thin film flows model

## C.1 Neutral stability analysis for thin films equations

We recall that the phase velocities of thin films equations (3.6) satisfy the equation :

$$\varepsilon(1-c_p)^2 - \frac{3i}{kRe}(1-c_p) - \varepsilon\left(\frac{cos\theta}{F^2} + \frac{2\lambda^2}{45} + \frac{k^2\varepsilon^2}{We}\right) - \frac{6i}{kRe} = 0$$

For lightness we introduce the positive constants :

$$a = \frac{\cos\theta}{F^2} + \frac{2\lambda^2}{45} + \frac{k^2\varepsilon^2}{We}, \quad b = \frac{3}{k\varepsilon Re} \tag{C.1}$$

and we consider the variable $X = 1 - c_p$. The stability condition becomes $Im(X) > 0$. Thus, under the previous notations, $X$ satisfies the equation :

$$X^2 - ibX - (a + 2ib) = 0 \tag{C.2}$$

In order to solve this second degree equation, we write

$$\Delta = -b^2 + 4(a + 2ib) \quad \Longrightarrow \quad X_{1,2} = \frac{-ic \pm \sqrt{\Delta}}{2} \tag{C.3}$$

Let us define a complex square root of $\Delta$ by taking :

$$\sqrt{\Delta} = p + iq, \quad \Delta = \sqrt{\Delta}^2 = p^2 - q^2 + 2ipq \tag{C.4}$$

so that the roots $X_{1,2}$ satisfy the equation :

$$2X_{1,2} = ib \pm (p + iq)$$

The stability condition $Im(X) > 0$ thus writes :

$$b \pm q > 0$$

Given that $b > 0$ then this condition is equivalent to :

$$b^2 > q^2$$

It remains to obtain an explicit expression of $q^2$. for that let us identify the real and imaginary parts in the equality $\Delta = \sqrt{\Delta}^2$:

$$\sqrt{\Delta}^2 = \Delta \Leftrightarrow \begin{cases} p^2 - q^2 = -b^2 + 4a \\ pq = 4b \ > 0 \Leftrightarrow p^2 = \frac{16b^2}{q^2} \end{cases}$$

Now we substitute $p^2 = 16b^2/q^2$ in the first equation and multiply by $q^2$ to obtain :

$$q^4 - (b^2 - 4a)q^2 - 16b^2 = 0 \Rightarrow q^2 = \frac{b^2 - 4a \pm \sqrt{(b^2 - 4a)^2 + 64b^2}}{2}$$

For $q \in \mathbb{R}$, we have $q^2 \geq 0$ and hence :

$$q^2 = \frac{b^2 - 4h_0 + \sqrt{(b^2 - 4ah_0)^2 + 64b^2u_0^2}}{2}$$

We substitute this expression into the stability condition $b^2 - q^2 > 0$ to obtain:

$$\begin{aligned} & b^2 + 4 - \sqrt{(b^2 - 4a)^2 + 64b^2} > 0 \\ \Leftrightarrow \quad & b^2 + 4a > \sqrt{(b^2 - 4a)^2 + 64b^2} > 0 \\ \Leftrightarrow \quad & (b^2 + 4a)^2 > (b^2 - 4a)^2 + 64b^2 \\ \Leftrightarrow \quad & 16b^2a > 64b^2 \Leftrightarrow a > 4 \end{aligned}$$

Finally:

$$\begin{aligned} a > 4 \Leftrightarrow \quad & \frac{cos\theta}{F^2} + \frac{2\lambda^2}{45} + \kappa\frac{k^2}{F^2} > 4 \\ \Leftrightarrow \quad & \frac{cos\theta}{F^2} + \frac{2}{15}\frac{Re\ \sin(\theta)}{F^2} + \kappa\frac{k^2}{F^2} > 4 \\ \Leftrightarrow \quad & cos\theta + \frac{2Re\ \sin(\theta)}{15} + \kappa k^2 > 4F^2 \\ \Leftrightarrow \quad & cos\theta + \frac{2Re\ \sin(\theta)}{15} + \kappa k^2 > \frac{4Resin\theta}{3} \\ \Leftrightarrow \quad & cotg\theta + \frac{\kappa k^2}{sin\theta} > \frac{6Re}{5} \end{aligned}$$

which concludes the development.

## C.2 Dispersion relation expansion

In this part, we show how to obtain the series expansions of the phase velocities of the augmented system for thin film flows. In order to perform an asymptotic analysis, the penalty parameters are set to be power functions of $\varepsilon$:

$$\alpha = \varepsilon^m, \quad \beta = \varepsilon^p, \qquad m, p \in \mathbb{N}^* \tag{C.5}$$

We will assume that $m \geq 1$ and $p \geq 1$, so that $\alpha$ and $\beta$ are $\mathcal{O}(\varepsilon)$. We recall that in this context, the characteristic polynomial writes :

$$\frac{1}{\alpha\beta\epsilon k^3} X Q_0(X) = 0, \quad Q_0(X) = A_0 + A_1 X + A_2 X^2 + A_3 X^3 + A_4 X^4 \tag{C.6}$$

with the coefficients :

$$A_0 = \frac{6i}{Re} + C\varepsilon k + \frac{6i}{ReWe} k^2 \varepsilon^{2+m} + \frac{1+\varepsilon^m C}{We} k^3 \varepsilon^3, \quad A_1 = \frac{3i}{Re} + \frac{3i}{ReWe} k^2 \varepsilon^{2+m}$$
$$A_2 = -\left( \frac{6i}{Re} \varepsilon^{m+p} k^2 + \left(1 + \varepsilon^p k^2 + C\varepsilon^{m+p} k^2\right) k\varepsilon + \frac{1}{We} k^3 \varepsilon^{3+m} \right), \quad A_3 = -\frac{3i\varepsilon^{m+p} k^2}{Re}, \quad A_4 = \varepsilon^{m+p+1} k^3$$

We disregard in what follows the trivial solution $X = 0$ and will focus on the roots of $Q_0(X)$. First, let us look for solutions that are compatible with a regular perturbation series of the form :

$$X = X_0 + \varepsilon X_1 + \varepsilon^2 X_2 + \varepsilon^m X_m + \varepsilon^{m+1} X_{m+1} + \varepsilon^{2m} X_{2m} + \varepsilon^p X_p + \varepsilon^{p+1} X_{p+1} + \varepsilon^{2p} X_{2p} + \varepsilon^{m+p} X_{m+p} + \mathcal{O}(\varepsilon^3) \tag{C.7}$$

Such an expansion only reveals the roots $X$ which remain bounded in the limit $\varepsilon \to 0$. Since we assumed that $m, p \geq 1$, all the combinations that may be of order $\mathcal{O}(\varepsilon^3)$ are used in the expansion. Replacing (C.7) in the characteristic polynomial (3.35) and setting to zero the

leading terms in $\varepsilon$ yields:

$$\begin{cases} \text{Order } 0 & : \quad \dfrac{3iX_0}{Re} + \dfrac{6i}{Re} = 0 \\ \text{Order } 1 & : \quad Ck - kX_0^2 + \dfrac{3iX_1}{Re} = 0 \\ \text{Order } 2 & : \quad -2kX_0X_1 + \dfrac{3iX_2}{Re} = 0 \\ \text{Order } m & : \quad \dfrac{3iX_m}{Re} = 0 \\ \text{Order } m+1 & : \quad -2kX_0X_m + \dfrac{3iX_{m1}}{Re} = 0 \\ \text{Order } 2m & : \quad \dfrac{3iX_2m}{Re} = 0 \\ \text{Order } p & : \quad \dfrac{3iX_p}{Re} = 0 \\ \text{Order } p+1 & : \quad -k^3X_0^2 - 2kX_0X_p + \dfrac{3i\text{Xp1}}{Re} = 0 \\ \text{Order } 2p & : \quad \dfrac{3iX_{2p}}{Re} = 0 \\ \text{Order } p+m & : \quad -\dfrac{3ik^2X_0^3}{Re} - \dfrac{6ik^2X_0^2}{Re} + \dfrac{3iX_{mp}}{Re} = 0 \end{cases} \tag{C.8}$$

Solving successively this set of equations gives the coefficients of the expansion as follows:

$$X_0 = -2, \quad X_1 = \frac{ikRe}{3}(C-4), \quad X_2 = -\frac{4}{9}k^2Re(C-4), \quad X_{p+1} = -\frac{4i}{3}Rek^3 \tag{C.9}$$

$$X_m = X_{2m} = X_{m+1} = X_p = X_{m+p} = X_{2p} = 0 \tag{C.10}$$

Thus, $X$ expands in a regular series as follows :

$$X = -2 + \frac{iRe}{3}\left(\frac{\cos\theta}{F^2} + \frac{2\lambda^2h^3}{45} - 4\right)k\varepsilon - \frac{4Re}{9}\left(\frac{\cos\theta}{F^2} + \frac{2\lambda^2h^3}{45} - 4\right)k^2\varepsilon^2 - \frac{4i}{3}Rek^3\varepsilon^{p+1} + \mathcal{O}(\varepsilon^3) \tag{C.11}$$

As the above computation confirms, the regular expansion (C.7) yields a unique root. Out of the coefficients that are dependent on $\alpha$ or $\beta$, only $X_{p+1}$ has a non-zero value among the concerned terms. The remaining terms correspond exactly to the series expansion of the phase velocity $c_{p_1}$ of the original system. Thus, in order keep this expansion valid up to first order, it suffices that the term $X_{p+1}\varepsilon^{p+1}$ is of second order, yielding a first constraint :

$$p \geq 1 \tag{C.12}$$

We shall now look for the remaining three roots. Trivially, $Q_0(X)$ degenerates into a first order polynomial in the limit $\varepsilon \to 0$. This means that the three remaining roots go to infinity in the same limit and only remains the obtained root of the previous expansion. These missed

solutions can be found by use of proper rescaling of the equation. In such a case we consider a new variable $Y(\varepsilon) = \varepsilon^d X$ where $Y(\varepsilon)$ can be expanded into a regular series. In order to obtain the singular expansion, one needs to find and balance the leading-order dominant terms with respect to $\varepsilon$. Substituting $X$ by $Y/\varepsilon^d$ leads to the polynomial:

$$Q_1(Y) = B_0 + B_1 Y + B_2 Y^2 + B_3 Y^3 + B_4 Y^4, \qquad B_n = \varepsilon^{-nd} A_n, \quad n \in \{0,1,2,3,4\} \tag{C.13}$$

where :

$$\begin{aligned}
B_0 &= \frac{6i}{ReWe} k^2 \varepsilon^{2+m} + \frac{C}{We} k^3 \varepsilon^{3+m} + \frac{6i}{Re} + Ck\varepsilon + \frac{k^3\varepsilon^3}{We} \\
B_1 &= \frac{3i}{ReWe} k^2 \varepsilon^{m+2-d} + \frac{3i}{Re}\varepsilon^{-d} \\
B_2 &= -k\varepsilon^{1-2d} - k^3\varepsilon^{p+1-2d} - \frac{1}{We} k^3 \varepsilon^{3-2d+m} - \frac{6i}{Re} k^2 \varepsilon^{-2d+m+p} - Ck^3\varepsilon^{1-2d+m+p} \\
B_3 &= -\frac{3i}{Re} k^2 \varepsilon^{m+p-3d} \\
B_4 &= k^3 \varepsilon^{1-4d+m+p}
\end{aligned}$$

Identifying the leading order in $\varepsilon$ in these coefficients yields :

$$B_0 = \mathcal{O}(1), \quad B_1 = \mathcal{O}(\varepsilon^{-d}), \quad B_2 = \mathcal{O}(\varepsilon^{1-2d}), \quad B_3 = \mathcal{O}(\varepsilon^{m+p-3d}), \quad B_4 = \mathcal{O}(\varepsilon^{m+p+1-4d}) \tag{C.14}$$

Balancing leading order dominant terms yields two possibilities :

- Balancing $B_1$ and $B_2$ as leading terms gives $d = 1$.
- Balancing $B_2$ and $B_4$ as leading terms gives $d = \frac{m+p}{2}$.

It is easy to check that all other combinations either lead to balancing non dominant terms or overlap with the previous cases. Thus, rescaling the equation by $\varepsilon^d$ for $d = 1$ and $d = \frac{m+p}{2}$ should reveal all the remaining roots.

Let us start with the case $d = 1$. This scaling is trivially reminiscent of $c_{p_2}$ of the original system, so one should intuitively recover its corresponding eigenvalue. Similarly to before we consider the following expansion :

$$\begin{aligned}
Y = {} & Y_0 + \varepsilon Y_1 + \varepsilon^2 Y_2 + \varepsilon^3 Y_3 + \varepsilon^m Y_m + \varepsilon^{m+1} Y_{m+1} + \varepsilon^{m+2} Y_{m+2} + \varepsilon^{2m} Y_{2m} + \varepsilon^{2m+1} Y_{2m+1} + \varepsilon^{3m} Y_{3m} + \varepsilon^p Y_p \\
& + \varepsilon^{p+1} Y_{p+1} + \varepsilon^{p+2} Y_{p+2} + \varepsilon^{2p} Y_{2p} + \varepsilon^{2p+1} Y_{2p+1} + \varepsilon^{3p} Y_{3p} + \varepsilon^{m+p} Y_{m+p} + \varepsilon^{m+p+1} Y_{m+p+1} + \mathcal{O}(\varepsilon^3)
\end{aligned}$$

There are more terms in this expansion compared with (C.7) since, in order to have a second order expansion in $X$, one needs to go to third order in $Y$. It would be best to replace $Y$ by this series in $\varepsilon Q(Y)$ to have order 0 as leading order. Setting to zero the leading terms in $\varepsilon$ and identifying, yields the expressions :

$$Y_0 = \frac{3i}{kRe}, \quad Y_1 = 2, \quad Y_2 = -\frac{i}{3} kRe(C-4), \quad Y_3 = \frac{4}{9} k^2 Re^2 (C-4) \tag{C.15}$$

$$Y_p = -\frac{3ik}{Re}, \quad Y_{p+2} = \frac{4}{3} ik^3 Re, \quad Y_{2p} = \frac{3ik^3}{Re}, \quad Y_{3p} = -\frac{3ik^5}{Re} \tag{C.16}$$

$$Y_m = Y_{m+1} = Y_{m+2} = Y_{2m} = Y_{3m} = Y_{p+1} = Y_{2p+1} = Y_{m+p} = Y_{m+p+1} = 0 \tag{C.17}$$

So that when replacing back $Y$ by $X/\varepsilon$ we obtain :

$$\begin{aligned} X =& \frac{3i}{Rek\varepsilon} + 2 - \frac{i}{3}Re\left(\frac{\cos\theta}{F^2} + \frac{2\lambda^2 h^3}{45} - 4\right)k\varepsilon + \frac{4}{9}Re^2\left(\frac{\cos\theta}{F^2} + \frac{2\lambda^2 h^3}{45} - 4\right)k^2\varepsilon^2 \\ & - \frac{3i}{Re}k\varepsilon^{p-1} + \frac{4}{3}iRek^3\varepsilon^{p+1} + \frac{3i}{Re}k^3\varepsilon^{2p-1} - \frac{3i}{Re}k^5\varepsilon^{3p-1} + \dots \end{aligned}$$

Thus, the same situation arises as before and we need to choose $p$ so that leading order in the additional terms is at least of second order. In this case, and $\forall p \geq 1$, the leading order is of order $\varepsilon^{p-1}$, which leads the constraint :

$$p \geq 3 \tag{C.18}$$

This concludes the asymptotic comparison of phase velocities. In conclusion, in order to have phase velocities which comply with the original model at least at first order, one needs to impose $p \geq 3$. Note that, the previous analysis does not yields any constraint on $\alpha$ besides the assumption we did in the beginning $m \geq 1$. This is due to the fact that in the augmented model, $\alpha$ mainly intervenes in the capillary terms which are taken proportional to $\varepsilon^2$, thus making them appear in later orders in the expansion (order 3 and higher). Even in this case where $\kappa$ is assumed $\mathcal{O}(1)$, one can show that the same analysis yields $m \geq 1$ through the asymptotic analysis, which is not any more restrictive than the initial assumption.

While, this fully addresses the comparison between the original system system phase velocities and their equivalents in the augmented system, this analysis did not yet investigate all of the eigenvalues. In fact, two eigenvalues are still missing and which are obtained by taking $d = \frac{m+p}{2}$ as a rescaling factor. While, the task of obtaining a precise expansion in this case is more tedious, it is not anymore rewarding, since stability analysis was already performed regardless of these values. Nevertheless, it remains beneficial to at least give the leading order behavior of these roots which write as follows :

$$X_{3,4} = \pm\frac{1}{k\sqrt{\epsilon^{p+m}}} + \cdots \tag{C.19}$$

such a form was expected in the sense that these velocities should go to infinity in the limit $\varepsilon \to 0$. To conclude this part, we obtained a necessary and sufficient condition for the leading order behavior of the phase velocity to match that of the original system. While the choice of $p = 3$ yields the best results in this regard, the resulting system requires both $\varepsilon$ and $k$ as independent variables, while in the original system, the analysis can be performed in terms of only $k\varepsilon$. It is possible to conserve such a structure for a unique choice of $\alpha$ and $\beta$ such that $\alpha \propto \varepsilon$ and $\beta \propto \varepsilon^2$. However, as proved by the previous computations, this does not guarantee accuracy nor does it maintain a good agreement stability-wise.

APPENDIX D

# Modified equations : Computing radii of convergence

In each of the following cases, we would like to check the convergence radius of the series :

$$\ln(S(\theta, \lambda_q, \Delta x) = \sum_{p=0}^{\infty} \frac{\alpha_p}{\lambda_q \Delta x^q} \theta^p = \sum_{p=0}^{\infty} a_p \theta^p \tag{D.1}$$

For that, we use the root convergence test, that is the radius of convergence $R$ is given by :

$$\lim_{p \to +\infty} |a_p|^{1/p} = \frac{1}{R} \tag{D.2}$$

## D.1 Centered scheme for the heat equation : $\lambda_2 = \frac{1}{2}$

We have :

$$S(\theta, \lambda_2, \Delta x) = 1 - 2\sin(\theta/2)^2 \tag{D.3}$$

In this case, it is possible to obtain an explicit expression of the $pth$ term of the sequence $a_p$ which is given by :

$$\begin{cases} a_0 = 0 & \\ a_{2p} = -\dfrac{(-4)^p E_{2p-1}(0)}{2(2p)!} & \forall p \geq 1 \\ a_{2p+1} = 0 & \forall p \geq 0 \end{cases} \tag{D.4}$$

where $E_p(x)$ denotes the $pth$ Euler polynomial function defined by :

$$\frac{2e^{xt}}{e^t + 1} = \sum_{n=0}^{\infty} E_n(x) \frac{t^n}{n!} \tag{D.5}$$

Since the series has only even order coefficients, it is equivalent and more convenient to check the convergence of the series :

$$\sum_{p=0}^{\infty} b_p \phi^p \tag{D.6}$$

where $\phi = \theta^2$ and $b_p = a_{2p}$. We proceed as follows in order to check the radius of convergence. We use Bernoulli's number [74]:

$$B_{2p} = \frac{-pE_{2p-1}(0)}{(2^{2p}-1)} \tag{D.7}$$

and plug this ansatz into $b_p$ to obtain :

$$|b_p| = \left|\frac{(4)^p E_{2p-1}(0)}{2(2p)!}\right| = \left|\frac{(4)^p(2^{2p}-1)}{2p(2p)!}B_{2p}\right| \underset{p\to\infty}{\sim} \left|\frac{2^{2p}}{p(\pi)^{2p}}\right|$$

Hence we can write :

$$|b_p| \underset{p\to\infty}{\sim} \frac{1}{p}\left(\frac{2}{\pi}\right)^{2p} \tag{D.8}$$

and so:

$$\lim_{p\to+\infty} |b_p|^{1/p} = \left(\frac{2}{\pi}\right)^2 \tag{D.9}$$

This means that the series (D.6) has a radius $R_\phi = \pi^2/4$ which implies that the radius of convergence of the series $G(\theta, 1/2, \Delta x)$ is $R = \pi/2$.

## D.2 Centered scheme for the heat equation : $\lambda_2 = \frac{1}{4}$

We have :

$$S(\theta, \lambda_2, \Delta x) = 1 - \sin(\theta/2)^2 \tag{D.10}$$

The expression of the $pth$ term $a_p$ is given by :

$$\begin{cases} a_0 = 0 & \\ a_{2p} = -\dfrac{(-1)^p E_{2p-1}(0)}{(2p)!} & \forall p \geq 1 \\ a_{2p+1} = 0 & \forall p \geq 0 \end{cases} \tag{D.11}$$

Using the same previous notations as in (D.6), we have :

$$|b_p| = \left|\frac{(E_{2p-1}(0)}{(2p)!}\right| = \left|\frac{(2^{2p}-1)}{p(2p)!}B_{2p}\right| \underset{p\to\infty}{\sim} \left|\frac{2}{p\pi^{2p}}\right| \tag{D.12}$$

which yields :

$$\lim_{p\to+\infty} |b_p|^{1/p} = \frac{1}{\pi^2}. \tag{D.13}$$

This gives the radius of convergence of the series $G(\theta, 1/4, \Delta x)$ is $R = \pi$.

## D.3 Proof of convergence for $\lambda_1 \leq 1$ of Upwind Euler for transport equation

Since $\Omega_c = \left\{(\lambda_q, \Delta x) \in \mathbb{R}^2 : \lambda_1 \leq \frac{1}{2}\right\}$, then the series $G(\theta, \lambda_1, \Delta x)$ is convergent $\forall \lambda_1 \leq 1/2$ and we have :

$$e^{\Delta t G(\theta, \lambda_1, \Delta x)} = S(\theta, \lambda_1, \Delta x)$$

Since stability only depends on the modulus of $S(\theta, \lambda_1, \Delta x)$ it is sufficient for our analysis to consider the equality :

$$e^{\Delta x \lambda_1 \mathrm{Re}(G(\theta, \lambda_1, \Delta x))} = |S(\theta, \lambda_1, \Delta x)| \tag{D.14}$$

Next, we show the following symmetry :

$$|S(\theta, 1/2 - \lambda_1, \Delta x)| = |S(\theta, 1/2 + \lambda_1, \Delta x)|$$

Indeed we have :

$$\begin{aligned}
S(\theta, 1/2 - \lambda_1, \Delta x) &= 1 - (1/2 - \lambda_1)\left(1 - \mathrm{e}^{-\mathrm{i}\theta}\right) = (1/2 + \lambda_1) + (1/2 - \lambda_1)e^{-\mathrm{i}\theta} \\
S(\theta, 1/2 + \lambda_1, \Delta x) &= 1 - (1/2 + \lambda_1)\left(1 - \mathrm{e}^{-\mathrm{i}\theta}\right) = (1/2 - \lambda_1) + (1/2 + \lambda_1)e^{-\mathrm{i}\theta} \\
&= \left((1/2 + \lambda_1) + (1/2 - \lambda_1)e^{\mathrm{i}\theta}\right)e^{-\mathrm{i}\theta} \\
&= e^{-\mathrm{i}\theta}\overline{S(\theta, 1/2 - \lambda_1, \Delta x)}
\end{aligned}$$

where the bar denotes the complex conjugate. Hence $|S(\theta, 1/2 - \lambda_1, \Delta x)| = |S(\theta, 1/2 + \lambda_1, \Delta x)|$. This implies through equality (D.14) that :

$$\Delta x(1/2 - \lambda_1)\mathrm{Re}(G(\theta, 1/2 - \lambda_1, \Delta x)) = \Delta x(1/2 + \lambda_1)\mathrm{Re}(G(\theta, 1/2 + \lambda_1, \Delta x)) \tag{D.15}$$

which also implies in terms of coefficients:

$$(1/2 - \lambda_1)\alpha_{2p}(1/2 - \lambda_1, \Delta x) = (1/2 + \lambda_1)\alpha_{2p}(1 + \lambda_1, \Delta x) \quad \forall p \geq 1 \tag{D.16}$$

Therefore, since the series is convergent for $0 \leq \lambda_1 \leq 1/2$, it is also convergent for $1/2 \leq \lambda_1 \leq 1$.