Università degli Studi di Pisa

Facoltà di Ingegneria
Corso di Dottorato in Ingegneria Aerospaziale
XXIII Ciclo

Tesi di Dottorato

Numerical Simulations of Barotropic Flows in Complex Geometries

Tutori
Prof. Maria Vittoria Salvetti
Dr. François Beux

Candidato
Marco Bilanceri

Direttore del Corso di Dottorato
Prof. Maria Vittoria Salvetti

Marzo 2011

To Sara,

for all the love, tenderness and affection
you constantly give me since the day we met.

To my Family,

for all the support they gave me
all over my life.

Abstract

The numerical simulation of barotropic flows in complex geometries is considered together with its application to cavitating flows and sediment transport problems. A general methodology based on different finite-volume methods applicable to unstructured grids is developed and validated. Second-order extension in space is obtained by using MUSCL-like methodologies. The time-advancing is carried out by using implicit linearised methods and the second-order accuracy in time is achieved by exploiting the Defect Correction technique.

As far as cavitating flows are concerned, a homogeneous fluid cavitation model is considered, which leads to a barotropic state equation. The continuity and momentum equations for compressible flows are discretized through a mixed finite-element/finite-volume approach, applicable to unstructured grids. P1 finite elements are used for the viscous terms, while finite volumes for the convective ones. The numerical fluxes are computed by shock-capturing schemes and ad-hoc preconditioning is used to avoid accuracy problems in the low-Mach regime. A HLL flux function for barotropic flows is proposed, in which an anti-diffusive term is introduced to counteract accuracy problems for contact discontinuities and viscous flows typical of this class of schemes.

For this HLL-like flux function two different time linearizations are considered: in the first one the upwind part of the flux function is frozen in time, while in the second one its time variation is taken into account. The proposed numerical ingredients are validated through the simulations of different flow configurations, namely the Blasius boundary layer, a Riemann problem, the quasi-1D cavitating flow in a nozzle, the flow around a hydrofoil mounted in a tunnel, both in non-cavitating and cavitating conditions The introduction of the RANS $k-\varepsilon$ model is validated by the simulation of the high-Reynolds number flow over a flat plate. Finally the simulation of the flow in a real three dimensional inducer in both non-cavitating and cavitating conditions is addressed.

Sediment transport problems are also considered. The problem is modeled through the Shallow-Water equations coupled with the Exner one to describe the time evolution of the bed profile. The system is closed by the Grass model which is used for the sediment transport fluxes. The governing equations are discretized by using two different finite-volume methods, the SRNH predictor-corrector scheme and a Modified Roe scheme for non conservative systems of equations. Starting from the explicit versions, the corresponding linearised implicit schemes are generated. The flux Jacobians are computed through automatic differentiation thus avoiding the difficult analytical differentiation of the numerical fluxes. Finally, the considered numerical ingredients are compared in terms of accuracy and computational time using different one dimensional and two dimensional problems, characterised by different time scales for the evolution of the bed and of the water flow.

Sommario

Nella presente tesi é considerata la simulazione di flussi barotropici all'interno di geometrie complesse e le sue applicazioni a flussi cavitanti e a problemi di trasporto di sedimenti. Un approccio generale basato su vari metodi ai volumi finiti e applicabile a griglie non strutturate, é stato sviluppato e testato. L'estensione al secondo ordine spaziale é ottenuta usando metodi tipo MUSCL. L'avanzamento temporale si basa su metodi impliciti linearizati e il secondo ordine temporale si basa sulla tecnica Defect Correction.

Per quanto riguarda i flussi cavitanti, é stato utilizzato un modello a flusso omogeneo che si basa su una equazione di stato barotropica. Il bilancio di massa e di quantitá di moto per flussi comprimibili sono discetizati tramite un metodo misto ai volumi finiti e agli elementi finiti. Gli elementi finiti P1 sono utilizzati per i termini viscosi mentre i volumi finiti per quelli convettivi. I flussi numerici sono calcolati utilizzando schemi in grado di calcolare soluzioni discontinue e una strategia di precondizionamento ad-hoc é stata utilizzata per risolvere i problemi di accuratezza che si riscontrano per bassi numeri di Mach. Una funzione di flusso di tipo HLL per flussi barotropici é stata proposta introdotta. In questa funzione di flusso é stato aggiunto un termine antidiffusivo che riduce i problemi di accuratezza che tipicamente si riscontrano per discontinuità di contatto e flussi viscosi quando si utilizzano schemi appartenenti a questa categoria.

Per questa funzione di flusso di classe HLL due differenti linearizzazioni temporali sono state considerate: nella prima la matrice di upwind della funzione di flusso é considerata constante, mentre nella seconda la sua variazione temporale viene tenuta in considerazione. Gli ingredienti numerici proposti sono stati quindi testati simulando varie tipologie di flussi, in particolare lo strato limite di Blasius, un problema di Riemann, il flusso quasi-1D in un ugello e il flusso di acqua intorno ad un profilo, sia in condizioni cavitanti che non cavitanti. Inoltre l'introduzione degli effetti della turbolenza tramite il modello RANS $k - \varepsilon$ é stata testata simulando il flusso ad alto numero di Reynolds su una lastra piana e, per finire, é stata affrontata la simulazione numerica di un induttore reale tridimensionale, sia in condizioni non cavitanti e cavitanti.

Oltre a quanto detto sono stati considerati anche problemi di trasporto di sedimenti. Il modello fisico di questo problema é basato sulle equazioni Shallow-Water a cui si aggiunge l'equazione di Exner per descrivere l'evoluzione temporale del profilo del fondale. In particolare é il flusso di sedimenti é stato descritto utilizzando il modello di Grass. Il sistema completo di equazioni é stato discretizato utilizzando due metodi ai volumi finiti, lo schema di tipo predittore-correttore SRNH e uno schema di Roe modificato per sistemi di equazioni in forma non conservativa. Partendo dalle versioni esplicite di questi schemi, sono stati sviluppati i corrispondenti metodi impliciti e, in particolare lo Jacobiano della funzione di flusso é stato calcolato utilizzando strumenti di differenziazione automatica. Questo approccio permette di non dover calcolare manualmente le complesse espressioni delle derivate della funzione di flusso. Questi metodi sono poi stati comparati in termini di accuratezza e costi computazionali utilizzando specifici problemi monodimensionali e bidimensionali caratterizzati da scale temporali diverse per l'evoluzione del fondale e del flusso d'acqua

Acknowledgements

First of all, I owe my deepest gratitude to my supervisors Prof. Maria Vittoria Salvetti, University of Pisa (Italy), and Dr. François Beux of Alta Spa who advised me during the three years of my Ph.D studies giving me full support in the development of my research and in my personal and professional growth.

I wish to express my sincere thanks to Hervé Guillard, Inria (Sophia-Antipolis, France), who introduced me into the field of sediment transport problems and gave me guidance and constructive suggestions during my visiting periods at Sophia-Antipolis.

I would like to acknowledge Prof. Luca d'Agostino (University of Pisa, Italy), for his advices and the precious suggestions he gave me to improve this work and for the interesting discussions on the topic of cavitating flows.

I am very grateful to Prof. Giorgio Rosatti (University of Trento, Italy) and Prof. Richard Saurel (University Aix Marseille I, France) who kindly agreed to dedicated their time to review my thesis.

I would like to thank the people of Alta Spa (Ospedaletto Pisa Italy), in particular Lucio Torre, Angelo Pasini e Angelo Cervone, for having provided the set of experimental data used for this work.

I wish to acknowledge the ESA project No. 20081/06/NL/IA *Scaling of thermal cavitation effects on cavitation-induced instabilities* and the EuroMéditerranée $3+3$ network MhyCoF which supported part of the present work.

Finally I also wish to thank the Italian Computer Centres CASPUR and CINECA for having provided computational resources and support.

Contents

Chapter 1

Introduction

The present thesis reports the progress made for the development of numerical tools for the simulation of barotropic flows in complex geometries. Many physical models can be classified as barotropic flows, the most popular of them being probably the Shallow-Water equations.

From a practical point of view barotropic models can be attractive. Indeed, while the standard systems of conservation laws include mass, momentum and energy balances, the barotropic formulation permits to decouple the energy balance from the rest of the governing equations and, thus, a reduced system can be considered. The computational resources required for the numerical simulations are therefore greatly reduced, making barotropic models interesting alternatives for engineering or environmental applications.

Two main categories of problems, which may be described by barotropic flows, are considered in this study, namely cavitating flows and sediment transport problems. In this context, the present work is part of a research activity aimed at developing a numerical set-up for the simulation of flows characterised by a generic barotropic equation of state [97, 102–104].

The numerical approaches developed and validated in the present thesis for these two classes of problems are different, but they share a few commnon points. First, space and time discretisation are kept separated ("method of lines") in order to simply switch between different space discretisation schemes and time-advancing techniques. The development of numerical tools suitable for the simulations of barotropic flows in real conditions usually requires to deal with complex geometries. Thus, the use of triangular or tetrahedral unstructured grids, which are easily adaptable to generic geometries, becomes particularly attractive. Since, for both cavitation and shallow-water applications, the convective part of the equation systems is hyperbolic, finite-volume discretizations applicable to unstructured grids are used. Furthermore, we are interested in *stiff* problems; the sources of stiffness for cavitating flows and sediment transport problems will be described in details in the following sections. To counteract the efficiency and stability problems due to the problem stiffness, implicit linearised approaches have been adopted and developed herein for time-advancing.

1.1 Cavitating flows

Cavitating flows occur in a number of engineering devices, as, for instance, rocket turbopumps, turbomachinery, hydrofoils, marine propellers, nozzles, etc. The prediction and characterisation of cavitating flows is hence of great importance, since cavitation has strong effects, usually negative, on performance and life of such devices.

From a physical view point cavitating flows are characterised by different phenomena interacting each other, such as change of phase, complex interactions between vapour and liquid, unsteadiness, not well defined vapour to liquid interfaces and turbulence. All this renders the modeling of cavitating flows a very complex task. Several models exist in the literature of different levels of complexity, see i.e. [3, 46, 65, 67, 77, 99, 100, 102] and the references there contained for a review. Among them are the so-called *one-fluid or equivalent-fluid* models, in which the cavitating flow is described in terms of a single fluid or mixture, whose properties are derived through suitable assumptions, and, more particularly, the *barotropic homogeneous fluid* models, in which the density and the pressure are linked each other through an equation of state, both for pure liquid and for the liquid-vapour mixture (see e.g. [37, 42, 80, 87]). The models of this kind, although they introduce rather strong simplifications and neglect a fine description of the local behaviour of cavitation, are attracting because of their simplicity and because they a-priori have the capability of describing the large-scale effects of cavitation, which are dominating in many applications of interest, such as for instance in the field of rocket propulsion. Examples of numerical simulations using this type of models can be found in [33–35,54,62,87,105,114]. In spite of the simplifying assumptions made and the apparent simplicity of such models, strong difficulties arise for numerical simulation, which are mainly due to the fact that the cavitating mixture is described as a highly compressible fluid, characterised by speed of sound values of several orders of magnitude lower that those of the pure liquid. Moreover, an abrupt transition from the wetted (incompressible) to the cavitating (highly supersonic) regimes occurs. In the present work, in the non-cavitating regions, a weakly-compressible liquid at constant temperature was considered, while for the cavitating regime, the homogeneous-flow model explicitly accounting for thermal cavitation effects and for the concentration of the active cavitation nuclei in the pure liquid proposed in [37] was adopted. For cavitating flow simulation, it is clear, on the basis of the previous considerations, that the numerical schemes must be designed in order to cope with nearly incompressible and highly-compressible regions coexisting in the flow. Two opposite approaches can be found in the literature: adaptation to compressible flows of methods developed for incompressible flows (e.g. [101, 114]) or adaptation to the incompressible limit, usually through ad-hoc preconditioning, of compressible flow solvers (e.g. [33, 54, 74]). The latter approach in particular, was used in a work [102] prior to the one presented here, which was our starting point. In the aforementioned work a preconditioned linearised implicit numerical method for the simulation of inviscid compressible barotropic flows on unstructured grids was developed, in which the spatial discretization of the convective terms is carried out through a finite-volume approach. The Roe numerical flux function [91] was adapted to barotropic flows while a Turkel-like preconditioning was considered to deal with low Mach number regime [102,104]. As for time advancing, a linearised implicit algorithm was defined by considering a Jacobian-free linearization of fluxes only relying on the properties of the Roe matrix [97, 102, 104]. The set-up numerical tool was tested for different types of barotropic equations of state and flow regimes, and this validated most of the used ingredients, as, for instance, the accuracy and efficiency in the low Mach regime [97,102]. In [104], an application to the simulation of the inviscid flow in a realistic configuration of a rocket turbopump inducer in non-cavitating conditions is described. However, for cavitating flows, the stability properties of the scheme were found to deteriorate dramatically and only very small time steps are allowed. This clearly increases the computational costs and, thus, makes difficult to afford the simulation of complex cavitating flows, as occur in many aerospace and industrial applications. A rather strong reduction of the CFL number allowed by numerical stability was also recently observed in [54] for a linearised implicit time-advancing scheme, when passing from non-cavitating to cavitating conditions.

Two new ingredients are introduced and investigated in the present thesis, in order to counteract with these efficiency limitations. First, a different numerical flux function is used for the convective fluxes. The starting point is the Rusanov numerical flux function [41, 95], which is the simplest scheme of the HLL family [107]. This scheme is known to have excellent robustness properties and is attracting because of its simplicity. A preliminary study [16] was carried out in this direction showing promising results. However, the Rusanov flux as well all the HLL schemes are also known for their excessive diffusive behaviour in presence of a contact discontinuity and they are, thus, not well suited for viscous flow simulations. To avoid this problem, more complex average-state approximate Riemann solvers have been introduced as, for instance, the HLLC scheme, proposed in [107], which involves two intermediate states in the approximate solution. Alternatively, an anti-diffusive term for the contact discontinuity can be directly added in the single-state HLL formulation as done in the HLLEM [44] and HLLE+ [84] schemes. In the present work, on the basis of the observation that in the 1D Riemann problem for barotropic flows density, velocity and pressure are continuous across the contact discontinuity, and thus, the presence of two different intermediate states is only due to the passive scalar (see [102] for details on the Riemann problem for barotropic flow), an anti-diffusive term only acting in the passive scalar equation is introduced. This leads to a low-diffusive HLL (LD-HLL) scheme, in which the first two equations, which are related to the acoustic waves, are unchanged with respect to the original Rusanov scheme. Thus, the simple structure of the Rusanov scheme is maintained in the LD-HLL one. Furthermore, on the basis of an asymptotic analysis in power of Mach, a preconditioning of the same kind of that used in [57] for perfect gases was adapted to barotropic flows and introduced in the upwind part of the numerical flux, in order to counteract the accuracy problems encountered in the low Mach regime. Since the preconditioning matrix multiplies the upwind part of the flux only, consistency in time is preserved.

The second point investigated herein is the linearization in time of the numerical flux function, which is needed in order to avoid the solution of a non-linear system at each time step of the implicit algorithm. A classical linearization consists in applying a first-order Taylor expansion in time but with a complete differentiation only for the centred part of the numerical flux function while the matrix in the upwind part is frozen at the previous time (see, e.g., [4, 119]). Even if this choice of linearization is in general a reasonable one, the time variation of the upwind part of the flux can be large in presence of huge variations of the flow velocity or when the speed of sound has a stiff change in magnitude. The latter is a typical situation in presence of cavitation. Hence, a more complete time linearization of the LD-HLL flux is proposed herein. Thanks to the simple structure of this scheme, the time variation of the upwind terms of the flux function is analytically derived while some simplifications are made on the basis of physical considerations suitable for the cavitating case. The remaining terms are numerically computed through centred finite differences.

Second-order accuracy in space is obtained through the MUSCL reconstruction technique [115]. Furthermore, the linearised implicit formulation is associated to a defect-correction technique to obtain second-order accuracy (both in time and space) at limited computational costs [103].

Finally, the implementation of viscous and turbulent effects is carried out in the present study through a mixed finite-volume/finite-element approach in which P1 finite elements are considered for the discretization of viscous terms. More in details, turbulence effects have been implemented in the code through a $k - \varepsilon$ RANS turbulence model [2, 76]: even if the considered turbulence model does not explicitly take into account cavitating effects, it is a well-known model and its positive features, as well as its drawbacks have been widely studied. Thus, it seems a reasonable starting point for the introduction of turbulence effects.

Several test-cases are considered to investigate the effects of the previously described new numerical and modelling ingredients. The anti-diffusive term introduced in the LD-HLL scheme is validated using a Riemann problem characterised by the presence of a contact discontinuity and the classical Blasius boundary layer over a flat plate. Comparison with exact solutions and with the results given by the Roe scheme are provided. The accuracy, robustness and efficiency properties of the Roe and the LD-HLL schemes in cavitating conditions are analysed for the quasi-1D steady cavitating flow in a convergent-divergent nozzle and for the flow around a NACA0015 hydrofoil mounted in a wind tunnel, for which experimental data are available [30]. For the LD-HLL scheme, the effects of the more complete time linearization are investigated. In particular, it is shown that this is a key point to largely improve the efficiency in cavitating conditions.

The implementation of turbulence effects is validated for the turbulent flow over a flat plate at zero angle of attack.

Finally, to conclude this part of the thesis, the flow in a real inducer is considered. This kind of application is really challenging due to the presence of both complex physical phenomena and complex geometries and, as a consequence, they require large computational resources. However since the experimental set-up is even more expensive, the interest in the development of numerical tools for these applications is growing, see e.g. [25, 36, 50, 63, 66, 69, 71, 82, 85, 86, 118], although 3D simulations in complete inducer geometries and in actual cavitating conditions are only a few [63, 86]. The inducer considered in this work is a three blade axial inducer. The geometry was provided by Alta Space, which also made available experimental data for both non-cavitating and cavitating flow conditions [31, 32, 38–40, 110–112]. Simulations in non-cavitating conditions at a specific value of the angular velocity ω_z and different values of the discharge have been done. Comparisons between numerical and experimental solutions are then performed. Simulations of the inducer flow in cavitating conditions have also been carried out. The obtained results are compared with experimental data and preliminary analyses on the possible presence of different kinds of instabilities related to cavitation are carried out.

1.2 Sediment transport problems

Let us focus on the second category of barotropic flows considered herein. The design and validation of numerical methods for the simulation of bedload sediment transport processes caused by the movement of a fluid in contact with the sediment layer has a significant interest for environmental and engineering problems. A few examples of such problems are beach profile changes due to severe climate waves, seabed response to dredging procedures or imposed structures, harbour siltation or transport in gravel-bed rivers.

The hydrodynamic part is usually modeled through the classical Shallow-Water equations coupled with an additional equation modeling the morphodynamical component. The system of Shallow-Water equations is used in a wide range of field, ranging from oceanographic to atmospheric applications. Moreover the standard Shallow-Water equations with constant bottom topography are equivalent to Euler equations for a barotropic flow. Thus, many numerical methods suitable for the simulation of inviscid compressible flows can be easily extended to this field of applications.

On the other hand, when the bottom topography is not constant, an additional source term appears in the system of equation. Indeed, a well known problem is that shallow water equations on non-flat topography have steady-state solutions in which the flux gradients are non-zero but are exactly balanced by the source terms. Standard numerical methods for the discretization of conservation laws may fail in correctly reproducing this balance and thus, specific methods have been developed to deal with this problem (C-property, well-balanced schemes, see e.g. [10, 51, 88, 89, 93, 108, 116, 120]).

Focusing on sediment transport problems, the additional morphodynamical component is usually modelled by continuity or Exner equation, expressing the conservation of the sediment volume, in which the solid transport discharge is provided by a closure model. According to [26] sediment transport can be classified into three main categories: bedload, saltation and suspension: in this thesis only the former is considered. Namely, in bedload sediment transport the sediment simply slides along the bed [26].

Many different models of solid transport discharge are available in the literature both deterministic as well as probabilistic ones (see, e.g., [26] for a review). As a first step, the Grass equation [56] is considered herein, which is one of the most popular and simple models.

A huge amount of work has been done in the last decades to develop numerical methods for the simulation of the previous system of equations (see, e.g., the references in [7, 8, 26, 28, 90, 109]). In particular the treatment of source terms and of the bed-load fluxes has received the largest attention. On the other hand, in Shallow-Water problems, time advancing has received much less attention since it is usually carried out by explicit schemes. Recently a semi-implicit method has been proposed [92] for Shallow-Water problems with fixed bed.

In this context, the focus of the present thesis is on the comparison between explicit and implicit schemes in the simulation of coupled shallow-water equations and sediment transport. Indeed, if the interaction of the water flow with the mobile bed is weak, the characteristic time scales of the flow and of the sediment transport can be very different introducing time stiffness in the global problem. For these cases, the stability properties of explicit schemes may significantly be deteriorated and, hence, it can be advantageous to use implicit schemes. On the other hand, since the considered problems are unsteady, attention must be paid for implicit schemes in the choice of the time step. Indeed, a too large time step could deteriorate the accuracy of the results and one issue is to investigate whether and for which conditions the use of implicit schemes is really convenient from a computational viewpoint. A first investigation of this issue is provided in the present thesis for 1D and 2D sediment-transport problems, involving different rates of bedload/water-flow interaction. We focus on flows over wet areas. The extension to the case of dry bed will the object of further studies.

As previously mentioned, the considered numerical formulation is based on the method of the lines, in which the spatial and temporal discretisation are kept separated. The computational domain is discretized by the use of unstructured triangular grids. Two different existing explicit numerical methods have been considered in this work, namely the SRNH scheme [5, 6, 8] and a Modified Roe one [26–28]. The second-order spatial accuracy is obtained by using MUSCL-like reconstruction techniques. Exploiting the independence between spatial and temporal discretisation, an implicit linearised formulation has been considered for the time-advancing of the aforementioned numerical methods and second-order order accuracy in time is obtained by using a backward differentiation formula coupled with the Defect Correction technique. Thus, the numerical setup for sediment transport problems shows several aspects (unstructured grids, finite-volume approach, MUSCL-like reconstruction, implicit time-advancing, DeC technique) in common with the one developed for cavitating flows.

A specific peculiarity of the numerical tool developed in this work for sediment transport problem concerns the linearised implicit time-advancing. The presence of source terms, of non conservative products and the specific structure of the considered numerical methods (e.g. the SRNH scheme is a predictor-corrector method), clearly make difficult to define a time linearisation, even approximated, for the numerical fluxes. The efforts required to compute the analytical Jacobian of the numerical fluxes are really remarkable, even when the simple Grass model is employed. In addition they are further increased by practical issues, as for instance the time required for writing and debugging a routine implementing the aforementioned linearisation procedure. In order to overcome these difficulties, an automatic differentiation tool (Tapenade, [59], http://www-sop.inria.fr/tropics/) has been used in this work. Given an explicit method, this tool can automatically generate a routine which computes the Jacobian of the considered function. Thus, starting from an explicit scheme, it is possible to quickly generate a routine which computes the Jacobian of the numerical fluxes and, by using this routine it is possible to straightforwardly define an implicit linearised method. Finally, it is work notice that the aforementioned numerical ingredients are easily adaptable to different numerical methods (thanks to the automatic differentiation procedure) and different physical models, as for instance different models for the solid transport discharge.

As previously mentioned, the different considered numerical schemes, viz. SRNH and Roe schemes with implicit or explicit time advancing and at first and second-order of accuracy, are applied to 1D and 2D benchmark problems. Comparisons in terms of accuracy and efficiency are provided.

1.3 Thesis Outline

The present thesis is organised in two main parts: in Part. I the study concerning cavitating flows is discussed, while Part. II deals with sediment transport problems.

1. Part. I: Numerical Simulations of Cavitating Flows

    - In Chap. 2 the adopted barotropic equations of state for cavitating flows are presented. In addition, several systems of governing equations of increasing level of complexity, are discussed. The main mathematical properties of the aforementioned systems are briefly described in order to set the ground for the development of the numerical discretisation.
    - In Chap. 3 the proposed 1D numerical discretisation is introduced. A general finite-volume discretization for hyperbolic systems of conservation laws is presented. A low-diffusive HLL scheme is

introduced and a linearised implicit time-advancing strategy, capable of overcome the severe CFL limitations encountered in cavitating conditions, is proposed. The extension to the second-order accuracy both in space and in time is achieved using the MUSCL reconstruction technique and the Defect Correction method. Finally the behaviour of the consider method dealing with nearly-incompressible flows is analysed and a suitable preconditioning strategy is proposed.

- In Chap. 4 the proposed numerical ingredients are validated using 1D benchmarks for which the exact solution is available. The considered test-cases involves a non-cavitating and a cavitating Riemann problem which, respectively, test the second-order extension and the diffusive properties of the numerical approaches. Two additional benchmarks on the stationary solution in a convergent-divergent nozzle are also considered to validate the preconditioning strategy and the efficiency of the proposed linearisation technique.
- In Chap. 5 the numerical ingredients proposed in the 1D case are extended to the 3D case and the issue of the discretisation of the 3D computational domain is discussed. Furthermore, the numerical discretisation of the laminar viscous fluxes and of the turbulent fluxes and source terms introduced by a RANS turbulence model is addressed. Finally, the aforementioned numerical ingredients are extendend to 3D rotating frame of reference.
- In Chap. 6 the 3D numerical ingredients are validated against some reference benchmark problems. More specifically, they include the flow over a flat plate, both in laminar (Blasius boundary layer) and turbulent conditions in order to validate the proposed discretisation for the viscous fluxes, to test the diffusive properties of the LD-HLL scheme and to validate the implementation of the turbulence model. In addition the water flow around a NACA0015 hydrofoil is considered, both in non-cavitating e cavitating conditions. In this latter case a quantitative appraisal, based on available experimental data is given and the efficiency of the proposed linearisation technique is analysed.
- In Chap. 7 the proposed 3D numerical method is exploited for the simulation of the flow in a real axial turbopump inducer. Several operational points, at different flow rates in non-cavitating as well as in cavitating conditions, have been considered. The numerical results are analysed on the ground of the available experimental data for this case. In addition a preliminary analysis of the in-

stabilities induced by cavitation in the aforementioned inducer is presented.

2. Part. II: Numerical Simulation of Sediment Transport Problems
   - In Chap. 8 the physical model considered for the description of sediment transport problems is introduced. A few alternative formulations and their mathematical properties are described in order to set the ground for the development of the numerical discretisation.
   - In Chap. 9 the proposed 2D numerical discretisation is addressed. Two explicit finite-volume numerical methods, namely the SNRH and the Modified Roe one are introduced and their main features are described. Starting from these explicit methods, their linearised implicit counterparts, which are generated by exploiting an automatic differentiation tool, are proposed. The extensions to second-order accuracy both in space and in time for both the explicit and implicit approaches are discussed.
   - In Chap. 10 the proposed numerical methods (explicit and implicit as well as first and second-order accurate both for the SNRH and Modified Roe schemes) are compared by using a 1D benchmark problem, in terms of efficiency and accuracy. Different conditions, characterised by different time scales for the evolution of the flow and of the sediment, are considered.
   - In Chap. 11 a 2D numerical test case is considered to validate the proposed 2D numerical methodologies and, in particular the linearised implicit time advancing automatically generated. Two test-cases, characterised by different ratios between the time scales of the evolution of the flow and of the sediment, are considered to compare the different numerical approaches.

Furthermore, Part. III contains some additional appendices. In appendices A and D the derivation of the exact solution for, respectively, the Riemann problem and the convergent-divergent nozzle are reported. Those exact solutions are the ones used for the appraisal of the 1D numerical approaches for cavitating flows. Appendix C presents the study of the preconditioning strategy proposed in this work and, finally, appendix B proves the equivalence between two different formulations for the Roe numerical scheme.

1.4 Related scientific publications and projects

Some of the topics described in this thesis have been documented through previous works, and in particular:

- The international publication [18]:
  M. Bilanceri, F. Beux, and M. V. Salvetti, *An Implicit Low-Diffusive HLL Scheme with Complete Time Linearization: Application to Cavitating Barotropic Flows*, Computer & Fluids, 39 (2010), pp. 1990–2006

- The INRIA research report [15]:
  M. Bilanceri, F. Beux, I. Elmahi, H. Guillard, and M. V. Salvetti, *Linearised implicit time-advancing applied to sediment transport simulations*, INRIA Rapport de Recherche 7492, 2010

- The proceedings of the international conferences [17, 19, 21]:
  M. Bilanceri, F. Beux, and M. V. Salvetti, *Investigation on Numerical Schemes in the Simulation of Barotropic Cavitating Flows*, in Proceedings of the 7^{th} International Symposium on Cavitation, Ann Arbor, Michigan, USA, 17–22 August 2009
  M. Bilanceri, F. Beux, and M. V. Salvetti, *An Implicit Low-Diffusive HLL Scheme for Cavitating Flow Simulation*, in V European Conference on Computational Fluid Dynamics, Lisbon, Portugal, 14–17 June 2010
  M. Bilanceri, H. Guillard, M. V. Salvetti, and F. Beux, *Implicit Simulations of Shallow-Water Equations with Mobile Bed*, in V European Conference on Computational Fluid Dynamics, Lisbon, Portugal, 14–17 June 2010

- The talks given at the international conferences and workshops [20, 22, 98]:
  M. V. Salvetti, F. Beux, M. Bilanceri, and E. Sinibaldi, *A Numerical Method for Barotropic Flow Simulation with Applications to Cavitation.* Presented by invitation at the Third Workshop "Micro-Macro Modelling and Simulation of Liquid-Vapour Flows", Strasbourg, France, 22–24 Jannuary 2008
  M. Bilanceri, E. Sinibaldi, F. Beux, and M. V. Salvetti, *A Preconditioned Second-Order Linearized Implicit Formulation For Barotropic Cavitating Flows.* Presented at the 8^{th} World Congress on Computational Mechanics (WCCM8), 5^{th} European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS 2008), Venice, Italy, 30 June– 5 July 2008
  M. Bilanceri, I. Elmahi, H. Guillard, M. V. Salvetti, and F. Beux, *Implicit Time-Advancing Applied to Shallow-Water Problems with Mobile Bed.* Presented at the Workshop on Numerical Method for

Interaction between Sediment & Water, Paris, France, 20–24 September 2010

- The internal publications of Department of Aerospace Engineering of the University of Pisa [16, 103]:
  E. Sinibaldi, F. Beux, M. Bilanceri, and M. V. Salvetti, *A Second-Order Linearised Implicit Formulation for Hyperbolic Conservation Laws, with Application to Barotropic Flows*, ADIA 2008-4, University of Pisa, June 2008
  M. Bilanceri, F. Beux, and M. V. Salvetti, *Investigation on Numerical Schemes, Preconditioning and Time Advancing in the Simulation of 1D Cavitating Flows*, ADIA 2008-7, University of Pisa, October 2008

- The technical reports of the ESA project No. 20081/06/NL/IA [11–13]:
  F. Beux, M. Bilanceri, and M. V. Salvetti, *TN5: Work Package WP3200 - 1D Cavitation Model Development*, technical Report for the ESA Project: *Scaling of thermal cavitation effects on cavitation-induced instabilities*
  F. Beux, M. Bilanceri, G. Pagano, and M. V. Salvetti, *TN6: Work Package WP3300 - 2D Cavitation Model Development*, technical Report for the ESA Project: *Scaling of thermal cavitation effects on cavitation-induced instabilities*
  F. Beux, M. Bilanceri, and M. V. Salvetti, *TN7: Work Package WP3400 - 3D Cavitation Model Development*, technical Report for the ESA Project: *Scaling of thermal cavitation effects on cavitation-induced instabilities*

Furthermore a part of the work presented has been developed as part of the following projects:

- ESA project No. 20081/06/NL/IA
- the EuroMéditerranée $3 + 3$ network MhyCoF

Part I

Numerical Simulation of Cavitating Flows

Chapter 2

Physical Model

In this chapter several physical models, of different level of complexity, are presented for the case of barotropic flows. They include the 1D governing equations for an inviscid barotropic flow, the 3D laminar viscous case as well as the classical $k-\varepsilon$ turbulence model. Note that the only specificity relative to cavitating flows is furnished by the choice of the cavitating equations and, thus, by the particular barotropic equation of state (EOS) adopted herein (see Sec. 2.2). By virtue of the barotropic EOS, the energy balance is decoupled from the mass and momentum balances and therefore, it is possible to consider a reduced set of governing equations.

First, the mathematical notation is described in Sec. 2.1 and the barotropic EOS under consideration is presented in Sec. 2.2. The 1D governing equations for an inviscid barotropic flow are presented in Sec. 2.3, they form the basis for the different numerical approaches considered in this work, as shown in Chap. 3. In Sec. 2.4 the 3D laminar viscous case is presented and, finally, the classical $k-\varepsilon$ turbulence model is added to the system of equations in Sec. 2.5.

2.1 Mathematical Notation

The adopted mathematical notation is briefly presented. Vectors are denoted using a bold character. Considering a reference frame in a 3D space, the position of a generic point P with respect to the origin is denoted by the vector $\mathbf{x}$, whose components x_1, x_2, x_3 are the Cartesian coordinates of P.

For a generic fluid system, ρ is density of the fluid, p its pressure and a the speed of sound. The symbol for the velocity vector is $\mathbf{u}$ and its components will be addressed either as u_1, u_2, u_3 or as u, v, w, depending on which of them makes easier the comprehension. The symbol $\mathbf{u}^T$ means the transpose of the vector $\mathbf{u}$. In some circumstances it is useful to consider a compact notation for the components of a vector: the symbol $(1, \mathbf{u}, 1)^T$ is a five-components vector whose first and last elements are 1 and the others are the components of $\mathbf{u}$, that is:

$$(1, \mathbf{u}, 1)^T = (1, u_1, u_2, u_3, 1)^T \tag{2.1}$$

The Einstein notation, summation over repeated indexes is used in this work, that is:

$$u_j v_j \Longrightarrow \sum_j u_j v_j \tag{2.2}$$

Finally the component of a generic vector $\mathbf{v}$ along a direction (a vector with unit norm) $\mathbf{n}$ is defined as:

$$\mathbf{v_n} = v_j n_j = \langle \mathbf{v}, \mathbf{n} \rangle \tag{2.3}$$

n_j being the j^{th} component of $\mathbf{n}$.

Some simplifications of the notation are considered when dealing with 1D systems. In this case the position of a point is defined by x and the only velocity component is u. A final remark is deserved to the symbols $\mathbf{W}$, F, V which denote, respectively, the vector of unknowns, the convective flux function and the viscous fluxes. Depending on the particular physical model investigated, 1D or 3D, laminar or turbulent, the number of unknowns and the definition of the fluxes are different. However, to avoid a heavy use of the notation, the symbols $\mathbf{W}$, F, V are always used for the different physical models. Time by time the specific system referred by those quantities will be clear from the context.

2.2 Constitutive equations

The physical formulation and the development of the numerical methods are based on a generic barotropic equation of state:

$$p = p(\rho) \tag{2.4}$$

The derivative $dp/d\rho$ is assumed to be strictly positive (a classical thermodynamic stability requirement for common fluids) and can be regarded as the square of the fluid sound speed $a(\rho)$:

$$a(\rho) \doteq \sqrt{\frac{\mathrm{d}p(\rho)}{\mathrm{d}\rho}} \tag{2.5}$$

In particular for cavitating flows a weakly-compressible liquid at constant temperature T_L is considered as working fluid. The liquid density ρ is allowed to locally fall below the saturation limit $\rho_{Lsat} = \rho_{Lsat}\,(T_L)$ thus originating cavitation phenomena. A regime-dependent (wetted/cavitating) constitutive relation is therefore adopted. As for the wetted regime ($\rho \geq \rho_{Lsat}$), the chosen model is of the form

$$p \;=\; p_{sat} \;+\; \frac{1}{\beta_{sL}}\;\ln\left(\frac{\rho}{\rho_{Lsat}}\right) \tag{2.6}$$

$p_{sat} = p_{sat}\,(T_L)$ and $\beta_{sL} = \beta_{sL}\,(T_L)$ being the saturation pressure and the liquid isentropic compressibility, respectively. Concerning the cavitating regime ($\rho < \rho_{Lsat}$), a homogeneous-flow model explicitly accounting for thermal cavitation effects and for the concentration of the active cavitation nuclei in the pure liquid has been adopted [37]:

$$\frac{p}{\rho}\frac{d\rho}{dp} = (1-\alpha)\left[(1-\varepsilon_L)\frac{p}{\rho_{Lsat}a^2_{Lsat}} + \varepsilon_L g^\star\left(\frac{p_c}{p}\right)^\eta\right] + \frac{\alpha}{\gamma_v} \tag{2.7}$$

where $g^\star$, η, γ_V and p_c are constant parameters depending on the particular substance considered, a_{Lsat} is the liquid sound speed at saturation, and α is the void fraction defined as

$$\alpha = \frac{\rho_{Lsat} - \rho}{\rho_{Lsat} - \rho_v} \simeq 1 - \frac{\rho}{\rho_{Lsat}} \tag{2.8}$$

ρ_v being the vapour density.
Finally, ε_L is the fraction of the liquid in thermal equilibrium with the vapour and is given by:

$$\varepsilon_L = \left\{\left[((1+\zeta)^3 - 1)\frac{\alpha}{1-\alpha}\right]^{-3} + 1\right\}^{-1/3} \tag{2.9}$$

ζ being a free model parameter accounting for thermal cavitation effects and, possibly, for the concentration of the active cavitation nuclei, see [37] for more details. Note that despite the model simplifications leading to a unified barotropic EOS, the transition between wetted and cavitating regimes is extremely abrupt. Indeed, the sound speed falls from values of order 10^3 m/s in the pure liquid down to values of order 10^{-1} m/s or 1 m/s in the mixture [37, 102, 104]. The corresponding Mach number variation makes this state equation very stiff from a numerical viewpoint. Note, also, that equations (2.6) and (2.7) for the cavitating case are a particular instance of (2.4).

As for the definition of the molecular viscosity, a simple model, which is linear in the cavitating regime, is considered:

$$\mu(\rho) = \begin{cases} \mu_L & \text{if} \quad \rho \geq \rho_{Lsat} \\ \mu_v & \text{if} \quad \rho \leq \rho_v \\ \alpha\mu_v + (1-\alpha)\mu_L & \text{otherwise} \end{cases} \tag{2.10}$$

in which μ_v and μ_L are the molecular viscosity of the vapour and of the liquid respectively, which, consistently with the assumptions made in the adopted cavitation model, are considered constant and computed at $T = T_L$.

2.3 1D Physical Model

Let us consider a 1D inviscid barotropic flow coupled with an additional substance ξ, which is simply advected with the flow. The evolution of the fluid variables can be described by the following system of equations:

$$\begin{cases} \dfrac{\partial \mathbf{W}}{\partial t} + \dfrac{\partial F(\mathbf{W})}{\partial x} = 0 \\ F(\mathbf{W}) = \left(\rho u, \rho u^2 + p, \rho u \xi\right)^T \\ \mathbf{W} = \left(\rho, \rho u, \rho \xi\right)^T \end{cases} \tag{2.11}$$

The first two equations of system (2.11) are, respectively, the mass and momentum balances for the fluid. The third equation equation describes the time-evolution of the substance ξ. The conservation law for ξ is decoupled from the rest of the system and, as a consequence, ξ is usually referred as a passive scalar. From a mathematical point of view, system (2.11) is hyperbolic. In order to verify this property, it is sufficient to explicitly write the Jacobian matrix $A^{(3)}$ of the flux function $F(\mathbf{W})$:

$$A^{(3)} \doteq \frac{\partial F(\mathbf{W})}{\partial \mathbf{W}} = \begin{pmatrix} 0 & 1 & 0 \\ a^2 - u^2 & 2u & 0 \\ -u\xi & \xi & u \end{pmatrix} \tag{2.12}$$

where the superscript (3) reminds the three components of the flux function. The eigenvalues of $A^{(3)}$ are:

$$\lambda_1^{(3)} = u - a, \quad \lambda_2^{(3)} = u, \quad \lambda_3^{(3)} = u + a \tag{2.13}$$

and the matrix $R^{(3)}$, whose columns correspond to the right eigenvectors of $A^{(3)}$ is easily computed as:

$$R^{(3)} = \begin{pmatrix} 1 & 0 & 1 \\ u - a & 0 & u + a \\ \xi & 1 & \xi \end{pmatrix} \tag{2.14}$$

Since the passive scalar equation is decoupled from the other equations, system (2.11) is easily generalisable to cases in which two or more passive scalars are included. Considering the presence of two passive scalars, ξ_1 and ξ_2, a system equivalent to (2.11) can be formulated as:

$$\begin{cases} \dfrac{\partial \mathbf{W}}{\partial t} + \dfrac{\partial F(\mathbf{W})}{\partial x} = 0 \\ F(\mathbf{W}) = (\rho u, \rho u^2 + p, \rho u \xi_1, \rho u \xi_2)^T \\ \mathbf{W} = (\rho, \rho u, \rho \xi_1, \rho \xi_2)^T \end{cases} \tag{2.15}$$

The Jacobian matrix $A^{(4)}$ is:

$$A^{(4)} \doteq \frac{\partial F(\mathbf{W})}{\partial \mathbf{W}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ a^2 - u^2 & 2u & 0 & 0 \\ -u\xi_1 & \xi_1 & u & 0 \\ -u\xi_2 & \xi_2 & 0 & u \end{pmatrix} \tag{2.16}$$

and the corresponding eigenvalues and right eigenvectors matrix $R^{(4)}$ are:

$$\lambda_1^{(4)} = u - a, \quad \lambda_2^{(4)} = \lambda_3^{(4)} = u, \quad \lambda_4^{(4)} = u + a \tag{2.17}$$

$$R^{(4)} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ u - a & 0 & 0 & u + a \\ \xi_1 & 1 & 0 & \xi_1 \\ \xi_2 & 0 & 1 & \xi_2 \end{pmatrix} \tag{2.18}$$

Comparing (2.16) with (2.12) and (2.18) with (2.14) the generalisation to an arbitrary number of passive scalar equations is straightforward. Note that in the following the superscripts $(\cdot)$ indicating the number of components of the system will be dropped, since it should be always clear what is the specific system considered.

2.4 3D Laminar Physical Model

The 3D laminar Navier-Stokes equations for a barotropic flow can be formulated in the following compact form, which is also the starting point for the derivation of the Galerkin formulation and of the discretization of the problem:

$$\frac{\partial \mathbf{W}}{\partial t} + \frac{\partial}{\partial x_j} F_j(\mathbf{W}) - \frac{\partial}{\partial x_j} \mu V_j(\mathbf{W}, \nabla \mathbf{W}) = 0 \tag{2.19}$$

The vector of the unknown variables $\mathbf{W}$ is defined as:

$$\mathbf{W} = (\rho,\ \rho u,\ \rho v,\ \rho w)^T \tag{2.20}$$

and the vectors $\mathbf{F} = (F_1, F_2, F_3)$ and $\mathbf{V} = (V_1, V_2, V_3)$ are, respectively, the convective fluxes and the diffusive ones defined as:

$$\begin{cases} F_1 = (& \rho u, & \rho u^2 + p, & \rho uv, & \rho uw &)^T \\ F_2 = (& \rho v, & \rho vu, & \rho v^2 + p, & \rho vw &)^T \\ F_3 = (& \rho w, & \rho wu, & \rho wv, & \rho w^2 + p &)^T \end{cases} \tag{2.21}$$

and

$$\begin{cases} V_1 = (& 0, & \sigma_{11}, & \sigma_{12}, & \sigma_{13} &)^T \\ V_2 = (& 0, & \sigma_{21}, & \sigma_{22}, & \sigma_{23} &)^T \\ V_3 = (& 0, & \sigma_{31}, & \sigma_{32}, & \sigma_{33} &)^T \end{cases} \tag{2.22}$$

where σ_{ij} is the generic component of the stress tensor:

$$\sigma_{ij} = -\frac{2}{3}\left(\frac{\partial u_k}{\partial x_k}\delta_{ij}\right) + \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \tag{2.23}$$

Let us consider the integral formulation of system (2.19) obtained by integrating over a general control volume $\mathcal{V}$:

$$\int_{\mathcal{V}} \frac{\partial \mathbf{W}}{\partial t} \mathrm{d}\mathcal{V} + \int_{\mathcal{V}} \frac{\partial}{\partial x_j} F_j(\mathbf{W}) \mathrm{d}\mathcal{V} - \int_{\mathcal{V}} \frac{\partial}{\partial x_j} \mu V_j(\mathbf{W}, \nabla \mathbf{W}) \mathrm{d}\mathcal{V} = 0 \tag{2.24}$$

Using the divergence theorem it is possible to reformulate (2.24) as follows

$$\int_{\mathcal{V}} \frac{\partial \mathbf{W}}{\partial t} \mathrm{d}\mathcal{V} + \int_{\mathcal{S}} n_j F_j(\mathbf{W}) \mathrm{d}\mathcal{S} - \int_{\mathcal{S}} \mu n_j V_j(\mathbf{W}, \nabla \mathbf{W}) \mathrm{d}\mathcal{S} = 0 \tag{2.25}$$

where $\mathcal{S}$ is the surface of the control volume $\mathcal{V}$ and $\mathbf{n} = (n_1, n_2, n_3)^T$ is the external normal unit vector to the surface. Considering now the convective part of system (2.19), the Jacobian matrix $A_\mathbf{n}$ of the convective flux $n_j F_j$ is:

$$A_\mathbf{n} = \begin{pmatrix} 0 & n_1 & n_2 & n_3 \\ a^2 n_1 - u u_\mathbf{n} & u_\mathbf{n} + u n_1 & u n_2 & u n_3 \\ a^2 n_2 - v u_\mathbf{n} & v n_1 & u_\mathbf{n} + v n_2 & v n_3 \\ a^2 n_3 - w u_\mathbf{n} & w n_1 & w n_2 & u_\mathbf{n} + w n_3 \end{pmatrix} \tag{2.26}$$

We can notice that, taken $\xi_1 = v$ and $\xi_2 = w$, the 1D and 3D formulations are consistent since the matrix $A_\mathbf{n}$ is equivalent to the matrix $A^{(4)}$ defined in (2.16) when the normal vector $\mathbf{n}$ is coincident with the x_1-axis. This feature will be further investigated in Chap. 5 where the 3D numerical methods are described.

2.4.1 3D laminar physical model in rotating frames

In the previous sections it was implicitly assumed that the reference frame was a Cartesian inertial frame. Instead in this section a frame rotating with constant angular velocity $\vec{\omega}$ is considered. When a rotating frame is considered, the system (2.19) must be modified as follows:

$$\frac{\partial \mathbf{W}}{\partial t} + \frac{\partial}{\partial x_j} F_j(\mathbf{W}) - \frac{\partial}{\partial x_j} \mu V_j(\mathbf{W}, \nabla \mathbf{W}) = \mathbf{S}\left(\vec{\omega}, \mathbf{x}, \mathbf{W}\right) \tag{2.27}$$

where $\mathbf{S}\left(\omega, \mathbf{x}, \mathbf{W}\right)$ is the source term taking into account non-inertial effects:

$$\begin{cases} \overline{\mathbf{S}} & = & -\left(2\vec{\omega} \wedge \rho \mathbf{u} + \rho \vec{\omega} \wedge \left(\vec{\omega} \wedge \mathbf{x}\right)\right) \\ \mathbf{S}\left(\vec{\omega}, \mathbf{x}, \mathbf{W}\right) & = & \left(0,\ \overline{\mathbf{S}}^T\right)^T \end{cases} \tag{2.28}$$

2.5 3D Turbulent Physical Model

In order to account for the turbulence effects, the RANS standard $k - \varepsilon$ model [76] has been used. The governing equations are the following:

$$\begin{aligned} \frac{\partial \mathbf{W}}{\partial t} + \frac{\partial}{\partial x_j} F_j(\mathbf{W}) + \frac{\partial}{\partial x_j} \widetilde{F}_j(\mathbf{W}) & \\ - \frac{\partial}{\partial x_j} \mu V_j(\mathbf{W}, \nabla \mathbf{W}) - \frac{\partial}{\partial x_j} \mu_t \widetilde{V}_j(\mathbf{W}, \nabla \mathbf{W}) &= \Omega(\mathbf{W}) \end{aligned} \tag{2.29}$$

In (2.29) the vector of unknowns $\mathbf{W}$ is defined as

$$\mathbf{W} = \left(\rho,\ \rho u,\ \rho v,\ \rho w,\ \rho k,\ \rho \varepsilon\right)^T \tag{2.30}$$

where k is the turbulent kinetic energy and ε is the turbulent dissipation. The other terms appearing in (2.29) are the convective fluxes F_j, the turbulence contribution to convective fluxes $\widetilde{F}_j$, the viscous laminar fluxes V_j, the viscous turbulent fluxes $\widetilde{V}_j$ and the source term Ω related to the $k - \varepsilon$ model. The molecular and turbulent viscosity are, respectively, μ defined by (2.10) and μ_t defined as follows:

$$\mu_t = C_\mu \frac{\rho k^2}{\varepsilon} \tag{2.31}$$

C_μ being an empirical parameter. The convective fluxes have the following expression:

$$\begin{cases} F_1 = (& \rho u, & \rho u^2 + p, & \rho u v, & \rho u w, & \rho u k, & \rho u \varepsilon &)^T \\ F_2 = (& \rho v, & \rho v u, & \rho v^2 + p, & \rho v w, & \rho v k, & \rho v \varepsilon &)^T \\ F_3 = (& \rho w, & \rho w u, & \rho w v, & \rho w^2 + p, & \rho w k, & \rho w \varepsilon &)^T \end{cases} \tag{2.32}$$

$$\begin{cases} \widetilde{F}_1 = \begin{pmatrix} 0, & \frac{2}{3}\rho k, & 0, & 0, & 0, & 0 \end{pmatrix}^T \\ \widetilde{F}_2 = \begin{pmatrix} 0, & 0, & \frac{2}{3}\rho k, & 0, & 0, & 0 \end{pmatrix}^T \\ \widetilde{F}_3 = \begin{pmatrix} 0, & 0, & 0, & \frac{2}{3}\rho k, & 0, & 0 \end{pmatrix}^T \end{cases} \tag{2.33}$$

The laminar and turbulent viscous fluxes are defined as follows:

$$\begin{cases} V_1 = \begin{pmatrix} 0, & \sigma_{11}, & \sigma_{12}, & \sigma_{13}, & \dfrac{\partial k}{\partial x_1}, & \dfrac{\partial \varepsilon}{\partial x_1} \end{pmatrix}^T \\ V_2 = \begin{pmatrix} 0, & \sigma_{21}, & \sigma_{22}, & \sigma_{23}, & \dfrac{\partial k}{\partial x_2}, & \dfrac{\partial \varepsilon}{\partial x_2} \end{pmatrix}^T \\ V_3 = \begin{pmatrix} 0, & \sigma_{31}, & \sigma_{32}, & \sigma_{33}, & \dfrac{\partial k}{\partial x_3}, & \dfrac{\partial \varepsilon}{\partial x_3} \end{pmatrix}^T \end{cases} \tag{2.34}$$

$$\begin{cases} \widetilde{V}_1 = \begin{pmatrix} 0, & \sigma_{11}, & \sigma_{12}, & \sigma_{13}, & \dfrac{1}{\sigma_k}\dfrac{\partial k}{\partial x_1}, & \dfrac{1}{\sigma_\varepsilon}\dfrac{\partial \varepsilon}{\partial x_1} \end{pmatrix}^T \\ \widetilde{V}_2 = \begin{pmatrix} 0, & \sigma_{21}, & \sigma_{22}, & \sigma_{23}, & \dfrac{1}{\sigma_k}\dfrac{\partial k}{\partial x_2}, & \dfrac{1}{\sigma_\varepsilon}\dfrac{\partial \varepsilon}{\partial x_2} \end{pmatrix}^T \\ \widetilde{V}_3 = \begin{pmatrix} 0, & \sigma_{31}, & \sigma_{32}, & \sigma_{33}, & \dfrac{1}{\sigma_k}\dfrac{\partial k}{\partial x_3}, & \dfrac{1}{\sigma_\varepsilon}\dfrac{\partial \varepsilon}{\partial x_3} \end{pmatrix}^T \end{cases} \tag{2.35}$$

in which σ_ε and σ_k are empirical parameters. Finally the formulation of the source term $\Omega(\mathbf{W})$ is:

$$\begin{cases} \mathcal{P} = -\left(\dfrac{2}{3}\rho k\delta_{ij} - \mu_t\left(\dfrac{\partial u_i}{\partial x_j} + \dfrac{\partial u_j}{\partial x_i} - \dfrac{2}{3}\dfrac{\partial u_k}{\partial x_k}\delta_{ij}\right)\right)\dfrac{\partial u_i}{\partial x_j} \\ \omega_k = -\rho\varepsilon + \mathcal{P} \\ \omega_\varepsilon = c_{\varepsilon 1}\dfrac{k}{\varepsilon}\mathcal{P} - c_{\varepsilon 2}\dfrac{\rho\varepsilon^2}{k} \\ \Omega(\mathbf{W}) = \begin{pmatrix} 0 & 0 & 0 & 0 & \omega_k & \omega_\varepsilon \end{pmatrix}^T \end{cases} \tag{2.36}$$

$\mathcal{P}$ being the production term of the turbulent kinetic energy. For all the empirical parameters contained in the $k-\varepsilon$ model, the standard values proposed in [76] are used herein. Similarly to the laminar case, the integral form of (2.29) can be recasted as:

$$\int_{\mathcal{V}} \frac{\partial \mathbf{W}}{\partial t}\mathrm{d}\mathcal{V} + \int_{\mathcal{S}} n_j F_j(\mathbf{W})\mathrm{d}\mathcal{S} + \int_{\mathcal{S}} n_j \widetilde{F}_j(\mathbf{W})\mathrm{d}\mathcal{S} - \int_{\mathcal{V}} \frac{\partial}{\partial x_j}\mu V_j(\mathbf{W}, \nabla\mathbf{W})\mathrm{d}\mathcal{V} \\ - \int_{\mathcal{V}} \frac{\partial}{\partial x_j}\mu_t \widetilde{V}_j(\mathbf{W}, \nabla\mathbf{W})\mathrm{d}\mathcal{V} = \int_{\mathcal{V}} \Omega(\mathbf{W})\mathrm{d}\mathcal{V} \tag{2.37}$$

Considering now the convective part of system (2.37), the Jacobian matrix $A_\mathbf{n}$ of the flux $n_j F_j$ is:

$$A_\mathbf{n} = \begin{pmatrix} 0 & n_1 & n_2 & n_3 & 0 & 0 \\ a^2 n_1 - u u_\mathbf{n} & u_\mathbf{n} + u n_1 & u n_2 & u n_3 & 0 & 0 \\ a^2 n_2 - v u_\mathbf{n} & v n_1 & u_\mathbf{n} + v n_2 & v n_3 & 0 & 0 \\ a^2 n_3 - w u_\mathbf{n} & w n_1 & w n_2 & u_\mathbf{n} + w n_3 & 0 & 0 \\ -k u_\mathbf{n} & k n_1 & k n_2 & k n_3 & u_\mathbf{n} & 0 \\ -\varepsilon u_\mathbf{n} & \varepsilon n_1 & \varepsilon n_2 & \varepsilon n_3 & 0 & u_\mathbf{n} \end{pmatrix} \tag{2.38}$$

Applying the same arguments considered for the laminar case, it is clear that using a suitable reference frame, the Jacobian matrix $A_\mathbf{n}$ is equivalent to a 1D hyperbolic system augmented with four passive scalar equations.

2.5.1 Wall treatment

As often done for the $k-\varepsilon$ model, a free-slip condition is imposed at a distance δ from the wall, which depends on the used computational grid; then, the shear stress due to the presence of the wall is obtained from the computed velocity by using a wall law. The Reichardt wall [61] law is used herein. Let u^+ and y^+ be the velocity and distance from the wall in wall units defined as follows:

$$u^+ = \frac{u_t}{u_f} \tag{2.39}$$

$$y^+ = \frac{\rho u_f}{\mu} y \tag{2.40}$$

in which u_t is the velocity component in the direction locally tangent to the wall, y the distance from the wall and u_f the friction velocity which can be computed from the wall shear stress, τ_w, as follows:

$$u_f = \left(\frac{\tau_w}{\rho}\right)^{\frac{1}{2}} \tag{2.41}$$

The Reichardt wall law can be written as follows (see [61] for more details):

$$u^+ = \frac{1}{k}\ln(1 + k y^+) + 7.8\left(1 - e^{-\frac{y^+}{11}} - \frac{y^+}{11} e^{-0.33 y^+}\right) \tag{2.42}$$

This wall-law has the advantage of describing once for all the three types of behaviour of the turbulent boundary layer, viz. the viscous layer, the logarithmic zone and the intermediate one, which is generally not accounted for in usual wall laws. For additional details we refer to [2].

2.5.2 3D Turbulent Physical Model in rotating frames

The extension of system (2.29) to a rotating frame of reference can be carried using the same approach considered for the laminar case:

$$\frac{\partial \mathbf{W}}{\partial t} + \frac{\partial}{\partial x_j} F_j(\mathbf{W}) + \frac{\partial}{\partial x_j} \widetilde{F}_j(\mathbf{W}) - \frac{\partial}{\partial x_j} \mu V_j(\mathbf{W}, \nabla \mathbf{W}) - \frac{\partial}{\partial x_j} \mu_t \widetilde{V}_j(\mathbf{W}, \nabla \mathbf{W}) = \Omega(\mathbf{W}) + \mathbf{S}\left(\vec{\omega}, \mathbf{x}, \mathbf{W}\right) \quad (2.43)$$

where $\mathbf{S}\left(\vec{\omega}, \mathbf{x}, \mathbf{W}\right)$ is the source term taking into account non-inertial effects:

$$\begin{cases} \overline{\mathbf{S}} & = & -\left(2\vec{\omega} \wedge \rho \mathbf{u} + \rho \vec{\omega} \wedge \left(\vec{\omega} \wedge \mathbf{x}\right)\right) \\ \mathbf{S}\left(\vec{\omega}, \mathbf{x}, \mathbf{W}\right) & = & \left(0,\ \overline{\mathbf{S}}^T, 0, 0\right)^T \end{cases} \quad (2.44)$$

From (2.44) and (2.43) it is clear that the formulation of the equations for k and ε is unchanged in a rotating frame of reference. This is a well-known feature of the $k - \varepsilon$ model see [23, 70, 117] and, as many authors point out, could lead to an underestimation of the turbulence generation due to the rotation [23]. In the literature some modification of the standard $k - \varepsilon$ model for rotating frame of reference can be found. However the physical modelling of turbulent cavitating flows is really challenging and, in the author knowledge, there is no physical model free from drawbacks or case-dependent assumptions.

On the other hand, the standard $k - \varepsilon$ model is commonly applied to a wide range of cases and its limitations are well-known. As a consequence, this model seems a good starting point for the introduction of turbulence in the physical model.

Chapter 3

1D Numerical Method

In this chapter is proposed a numerical formulation for the simulation of the barotropic flows which are described by the physical models introduced in Chap. 2. The spatial discretization of the governing equations and, more specifically, of the convective fluxes, is based on a finite-volume formulation. Time advancing is carried out through an implicit linearised algorithm. In this chapter the main numerical ingredients are presented in details in the 1D inviscid case, while the generalisations to 3D viscous and turbulent flow equations are discussed in Chap. 5.
First, from Sec. 3.1 to Sec. 3.4 a general finite volume formulation for 1D hyperbolic system of conservation laws is presented. The proposed formulation, which is independent from any specific structure of the hyperbolic system, is then specialised in the sections from 3.5 to 3.8 for the specific case under consideration, i.e. barotropic flows.

3.1 A linearised implicit formulation for a generic hyperbolic system of conservation laws

Let us consider a generic hyperbolic system of conservation laws,

$$\begin{cases} \dfrac{\partial \mathbf{W}}{\partial t} + \dfrac{\partial F(\mathbf{W})}{\partial x} = 0 \\ \mathbf{W}(x,0) = \mathbf{W}_0(x) \end{cases} \tag{3.1}$$

of which, systems (2.11) and (2.15) are particular instances.
A semi-discrete finite-volume formulation for (3.1) can be expressed as

$$\begin{cases} \dfrac{d}{dt}\mathbf{W}_h + \Psi_h^{(p)}(\mathbf{W}_h) = 0 \\ \mathbf{W}_h(0) = \mathbf{W}_h^0 \end{cases} \tag{3.2}$$

where $\mathbf{W}_h = \mathbf{W}_h(t)$ is the state vector representing the semi-discrete solution and $\Psi_h^{(p)}(\,\cdot\,)$ corresponds to the discretization with an accuracy of order p of the spatial differential operator.

In order to avoid severe Courant-Friedrichs-Lewy (CFL) restrictions typical of explicit approaches, an implicit formulation can be used for time integration. In particular, the whole discretization of (3.1) can be achieved through the use of a Backward Differentiation Formula (hereafter BDF), q being the order of accuracy of the method, which yields the following non-linear system:

$$\mathcal{L}_p(\mathbf{W}_h^{n+1}) = 0 \quad \text{with} \quad \mathcal{L}_p(\mathbf{U}) = \frac{1}{\Delta^n t}\left(a_{q,0}^n \mathbf{U} - \sum_{i=1}^{q} a_{q,i}^n \mathbf{W}_h^{n+1-i}\right) + \Psi_h^{(p)}(\mathbf{U}) \tag{3.3}$$

where $\mathbf{W}_h^n$ is an approximation of $\mathbf{W}_h$ in $t = t^n$, $\Delta^n(\,\cdot\,) = (\,\cdot\,)^{n+1} - (\,\cdot\,)^n$ and $a_{q,i}^n$ are the coefficients of the BDF method.

3.1.1 A linearised implicit DeC approach

A possible way to proceed is to define an adequate linearisation of the non linear term in (3.3). Usually linearisation is obtained via differentiation, for instance through the following approximation:

$$\Psi_h^{(p)}(\mathbf{W}_h^{n+1}) \;\simeq \Psi_h^{(p)}(\mathbf{W}_h^n) + \frac{\partial \Psi_h^{(p)}}{\partial \mathbf{W}_h}(\mathbf{W}_h^n)\; \Delta^n \mathbf{W}_h \tag{3.4}$$

Thus, using (3.4) in (3.3), which just corresponds to the application of one Newton iteration to (3.3) starting from $\mathbf{W}_h^n$, the following linear system is obtained:

$$\left[\frac{a_{q,0}^n}{\Delta^n t} I + \frac{\partial \Psi_h^{(p)}}{\partial \mathbf{W}_h}(\mathbf{W}_h^n)\right] \Delta^n \mathbf{W}_h = \frac{1}{\Delta^n t}\left[\sum_{i=1}^{q} a_{q,i}^n \mathbf{W}_h^{n+1-i} - a_{q,0}^n \mathbf{W}_h^n\right] - \Psi_h^{(p)}(\mathbf{W}_h^n). \tag{3.5}$$

In practise, it is not always possible nor convenient to exactly compute the Jacobian matrix of $\Psi_h^{(p)}$. Indeed, it is not unusual to have some lack of differentiability of the numerical fluxes and, moreover, even when the existence of the Jacobian is guaranteed, it can be a rather tedious task to define its exact expression. Thus, the exact Jacobian matrix is commonly replaced by an approximate one or, more generally, by a matrix $M_h^{(p)}$ such that an approximation like (3.4) is still verified. More precisely, adopting the following approximate linearisation

$$\Psi_h^{(p)}(\mathbf{W}_h^{n+1}) \;\simeq \Psi_h^{(p)}(\mathbf{W}_h^n) + M_h^{(p)}(\mathbf{W}_h^n)\left(\mathbf{W}_h^{n+1} - \mathbf{W}_h^n\right) \tag{3.6}$$

the resulting linear system, which should be solved instead of (3.5) at each time step is expressed as:

$$\left[\frac{a^n_{q,0}}{\Delta^n t} I + M_h^{(p)}(\mathbf{W}_h^n)\right] \Delta^n \mathbf{W}_h = \frac{1}{\Delta^n t}\left[\sum_{i=1}^{q} a^n_{q,i}\mathbf{W}_h^{n+1-i} - a^n_{q,0}\mathbf{W}_h^n\right] - \Psi_h^{(p)}(\mathbf{W}_h^n) \tag{3.7}$$

Let us now specialise to first-order accuracy in space and in time, namely taking $p = q = 1$, then (3.7) reduces to:

$$\left[\frac{1}{\Delta^n t} I + M_h^{(1)}(\mathbf{W}_h^n)\right] (\mathbf{W}_h^{n+1} - \mathbf{W}_h^n) = -\Psi_h^{(1)}(\mathbf{W}_h^n). \tag{3.8}$$

In this way, a first-order accurate approximation of a time-dependent solution is obtained at the cost of solving one linear system at each time step.
Increasing the order of accuracy can lead to additional difficulties: the computational effort required for the matrix inversion increases with the order of the spatial accuracy p (the matrix bandwidth and condition number becoming larger). Avoiding the use of an exact Jacobian is not in general sufficient to cure this problem in a satisfactory way. A possible approach to cope with this problem is to introduce Defect-Correction (DeC) techniques, in which a numerical discretization of lower accuracy degree is involved for the implicit part. More precisely, instead of directly solving (3.3), simpler non linear problems are iteratively considered through the following fixed-point like algorithm with $1 \leq m < p$:

$$\begin{cases} \mathcal{W}^0 \text{ given} \\ \mathcal{L}_m(\mathcal{W}^{s+1}) = \mathcal{L}_m(\mathcal{W}^s) - \mathcal{L}_p(\mathcal{W}^s) & \text{for } \ s = 0, \cdots, s_M - 1. \\ \mathbf{W}_h^{n+1} = \mathcal{W}^{s_M} \end{cases} \tag{3.9}$$

Then, adopting the approximate linearisation (3.6), applied to $\Psi_h^{(m)}(\,\cdot\,)$ the resulting linear system, which should be solved at the s-th DeC iteration, is expressed as:

$$\left[\frac{a^n_{q,0}}{\Delta^n t} I + M_h^{(m)}(\mathcal{W}^s)\right] \Delta^s \mathcal{W} = -\frac{1}{\Delta^n t}\left(a^n_{q,0}\mathcal{W}^s - \sum_{i=1}^{q} a^n_{q,i}\mathbf{W}_h^{n+1-i}\right) - \Psi_h^{(p)}(\mathcal{W}^s) \tag{3.10}$$

where Δ^s is clearly understood.
Let us now specialise to second-order accuracy in space, which can already be considered as a high-order accuracy in the context of industrial CFD applications. A formulation as (3.2) is often used also for solving steady hyperbolic problems through a pseudo-time marching procedure. In this case, a first-order accuracy in time is, in general, sufficient since we only need a fast

convergence to a steady solution. Then, one DeC iteration (symbol t is also kept for the pseudo-time, for simplicity) starting from $\mathcal{W}^0 = \mathbf{W}_h^n$ reduces to:

$$\left[\frac{1}{\Delta^n t} I + M_h^{(1)}(\mathbf{W}_h^n)\right] (\mathbf{W}_h^{n+1} - \mathbf{W}_h^n) = -\Psi_h^{(2)}(\mathbf{W}_h^n).$$

In this way, a second-order accurate approximation of the steady solution is obtained at convergence by only solving simplified linear systems. We refer to [43] for a study of the convergence properties. For unsteady cases, this second-order accuracy is not a priori guaranteed. First of all, the time accuracy should also be increased to the second-order yielding the following s-th DeC iteration:

$$\begin{aligned}\left(\frac{(1+2\tau)}{\Delta^n t(1+\tau)} I + M_h^{(1)}(\mathcal{W}^s)\right) \Delta^s \mathcal{W} = \\ -\left(\frac{(1+2\tau)\mathcal{W}^s - (1+\tau)^2 \mathbf{W}_h^n + \tau^2 \mathbf{W}_h^{n-1}}{(1+\tau)\Delta^n t} + \Psi_h^{(2)}(\mathcal{W}^s)\right) \quad (3.11)\end{aligned}$$

where $\tau = \frac{\Delta^n t}{\Delta^{n-1} t}$ is the ratio between two consecutive time steps. Then, a second-order approximation should be recovered asymptotically when a full convergence of the DeC iterations is obtained. Fortunately, a finite termination property of DeC methods ensures that only few DeC iterations are necessary to reach the higher-order of accuracy. Indeed, it has been proved in [81] that $s_M \geq p/(m+1)$ is a sufficient condition to obtain a p^{th}-order accuracy for $q = p$, assuming a linear relationship between time and space steps and taking $\mathcal{W}^0 = \mathbf{W}_h^n$. In particular, one DeC iteration of (3.11) should be sufficient to reach a second-order accurate approximation. In this case, the computation of one time iteration is reduced to the solution of the following linear system:

$$\begin{aligned}\left(\frac{(1+2\tau)}{\Delta^n t(1+\tau)} I + M_h^{(1)}(\mathbf{W}_h^n)\right) (\mathbf{W}_h^{n+1} - \mathbf{W}_h^n) = \\ -\left(\frac{\tau^2}{(1+\tau)} \frac{\mathbf{W}_h^{n-1} - \mathbf{W}_h^n}{\Delta^n t} + \Psi_h^{(2)}(\mathbf{W}_h^n)\right) \quad (3.12)\end{aligned}$$

Note that the theoretical results presented in [81] are obtained by assuming a sufficient regularity and by using, in particular, the exact Jacobians. Nevertheless, this result can be generalised in order to include the complete formulation, i.e. considering also the error due to the linearisation. In particular the second-order accuracy of the one-iteration DeC approach (3.12), in which the first-order linearisation is used, can be checked as shown in Sec. 3.1.2. Furthermore, from a practical point of view, it has been observed that the use of few additional DeC iterations generally increase the robustness of the numerical approach. As a consequence, in 3D industrial applications it is common practise to use several DeC iterations.

3.1.2 An appraisal of the accuracy order for the one-iteration DeC

The linearisations proposed in this thesis and defined in Sec. 3.3 introduce an error which is formally of $O(\nu^2)$ where $\nu^2 = \max((\Delta t)^2, \Delta t \Delta x)$ (linearisations (3.26) and (3.32)) or $\nu^2 = (\Delta t)^2$ (linearisation (3.52)). This kind of linearisation can be applied for the extension to second-order accuracy in space, keeping the same order for the truncation error. Hence, it means that for $m = 1, 2$:

$$\Psi_h^{(m)}(W_h^{n+1}) = \Psi_h^{(m)}(W_h^n) + M_h^{(m)}(W_h^n)\ \Delta^n W_h + O(\nu^2)$$

Let us define $D_\Psi = \Psi_h^{(2)} - \Psi_h^{(1)}$; then we have at least:

$$D_\Psi(W_h^{n+1}) - D_\Psi(W_h^n) = [M_h^{(2)}(W_h^n) - M_h^{(1)}(W_h^n)]\ \Delta^n W_h + O(\nu^2) \quad (3.13)$$

Since D_Ψ corresponds to the difference between a first- and a second-order spatial discretisation of the same differential equation, both $D_\Psi(U^{n+1})$ and $D_\Psi(U^n)$ should be $O(\Delta x)$. Moreover, if D_Ψ is regular enough (e.g. Lipschitzian), $D_\Psi(U^{n+1}) - D_\Psi(U^n)$ should also be $O(\Delta t)$ and, therefore, the following relation is formally introduced:

$$D_\Psi(U^{n+1}) - D_\Psi(U^n) = O(\Delta t \Delta x)$$

Consequently, from (3.13) we obtain that

$$[M_h^{(2)}(W_h^n) - M_h^{(1)}(W_h^n)]\ \Delta^n W_h = O\left((\Delta t)^2, \Delta x \Delta t\right)$$

and thus $M_h^{(1)}$ can be used instead of $M_h^{(2)}$, without loss of accuracy. In particular, we obtain:

$$\Psi_h^{(2)}(W_h^{n+1}) = \Psi_h^{(2)}(W_h^n) + M_h^{(1)}(W_h^n)\ \Delta^n W_h + O((\Delta t)^2, \Delta x \Delta t). \quad (3.14)$$

Let us consider a second-order BDF time discretisation scheme, associated with a second-order spatial discretisation, namely:

$$\frac{3}{2\Delta t} I\left(U^{n+1} - U^n\right) + \frac{U^{n-1} - U^n}{2\Delta t} + \Psi_h^{(2)}(U^{n+1}) = O\left((\Delta t)^2, (\Delta x)^2\right).$$

Thus, by introducing (3.14) we obtain:

$$\begin{aligned}\left(\frac{3}{2\Delta t} + M_h^{(1)}(W_h^n)\right)\Delta^n W_h = \\ -\left(\frac{W_h^{n-1} - W_h^n}{2\Delta t} + \Psi_h^{(2)}(W_h^n)\right) + \mathrm{O}\left((\Delta x)^2, \Delta x \Delta t, (\Delta t)^2\right) \quad (3.15)\end{aligned}$$

Thus, the resulting scheme can be viewed as a second-order accurate approach, since the corresponding error is $\mathrm{O}\left(\eta^2\right)$, with $\eta = \max\left(\Delta x, \Delta t\right)$.

3.2 First-order spatial discretization

In order to completely specify the previously described numerical formulation, a spatial discretisation $\Psi_h^{(p)}(\,\cdot\,)$, as well as a specific linearization $M_h^{(q)}(\,\cdot\,)$ should be defined. Concerning the spatial discretisation, let us consider, in a first step, the first-order accuracy, i.e. $\Psi_h^{(1)}(\,\cdot\,)$. Thus the semi-discrete formulation can be expressed as:

$$\frac{d}{dt}\mathbf{W}_i \;+\; \frac{1}{\delta x_i}\Big(\Phi_{i,i+1} - \Phi_{i-1,i}\Big) = 0 \tag{3.16}$$

where $\mathbf{W}_i$ represents the mean integral value of $\mathbf{W}$ on the i-th cell centred on x_i, δx_i is the length of i-th cell i.e. $\delta x_i = (h_i + h_{i+1})/2$, with $h_i = x_i - x_{i-1}$ and $\Phi_{l,r}$ is the numerical flux function between the l-th (left) cell and the r-th (right) one.

For a generic system of conservation laws several choices are possible. In the following some families of numerical schemes and their main features will be addressed.

3.2.1 Godunov-like schemes

In this work all the considered numerical methods are based on Godunov-type schemes. In the original work of Godunov [53], a Riemann problem is solved at each cell interface and the exact solution is used to define the numerical flux function. However the exact solution of the Riemann problem is in general non-linear (see i.e. appendix A for the solution of the Riemann problem for a convex barotropic state law) thus greatly increasing the computational cost of the numerical method. As a consequence, many Godunov-type schemes, in which the numerical flux function is defined by only approximately solving a Riemann problem at each cell interface, can be found in the literature.

3.2.2 The Roe scheme

One of the most popular choices for $\Phi_{l,r}$ is the so-called Roe numerical flux function [91], which defines a first-order (in space) upwind scheme. We recall here its expression:

$$\Phi_{l,r} = \Phi(\mathbf{W}_l, \mathbf{W}_r) = \frac{F(\mathbf{W}_l) + F(\mathbf{W}_r)}{2} - \frac{1}{2}|\tilde{J}(\mathbf{W}_l, \mathbf{W}_r)|(\mathbf{W}_r - \mathbf{W}_l) \tag{3.17}$$

where the matrix $|\tilde{J}|$ is given by a suitable matrix $\tilde{J}$, called “Roe matrix”, verifying the following conditions for any $(\mathbf{W}_l, \mathbf{W}_r)$:

1. $\tilde{J}(\mathbf{W}_l, \mathbf{W}_r)$ is diagonalizable with real eigenvalues;
2. $\lim\limits_{\mathbf{W}_l, \mathbf{W}_r \to \mathbf{W}^*} \tilde{J}(\mathbf{W}_l, \mathbf{W}_r) = \dfrac{\partial F}{\partial \mathbf{W}}(\mathbf{W}^*)$;
3. $\tilde{J}(\mathbf{W}_l, \mathbf{W}_r)\ (\mathbf{W}_r - \mathbf{W}_l) = F(\mathbf{W}_r) - F(\mathbf{W}_l)$.

The matrix $|\tilde{J}|$ in (3.17) is defined as $|\tilde{J}| = T\,|D|\,T^{-1}$ where T is the matrix of the right eigenvectors associated with $\tilde{J}$ and $|D|_{km} = |\lambda_k|\delta_{km}$, λ_k and δ_{km} respectively being the k-th eigenvalue and the Kronecker symbol. It is well-known that this scheme suffers of some limitations with the possibility to encounter non-physical solutions or shock instabilities (see Chap. 4). These problems can be partially solved by introducing an entropy fix correction. Clearly, the Roe matrix, originally defined in [91] for the case of the Euler equations with an ideal-gas state law, depends on the specific hyperbolic problem under consideration.

3.2.3 The Rusanov scheme

Another possible choice for $\Phi_{l,r}$ is the Rusanov scheme [95], a simple average-state approximate Riemann Solver, which can be expressed as follows:

$$\Phi_{l,r} = \Phi(\mathbf{W}_l, \mathbf{W}_r) = \frac{F(\mathbf{W}_l) + F(\mathbf{W}_r)}{2} - \frac{1}{2} S(\mathbf{W}_l, \mathbf{W}_r)(\mathbf{W}_r - \mathbf{W}_l) \tag{3.18}$$

where

$$S(\mathbf{W}_l, \mathbf{W}_r) = \lambda \mathrm{I} \tag{3.19}$$

and I denotes the identity matrix. Besides, the parameter $\lambda = \lambda\,(\mathbf{W}_l \mathbf{W}_r)$ appearing in (3.19) is an upper bound for the fastest signal velocities of the Riemann problem [107]. It is well known that the Rusanov scheme is more dissipative than the Roe one and, in particular, it is exceedingly dissipative for contact discontinuities (see [107] for the analytical treatment). This means that the Rusanov scheme is not well suited for viscous computations, since a contact discontinuity is numerically equivalent to a limiting case of a viscous boundary layer [84].

On the other hand, as far as the choice of λ is a reasonable one, the stability properties of this scheme are usually better than those of Roe. As for the Roe scheme, the Rusanov scheme depends on the specific hyperbolic problem under consideration. A slightly more general formulation can be to consider a matrix S with different diagonal coefficients, i.e. in which each coefficient can be written in the following form:

$$S_{km} = \lambda_k \delta_{km} \tag{3.20}$$

Indeed this formulation will be useful in the following in order to take into account more sophisticated schemes, as it is shown in Sec. 3.6 and Sec. 3.5.3.

3.2.4 The Q schemes

We can notice that as long as the Rusanov scheme, the Roe scheme or, in general, a scheme defined by (3.18) are concerned, the numerical flux function can be expressed as follows:

$$\Phi\left(\mathbf{W}_l, \mathbf{W}_r\right) = \Phi_{lr} = \Phi_{lr,c} + \Phi_{lr,u} \tag{3.21}$$

in which the centred part $\Phi_{lr,c}$ and the upwind one $\Phi_{lr,u}$ are defined as:

$$\Phi_{lr,c} = \frac{F(\mathbf{W}_r) + F(\mathbf{W}_l)}{2}, \quad \Phi_{lr,u} = -\frac{1}{2}\, Q_{lr}\left(\mathbf{W}_r - \mathbf{W}_l\right) \tag{3.22}$$

where the matrix $Q_{lr} = Q\left(\mathbf{W}_l, \mathbf{W}_r\right)$ is a generic matrix which characterises the particular scheme. Indeed, $Q = |\tilde{J}(\mathbf{W}_l, \mathbf{W}_r)|$ for the Roe scheme (3.17) while $Q = \lambda \mathrm{I}$ for the Rusanov scheme (3.18). However, it is also important to notice that contrary to the Rusanov scheme or, generally speaking, to its extension (3.20), the upwind matrix of the Roe scheme is not diagonal.

3.3 Flux linearisation

Let us consider the flux linearisation, still for the first-order of accuracy, i.e. the matrix $M_h^{(1)}(\,\cdot\,)$. A fully discrete formulation of (3.16) can be obtained by applying a backward Euler scheme for the time discretization of :

$$\frac{\delta x_i}{\Delta^n t}\, \Delta^n \mathbf{W}_i + \Delta^n \Phi_{i,i+1} - \Delta^n \Phi_{i-1,i} = -\left(\Phi^n_{i,i+1} - \Phi^n_{i-1,i}\right).$$

Then, through a classical linearisation via differentiation, the following block tridiagonal linear system, which corresponds to (3.8) is obtained:

$$-H^{(1)n}_{i-1,i}\, \Delta^n \mathbf{W}_{i-1} + \mathcal{H}^n_i \Delta^n \mathbf{W}_i + H^{(2)n}_{i,i+1}\, \Delta^n \mathbf{W}_{i+1} = -\left(\Phi^n_{i,i+1} - \Phi^n_{i-1,i}\right) \tag{3.23}$$

in which:

$$\begin{cases} H^{(1)n}_{i,j} = \dfrac{\partial \Phi_{ij}}{\partial \mathbf{W}_i}(\mathbf{W}^n),\; H^{(2)n}_{i,j} = \dfrac{\partial \Phi_{ij}}{\partial \mathbf{W}_j}(\mathbf{W}^n) \\ \mathcal{H}^n_i = \left(\dfrac{\delta x_i}{\Delta^n t} I + H^{(1)n}_{i,i+1} - H^{(2)n}_{i-1,i}\right) \end{cases} \tag{3.24}$$

As previously pointed out, the exact computation of the Jacobian matrices is not always possible since the numerical flux function is generally not differentiable, as it is indeed the case for the Roe and the Rusanov schemes. On the other hand, simplified Jacobian-free linearisations, which maintain the sparse structure of the final linear system, can be obtained by finding two matrices D_1 and D_2 such that

$$\Delta^n \Phi_{ij} \simeq D_1(\mathbf{W}^n_i, \mathbf{W}^n_j)\, \Delta^n \mathbf{W}_i + D_2(\mathbf{W}^n_i, \mathbf{W}^n_j)\, \Delta^n \mathbf{W}_j. \tag{3.25}$$

In particular, this could be achieved by exploiting the first-order homogeneity of the flux function F (this property being verified, for instance, for the Euler equations for perfect gases); see, e.g. [49]. However, as previously mentioned, in this study we are interested in linearisations which do not require particular properties for the flux function.

3.3.1 A linearisation for the Roe scheme

Let us consider the following relation which is satisfied for any Roe numerical flux function:

$$\Delta^n \Phi_{ij} = \tilde{J}^{+}(\mathbf{W}_i^n, \mathbf{W}_j^n)\, \Delta^n \mathbf{W}_i + \tilde{J}^{-}(\mathbf{W}_i^n, \mathbf{W}_j^n)\, \Delta^n \mathbf{W}_j + \mathcal{R}_{ij}^{n,n+1} \tag{3.26}$$

in which, if $\tilde{J}^{\pm} = \dfrac{1}{2}\left(\tilde{J} \pm |\tilde{J}|\right)$ are sufficiently regular, $\mathcal{R}_{ij}^{n,n+1}$ is such that $\|\mathcal{R}_{ij}^{n,n+1}\| = O\left((\Delta t)^2, \Delta t \Delta x\right)$ (see [102, 103] for a proof) and thus can be neglected in (3.26), yielding the following block tridiagonal linear system in place of (3.23):

$$B_{-1}^{i,n}\, \Delta^n \mathbf{W}_{i-1} + B_0^{i,n}\, \Delta^n \mathbf{W}_i + B_1^{i,n} \Delta^n \mathbf{W}_{i+1} = -\left(\Phi_{i,i+1}^n - \Phi_{i-1,i}^n\right) \tag{3.27}$$

where:

$$\begin{cases} B_{-1}^{i,n} = -\tilde{J}^{+}(\mathbf{W}_{i-1}^n, \mathbf{W}_i^n) \\ \\ B_0^{i,n} = \dfrac{\delta x_i}{\Delta^n t}\, I + \tilde{J}^{+}(\mathbf{W}_i^n, \mathbf{W}_{i+1}^n) - \tilde{J}^{-}(\mathbf{W}_{i-1}^n, \mathbf{W}_i^n) \\ \\ B_1^{i,n} = \tilde{J}^{-}(\mathbf{W}_i^n, \mathbf{W}_{i+1}^n) \end{cases} \tag{3.28}$$

It may be worth remarking that the linearisation has been derived by exploiting only the algebraic properties of the Roe matrix and, therefore, it does not depend on the specific equation of state which completes the mathematical problem.

3.3.2 A simple linearisation for the Q-schemes

Due to its particular construction, the linearisation (3.26) can be used only for the Roe scheme. More generally, let us now consider a generic Q-scheme. Without assuming any particular algebraic property for the matrix Q, let us express $\Delta^n \Phi_{ij}$ as follows:

$$\Delta^n \Phi_{ij} = \Delta^n \Phi_{ij,c} + \Delta^n \Phi_{ij,u}$$

where $\Phi_{ij,c}$ and $\Phi_{ij,u}$ are defined by (3.22). Since the centred term is a differentiable function of $\mathbf{W}$, its variation can be easily approximated by a

Taylor expansion as:

$$\Delta^n \Phi_{ij,c} = \frac{1}{2}\left(A(\mathbf{W}_i^n)\Delta^n \mathbf{W}_i + A(\mathbf{W}_j^n)\Delta^n \mathbf{W}_j\right) + O(\Delta t^2) \tag{3.29}$$

in which A is the Jacobian matrix of F, i.e. $A \doteq \dfrac{\partial F}{\partial \mathbf{W}}$.

It is more difficult to estimate the variation of the upwind term because it is generally non differentiable. However, it is possible to rewrite this term as:

$$\Delta^n \Phi_{ij,u} = -\frac{Q_{ij}^n}{2}(\Delta^n \mathbf{W}_j - \Delta^n \mathbf{W}_i) - \frac{\Delta^n Q_{ij}}{2}(\mathbf{W}_j^{n+1} - \mathbf{W}_i^{n+1}) \tag{3.30}$$

Only the second term of the right hand side of (3.30) involves quantities at time $n+1$ and, as a consequence, several approximations of $\Delta^n \Phi_{ij,u}$ can be obtained simply varying the way in which this implicit term is approximated. A first linearisation can be derived by assuming that the solution is regular enough to satisfy:

$$\mathbf{W}_j^{n+1} - \mathbf{W}_i^{n+1} \propto O(\Delta x) \quad \text{and} \quad \Delta^n Q_{ij} \propto O(\Delta t) \tag{3.31}$$

which implies that

$$\Gamma_{ij}^{n,n+1} \doteq \frac{\Delta^n Q_{ij}}{2}(\mathbf{W}_j^{n+1} - \mathbf{W}_i^{n+1}) \propto O(\Delta t \Delta x)$$

Then, an approximation of $\Delta^n \Phi_{ij}$ can be obtained by neglecting the term $R_{ij}^{n,n+1}$ in the following expression:

$$\begin{aligned}\Delta^n \Phi_{ij} =& \frac{1}{2}\left(A(\mathbf{W}_i^n)\Delta^n \mathbf{W}_i + A(\mathbf{W}_j^n)\Delta^n \mathbf{W}_j\right) \\ & - \frac{Q(\mathbf{W}_i^n, \mathbf{W}_j^n)}{2}(\Delta^n \mathbf{W}_j - \Delta^n \mathbf{W}_i) + R_{ij}^{n,n+1}\end{aligned} \tag{3.32}$$

where $R_{ij}^{n,n+1}$ corresponds to $\Gamma_{ij}^{n,n+1}$ plus the term in $O(\Delta t^2)$ of (3.29). If (3.31) holds true, $R_{ij}^{n,n+1}$ is such that $\|R_{ij}^{n,n+1}\| = O\left((\Delta t)^2, \Delta t \Delta x\right)$. We can remark that $\Delta^n \Phi_{ij,u}$, is already a term in $O(\Delta t \Delta x)$, and thus, since a contribution of the upwind part in the approximation of $\Delta^n \Phi_{ij}$ is suitable, a term of the same order of $R_{ij}^{n,n+1}$ is also present in (3.32). It corresponds on the non neglected part of $\Delta^n \Phi_{ij,u}$. On the other hand, by comparing (3.32) with (3.26), it is clear that this approach leads to a truncation error analogous to that of the linearisation for the Roe scheme. It is worth noting that this approach, which can be used as long as (3.31) is satisfied, is quite general since it does not require any particular property for the matrix Q_{ij}. Moreover, it corresponds to the frozen Jacobian method described e.g. in [4]. Note that in [4] this approach has been used for both the Roe and the Rusanov schemes.

3.3.3 A complete linearisation for diagonal Q-schemes

Hypothesis (3.31) is a reasonable one but there are situations of practical interest in which it is not satisfied, as it is pointed out in [4]. If a discontinuity is present (i.e. a shock wave), the magnitude of the term $\mathbf{W}_j^{n+1} - \mathbf{W}_i^{n+1}$ can be large, independently of the size of Δx. Moreover, the term $\Delta^n Q_{ij}$ can be not so small. This can happen, in particular, when the speed of sound has a stiff change in magnitude or in presence of huge variations of the flow velocity.

Let us focus now on diagonal Q-schemes, i.e. when the upwind matrix Q is a diagonal matrix S as, for instance, for the Rusanov scheme. To estimate the second part of (3.30), i.e. $\Gamma_{ij}^{n,n+1}$, let us consider, in a first step, one diagonal coefficient of S, λ_k. Indeed, due to the particular expression of the matrix S, approximating $\Delta^n S$ is equivalent to find an approximation for each term $\Delta^n \lambda_k$. Moreover, we assume here that λ_k is a composite function of two variables, a and u, i.e.

$$\lambda_k = \lambda_k \left(u \left(\mathbf{W}_i(t), \mathbf{W}_j(t) \right), a \left(\mathbf{W}_i(t), \mathbf{W}_j(t) \right) \right) \tag{3.33}$$

Note that this particular dependency on the flow variables is always verified for the numerical schemes proposed in this study (see section 3.5).

In the following we assume that the solution is differentiable with respect to time and it is regular enough so that the following Taylor expansion holds true:

$$\Delta^n \lambda_k = \left. \frac{\mathrm{d}\lambda_k}{\mathrm{d}t} \right|_{t^n} \Delta^n t \, + \, O\left((\Delta^n t)^2 \right) \tag{3.34}$$

it is possible to recast the term $\dfrac{\mathrm{d}\lambda_k}{\mathrm{d}t}$ in (3.34) as:

$$\begin{aligned} \frac{\mathrm{d}\lambda_k}{\mathrm{d}t} = \frac{\partial \lambda_k}{\partial u} & \left(\left\langle \frac{\partial u}{\partial \mathbf{W}_i}, \frac{\mathrm{d}\mathbf{W}_i}{\mathrm{d}t} \right\rangle + \left\langle \frac{\partial u}{\partial \mathbf{W}_j}, \frac{\mathrm{d}\mathbf{W}_j}{\mathrm{d}t} \right\rangle \right) \\ + \frac{\partial \lambda_k}{\partial a} & \left(\left\langle \frac{\partial a}{\partial \mathbf{W}_i}, \frac{\mathrm{d}\mathbf{W}_i}{\mathrm{d}t} \right\rangle + \left\langle \frac{\partial a}{\partial \mathbf{W}_j}, \frac{\mathrm{d}\mathbf{W}_j}{\mathrm{d}t} \right\rangle \right) \end{aligned} \tag{3.35}$$

where the notation $\langle .,. \rangle$ represents the Euclidean scalar product in $\mathbb{R}^\nu$, ν being the number of components of $\mathbf{W}_i$. Moreover, considering also a Taylor expansion in time at $\mathbf{W}_i$, we obtain:

$$\begin{aligned} \left. \frac{\mathrm{d}\lambda_k}{\mathrm{d}t} \right|_{t^n} = \left. \frac{\partial \lambda_k}{\partial u} \right|_{t^n} & \left(\left\langle \left. \frac{\partial u}{\partial \mathbf{W}_i} \right|_{t^n}, \frac{\Delta^n \mathbf{W}_i}{\Delta^n t} \right\rangle + \left\langle \left. \frac{\partial u}{\partial \mathbf{W}_j} \right|_{t^n}, \frac{\Delta^n \mathbf{W}_j}{\Delta^n t} \right\rangle \right) + \\ \left. \frac{\partial \lambda_k}{\partial a} \right|_{t^n} & \left(\left\langle \left. \frac{\partial a}{\partial \mathbf{W}_i} \right|_{t^n}, \frac{\Delta^n \mathbf{W}_i}{\Delta^n t} \right\rangle + \left\langle \left. \frac{\partial a}{\partial \mathbf{W}_j} \right|_{t^n}, \frac{\Delta^n \mathbf{W}_j}{\Delta^n t} \right\rangle \right) + O\left(\Delta^n t \right) \end{aligned} \tag{3.36}$$

and consequently, by substituting (3.36) into (3.34) the following expression is achieved:

$$\Delta^n\lambda_k = \left.\frac{\partial\lambda_k}{\partial u}\right|_{t^n}\left(\left\langle \left.\frac{\partial u}{\partial \mathbf{W}_i}\right|_{t^n}, \Delta^n\mathbf{W}_i\right\rangle + \left\langle \left.\frac{\partial u}{\partial \mathbf{W}_j}\right|_{t^n}, \Delta^n\mathbf{W}_j\right\rangle\right) + \left.\frac{\partial\lambda_k}{\partial a}\right|_{t^n}\left(\left\langle \left.\frac{\partial a}{\partial \mathbf{W}_i}\right|_{t^n}, \Delta^n\mathbf{W}_i\right\rangle + \left\langle \left.\frac{\partial a}{\partial \mathbf{W}_j}\right|_{t^n}, \Delta^n\mathbf{W}_j\right\rangle\right) + O\left((\Delta^n t)^2\right) \quad (3.37)$$

(3.37) are independent from the index k, that is from the particular term of matrix S into consideration.
Furthermore, $\Delta^n S_{ij}$ can be directly written as a function of the terms $\Delta^n\lambda_k$ since the coefficients of this matrix can be expressed as:

$$\left(\Delta^n S_{ij}\right)_{km} = \delta_{km}\Delta^n\lambda_k$$

Thus, using (3.37) the following expression for $\Delta^n S_{ij}$ is obtained:

$$\Delta^n S_{ij} = \left\langle \left.\frac{\partial u}{\partial \mathbf{W}_i}\right|_{t^n}, \Delta^n\mathbf{W}_i\right\rangle \left.\frac{\partial S}{\partial u}\right|_{t^n} + \left\langle \left.\frac{\partial a}{\partial \mathbf{W}_i}\right|_{t^n}, \Delta^n\mathbf{W}_i\right\rangle \left.\frac{\partial S}{\partial a}\right|_{t^n} + \left\langle \left.\frac{\partial u}{\partial \mathbf{W}_j}\right|_{t^n}, \Delta^n\mathbf{W}_j\right\rangle \left.\frac{\partial S}{\partial u}\right|_{t^n} + \left\langle \left.\frac{\partial a}{\partial \mathbf{W}_j}\right|_{t^n}, \Delta^n\mathbf{W}_j\right\rangle \left.\frac{\partial S}{\partial a}\right|_{t^n} + O\left((\Delta^n t)^2\right) \quad (3.38)$$

where the matrices $\frac{\partial S}{\partial u}$ and $\frac{\partial S}{\partial a}$ are defined as follows:

$$\text{for } k \text{ and } m = 1,\nu \qquad \left(\frac{\partial S}{\partial u}\right)_{km} \doteq \delta_{km}\frac{\partial\lambda_k}{\partial u} \quad \text{and} \quad \left(\frac{\partial S}{\partial a}\right)_{km} \doteq \delta_{km}\frac{\partial\lambda_k}{\partial a} \quad (3.39)$$

Finally we can recast (3.38) in the form

$$\Delta^n S_{ij} = \underbrace{H_1\left(\mathbf{W}_i^n, \mathbf{W}_j^n, \Delta^n\mathbf{W}_i\right)}_{\mathcal{H}_{1,ij}} + \underbrace{H_2\left(\mathbf{W}_i^n, \mathbf{W}_j^n, \Delta^n\mathbf{W}_j\right)}_{\mathcal{H}_{2,ij}} + O\left((\Delta^n t)^2\right) \quad (3.40)$$

Using (3.40) it is now possible to approximate the term $\Delta^n S_{ij}(\mathbf{W}_j^{n+1} - \mathbf{W}_i^{n+1})$. Firstly let us decompose $\mathbf{W}_j^{n+1} - \mathbf{W}_i^{n+1}$ as:

$$\mathbf{W}_j^{n+1} - \mathbf{W}_i^{n+1} = \mathbf{W}_j^n - \mathbf{W}_i^n + \Delta^n\mathbf{W}_j - \Delta^n\mathbf{W}_i \quad (3.41)$$

then we can write

$$\begin{aligned}
\Delta^n S_{ij}(\mathbf{W}_j^{n+1} - \mathbf{W}_i^{n+1}) = & \left(\mathcal{H}_{1,ij} + \mathcal{H}_{2,ij} + O\left((\Delta^n t)^2\right)\right)\left(\mathbf{W}_j^{n+1} - \mathbf{W}_i^{n+1}\right) \\
= & \left(\mathcal{H}_{1,ij} + \mathcal{H}_{2,ij}\right)\left(\mathbf{W}_j^{n} - \mathbf{W}_i^{n}\right) \\
& + \left(\mathcal{H}_{1,ij} + \mathcal{H}_{2,ij}\right)\left(\Delta^n \mathbf{W}_j - \Delta^n \mathbf{W}_i\right) \\
& + O\left((\Delta^n t)^2\right)\left(\mathbf{W}_j^{n+1} - \mathbf{W}_i^{n+1}\right)
\end{aligned} \tag{3.42}$$

Then, we obtain the following expression:

$$\Delta^n S_{ij}(\mathbf{W}_j^{n+1} - \mathbf{W}_i^{n+1}) = \left(\mathcal{H}_{1,ij} + \mathcal{H}_{2,ij}\right)\left(\mathbf{W}_j^{n} - \mathbf{W}_i^{n}\right) + \bar{R}_{ij}^{n,n+1} \tag{3.43}$$

where $\bar{R}_{ij}^{n,n+1}$, defined as follows, should be neglected:

$$\begin{aligned}
\bar{R}_{ij}^{n,n+1} = & \left(\mathcal{H}_{1,ij} + \mathcal{H}_{2,ij}\right)\left(\Delta^n \mathbf{W}_j - \Delta^n \mathbf{W}_i\right) \\
& + O\left((\Delta^n t)^2\right)\left(\mathbf{W}_j^{n+1} - \mathbf{W}_i^{n+1}\right)
\end{aligned} \tag{3.44}$$

To estimate the order of magnitude of $\|\bar{R}_{ij}^{n,n+1}\|$ we firstly assume that the solution is sufficiently regular both in space and in time in order to have:

$$\begin{cases}
\mathbf{W}_j^{n+1} - \mathbf{W}_i^{n+1} & \propto & O\left(\Delta x\right) \\
\Delta^n \mathbf{W}_j - \Delta^n \mathbf{W}_i & \propto & O\left(\Delta x \Delta t\right) \\
\mathcal{H}_{1,ij}^n + \mathcal{H}_{2,ij}^n & \propto & O\left(\Delta t\right)
\end{cases} \tag{3.45}$$

where the last one is a direct consequence that $\mathcal{H}_{1,ij}^n + \mathcal{H}_{2,ij}^n$ is an approximation of $\frac{\mathrm{d}S}{\mathrm{d}t}\Delta^n t$. Substituting (3.45) into (3.44) yields:

$$\|\bar{R}_{ij}^{n,n+1}\| = O\left((\Delta^n t)^2 \Delta x\right) \tag{3.46}$$

Since we are trying to approximate terms which have been previously neglected in the first linearisation (3.32), an increase in accuracy is expected. Actually, an improvement for $\|\bar{R}_{ij}^{n,n+1}\|$ of one order of magnitude (more precisely of order of $\max(\Delta t, \Delta x)$) is achieved when the solution is regular enough. Nevertheless, as previously pointed out, we are interested here in less regular solutions. This second linearisation is of major interest in this case, since for regular solutions the frozen Jacobian approach is already accurate enough with respect to the global order of accuracy of the numerical method. Indeed, we consider here the formulation (3.8) which is first-order accurate in time and space, and moreover, even with a higher accurate approach, i.e. with (3.7) with $p > 1$ and $q > 1$, we are limited to second order in time as long as approximation (3.29) is used. On the contrary, it seems crucial to take into account an approximation of $\Gamma_{ij}^{n,n+1}$ when $\Delta^n S_{ij}$ and $\mathbf{W}_j^{n+1} - \mathbf{W}_i^{n+1}$ are independent from the time and space discretization as for instance, in presence of a shock where no assumption on the regularity in space of the solution can be done. Nevertheless, in this case, a second-order accuracy is still obtained if we suppose that the solution is regular enough in time to satisfy (3.34). Indeed, it is then possible to write

$$\begin{cases} \Delta^n\mathbf{W}_j - \Delta^n\mathbf{W}_i & \propto \ O\left(\Delta t\right) \\ \mathcal{H}_{1,ij}^n + \mathcal{H}_{2,ij}^n & \propto \ O\left(\Delta t\right) \end{cases} \tag{3.47}$$

and substituting (3.47) into (3.44) yields:

$$\|\bar{R}_{ij}^{n,n+1}\| = O\left((\Delta^n t)^2\right) \tag{3.48}$$

As done for the previous linearisations, i.e. (3.26) and (3.32), we are interested in obtaining an approximation of $\Delta^n\Phi_{ij}$ as proposed in equation (3.25), that is with a linear dependency on the unknowns $\Delta^n\mathbf{W}_i$ and $\Delta^n\mathbf{W}_j$. However, from (3.43) we have apparently lost this kind of expression. Taking into account (3.43), and using (3.38), (3.40) and (3.43), the term depending on $\Delta^n\mathbf{W}_i$ can be expressed as follows:

$$\left[\left\langle \frac{\partial u}{\partial \mathbf{W}_i}^n, \Delta^n\mathbf{W}_i\right\rangle \frac{\partial S}{\partial u}^n + \left\langle \frac{\partial a}{\partial \mathbf{W}_i}^n, \Delta^n\mathbf{W}_i\right\rangle \frac{\partial S}{\partial a}^n\right]\left(\mathbf{W}_j^n - \mathbf{W}_i^n\right)$$

Let us consider only one of the previous terms, i.e. let us define the following vector:

$$T_{ij}^n = \left\langle \frac{\partial v}{\partial \mathbf{W}_i}^n, \Delta^n\mathbf{W}_i\right\rangle \frac{\partial S}{\partial v}^n \left(\mathbf{W}_j^n - \mathbf{W}_i^n\right)$$

in which v can be either u or a. Then, developing both the scalar product and the product matrix by vector, we obtain the following expression for the k-th component of T_{ij}^n:

$$(T_{ij}^n)_k = \left[\sum_{m=1}^{\nu} \frac{\partial v^n}{\partial \mathbf{W}_{i,m}} \Delta^n\mathbf{W}_{i,m}\right] \sum_{s=1}^{\nu} \left(\frac{\partial S}{\partial v}^n\right)_{ks} \left(\mathbf{W}_{j,s}^n - \mathbf{W}_{i,s}^n\right)$$

$\mathbf{W}_{i,s}^n$ being the s-th component of $\mathbf{W}_i^n$. Then, the previous expression can be simplified using (3.39) as:

$$\begin{aligned} (T_{ij}^n)_k &= \left[\sum_{m=1}^{\nu} \frac{\partial v^n}{\partial \mathbf{W}_{i,m}} \Delta^n\mathbf{W}_{i,m}\right] \frac{\partial \lambda_k}{\partial v}^n \left(\mathbf{W}_{j,k}^n - \mathbf{W}_{i,k}^n\right) \\ &= \sum_{m=1}^{\nu} \left[\frac{\partial \lambda_k}{\partial v}^n \frac{\partial v^n}{\partial \mathbf{W}_{i,m}} \left(\mathbf{W}_{j,k}^n - \mathbf{W}_{i,k}^n\right)\right] \Delta^n\mathbf{W}_{i,m} \end{aligned} \tag{3.49}$$

Consequently, an adequate dependency on $\Delta^n \mathbf{W}_i$ has been obtained and (3.43) can be rewritten as follows:

$$\Delta^n S_{ij}\left(\mathbf{W}_j^{n+1} - \mathbf{W}_i^{n+1}\right) = K_{ij}^{+}\Delta^n \mathbf{W}_j + K_{ij}^{-}\Delta^n \mathbf{W}_i + \bar{R}_{ij}^{n,n+1} \tag{3.50}$$

in which $\bar{R}_{ij}^{n,n+1}$ is the one defined in (3.44) while the matrices K_{ij}^{+} and K_{ij}^{-} are defined by:

$$\left\{\begin{array}{lcl} (K_{ij}^{+})_{km} & = & \left(\dfrac{\partial \lambda_k}{\partial u}\dfrac{\partial u}{\partial \mathbf{W}_{j,m}} + \dfrac{\partial \lambda_k}{\partial a}\dfrac{\partial a}{\partial \mathbf{W}_{j,m}}\right)\left(\mathbf{W}_{j,k}^n - \mathbf{W}_{i,k}^n\right) \\ \\ (K_{ij}^{-})_{km} & = & \left(\dfrac{\partial \lambda_k}{\partial u}\dfrac{\partial u}{\partial \mathbf{W}_{i,m}} + \dfrac{\partial \lambda_k}{\partial a}\dfrac{\partial a}{\partial \mathbf{W}_{i,m}}\right)\left(\mathbf{W}_{j,k}^n - \mathbf{W}_{i,k}^n\right) \end{array}\right. \tag{3.51}$$

Finally, using (3.50) the following more complete approximation of the time variation of the numerical flux function can be obtained:

$$\begin{aligned} \Delta^n \Phi_{ij} \simeq & \left(\frac{1}{2}\left[A(\mathbf{W}_i^n) + S(\mathbf{W}_i^n, \mathbf{W}_j^n)\right] + K_{ij}^{-}\right)\Delta^n \mathbf{W}_i \\ & + \left(\frac{1}{2}\left[A(\mathbf{W}_j^n) - S(\mathbf{W}_i^n, \mathbf{W}_j^n)\right] + K_{ij}^{+}\right)\Delta^n \mathbf{W}_j \end{aligned} \tag{3.52}$$

If we use now this linearisation in the formulation (3.8), as for the case of the Roe scheme, we obtain a block tridiagonal linear system for a general diagonal Q-scheme. More precisely:

$$C_{-1}^{i,n}\,\Delta^n \mathbf{W}_{i-1} + C_0^{i,n}\,\Delta^n \mathbf{W}_i + C_1^{i,n}\Delta^n \mathbf{W}_{i+1} = -\left(\Phi_{i,i+1}^n - \Phi_{i-1,i}^n\right) \tag{3.53}$$

where:

$$\left\{\begin{array}{l} C_{-1}^{i,n} = \dfrac{-A(\mathbf{W}_{i-1}^n) - S_{i-1,i}^n + K_{i-1,i}^{-}}{2} \\ \\ C_0^{i,n} = \dfrac{\delta x_i}{\Delta^n t} I + \dfrac{S_{i-1,i}^n + S_{i,i+1}^n + K_{i-1,i}^{+} - K_{i,i+1}^{-}}{2} \\ \\ C_1^{i,n} = \dfrac{A(\mathbf{W}_{i+1}^n) - S_{i,i+1}^n - K_{i,i+1}^{+}}{2} \end{array}\right. \tag{3.54}$$

Note that if K^+ and K^- are null matrices, this linearisation reduces to the previous one obtained from (3.32).

3.4 Extension to second-order accuracy

In this section the numerical techniques used to reach the second-order of accuracy both in space and in time are described. First, in Sec. 3.4.1 the numerical formulation proposed in sections 3.2 and 3.3 is extended to the second-order accuracy in space, then in Sec. 3.4.2 the second-order accuracy in time is achieved.

3.4.1 Spatial second-order accuracy

The extension to second-order accuracy in space can be achieved by using a classical MUSCL technique [115], in which instead of $\Phi_{i,i+1} = \Phi(\mathbf{W}_i, \mathbf{W}_{i+1})$, the numerical flux $\Phi_{i+\frac{1}{2}} = \Phi(\mathbf{W}^-_{i+\frac{1}{2}}, \mathbf{W}^+_{i+\frac{1}{2}})$ is considered at the cell interface $x_{i+\frac{1}{2}}$. The considered values $\mathbf{W}^{\pm}_{i+\frac{1}{2}}$ are defined by piecewise linear reconstruction of the solution, and can be expressed as follows in a β-scheme form (see, e.g. [58]):

$$\begin{cases} \mathbf{W}^-_{i+\frac{1}{2}} = \mathbf{W}_i + \dfrac{h_{i+1}}{2}\left[(1-\beta)\dfrac{\mathbf{W}_{i+1}-\mathbf{W}_i}{h_{i+1}} + \beta\dfrac{\mathbf{W}_i-\mathbf{W}_{i-1}}{h_i}\right] \\[2ex] \mathbf{W}^+_{i+\frac{1}{2}} = \mathbf{W}_{i+1} - \dfrac{h_{i+1}}{2}\left[(1-\beta)\dfrac{\mathbf{W}_{i+1}-\mathbf{W}_i}{h_{i+1}} + \beta\dfrac{\mathbf{W}_{i+2}-\mathbf{W}_{i+1}}{h_{i+2}}\right] \end{cases} \tag{3.55}$$

where β is a given parameter.

Formulation for second-order schemes with limiters

The slopes used for the MUSCL extrapolation have to be limited in order to avoid overshoots, specially when strong gradients and/or discontinuities are present in the solution (see, e.g. [78]). More precisely, the average values $\mathbf{W}^{\pm}_{i+\frac{1}{2}}$ defined in (3.55) should be substituted by the following limited values:

$$\begin{cases} \mathbf{W}^-_{i+\frac{1}{2}} = \mathbf{W}_i + \dfrac{h_{i+1}}{2}\left(\Lambda^-_{i+\frac{1}{2}}\dfrac{\mathbf{W}_{i+1}-\mathbf{W}_i}{h_{i+1}} + \Lambda^+_{i-\frac{1}{2}}\dfrac{\mathbf{W}_i-\mathbf{W}_{i-1}}{h_i}\right) \\[2ex] \mathbf{W}^+_{i+\frac{1}{2}} = \mathbf{W}_{i+1} - \dfrac{h_{i+1}}{2}\left(\Lambda^+_{i+\frac{1}{2}}\dfrac{\mathbf{W}_{i+1}-\mathbf{W}^n_i}{h_{i+1}} + \Lambda^-_{i+\frac{3}{2}}\dfrac{\mathbf{W}^n_{i+2}-\mathbf{W}_{i+1}}{h_{i+2}}\right) \end{cases} \tag{3.56}$$

in which $\Lambda^{\pm}_l$ are diagonal matrices introducing the (non linear) limiter function. Several classical choices can be used to obtain a total variation diminishing scheme [107]. For instance, in the numerical experiments presented in section 4, the slope limiter of van Albada [113] has been used. Then, in particular, matrix $\Lambda^-_{i+\frac{1}{2}}$ depends on $\mathbf{W}_{i-1}, \mathbf{W}_i$ and $\mathbf{W}_{i+1}$ while $\mathbf{W}_i, \mathbf{W}_{i+1}$ and $\mathbf{W}_{i+2}$ are involved in the definition of $\Lambda^+_{i+\frac{1}{2}}$. Thus, the introduction of limiters adds additional levels of non linearity: the limiter by itself is not linear (nor differentiable) so as the products like $\Lambda_l^{\pm}\mathbf{W}_j$. Consequently, in this case a DeC approach seems to be particularly suitable since it requires only the linearisation of the first order numerical flux function, where this additional non-linearity is not present.

3.4.2 DeC second-order formulation

Let us consider the algorithm (3.11) associated with the linearisation (3.25). In this case, the following space and time second-order accurate implicit formulation is obtained:

$$\begin{cases} \mathbf{W}^0 = \mathbf{W}^n \\ \\ \mathcal{B}_{-1}^{i,s}\Delta^s\mathbf{W}_{i-1} + \mathcal{B}_0^{i,s}\Delta^s\mathbf{W}_i + \mathcal{B}_1^{i,s}\Delta^s\mathbf{W}_{i+1} = \mathcal{S}_i^s \quad s = 0,\cdots,s_M-1 \\ \\ \mathbf{W}^{n+1} = \mathbf{W}^{s_M} \end{cases} \tag{3.57}$$

in which:

$$\begin{cases} \mathcal{B}_{-1}^{i,s} = -D_1(\mathbf{W}_{i-1}^s, \mathbf{W}_i^s) \\ \\ \mathcal{B}_0^{i,s} = \dfrac{\delta x_i(1+2\tau)}{(1+\tau)\Delta^n t} I + D_1(\mathbf{W}_i^s, \mathbf{W}_{i+1}^s) - D_2(\mathbf{W}_{i-1}^s, \mathbf{W}_i^s) \\ \\ \mathcal{B}_1^{i,s} = D_2(\mathbf{W}_i^s, \mathbf{W}_{i+1}^s) \\ \\ \mathcal{S}_i^s = -\left(\delta x_i \dfrac{(1+2\tau)\mathcal{W}^s - (1+\tau)^2\mathbf{W}_h^n + \tau^2\mathbf{W}_h^{n-1}}{(1+\tau)\Delta^n t} + \Phi_{i+\frac{1}{2}}^s - \Phi_{i-\frac{1}{2}}^s\right) \end{cases} \tag{3.58}$$

D_1 and D_2 being the generic matrices of the approximation (3.25) which, for the specific case of the Roe and diagonal Q-schemes are defined by, respectively, (3.26) and (3.52).

3.5 Barotropic flows

The formulation defined in the previous sections for a hyperbolic system of conservation laws is rather general since it can be applied to a generic flux function F. However, as previously pointed out, the complete definition of the aforementioned numerical methods, depends on the specific hyperbolic problem under consideration. Hence, hereafter the 1D system (2.11) closed by a convex barotropic state law is considered.

The numerical flux functions considered in this section are constructed through an approximation of the Riemann problem, by considering a linearised problem, as in the case of the Roe scheme, or an average-state approximation, as in the case of HLL schemes. The solution for the Riemann problem with a generic convex barotropic equation of state has been described in [102,103] and it is reported in appendix A. It is composed by three waves: the intermediate one is always a contact discontinuity, the others can be shock or rarefaction waves. This simple three-waves configuration is obtained in particular for the barotropic EOS described in section 2.2, i.e. Eqs. (2.6) and (2.7). Indeed, the convexity is trivially obtained for the wetted regime, while, even if it has not be rigorously proved due to the complexity of (2.7), this assumption has been also systematically verified by an analytical/numerical analysis for all the considered cavitating mixtures (we refer to [102] for details).

3.5.1 The Roe scheme

The generic Roe scheme described in (3.2.2) is completely defined once a suitable Roe matrix is derived for the specific hyperbolic problem under consideration. For a generic barotropic state equation and the 1D hyperbolic system (2.11), the following Roe matrix has been derived in [102,103]:

$$\tilde{J}_{ij} = \begin{pmatrix} 0 & 1 & 0 \\ \tilde{a}_{ij}^2 - \tilde{u}_{ij}^2 & 2\tilde{u}_{ij} & 0 \\ -\tilde{u}_{ij}\tilde{\xi}_{ij} & \tilde{\xi}_{ij} & \tilde{u}_{ij} \end{pmatrix} \tag{3.59}$$

in which $\tilde{u}_{ij}$ and $\tilde{\xi}_{ij}$ correspond to the well-known "Roe averages" for the states $\mathbf{W}_i$ and $\mathbf{W}_j$ of u and ξ respectively, i.e.:

$$\tilde{u}_{lr} = \frac{\sqrt{\rho_i}u_i + \sqrt{\rho_j}u_j}{\sqrt{\rho_i} + \sqrt{\rho_j}} \quad \text{and} \quad \tilde{\xi}_{lr} = \frac{\sqrt{\rho_i}\,\xi_i + \sqrt{\rho_j}\,\xi_j}{\sqrt{\rho_i} + \sqrt{\rho_j}}$$

whereas $\tilde{a}_{ij}$, which can be considered as a Roe average for the sound speed, is defined as:

$$\tilde{a}_{ij} = \begin{cases} \sqrt{\dfrac{p(\rho_j) - p(\rho_i)}{\rho_j - \rho_i}} & \text{if} \quad \rho_j \neq \rho_i \\[2ex] a\left(\rho_\star, p(\rho_\star)\right) & \text{if} \quad \rho_i = \rho_j = \rho_\star \end{cases} \tag{3.60}$$

Note that as far as the numerical implementation is concerned, $\tilde{a}_{lr}$ should be defined as follows:

$$\tilde{a}_{lr}^2 = \begin{cases} \dfrac{p(\rho_r) - p(\rho_l)}{\rho_r - \rho_l} & \text{if} \mid \rho_r - \rho_l \mid > \epsilon \\[2ex] a^2(\bar{\rho}_{lr}, p(\bar{\rho}_{lr})) & \text{if} \mid \rho_r - \rho_l \mid < \epsilon \end{cases}$$

where ϵ is a suitable numerical threshold and $\bar{\rho}_{lr}$ is an average value such that: $\lim_{W_l, W_r \to W^\star} \bar{\rho}_{lr} = \rho_\star$. Besides, the eigenvalues of the matrix $\tilde{J}_{ij}$ have the same formal expression of whose of the analytical Jacobian (2.12):

$$\tilde{\lambda}_1 = \tilde{u}_{ij} + \tilde{a}_{ij}\ , \ \tilde{\lambda}_2 = \tilde{u}_{ij} - \tilde{a}_{ij} \ \text{and} \ \tilde{\lambda}_3 = \tilde{u}_{ij} \tag{3.61}$$

3.5.2 A simple HLL scheme

The HLL schemes [58], which assume one intermediate wave state between two acoustic waves in the approximate Riemann problem, are widely used due to their simplicity, reliability and robustness. The choice of s_L and s_R, the acoustic wave speed estimates, fully determines each particular scheme. The simplest approach, known as the Rusanov scheme [41, 95], is obtained by considering $s_L = -s_R$, and thus, can be expressed as

$$\Phi_{i,j} = \frac{F(\mathbf{W}_i) + F(\mathbf{W}_j)}{2} - \frac{\lambda_{ij}}{2}(\mathbf{W}_j - \mathbf{W}_i) \tag{3.62}$$

where $\lambda_{ij} = s_R$ is an upper bound for the fastest wave speed. A classical choice for λ_{ij} is the largest absolute value of the Roe matrix eigenvalues.

The HLL schemes are trivially generalisable to the case of barotropic EOS. Indeed, both the acoustic wave speeds of the exact Riemann problem and the corresponding Roe matrix eigenvalues are formally identical in the barotropic and non-barotropic cases even if the definition of the sound speed is different. Thus, a Rusanov scheme for barotropic EOS is proposed, here, considering the numerical flux function (3.62) associated with

$$\lambda_{ij} = \max_p(|\tilde{\lambda}_p|) = |\tilde{u}_{ij}| + \tilde{a}_{ij} \tag{3.63}$$

in which $\tilde{a}_{ij}$ is obtained by (3.60) and λ_p is defined by (3.61). As already mentioned, note that equation (3.62) is a particular instance of the family of schemes defined in section 3.2.3. Equations (3.63) is not the only possible choice for λ_{ij}, another common approach is to consider:

$$\lambda_{ij} = \max(|u_i| + a_i, |u_j| + a_j) \tag{3.64}$$

Some numerical experiments carried out in [16] (see also Sec. 4.3.2) proved that the choices (3.64) and (3.63) can be considered equivalent since they computed the same results. Thus, in this work, except when differently stated, only the first choice, (3.63) is considered.

3.5.3 A modified HLL scheme with an anti-diffusive term

The HLL schemes are known for their excessive diffusive behaviour in presence of a contact discontinuity. Indeed, due to the assumption of a two-wave configuration with only one intermediate state, the two intermediate states of the exact Riemann solution are averaged, smearing, in this way, the contact discontinuity. To avoid this problem, more complex average-state approximate Riemann solvers have been introduced as, for instance, the HLLC scheme, proposed in [107], which involves two intermediate states in the approximate solution. Alternatively, an anti-diffusive term for the contact discontinuity can be directly added in the single-state HLL formulation. This kind of approaches initially proposed in [44] with the HLLEM scheme, has been successively improved considering the HLLE+ scheme [84]. In this context, the following unified expression of the flux function can be used to represent HLL (taking $\bar{\delta} = 0$), HLLEM, HLLE+ as well as Roe fluxes:

$$\Phi(\mathbf{W}_l, \mathbf{W}_r) = \frac{b^+ F(\mathbf{W}_l) - b^- F(\mathbf{W}_r)}{b^+ - b^-} + \frac{b^+ b^-}{b^+ - b^-} \left(\mathbf{W}_r - \mathbf{W}_l - \sum_{p \in S_{cd}} \bar{\delta} \alpha_p r_p \right) \tag{3.65}$$

in which b^+, b^- and $\bar{\delta}$ are parameters depending on the states $\mathbf{W}_l$ and $\mathbf{W}_r$, while r_p are the right eigenvectors of the Jacobian matrix of the flux function F evaluated at a particular intermediate state and α_p the components of $\mathbf{W}_r - \mathbf{W}_l$ in the right-eigenvector basis. S_{cd} is the set of indexes for which r_p is associated with the eigenvalue u of the flux Jacobian matrix, i.e. related to the contact discontinuity. Thus, the last term in the right hand side of (3.65) represents an anti-diffusive term for the contact discontinuity. However, for the classical case of an ideal-gas EOS, any modification of this term acts on the whole system of equations, and consequently, the choice of an adequate parameter $\bar{\delta}$ is rather critical (see [84]). The situation is different for the barotropic case since density, velocity and pressure are continuous across the contact discontinuity, and thus, the presence of two different intermediate states is only due to the passive scalar (see [102, 103]). Moreover, only one eigenvalue is associated with the contact discontinuity while the corresponding eigenvector r_3 can be always chosen to be $(0\ 0\ 1)^T$ whatever is the particular intermediate state. Consequently, the anti-diffusive term in (3.65) affects only the third equation of the system. Due to this decoupling from mass and momentum equations, there are less restrictions on the choice of $\bar{\delta}$ in the barotropic case. Clearly, it should correspond to an anti-diffusive term in order to avoid the smearing of the contact discontinuity of the HLL scheme, and on the other hand, the estimate of the wave speeds should remain as close as possible to the real ones. The simplicity of the scheme is privileged here for the choice of the modified term in view of its incorporation in the complete formulation including preconditioning for low Mach number and linearised implicit approach (see sections 3.6 and 3.3). Consequently, starting from the simplest HLL scheme, i.e. the Rusanov one, the following modified scheme, called in the following LD-HLL (Low-Diffusive HLL), is proposed:

$$\Phi_{i,j} = \frac{F(\mathbf{W}_i) + F(\mathbf{W}_j)}{2} - \frac{1}{2} \begin{pmatrix} \lambda_{ij} & 0 & 0 \\ 0 & \lambda_{ij} & 0 \\ 0 & 0 & |\tilde{u}_{ij}| \end{pmatrix} (\mathbf{W}_j - \mathbf{W}_i) \tag{3.66}$$

With respect to the original Rusanov scheme, the first two equations, which are related to the acoustic waves, are unchanged, while for the third equation, i.e. the one directly related to the contact discontinuity, the diffusive part of the scheme has been reduced. This can be more clearly shown by rewriting (3.66) as follows:

$$\Phi_{i,j} = \frac{F(\mathbf{W}_i) + F(\mathbf{W}_j)}{2} - \frac{1}{2}\lambda_{ij} \left((\mathbf{W}_j - \mathbf{W}_i) - \frac{\tilde{a}_{ij}}{\tilde{a}_{ij} + |\tilde{u}_{ij}|} (\rho_j \xi_j - \rho_i \xi_i) \, r_3 \right) \tag{3.67}$$

reminding that $r_3 = (0\ 0\ 1)^T$.

Furthermore, this scheme can also be expressed in terms of the unified Godunov-type formulation (3.65) by considering the following parameters:

$$b^- = -\lambda_{ij}, \, b^+ = \lambda_{ij} \quad \text{and} \quad \bar{\delta} = \frac{\tilde{a}_{ij}}{\tilde{a}_{ij} + |\tilde{u}_{ij}|} \left(1 + \frac{\tilde{\xi}_{ij} (\rho_j - \rho_i)}{\tilde{\rho}_{ij} (\xi_j - \xi_i)} \right) \tag{3.68}$$

in which we have chosen to express $\mathbf{W}_j - \mathbf{W}_i$ in the basis of the right eigenvectors of the Roe matrix. It is also interesting to investigate the behaviour of this scheme in the asymptotic limits. In particular, when the flow is nearly incompressible, i.e. $\rho_j \simeq \rho_i$, $|\tilde{u}_{ij}| \ll \tilde{a}_{ij}$ and consequently $\lambda_{ij} \simeq \tilde{\lambda}_1 = -\tilde{\lambda}_2$, the parameters defined in (3.68) tend to the following values:

$$b^- = \min\left(\tilde{\lambda}_2, 0\right), \, b^+ = \max\left(\tilde{\lambda}_1, 0\right) \quad \text{and} \quad \bar{\delta} = \frac{\tilde{a}_{ij}}{\tilde{a}_{ij} + |\tilde{u}_{ij}|} \tag{3.69}$$

The scheme associated with these parameters in the unified formulation (3.65) coincides with the Roe scheme. This is a classical result for ideal-gas EOS (see [84]) but it is also true for the barotropic case (the proof is reported in appendix B). Thus, in the limit of incompressible flows and, as a consequence, for low Mach number flows, the LD-HLL and Roe schemes are equivalent. Conversely, in the limit of highly compressible flows since $|\tilde{u}_{ij}| \gg \tilde{a}_{ij}$ we have $\lambda_{ij} \simeq |\tilde{u}_{ij}|$, and thus, it follows that the scheme (3.66) is equivalent to the standard Rusanov one.

3.6 Preconditioning for low Mach number flows

For the cavitating flow problem, a large part of the flow is characterised by very low Mach numbers since we have to deal with a weakly-compressible liquid. Compressible solvers encounter accuracy problems when dealing with nearly-incompressible flows [57]. In order to counteract this difficulty, some preconditioning must be applied.

3.6.1 A preconditioned Roe scheme

A Turkel-like preconditioning has been proposed in [102, 104] for the Roe flux function associated with a barotropic EOS using a similar formulation as proposed in [57] for a perfect-gas state equation. This preconditioning is acting only on the upwind part of the numerical flux function; more precisely, $Q_{ij} = |\tilde{J}(\mathbf{W}_i, \mathbf{W}_j)|$ is replaced by:

$$Q_{ij} = P_{ij}^{-1} |P_{ij} \tilde{J}(\mathbf{W}_i, \mathbf{W}_j)|$$

in which P_{ij} is the preconditioning matrix defined by

$$P_{ij} = \left. \frac{\partial \mathbf{W}_p}{\partial \mathbf{W}} \right|_{ij} \Lambda_p \left. \frac{\partial \mathbf{W}}{\partial \mathbf{W}_p} \right|_{ij}$$

where $\mathbf{W}_p$ is the vector of the primitive variables and $\Lambda_p = diag(\theta^2, 1, 1)$, the parameter θ being proportional to a reference Mach number, which is herein defined as:

$$\theta = \theta(M) = \begin{cases} 10^{-6} & \text{if } M \leq 10^{-6} \\ \min(M, 1) & \text{otherwise} \end{cases} \tag{3.70}$$

in which M represents a local Mach number as function of the Roe average values, i.e. $M = \dfrac{|\tilde{u}|}{\tilde{a}}$. It has been shown in [102] that, with this choice of P_{ij}, the preconditioned semi-discrete solution recovers the same asymptotic behaviour (with an expansion in power of a reference Mach number M) of the analytical one in the low Mach number limit. This theoretical result has been also corroborated by numerical experiments which show that the preconditioned formulation does not present accuracy problems for low Mach number flows. We refer to [97, 102, 104] for more details.

3.6.2 A preconditioned HLL-like scheme

By carrying out an asymptotic analysis reported in appendix C, it has been found that the Rusanov scheme also encounters accuracy problems in the low Mach number limit [16]. However, the correct asymptotic behaviour of the analytical solution can be recovered acting directly on the acoustic diagonal terms of the matrix S_{ij}. More precisely, it is possible to consider the following "preconditioned" matrix:

$$S_{ij} = \begin{pmatrix} \lambda_{ij,1}^p & 0 & 0 \\ 0 & \lambda_{ij,2}^p & 0 \\ 0 & 0 & \lambda_{ij,3}^p \end{pmatrix} \tag{3.71}$$

For the preconditioned Rusanov scheme, the “preconditioned” eigenvalues $\lambda^p_{ij,1}$ appearing in (3.71) are defined as:

$$\lambda^p_{ij,1} = \theta^{-1}\lambda_{ij},\ \lambda^p_{ij,2} = \theta\lambda_{ij},\ \lambda^p_{ij,3} = \lambda_{ij} \tag{3.72}$$

where the parameter θ is the same previously defined for the preconditioned Roe scheme (see (3.70)). With this simple preconditioning procedure, the correct asymptotic behaviour of the analytical solution is recovered (for details see appendix C).

Note that since the preconditioning affects only the mass and momentum balances and not the passive scalar field, the analysis is also valid for the LD-HLL scheme (3.66) and thus, the “preconditioned matrix” for this scheme is:

$$S_{ij} = \begin{pmatrix} \lambda_{ij}\theta^{-1} & 0 & 0 \\ 0 & \lambda_{ij}\theta & 0 \\ 0 & 0 & |\tilde{u}_{ij}| \end{pmatrix} \tag{3.73}$$

Note that (3.73) is a particular instance of (3.71) when the “preconditioned” eigenvalues are defined as:

$$\lambda^p_{ij,1} = \theta^{-1}\lambda_{ij},\ \lambda^p_{ij,2} = \theta\lambda_{ij},\ \lambda^p_{ij,3} = |\tilde{u}_{ij}| \tag{3.74}$$

3.7 Simplified expression of the matrices for the complete linearisation

To deal with the complete linearisation in 3.3.2 it is necessary to write the matrices K^+ and K^- for the particular case under consideration. For this purpose it is worth noting that the elements of the matrix S, even in the preconditioned case, λ_k, are in the form $f(M)(|u| + a)$ where $M = \dfrac{|u|}{a}$ and, as a consequence, they satisfy (3.33).

Independently of the way $|u|$ and a are chosen, λ_k is always a differentiable function of a (a being always > 0) and is differentiable with respect to u when u is different from 0. Furthermore, a and u are of the form:

$$a = a(\rho_j, \rho_i) \quad u = u(\rho_j, \rho_i, (\rho u)_j, (\rho u)_i)$$

Both u and a are differentiable functions of ρ_j, ρ_i but u is not always differentiable with respect to $(\rho u)_j$ and $(\rho u)_i$. However, excepting when the flow is highly supersonic, we have that the flow derivatives of u are far smaller than the ones of a. Consequently, it seems reasonable to neglect the variation in u for the expression of K^+ and K^- in (3.51), and then, to approximate the variation of the matrix S only through the variation in a. Moreover, it is important to notice that the matrices K^+ and K^- introduce a term that is second-order accurate except in presence of a discontinuity. Due to the particular state law under consideration, a shock wave is nearly constantly associated with a transition from vapour to liquid. This transition involves a step-like change of the speed of sound and, as a consequence, set straight the choice to neglect the variation in u. Then, with this assumption and using the fact that a only depends on the density, the matrices K^+ and K^- previously defined in (3.51) reduce here to:

$$K_{ij}^{\pm} = \begin{pmatrix} \dfrac{\partial \lambda_1}{\partial a} \dfrac{\partial a}{\partial \rho_\pm} (\rho_r^n - \rho_l^n) & 0 & 0 \\ \dfrac{\partial \lambda_2}{\partial a} \dfrac{\partial a}{\partial \rho_\pm} \Big((\rho u)_r^n - (\rho u)_l^n \Big) & 0 & 0 \\ \dfrac{\partial \lambda_3}{\partial a} \dfrac{\partial a}{\partial \rho_\pm} \Big((\rho \xi)_r^n - (\rho \xi)_l^n \Big) & 0 & 0 \end{pmatrix} \tag{3.75}$$

in which $\rho_\pm$ is equal to ρ_j and ρ_i for K^+ and K^- respectively.

3.7.1 A numerical approximation of the eigenvalues derivatives

Matrix K_{ij}^+ and K_{ij}^- require the computation of the eigenvalues derivatives with respect to density, but considering only the variation in a, i.e. terms $\partial_\rho \lambda_k$ defined by:

$$\partial_\rho \lambda_k = \frac{\partial \Lambda_k}{\partial \rho} \quad \text{where} \quad \Lambda_k = \Lambda_k(\rho_j, \rho_i) = \lambda_k \Big(u, a(\rho_j, \rho_i) \Big)$$

The analytical expression of those terms can be very complex, especially in the preconditioned case when λ_k is chosen as in (3.74). Also, $\partial_\rho \lambda_k$ could be discontinuous at the saturation point for cavitating flows. While the last point could be solved using some "average" between the right and left derivatives, the analytical expression of $\partial_\rho \lambda_k$ is expensive from a computational point of view. A possible way to overcome this problem is to consider the following approximation:

$$\begin{cases} \partial_{\rho_j} \lambda_k \simeq \dfrac{\lambda_k(u, a(\rho_j + \epsilon, \rho_i)) - \lambda_k(u, a(\rho_j - \epsilon, \rho_i))}{2\epsilon} \\[2ex] \partial_{\rho_i} \lambda_k \simeq \dfrac{\lambda_k(u, a(\rho_j, \rho_i + \epsilon)) - \lambda_k(u, a(\rho_j, \rho_i - \epsilon))}{2\epsilon} \end{cases} \tag{3.76}$$

When the analytical derivative exists, this approximation is second-order accurate, but (3.76) is always defined and near a discontinuity its value is an average of the right and left limits of the derivative.

3.8 Summary of the 1D numerical methods

In this study different linearised implicit numerical methods have been considered for barotropic flows. For the second-order accurate in space and time schemes, the general formulation of those approaches can be summarised in system (3.57)-(3.58). The expression of the matrices D_1 and D_2 appearing in (3.58) depends on the specific numerical flux function considered, (Roe, Rusanov, LD-HLL) as well as on the presence of a preconditioner and on the chosen flux linearisation. For the sake of clarity, a summary of the possible expressions of D_1 and D_2 is reported in table 3.1. Similarly, it is possible to define a general formulation for the first-order schemes, the only difference with the formulation (3.57)-(3.58) is that the time derivative is approximated with a backward Euler scheme instead that with a second-order BDF. Note in particular that even using the preconditioning both the Rusanov and the LD-HLL schemes are diagonal Q-schemes Thus, the complete linearisation defined by (3.52) in Sec. 3.3.3 can be applied also for these schemes.

	Roe	Rusanov	LD-HLL
D_1	$Q^{-1}\left[Q\tilde{J}_{ij}\right]^{+}$	$\frac{1}{2}(A(W_i^n) + S_{ij}^n - K_{ij})$	
D_2	$Q^{-1}\left[Q\tilde{J}_{ij}\right]^{-}$	$\frac{1}{2}(A(W_j^n) - S_{ij}^n + K_{ji})$	
no precond.	$Q = I$	$S_{ij}^n = \lambda_{ij} I$	S_{ij}^n from (3.66)
precond.	$Q = P_{ij}$	S_{ij}^n from (3.71)	S_{ij}^n from (3.73)
linearisation	defined in (3.26)	standard (from (3.32)): $K_{ij} = 0$ complete (from (3.52)): K_{ij} from (3.51)	

Table 3.1: Expressions of D_1 and D_2 for the different choices of linearisation and numerical flux function (with or without preconditioning).

Chapter 4

1D Numerical Experiments

In this chapter, the previously defined numerical approaches are validated and compared by considering different test-cases, for which the exact solution is available. They include, in particular, both steady and unsteady problems as well as different kinds of solutions (smooth, continuous and discontinuous ones) and regimes (from low Mach number to supersonic flows).

First a brief description of the considered test-case is given at the begin of each of the sections 4.2 and 4.3. Then the proposed methodologies for the second-order extension and the low diffusive modification of the Rusanov scheme are validated in section 4.2.1 and 4.2.2 respectively. The preconditioning strategy is addressed in section 4.3.2 and, finally the complete linearization proposed in section 3.3 is analysed in 4.3.3.

4.1 A barotropic state law model

Since the main purpose of this part of the thesis is the simulation of cavitating flows, most of the test-cases presented in this chapter have been performed considering the cavitating EOS (2.6)-(2.7). Nevertheless, less specific test-cases have also been considered in order to independently study the different numerical ingredients described in Chap. 2. Thus, the following barotropic state law has been also considered for some of the 1D test-cases:

$$p = \kappa\,\rho^{\varkappa} + \gamma \tag{4.1}$$

with $\kappa > 0$, $\varkappa > 0$ and γ given (real) constants. Note that this family of barotropic equations of state includes several classical cases as, for instance, the homogeneous shallow water equations, of course, augmented by the advection of the passive scalar ($\kappa = 2^{-1}$, $\varkappa = 2$ and $\gamma = 0$) or the Tait law, which is used for describing the isentropic flow of compressible liquids ($\kappa = \varepsilon\,\rho_0^{-\varkappa}$ and $\gamma = -\varepsilon$, with $\varepsilon > 0$ and $\rho_0 > 0$).

4.2 Riemann problems for convex barotropic law

Riemann problems, which are characterised by the convection of continuous and discontinuous waves, are classical benchmarks to test numerical schemes for time-dependent problems, as well as to investigate the behaviour of the schemes in presence of discontinuities (see e.g. [107]). The initial condition for this problem consists of constant right and left states separated by a discontinuity. The exact solution of the Riemann problem for (2.11) closed by a convex EOS, i.e. (2.4), is described in [102] and is reported in appendix A. For a generic convex barotropic equation, the solution for the Riemann problem is composed by three waves: the intermediate one is always a contact discontinuity, the others can be shock or rarefaction waves [102]. Moreover, the presence of contact discontinuities in the aforementioned solution to the Riemann problem permits to compare the results of the Rusanov scheme, which is known to have poor accuracy for this kind of discontinuities, with those of the Roe scheme.

4.2.1 Non-cavitating Riemann problem: study of DeC approaches

For the present test-case only one numerical scheme, namely the Roe one, has been considered. Indeed, the study is centred on the validation of the second-order DeC formulation (3.11)-(3.12). In particular, the actual increase of accuracy of the second-order approach, as well as the practical impact on the solution of using more than one DeC iteration is shown. In the present Riemann problem, the equation of state (4.1) is exploited using $\kappa = 10^6$, $\varkappa = 1$ and $\gamma = 0$. The initial conditions are the following: $\rho_L = 1\ \mathrm{kg/m^3}, \rho_R = 1.25\ \mathrm{kg/m^3}$, $u_L = -90\ \mathrm{m/s}$, $u_R = 100\ \mathrm{m/s}$, $\xi_L = -2$ and $\xi_L = 2$, where the subscripts L and R denote the left and right states respectively. In this case, the unsteady solution corresponds to the progressive separation of two waves (one shock and one rarefaction respectively) for pressure, density and velocity (and the evolution of a contact discontinuity for the passive scalar). The study is focused here on the second-order defect correction approaches considering different levels of refinement both in space and time (Δx from 0.5 to 500 and from Δt from $0.5 \cdot 10^{-4}$ to 0.05, respectively). The results computed by the first-order version of the Roe scheme and the second-order one with 1, 2 and 3 DeC iterations are shown in the following. All the solutions are taken at the time $t = 1$. Figs. 4.1 and 4.2 show the convergence with respect to space refinement with, as expected, an important diffusivity of the first-order approach near strong gradients and discontinuities. One iteration of defect correction is already sufficiently accurate since the cases with $S_M = 2$ or $S_M = 3$ provide very similar results.

Figs. 4.1 and 4.3 correspond to different numerical solutions obtained by varying the refinement, but now taking $\Delta x \Delta t$ constant. For the second-order approaches, the best solution is obtained by an intermediate set of values of $(\Delta x, \Delta t)$ (Fig. 4.3a); this appears in agreement with the accuracy appraisal of $\mathrm{O}\left((\Delta\eta)^2\right)$, $(\Delta\eta)^2 = \max\left((\Delta x)^2, \Delta x \Delta t, (\Delta t)^2\right)$, proposed in section 3.1.2. Indeed, for large Δx (Fig. 4.1), the term of order $(\Delta x)^2$ should be predominant and the grid coarseness deteriorates the solution accuracy and, similarly, too large Δt ($\Delta t = 10^{-2}$) yields significant errors. In order to be more quantitative, a scaling issue can be introduced in the present case. Indeed, the self-similar character of the solution of the Riemann problem intrinsically introduces a space/time ratio which, for low Mach number flows like the one considered, is basically associated with the fluid sound speed (equal to 10^3). By observing the exact solution in Figs. 4.1 or 4.3 (at $t = 1$), it appears that the characteristic length scale, given by the distance between the waves, is roughly equal to $L_x = 2000$. Hence, by comparing $\Delta x / L_x$ and Δt, it is not surprising that the best results amongst those shown in Figs. 4.1 and 4.3 are associated with $\Delta t = 10^{-3}$. Indeed, in this case, the chosen settings tend to approach the optimal condition $\Delta x / L_x = \Delta t$ minimising the non-dimensional counterpart of $\mathrm{O}\left((\Delta x)^2, \Delta x \Delta t, (\Delta t)^2\right)$.

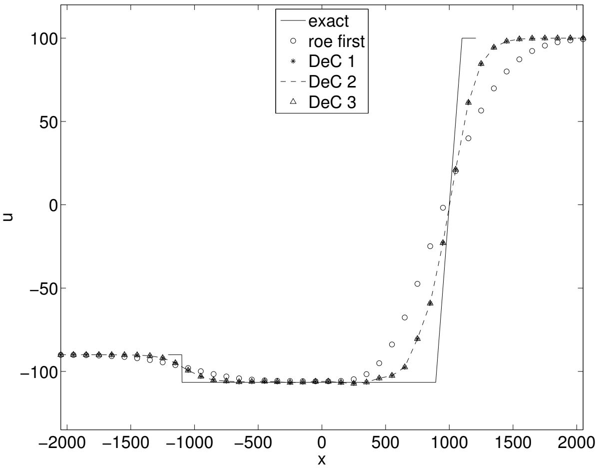
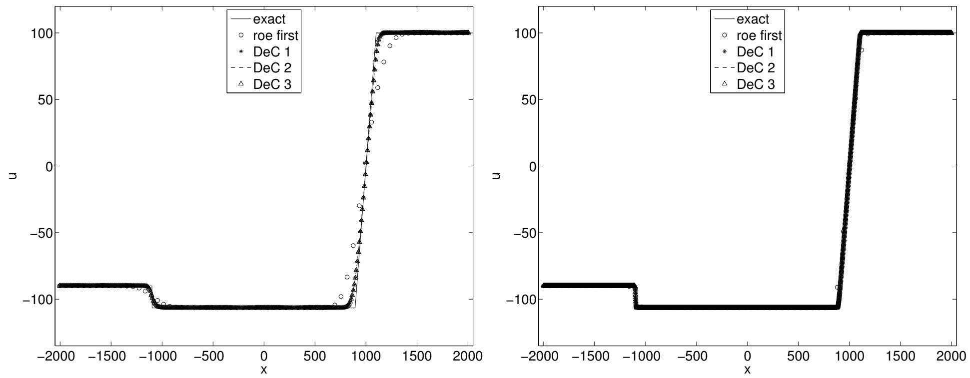


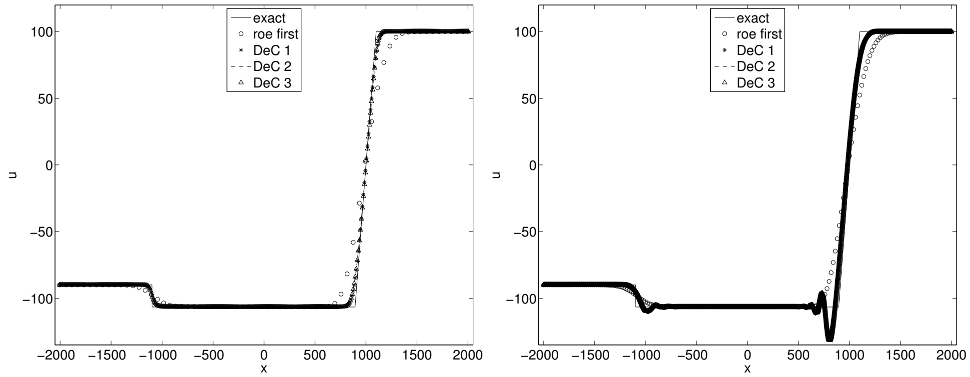
Figure 4.1: Non cavitating Riemann problem (Roe scheme), comparison between the first-order implicit scheme and second-order DeC approaches: velocity field for $\Delta t = 10^{-4}$ and $\Delta x = 100$.

In this spirit, the oscillations appearing in Fig. 4.3b, even if limiters are used, are probably due to the coarsening of Δt with respect to Δx (Δt here is 20 times larger than $\Delta x / L_x$). In consideration of this difference, it seems reasonable that applying several defect-correction steps does not significantly improve the numerical solution. On the other hand, by increasing Δx to 10



(a) $\Delta x = 10$ **(b)** $\Delta x = 1$

Figure 4.2: Non cavitating Riemann problem (Roe scheme), comparison between the first-order implicit scheme and second-order DeC approaches: velocity field for $\Delta t = 10^{-4}$.



(a) $(\Delta t, \Delta x) = (10^{-3}, 10)$ **(b)** $(\Delta t, \Delta x) = (10^{-2}, 1)$

Figure 4.3: Non cavitating Riemann problem (Roe scheme), comparison between the first-order implicit scheme and second-order DeC approaches: velocity field with $\Delta t \Delta x = 10^{-2}$.

(thus reducing the gap between time and space discretisation), an improved numerical solution is obtained by increasing the number of defect-correction steps, as shown in Fig. 4.4. More in detail, two defect-correction iterations seem to be enough to damp the oscillations as additional iterations do not refine the solution. Finally, it may be worth remarking that only implicit schemes permit to approach the aforementioned optimal condition for the space/time discretisation settings, being free of CFL-like constraints.

Let $x_1, \cdots, x_N$ be the nodes corresponding to a particular spatial dis-

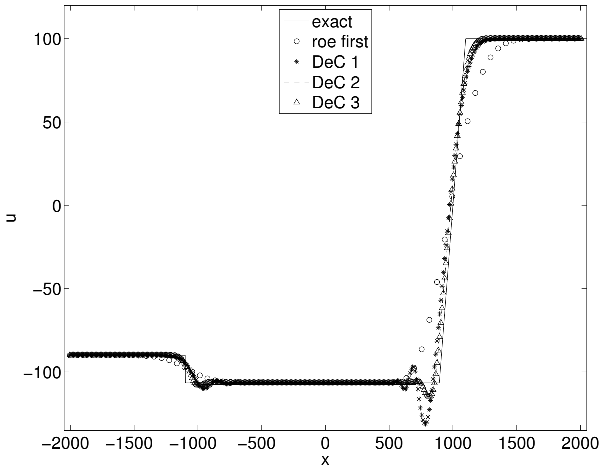


Figure 4.4: Non cavitating Riemann problem (Roe scheme), comparison between the first-order implicit scheme and second-order DeC approaches: velocity field with $(\Delta t, \Delta x) = (10^{-2}, 10)$.

cretisation, it is possible to define the error of the numerical simulation with respect to the exact solution, as $e(\varphi) = \|\vec{\varphi} - \vec{\varphi}_e\| / \|\vec{\varphi}_e\|$, where $\|.\|$ is the Euclidean norm in $I\!R^N$, φ and φ_e are the numerical and the exact solutions respectively and $\vec{\varphi} = (\varphi(x_1), \cdots, \varphi(x_N))^T$. Fig. 4.5 shows the evolution of the density error with refinement. The refinement is made simultaneously in space and time, i.e. with a constant ratio $\Delta x / \Delta t$ (corresponding to the case of Fig. 4.3a). The asymptotic slopes of the error curve are about 0.42 for first-order and 0.71 for the DeC approaches (DeC1, DeC2 and DeC3) respectively. The presence of discontinuities and strong gradients where the second-order is lost and which yield a major contribution to $e(u)$ prevents us from expecting the classical asymptotic error slopes of 1 and 2 for, respectively, the first and second order formulation. Nevertheless, it must be noticed that in the present case the difference of accuracy between the first-order approach and the DeC ones is clearly illustrated by a slope ratio around 2 (namely 1.7) thus confirming the effectiveness of the proposed methodology for second-order extensions. Furthermore, the error behaviour for a more regular solution will be shown in Sec. 4.3.1.

4.2.2 A Riemann problem for cavitating barotropic flow

As previously mentioned, the presence of the contact discontinuity permits to compare the results of the Rusanov scheme, which is known to encounter accuracy problems for this kind of discontinuities, with those of the Roe scheme and of the proposed LD-HLL scheme. The working fluid is water at

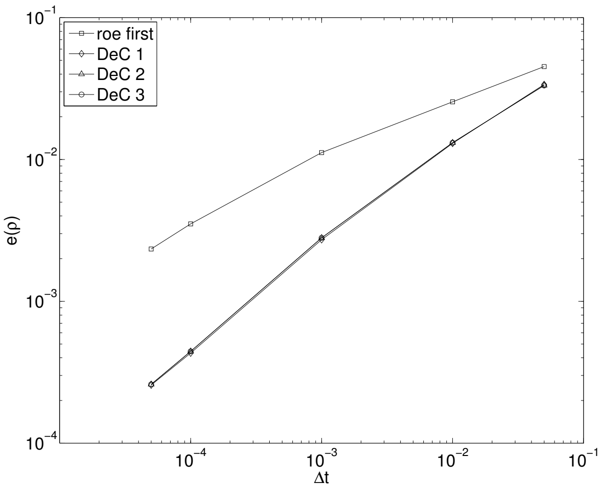
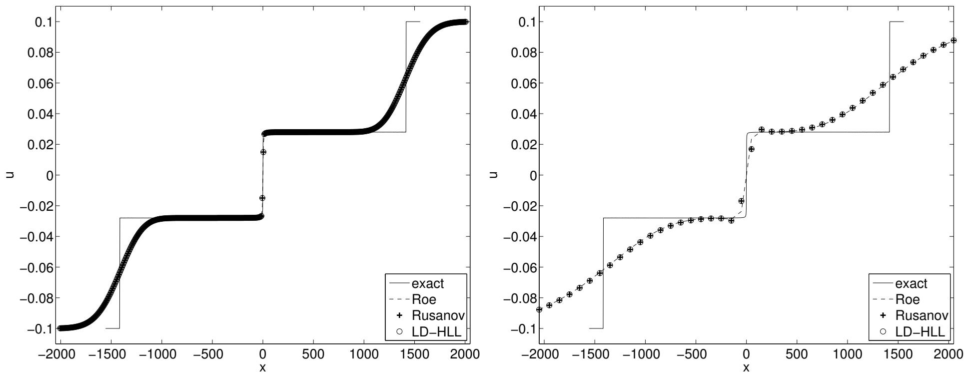


Figure 4.5: Non cavitating Riemann problem (Roe scheme), comparison between the first-order implicit scheme and second-order DeC approaches: density error with respect to Δt ($\Delta x/\Delta t$ frozen)

a temperature of $T = 293.16$ K and the equations of state are those given by the barotropic cavitation model (2.6)-(2.7), in which $T_L = 293.16$ K while the free model parameter ζ (involved in the definition of ε_L in Eq. (2.9)) is chosen equal to 0.1. The initial conditions are the following: $\rho_L = \rho_R = 998$ kg/m^3, $u_L = -0.1$ m/s, $u_R = 0.1$ m/s, $\xi_L = 2$ and $\xi_L = 4$, were the subscripts L and R denote the left and right states respectively. With these initial conditions, the exact solution is characterised by the formation of two rarefaction waves. This leads to an intermediate region which is cavitating ($\rho \approx 960.47 < \rho_{Lsat}$). The sound speed in the liquid is $a_L = a_R \approx 1415.63$ m/s while in the cavitating region falls down to approximately $a_{cav} \approx 0.37$ m/s.

Simulations have been carried out with two different spatial grids, with Δx equal to 100 and 10, and using a constant CFL for all the schemes, equal to 1.4, which is the maximum allowable for the stability of the Roe scheme. The CFL number is defined herein by taking the local value of the speed of sound as a reference velocity. Consequently, given Δx and CFL, the time step is set equal to $\Delta t = CFL\, \Delta x/(\max_i a_i)$, a_i being the value of the speed of sound in cell i.

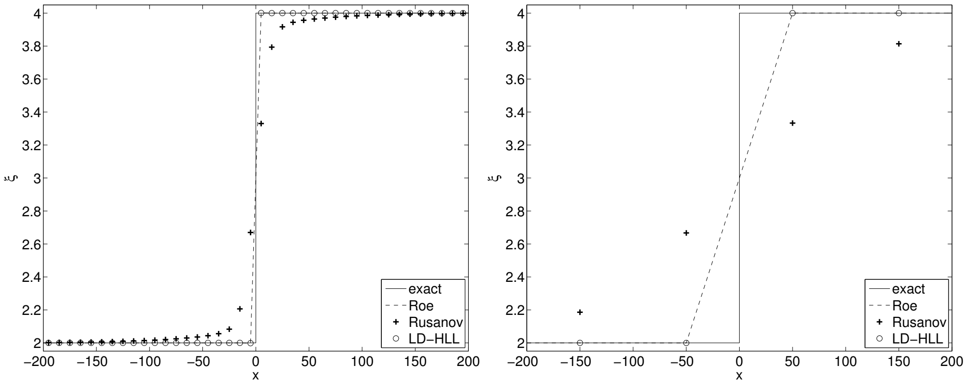
A comparison between the first-order version of the different schemes, namely Roe, Rusanov and LD-HLL has been performed. Figs. 4.6 show the velocity obtained with the different schemes for the two grid resolutions, compared against the exact solution. All the schemes give practically the same results for velocity, with grid refinement leading to the expected im-



(a) $\Delta x = 10$ **(b)** $\Delta x = 100$

Figure 4.6: Velocity for the cavitating Riemann problem.

provement of accuracy and decrease of numerical dissipation. The analysis of the behaviour of the passive scalar in Fig. 4.7 is particularly interesting. Indeed, it is clear that, as expected, the Rusanov scheme introduces a much larger dissipation than the Roe one. Conversely the LD-HLL scheme recovers the same accuracy as the Roe one and, thus, these results confirm that the proposed LD modification actually improves the accuracy for contact discontinuities. It may be surprising that, for a fixed grid resolution and numerical

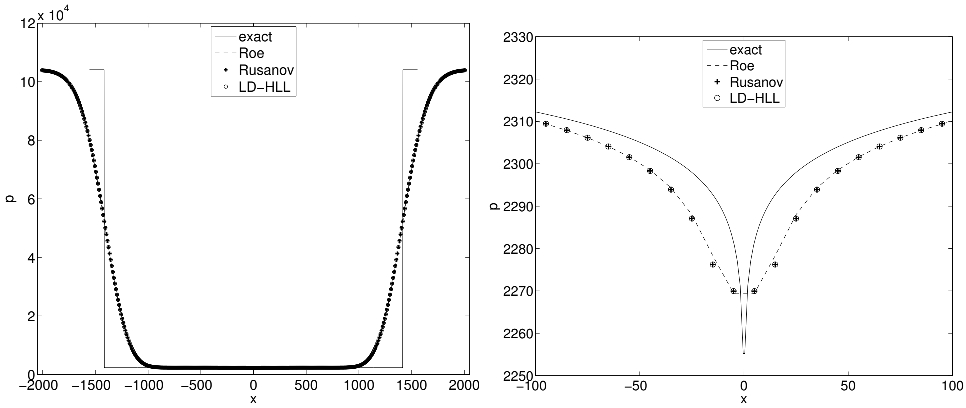


(a) $\Delta x = 10$ **(b)** $\Delta x = 100$

Figure 4.7: Passive scalar for the cavitating Riemann problem.

scheme, the contact discontinuity appears to be better captured than the rarefaction waves; this for all the considered schemes, included the Rusanov one. This behaviour is mainly due to the specific character of the barotropic state law in presence of cavitation and to the low Mach number of this test case $M \simeq 10^{-5}$. Indeed, in all the numerical schemes the upwind part of mass and momentum fluxes is associated with the acoustic eigenvalues and, as a consequence, the introduced numerical diffusivity is related to the speed of sound. Thus, all the schemes give the same numerical viscosity for the two rarefaction waves, which is rather large since they occur in zones of pure water, for which the speed of sound is, as previously said, of the order of 1400 m/s. Conversely, the upwind part of the flux in the passive scalar equation for the Roe and LD-HLL schemes is associated with the flow velocity which is 0.1 m/s, while for the Rusanov one is again related to the speed of sound. However, in this particular test case, the contact discontinuity is inside a cavitating region where the speed of sound is much lower than in pure liquid (as previously said it falls down to approximately 0.37 m/s), and thus, even if the numerical diffusivity introduced at the contact discontinuity by the Rusanov scheme is significantly larger than the one given by the other two schemes, it remains lower than the one introduced by all the schemes at the rarefaction waves.

Finally, in order to give an idea of the peculiarity of the Riemann problem solution due to cavitation, the pressure behaviour is shown in Fig. 4.8. As can be seen, most of the pressure variation occurs at the rarefaction waves, leading to values lower than the saturation one ($p_{sat} = 2340$ Pa). Furthermore, a spike-like pressure variation takes place in a narrow region centred at $x = 0$ (see the zoom in Fig. 4.8), which corresponds to the zone in which the sound speed reaches the value of a_{cav}. Clearly to accurately capture this behaviour



(a) whole domain **(b)** detail of the cavitation region

Figure 4.8: Pressure for the Riemann problem and $\Delta x = 10$.

a more refined grid would be required; nevertheless, also for pressure all the schemes give practically the same results.

4.3 Barotropic flows in a 1D convergent-divergent nozzle

Here an inviscid flow in a quasi 1D convergent-divergent nozzle is considered. The governing equations are the reduced 1D Euler equations (without passive scalar) with a source term, which accounts for variations of the cross-sectional area S of the nozzle. More precisely, the following system of differential equations is obtained:

$$\frac{\partial W}{\partial t} + \frac{\partial F(W)}{\partial x} = -\frac{1}{S(x)}\frac{dS}{dx}G(W) \tag{4.2}$$

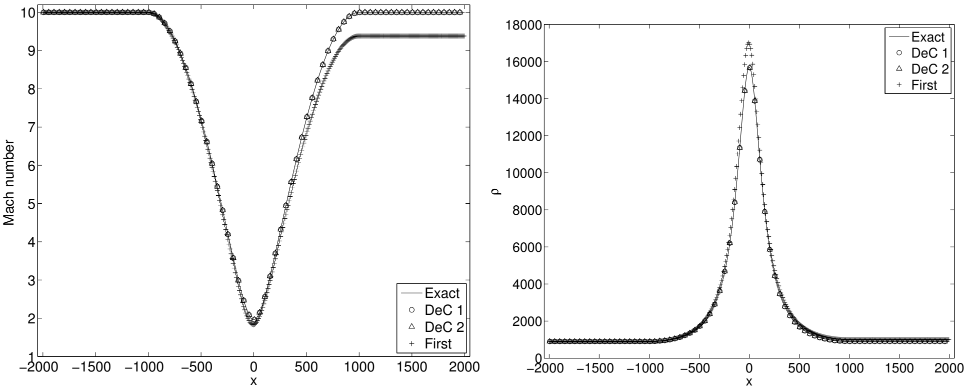
in which W and $F(W)$ corresponds to the first two components of the same terms in Eq. (2.11) and $G(W) = \left(\rho u, \rho u^2\right)^T$. The steady exact solution for a barotropic flow governed by Eq. (4.2) is derived in appendix D and is used here as a reference. Note, in particular, that since in this test-case the flow is inviscid and the passive scalar is not taken into account, the Rusanov scheme and the LD-HLL one are equivalent.

4.3.1 A supersonic non-cavitating polytropic flow

In this test-case, a fully supersonic flow, characterised by high Mach number values is obtained by considering the following inlet conditions $\rho_\infty = 900$, $u_\infty = 300$ and nozzle area ratio $S_{min}/S_\infty = 0.07$. S_{min} and S_∞ being the cross-section area at the nozzle throat and inlet (or outlet) respectively. Indeed, the Mach number, which is taken equal to 10 at the inlet, decreases to about 2 in the nozzle throat (see Fig. 4.9a).

For this problem different formulations of the Roe scheme have been compared; more specifically, the first-order linearised implicit scheme (3.27), and the DeC second-order formulation, with $s_M = 1$ and $s_M = 2$ (DeC1 and DeC2 respectively). The steady solution has been obtained after very few pseudo-time iterations (about 3) for all the linearised implicit approaches. Note that the exact steady solution can be obtained as described in Appendix D by using the barotropic EOS (4.1) with $\kappa = 0.5$, $\varkappa = 2$ and $\gamma = 0$.

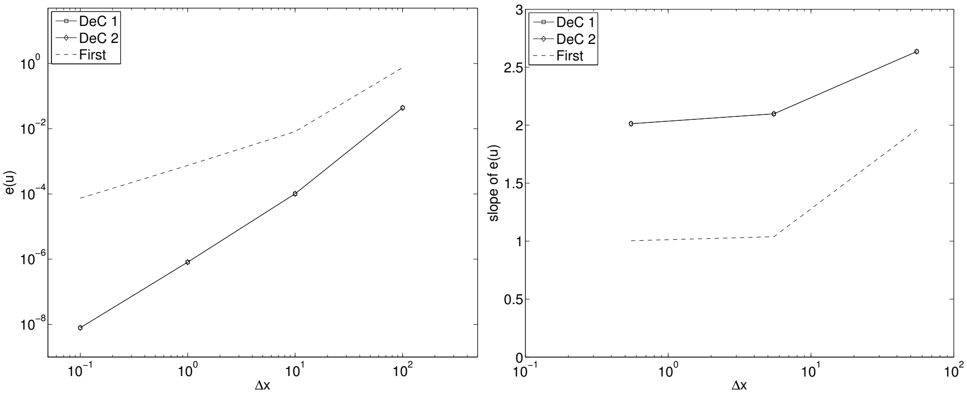
The comparison between the different steady solutions clearly shows a significant difference between the first-order approach and the second-order ones, which give very similar results. This is illustrated in Fig. 4.9, where at the considered level of grid refinement ($\Delta x = 10$), the second-order approaches are already close to the exact solution, while the first-order one still presents large discrepancies. Furthermore, Fig. 4.10 shows $e(u)$, the error on the velocity field (defined in Sec. 4.2.1), at different levels of refinement in



(a) Mach number **(b)** density field

Figure 4.9: Non cavitating supersonic flow, comparison between several implicit formulations, $\Delta x = 10$.

space (from 40 to 4 × 10^4 cells corresponding to a space discretisation step Δx from 100 to 0.1 respectively). For the first-order approach the slope of the error curve tends to 1 as Δx tends to zero while the second-order accuracy is recovered for all the other approaches since the slope tends asymptotically to 2, as shown in Fig. 4.10b. Thus, the second-order of accuracy is maintained by the Defect correction technique and, in particular, according to Sec. 3.1.2, one iteration of DeC (DeC1) is enough to provide a second-order accuracy.

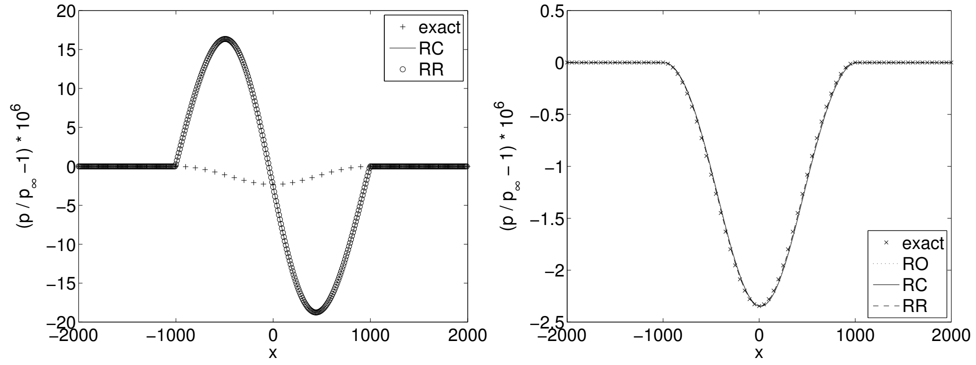


(a) error, log-log scale **(b)** slope of the error curve

Figure 4.10: Non cavitating supersonic flow, comparison between different implicit formulations, for the velocity error $e(u)$

4.3.2 A subsonic non-cavitating polytropic flow

In this test-case, a fully subsonic flow, characterised by a low Mach number value (approximately 10^{-3}), is obtained by considering the following inlet conditions $\rho_\infty = 900$, $u_\infty = 0.03$ and nozzle area ratio $S_{min}/S_\infty = 0.9$. Moreover, the barotropic law (4.1) with $\kappa = 0.5$, $\varkappa = 2$ and $\gamma = 0$ is used as equation of state. Similar test cases have been used in previous works [97,102] to investigate the behaviour of the numerical solution in the low Mach number regime, and, in particular, to validate the preconditioning defined for the Roe scheme. Thus, the purpose, here, is to verify that the Rusanov scheme requires a preconditioning strategy, and successively, to validate the technique proposed in (3.71)-(3.70). Fig. 4.11 shows the comparison between the preconditioned and the non-preconditioned simulations, as well as between the Roe and the Rusanov schemes. It turns out that a preconditioner is actually needed and the proposed one seems effective. For this test-case two different versions of the Rusanov scheme are considered, namely RC-approach which is based on equation (3.64) for the definition of the upwind matrix, and the RR one, which is based on (3.63). Note that, only the first-order formulations of the considered schemes have been used in these simulations. As previously pointed out, both the RC and RR compute the same steady solution which appears slightly less accurate than the one obtained using the Roe scheme, as shown in Fig. 4.11(b). This confirms that the RC and RR formulations are equivalent (see also [16] for a detailed investigation). Thus, this is the only test-case in which the Rusanov scheme based on formulation (3.64) is considered and in the following the formulation (3.63) will always be used



(a) Without preconditioning **(b)** With preconditioning

Figure 4.11: Comparison of the pressure field between the Roe (RO) and the Rusanov flux functions RR and RC which are, respectively based on the formulations (3.63) and (3.64), $\Delta x = L/4000$.

4.3.3 A Quasi-1D nozzle cavitating flow with shock

This test-case is characterised by inlet conditions $\rho_\infty \simeq 997.94994$, $u_\infty \simeq 1.0005$ and $S_{min}/S_\infty = 0.5$. The equations of state (2.6)-(2.7) for cavitating flows are used by taking $T_L = 293.16K$ and $\zeta = 0.1$. For the considered conditions, a subsonic/supersonic cavitating flow with a steady shock wave is obtained. Indeed, at the inlet the flow is characterised by a very low Mach number ($M_\infty = 1.4 \cdot 10^{-3}$), in the convergent zone the liquid expands and transitions to the vapour phase immediately upstream the throat, where Mach equal to 1 is reached. Afterward the flow is expanded until Mach 14 and then, through a shock wave, the flow reverts to subsonic and to the liquid phase (see e.g. Fig. 4.13b). This test-case contains all the difficulties typically encountered in cavitating conditions: two smooth flow regions divided by a shock wave, huge variations of the Mach number value with both highly compressible and incompressible areas and liquid/vapour and vapour/liquid transitions. This test-case is aimed at comparing the different schemes proposed in section 3.5 in a cavitating case as well as at evaluating their behaviour in presence of discontinuities.

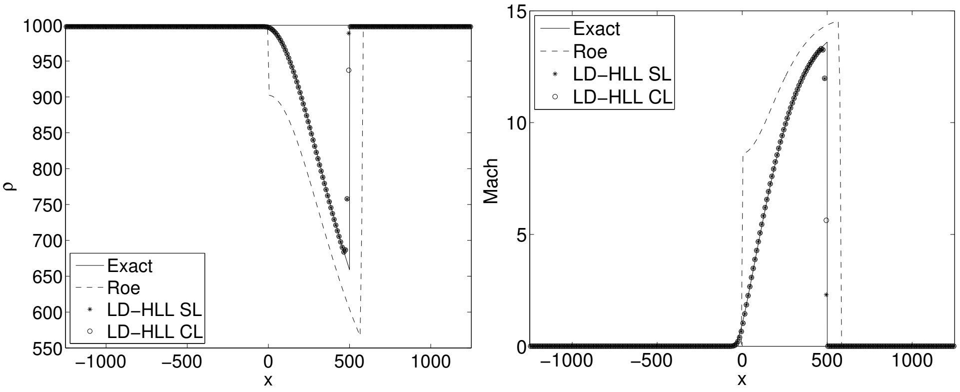
It has been shown in previous studies [97,102] that strong stability limitations appear when cavitation occurs for the case of a preconditioned linearised Roe formulation. Thus, by taking a fixed spatial grid ($\Delta x = L/250$, L being the computational domain length), the stability properties of the different approaches have been appraised. The maximum values of the CFL allowed by numerical stability obtained by this analysis are reported in Tab. 4.1. In accordance with the previous studies, the Roe scheme is constrained by a stability condition which limits the maximum CFL coefficient to a small value, here of about 0.01. The same limitation is also observed for the LD-HLL scheme when the standard approximated time linearization (3.32) is used. Conversely, by adopting the more complete linearization in time, i.e. (3.52)-(3.75), this limit is increased up to 1400, both for the first- and second-order schemes, as it is shown in Table 4.1. Thus, the LD-HLL scheme with the complete linearization can be considered a major improvement in term of robustness, since an increase of five orders of magnitude is obtained for

	1^{st} order Roe	1^{st} order LD-HLL SL	1^{st} order LD-HLL CL	2^{nd} order LD-HLL CL
CFL_{max}	0.01	0.01	1400	1400

Table 4.1: Maximum CFL values for the quasi-1D cavitating nozzle flow. SL stands for the standard time linearization (3.32) and CL for the more complete one (3.52)-(3.75).

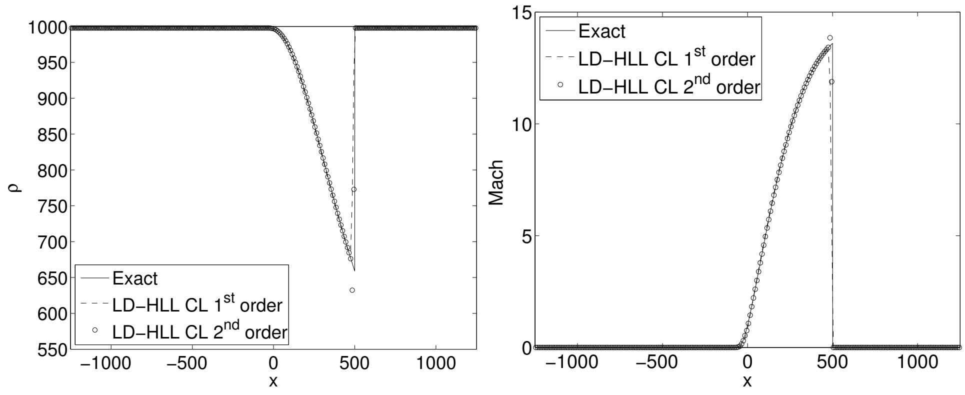
the CFL coefficient. These results confirm that the more complete linearization takes into account terms that become important when a subsonic phase transition or a shock wave (or both) are present.

Moreover, Fig. 4.12 shows that all the first-order accurate formulations based on the LD-HLL scheme provide practically the same results, which are in good agreement with the exact solution. Conversely, the Roe scheme introduces an unphysical expansion shock wave and, as a consequence, the solution given by the numerical scheme is not accurate: this is the well-know problem of the entropy violation of the Roe scheme and would require an entropy fix (see, e.g., [107]). Finally Fig. 4.13 shows the comparison between the results for the first- and second-order LD-HLL scheme with the complete linearization. For the considered grid resolution, the second-order scheme is slightly less diffusive than the first-order one. However, even with the use of a limiter, near the shock wave there is a small overshoot of the solution. We recall that vapour to liquid phase transition occurs at the same location as the shock wave, this adding to the problem stiffness.



(a) density **(b)** Mach number

Figure 4.12: Cavitating quasi-1D nozzle flow: comparison between different 1^{st}-order schemes.



(a) density **(b)** Mach number

Figure 4.13: Cavitating quasi-1D nozzle flow: comparison between 1st- and 2nd-order LD-HLL scheme with complete time linearization.

Chapter 5

3D numerical formulation

In this chapter a 3D numerical method for the simulation of cavitating flows is presented. The proposed method is based on a mixed finite-volume/finite-element formulation. The convective fluxes are defined by generalising the finite-volume formulation proposed in Chap. 3 while a finite-element approach is considered for the discretization of the viscous fluxes.

First the generation of a suitable discretisation for the computational domain is addressed in Sec. 5.1. Then the extension to the 3D case of the discretization of the convective fluxes and of the MUSCL reconstruction techniques are described in Sec. 5.3 and Sec. 5.4 respectively. Finally, the treatment of the viscous fluxes and of the source terms are presented in Sec. 5.5 and Sec. 5.6.

5.1 Mixed finite volume/finite elements approximation

The considered spatial approximation of the computational domain is based on a mixed finite-volume/finite-element approach. The same approach considered in [102] is used to discretise the computational domain. An unstructured grid is used to discretize the computational domain $\mathcal{D}$ in N_t tetrahedrons. The h^{th} tetrahedron is denoted by T_h and $i(T_h)$ is the set of indexes marking the vertexes $\mathbf{P}_r$ belonging to T_h. The generic tetrahedrons T_h will be also the basic geometric element for the finite element formulation described in section 5.5.

Starting from a tetrahedral grid, it is possible to associate a finite volume cell C_i to each vertex $\mathbf{P}_i$ by simply defining:

$$C_i = \bigcup_{h \in t(i)} C_i^{(h)} \tag{5.1}$$

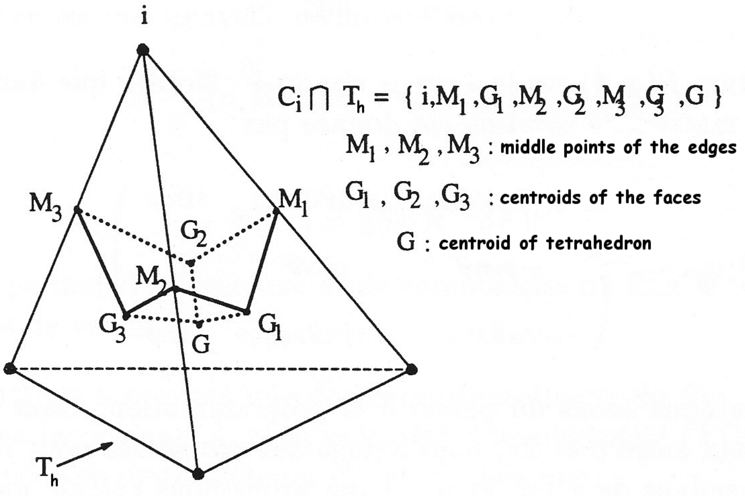


Figure 5.1: Example of the construction of a 3D finite volume cell starting from triangular elements.

where $C_r^{(h)}$ represents the subset of T_h which is defined by further dividing T_h into 24 sub-tetrahedrons by means of its median planes and subsequently considering those 6 sub-tetrahedrons which share $\mathbf{P}_r$ as a vertex. The set $t(i)$ is the set of indexes marking those tetrahedrons which share $\mathbf{P}_i$ as a vertex. The boundary ∂C_i of the cell can be defined as:

$$\partial C_i = \bigcup_{j \in \mathcal{N}(i)} \partial C_{ij} \tag{5.2}$$

where $\mathcal{N}(i)$ is the set of the neighbours of the i-th vertex and ∂C_{ij} is the common interface between the i-th cell the j one.

Clearly, there is a finite volume cell for each vertex and it is sometimes referred to as a "dual mesh" (see e.g. [52]), by virtue of the specific procedure which is adopted in order to build the cells by starting from the tetrahedrons. An example of the construction of the finite volume cells is shown in Fig. 5.1.

5.2 3D Spatial discretization

Let us specialise equation (2.37) for the case in which the control volume $\mathcal{V}$ is defined by (5.1):

$$\int_{C_i} \frac{\partial \mathbf{W}}{\partial t} \mathrm{d}\mathcal{V} + \int_{\partial C_i} n_k \left(F_k + \widetilde{F}_k \right) \mathrm{d}\mathcal{S} - \int_{C_i} \frac{\partial}{\partial x_k} \left(\mu V_k + \mu_t \widetilde{V}_k \right) \mathrm{d}\mathcal{V} = \int_{\mathcal{C}_i} \Omega \mathrm{d}\mathcal{V} \tag{5.3}$$

Thus, it is possible to reformulate (5.3) in the semi-discrete form:

$$\mathcal{C}_i \frac{\mathrm{d}\mathbf{W}_i}{\mathrm{d}t} + \sum_{j \in N(i)} \Phi_{ij} + \Upsilon_i = \mathbf{\Omega}_i \tag{5.4}$$

where Φ_{ij}, Υ_i, and $\mathbf{\Omega}_i$ are numerical approximation for, respectively, the convective fluxes, the viscous ones and the source term. $\mathcal{C}_i$ is the cell volume and $\mathbf{W}_i$ is the semi-discrete unknown associated with C_i, that is:

$$\mathbf{W}_i = \frac{1}{\mathcal{C}_i} \int_{C_i} \mathbf{W} \mathrm{d}\mathcal{V} \tag{5.5}$$

The semi-discrete formulation (5.4) is completely defined once the definition of Φ_{ij}, Υ_i, and $\mathbf{\Omega}_i$ are given.

5.3 3D Convective fluxes

Comparing equations (5.3) and (5.4), it follows that the numerical flux function Φ_{ij} must satisfy the relation:

$$\Phi_{ij} \simeq \int_{\partial C_i} n_k \left(F_k + \widetilde{F}_k \right) \mathrm{d}\mathcal{S} \tag{5.6}$$

For the sake of clarity, initially only the laminar convective fluxes will be considered, that is:

$$\Phi_{ij} \simeq \int_{\partial C_i} n_k F_k \mathrm{d}\mathcal{S} \tag{5.7}$$

then the required modifications needed to take into account $\widetilde{F}_k$ will be briefly outlined.

In section 2.4 it has been pointed out that F defined in (2.15) is a particular instance of $n_k F_k$, corresponding to the case $\mathbf{n} = (1, 0, 0)^T$. Considering two different Cartesian frames, $\mathbf{O}$ and $\mathbf{O}'$, sharing the same origin, let $\mathbf{R}$ denote the matrix associated with the rotation from $\mathbf{O}'$ to $\mathbf{O}$ (note that $\mathbf{R}^{-1} = \mathbf{R}^T$) The formal expression of $\mathbf{R}$ is:

$$\mathbf{R} = \begin{pmatrix} e'_{1,1} & e'_{2,1} & e'_{3,1} \\ e'_{1,2} & e'_{2,2} & e'_{3,2} \\ e'_{1,3} & e'_{2,3} & e'_{3,3} \end{pmatrix} \tag{5.8}$$

where $e'_{i,j}$ is the component of the j-th coordinate axis of the Cartesian frame $\mathbf{O}'$ along the i-th axis of the $\mathbf{O}$ one. Considering the state vector $\mathbf{W}$, its first component, the density, is a scalar which is invariant with respect to a frame change, while the others are the components of the momentum vector. As a consequence, the expressions of the state vector in the reference frame $\mathbf{O}$ ($\mathbf{W}$) and in the $\mathbf{O}'$ one ($\mathbf{W}'$) are related by the following transformation:

$$\mathbf{W} = \overline{\mathbf{R}} \cdot \mathbf{W}' \tag{5.9}$$

where $\overline{\mathbf{R}}$ is defined as:

$$\overline{\mathbf{R}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & e'_{1,1} & e'_{2,1} & e'_{3,1} \\ 0 & e'_{1,2} & e'_{2,2} & e'_{3,2} \\ 0 & e'_{1,3} & e'_{2,3} & e'_{3,3} \end{pmatrix} \tag{5.10}$$

Also, the flux function must be invariant by rotation ([52] and [107]), that is:

$$\Phi_{ij} = \overline{\mathbf{R}} \cdot \Phi'_{ij} \tag{5.11}$$

Applying (5.9) and (5.11) to the flux function (3.21)-(3.22) the transformation of the flux function reads:

$$\begin{cases} \Phi_{ij} & = & \Phi_{ij,c} + \Phi_{ij,u} \\ \Phi_{ij,c} & = & \dfrac{\overline{\mathbf{R}}F(\mathbf{W}'_i) + \overline{\mathbf{R}}F(\mathbf{W}'_j)}{2} \\ \Phi_{ij,u} & = & -\dfrac{1}{2}\overline{\mathbf{R}}Q\left(\mathbf{W}'_i, \mathbf{W}'_j\right)\overline{\mathbf{R}}^{-1}(\mathbf{W}_j - \mathbf{W}_i) \end{cases} \tag{5.12}$$

Once (5.12) is given, it is straightforward to generalise all the numerical flux functions described in Chap. 3 to the 3D case. To ease the comprehension, let consider for instance the following preconditioned S matrix for the LD-HLL scheme, i.e.S_{ij} defined by:

$$S_{ij} = \begin{pmatrix} \lambda_1^p & 0 & 0 & 0 \\ 0 & \lambda_2^p & 0 & 0 \\ 0 & 0 & \lambda_3^p & 0 \\ 0 & 0 & 0 & \lambda_3^p \end{pmatrix} \tag{5.13}$$

where λ_1^p, λ_2^p and λ_3^p are defined by (3.74).

The normal vector $\mathbf{n}_{ij} = (n_{ij,1}, n_{ij,2}, n_{ij,3})^T$ to the interface between the i-th cell and j-th is defined as:

$$\mathbf{n}_{ij} = \int_{\partial C_{ij}} \mathbf{n} \mathrm{d}\mathcal{S} \tag{5.14}$$

The 1D numerical flux function is a particular instance of the 3D case when $\mathbf{n}_{ij}$ is coincident with the first Cartesian axis of the reference frame, that is when:

$$\mathbf{n}_{ij} \equiv \mathbf{e}'_1 \tag{5.15}$$

First, let us decompose S_{ij} as follows:

$$S_{ij} = \begin{pmatrix} \lambda_1^p & 0 & 0 & 0 \\ 0 & \lambda_3^p & 0 & 0 \\ 0 & 0 & \lambda_3^p & 0 \\ 0 & 0 & 0 & \lambda_3^p \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \Delta^{32}\lambda^p & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.16}$$

where $\Delta^{32}\lambda^p = \lambda_2^p - \lambda_3^p$. Then, by applying (5.10) and (5.12) to (5.13) in the specific case (5.15), the following equality is obtained by using also the orthonormality of the base $(\mathbf{e}_1', \mathbf{e}_2', \mathbf{e}_3')$:

$$\overline{\mathbf{R}}\, S_{ij}\, \overline{\mathbf{R}}^{-1} = \begin{pmatrix} \lambda_1^p & 0 & 0 & 0 \\ 0 & \lambda_3^p & 0 & 0 \\ 0 & 0 & \lambda_3^p & 0 \\ 0 & 0 & 0 & \lambda_3^p \end{pmatrix} + \overline{\mathbf{R}} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \Delta^{32}\lambda^p & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \overline{\mathbf{R}}^T \tag{5.17}$$

and, finally, the resulting 3D upwind matrix for the LD-HLL scheme is:

$$S_{ij} = \begin{pmatrix} \lambda_1^p & 0 & 0 & 0 \\ 0 & \Delta^{32}\lambda^p\, n_{ij,1}^2 + \lambda_3^p & \Delta^{32}\lambda^p\, n_{ij,1}\, n_{ij,2} & \Delta^{32}\lambda^p\, n_{ij,1}\, n_{ij,3} \\ 0 & \Delta^{32}\lambda^p\, n_{ij,2}\, n_{ij,1} & \Delta^{32}\lambda^p\, n_{ij,2}^2 + \lambda_3^p & \Delta^{32}\lambda^p\, n_{ij,2}\, n_{ij,3} \\ 0 & \Delta^{32}\lambda^p\, n_{ij,3}\, n_{ij,1} & \Delta^{32}\lambda^p\, n_{ij,3}\, n_{ij,2} & \Delta^{32}\lambda^p\, n_{ij,3}^2 + \lambda_3^p \end{pmatrix} \tag{5.18}$$

Using the same approach the definition of the upwind matrix for the Roe and Rusanov schemes is straightforward.

5.3.1 Turbulent convective fluxes

The generalisation to the 3D case of the turbulent convective fluxes can be easily obtained considering the sum of the turbulent and laminar convective fluxes $n_k \left(F_k + \widetilde{F}_k \right)$ instead of simply the laminar ones. However, the equation of the turbulent kinetic energy k is not decoupled from the mass and momentum balances and, thus, a specific analysis is required for this case.
For the sake of simplicity let us consider a reduced 1D turbulent system containing only the unknowns ρ, ρu and ρk:

$$\begin{cases} \dfrac{\partial \hat{\mathbf{W}}}{\partial t} + \dfrac{\partial \hat{F}(\hat{\mathbf{W}})}{\partial x} = 0 \\ \hat{F}(\hat{\mathbf{W}}) = \left(\rho u, \rho u^2 + p + \dfrac{2}{3}\rho k, \rho u k \right)^T \\ \hat{\mathbf{W}} = (\rho, \rho u, \rho k)^T \end{cases} \tag{5.19}$$

The Jacobian matrix of $\hat{F}(\hat{\mathbf{W}})$ in (5.19) is different from the one of (2.11), due to the coupling term between the momentum balance and the equation for ρk:

$$\hat{A} \doteq \frac{\partial \hat{F}(\hat{\mathbf{W}})}{\partial \hat{\mathbf{W}}} = \begin{pmatrix} 0 & 1 & 0 \\ a^2 - u^2 & 2u & \frac{2}{3} \\ -uk & k & u \end{pmatrix} \tag{5.20}$$

However, by defining a “turbulent” speed of sound $\hat{a}$:

$$\hat{a} \doteq \sqrt{a^2 + \frac{2}{3}}$$

the eigenvalues of $\hat{A}$ can be formulated in formal analogy with (2.13) as:

$$\hat{\lambda}_1 = u - \hat{a}, \quad \hat{\lambda}_2 = u, \quad \hat{\lambda}_3 = u + \hat{a} \tag{5.21}$$

Let us focus on the consequences of (5.21) for the Rusanov and LD-HLL schemes. In the Rusanov scheme, the diagonal elements of the upwind matrix S should be an upper bound for the fastest signal velocity in the Riemann problem. However, considering the barotropic EOS described in section 2.2, it is almost always verified that:

$$a^2 \gg \frac{2}{3} \tag{5.22}$$

and, as consequence:

$$\hat{\lambda}_1 \simeq u - a, \quad \hat{\lambda}_2 = u, \quad \hat{\lambda}_3 \simeq u + a \tag{5.23}$$

Thus, when (5.22) is verified, it is reasonable to approximate the eigenvalues of the turbulent system (5.20) with the ones of (2.11). Even when (5.22) is not satisfied, it is possible to use (5.23) to approximate the eigenvalues of the system. Indeed, for the Rusanov and LD-HLL schemes, the choice of the eigenvalues only affects the upwind part of the numerical method so that a modification of the upwind term does not affect the consistency of the numerical scheme. As a consequence, when (5.22) is not satisfied, (5.23) is an underestimation of the true eigenvalues. However, in view of the previous considerations, this could affect only the stability properties of the numerical method, not its consistency. Using (5.23) it is possible to extend the Rusanov method to the 3D turbulent case without any additional modification. Instead, more attention is required for the LD-HLL. The low diffusive modification of the Rusanov scheme is based on the structure of the eigenvectors of system (2.11), where the eigenvector associated with u is in the form $(0\ 0\ 1)^T$. This is not true for system (5.19), where the right eigenvector matrix can be formulated as:

$$\hat{R} = \begin{pmatrix} 1 & -\dfrac{2}{3}\dfrac{1}{a^2} & 1 \\[2ex] u - \hat{a} & -\dfrac{2}{3}\dfrac{u}{a^2} & u + \hat{a} \\[2ex] k & 1 & k \end{pmatrix} \tag{5.24}$$

However, when (5.22) is satisfied, it is reasonable to approximate $\hat{R}$ as (5.23):

$$\hat{R} \simeq \begin{pmatrix} 1 & 0 & 1 \\ u - a & 0 & u + a \\ k & 1 & k \end{pmatrix} \tag{5.25}$$

and, thus, the low diffusive modification of the Rusanov scheme is appliable even to this case. On the contrary, if (5.22) is not verified, the use of (5.25) instead of (5.24) could results in a “partial” low diffusive modification. The consistency of the resulting scheme is not affected since the aforementioned modification only concerns the upwind part of the numerical method. On the other hand, the accuracy of the numerical method could be deteriorated. Nevertheless, as previously pointed out, the considered turbulence model does not explicitly take into account cavitation and no specific modification has been included in this work. Thus, since also the considered turbulent physical model can be questionable when (5.22) is not satisfied, it seems reasonable to always use the approximation (5.25), at least at this stage.

In view of the aforementioned considerations, remembering that v, w, ε can be regarded as passive scalars, the upwind matrix S_{ij} for the Rusanov or the LD-HLL in the 3D turbulent case can be formulated as:

$$\begin{pmatrix} \lambda_1^p & 0 & 0 & 0 & 0 & 0 \\ 0 & \Delta^{32}\lambda^p\, n_{ij,1}^2 + \lambda_3^p & \Delta^{32}\lambda^p\, n_{ij,1}\, n_{ij,2} & \Delta^{32}\lambda^p\, n_{ij,1}\, n_{ij,3} & 0 & 0 \\ 0 & \Delta^{32}\lambda^p\, n_{ij,2}\, n_{ij,1} & \Delta^{32}\lambda^p\, n_{ij,2}^2 + \lambda_3^p & \Delta^{32}\lambda^p\, n_{ij,2}\, n_{ij,3} & 0 & 0 \\ 0 & \Delta^{32}\lambda^p\, n_{ij,3}\, n_{ij,1} & \Delta^{32}\lambda^p\, n_{ij,3}\, n_{ij,2} & \Delta^{32}\lambda^p\, n_{ij,3}^2 + \lambda_3^p & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_3^p & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_3^p \end{pmatrix} \tag{5.26}$$

where $\Delta^{32}\lambda^p = \lambda_2^p - \lambda_3^p$ and λ_1^p, λ_2^p and λ_3^p are approximated using (5.23).

5.4 Second-order extension for the 3D case

The second-order extension of the convective fluxes described in Sec. 5.3 is carried on using a MUSCL reconstruction technique [115] similar to the one described for the 1D case in section 3.4. In particular the flux function Φ_{ij}

is not computed using the values $\mathbf{W}_i$ and $\mathbf{W}_j$ but considering two suitable states which, except in presence of discontinuities, approximate with second-order accuracy the solution at the boundary between the two cells, that is:

$$\Phi_{ij} = \Phi_{ij}\left(\mathbf{W}_{ij}^{-}, \mathbf{W}_{ij}^{+}\right) \tag{5.27}$$

where

$$\mathbf{W}_{ij}^{-} = \mathbf{W}_i + \frac{1}{2}\nabla\mathbf{W}_i \cdot \mathbf{d}_{ij}, \quad \mathbf{W}_{ij}^{+} = \mathbf{W}_j - \frac{1}{2}\nabla\mathbf{W}_j \cdot \mathbf{d}_{ij} \tag{5.28}$$

In (5.28) $\mathbf{d}_{ij}$ is the vector joining the i-th node with the j-th one and $\nabla\mathbf{W}_i$ is an approximation of the gradient in the i-th cell, possibly taking into account slope limiters. The definition of $\nabla\mathbf{W}_j \cdot \mathbf{d}_{ij}$ in this work is based on the same approach considered in [1, 24] when limiters are not required and on the works [45, 48] when limiters have to be included. First, a linear approximation $\nabla\mathbf{W}_{|T_h}$ of the gradient in each tetrahedron T_h is considered:

$$\nabla\mathbf{W}_{|T_h} = \sum_{k\in i(T_h)} \mathbf{W}_k \nabla\mu_{T_h}^{k} \tag{5.29}$$

where $\mu_{T_h}^{k}$ is the barycentric coordinates associated with the k^{th}-vertex of T_h. Once $\nabla\mathbf{W}_{|T_h}$ is available, a centred approximation of the gradient in the i-th cell is derived as follows:

$$\nabla\mathbf{W}_i^{c} = \frac{\displaystyle\sum_{h\in t(i)} \mathcal{V}_{T_h}\nabla\mathbf{W}_{|T_h}}{\displaystyle\sum_{h\in t(i)} \mathcal{V}_{T_h}} \tag{5.30}$$

$\mathcal{V}_{T_h}$ being the volume of the h-tetrahedron. Following [45, 48], it is now possible to obtain a second-order MUSCL reconstruction operator with slope limiters by defining $\nabla\mathbf{W}_i \cdot \mathbf{d}_{ij}$ and $\nabla\mathbf{W}_j \cdot \mathbf{d}_{ij}$ by:

$$\left\{\begin{array}{lcl} \nabla\mathbf{W}_i \cdot \mathbf{d}_{ij} & = & \mathrm{minmod}\{\mathbf{W}_j - \mathbf{W}_i, 2\nabla\mathbf{W}_i^{c} \cdot \mathbf{d}_{ij} - (\mathbf{W}_j - \mathbf{W}_i)\} \\ \nabla\mathbf{W}_j \cdot \mathbf{d}_{ij} & = & \mathrm{minmod}\{\mathbf{W}_j - \mathbf{W}_i, 2\nabla\mathbf{W}_j^{c} \cdot \mathbf{d}_{ij} - (\mathbf{W}_j - \mathbf{W}_i)\} \end{array}\right. \tag{5.31}$$

where minmod$\{a, b\}$ is defined as follows

$$\mathrm{minmod}\{a, b\} = \left\{\begin{array}{ll} \min\{a, b\} & \text{if } a > 0, b > 0 \\ \max\{a, b\} & \text{if } a < 0, b < 0 \\ 0 & \text{otherwise} \end{array}\right. \tag{5.32}$$

5.5 Diffusive fluxes

To correctly discretize (5.4), Υ_i must approximate the contribution of the viscous fluxes, that is:

$$\Upsilon_i \simeq \int_{C_i} \frac{\partial}{\partial x_k} \left(\mu V_k + \mu_t \widetilde{V}_k \right) \mathrm{d}\mathcal{V} \tag{5.33}$$

For the sake of simplicity, only the laminar diffusive fluxes will be considered since the extension to the turbulent fluxes is straightforward.

The discretization of the viscous fluxes is carried out using a P1 finite-element approach. The aforementioned formulation for the diffusive terms is obtained by multiplying the diffusive terms by a basis test function. First, let us focus on a single generic tetrahedron T_h: in the following the P1 finite-element basis function associated to the i^{th} node restricted to the tetrahedron T_h is denoted by $\phi^{(i,T_h)}$. In particular $\phi^{(i,T_h)}$ is assumed to be of unit value on the node i and to vanish linearly at the remaining vertexes of T_h. Thus, by integrating by parts the viscous fluxes multiplied by the basis test function on the T_h tetrahedron we obtain:

$$\int_{T_h} \frac{\partial \mu V_j}{\partial x_j} \phi^{(i,T_h)} \, \mathrm{d}\mathcal{V} = \int_{\partial T_h} \mu V_j \phi^{(i,T_h)} n_j^{T_h} \, \mathrm{d}\mathcal{S} - \int_{T_h} \mu V_j \frac{\partial \phi^{(i,T_h)}}{\partial x_j} \, \mathrm{d}\mathcal{V} \tag{5.34}$$

where $\mathbf{n}^{T_h}$ is the external unit normal vector to tetrahedron T_h. In the P1 formulation for the finite-element method, the test functions, $\phi^{(i,T_h)}$, are linear functions on the element T_h and so their gradient is constant. Moreover, in this formulation also the unknown variables are approximated by their projection on the P1 basis function. That is, considering the generic term σ_{ij} defined in (2.23), the corresponding P1-approximation on the tetrahedron T_h is given by:

$$\begin{aligned} \sigma_{ij}^{T_h} = \Bigg(\sum_{L \in i(T_h)} & -\frac{2}{3} \left(u_{L,k} \frac{\partial \phi^{(L,T_h)}}{\partial x_k} \right) \delta_{ij} \\ & + \left(u_{L,i} \frac{\partial \phi^{(L,T_h)}}{\partial x_j} + u_{L,j} \frac{\partial \phi^{(L,T_h)}}{\partial x_i} \right) \Bigg) \end{aligned} \tag{5.35}$$

where $u_{L,i}$ represents the i^{th} components of the velocity of the L^{th} node. Using (5.35) it is possible to define the following approximation of the viscous fluxes on each tetrahedron:

$$\begin{cases} V_1^{T_h} = \begin{pmatrix} 0, & \sigma_{11}^{T_h}, & \sigma_{12}^{T_h}, & \sigma_{13}^{T_h} \end{pmatrix}^T \\ V_2^{T_h} = \begin{pmatrix} 0, & \sigma_{21}^{T_h}, & \sigma_{22}^{T_h}, & \sigma_{23}^{T_h} \end{pmatrix}^T \\ V_3^{T_h} = \begin{pmatrix} 0, & \sigma_{31}^{T_h}, & \sigma_{32}^{T_h}, & \sigma_{33}^{T_h} \end{pmatrix}^T \end{cases} \tag{5.36}$$

Starting from (5.36), the approximation of the second term of the right hand side of (5.34) is done as follows:

$$\int_{T_h} \mu V_j \frac{\partial \phi^{(i,T_h)}}{\partial x_j} \, \mathrm{d}\mathcal{V} \simeq V_j^{T_h} \frac{\partial \phi^{(i,T_h)}}{\partial x_j} \int_{T_h} \mu \mathrm{d}\mathcal{V} \tag{5.37}$$

where $V_j^{T_h}$ and $\phi^{(i,T_h)}$ have been taken out of the integral since they are constant. Consistently with the P1 formulation, even the viscosity μ is a linear function on the tetrahedron T_h and, thus:

$$\int_{T_h} \mu \mathrm{d}\mathcal{V} = \mathcal{V}_{T_h} \mu^{T_h} \quad \text{where} \quad \mu^{T_h} = \sum_{L \in i(T_h)} \frac{\mu_L}{4} \tag{5.38}$$

and $\mathcal{V}_{T_h}$ is the volume of the tetrahedron T_h. As a consequence, substituting (5.38) in (5.37) the following approximation of the viscous fluxes on the T_h tetrahedron is obtained:

$$\int_{T_h} \mu V_j \frac{\partial \phi^{(i,T_h)}}{\partial x_j} \, \mathrm{d}\mathcal{V} \simeq \mathcal{V}_{T_h} \, \mu^{T_h} \, V_j^{T_h} \frac{\partial \phi^{(i,T_h)}}{\partial x_j} \tag{5.39}$$

In order to build the fluxes for the node i consistently with the finite-volume formulation, the contribution of all the elements having i as a vertex needs to be summed together as follows [47, 75, 94]:

$$\int_{C_i} \frac{\partial}{\partial x_j} \mu V_j \mathrm{d}\mathcal{V} \simeq \sum_{T_h \in t(i)} \left(\int_{\partial T_h} \mu V_j \phi^{(i,T_h)} n_j^{T_h} \, \mathrm{d}\mathcal{S} - \int_{T_h} \mu V_j \frac{\partial \phi^{(i,T_h)}}{\partial x_j} \, \mathrm{d}\mathcal{V} \right) \tag{5.40}$$

Considering (5.40) and (5.39), except for an additional contribution due to the boundary conditions, the following approximation holds for the viscous fluxes:

$$\Upsilon_i = - \sum_{T_h \in t(i)} \mathcal{V}_{T_h} \, \mu^{T_h} \, V_j^{T_h} \frac{\partial \phi^{(i,T_h)}}{\partial x_j} \tag{5.41}$$

Note that (5.41) has been derived for the laminar viscous fluxes. However, due to the formal analogy in the expression of the laminar and turbulent viscous fluxes, the extension to the turbulent case is straightforward.

5.6 Source Terms

In this section the numerical discretization of the source term appearing in (5.1) is addressed. Two contributions appear in the source term, the first one is the turbulence generation while, in rotating frames of reference, also the non inertial terms must be taken into account.

5.6.1 Turbulence generation

The discretization of the turbulence term appearing in (2.36) can be carried out using the same approach described in the previous sections. In particular for the discretization of the term $\mathcal{P}$, the contribution due to the viscous stress is computed as described in section 5.5.

5.6.2 Rotating frames

The non inertial terms due to a rotating frame of reference are addressed. Considering the boundary of the computational domain, $\partial\mathcal{D}$, only boundaries symmetric with respect to the rotation axis will be addressed in this work. Under this hypothesis, the frontier behaves like a fixed one in the body-frame and therefore it is possible to discretize the system of equation (5.3) without dealing with moving computational grids [102]. Let $\mathbf{g}_i$ denote the centroid associated with C_i:

$$\mathbf{g}_i := \frac{1}{\mathcal{C}_i} \int_{C_i} \mathbf{x}\, \mathrm{d}\mathcal{V} \tag{5.42}$$

Moreover, let $\mathbf{r}_i$ denote the projection of $\mathbf{g}_i$ on the rotation axis:

$$\mathbf{r}_i := -\,\hat{\vec{\omega}} \wedge \left(\hat{\vec{\omega}} \wedge \mathbf{g}_i\right)$$

where $\hat{\vec{\omega}}$ is the unit vector associated with $\vec{\omega}$:

$$\hat{\vec{\omega}} := \frac{\vec{\omega}}{\|\vec{\omega}\|}$$

Then it is possible to approximate the source term (2.44) as:

$$\int_{C_i} \mathbf{S}\, \mathrm{d}\mathcal{V} \approx \mathcal{C}_i\, \mathbf{S}_i \tag{5.43}$$

where

$$\mathbf{S}_i := \|\vec{\omega}\| \begin{pmatrix} 0 \\ -\,2\,\hat{\vec{\omega}} \wedge \rho_i \mathbf{u}_i + \rho_i\, \|\vec{\omega}\|\, \mathbf{r}_i \end{pmatrix} \tag{5.44}$$

5.7 Boundary conditions

The boundary conditions are introduced in a weak way, i.e. they are not forced directly on the solution vector at each time step. Firstly, the real boundary $\partial\mathcal{D}$ is approximated by a polygonal boundary Γ; this can be split in four parts:

$$\Gamma = \Gamma_{\mathrm{I}} \cup \Gamma_{\mathrm{O}} \cup \Gamma_{\mathrm{s}} \cup \Gamma_{\mathrm{b}} \tag{5.45}$$

in which the term Γ_{I} represents the inflow far-field boundary, Γ_O is the outflow far-field boundary, and Γ_s and Γ_b are, respectively, the boundaries at which free-slip and no-slip conditions are imposed.

Considering the convective fluxes, the treatment of the inflow and outflow boundary conditions is based on the conventional ghost-cell approach. More precisely, the numerical flux function at the aforementioned boundaries is computed through the same numerical flux function as for the internal cells by introducing fictitious state vectors. Namely, the flux $\mathbf{\Phi}_{i\mathrm{I}}$ between the inflow boundary and the generic boundary i^{th} cell is computed as follows:

$$\Phi_{i\mathrm{I}} = \Phi\left(\mathbf{W}_i, \mathbf{W}_\mathrm{I}, \mathbf{n}_{i\mathrm{I}}\right) \tag{5.46}$$

where $\mathbf{n}_{i\mathrm{I}}$ is the normal to the inflow boundary and $\mathbf{W}_\mathrm{I} = \mathbf{W}_\mathrm{I}\left(\mathbf{W}_i\right)$ is the inflow state vector associated to the i^{th} cell. The number and type of conditions at a boundary are defined by the eigenvalues of the Jacobian of the flux function [72, 73]. In particular at each boundary only the waves which are propagating inside the computational domain should be defined by the fictitious state vectors while the outgoing waves should be undisturbed in order to avoid reflections at the boundary. However, the definition of time-dependent non reflecting boundary conditions could require the solution of an additional equation for the definition of $\mathbf{W}_\mathrm{I}$ [73, 106].

A simpler approach, commonly used (see i.e. [60, 68]), is to assign a number of components of the fictitious state vector equal to the number of incoming waves and to extrapolate the others. In particular, by considering a subsonic inflow, the number of positive eigenvalues of the flux function is 3 and only one eigenvalue is negative. Thus 3 components of the state vector $\mathbf{W}_\mathrm{I}$, namely the 3 componets of the momentum vector ρu_I, ρu_I and ρu_I, are assigned while the last one, the density is extrapolated, that is:

$$\mathbf{W}_\mathrm{I}\left(\mathbf{W}_i\right) = \left(\rho_i,\ \rho u_\mathrm{I},\ \rho v_\mathrm{I},\ \rho w_\mathrm{I}\right)^T \tag{5.47}$$

The same approach is considered for the outflow boundary: in this case there is only one incoming wave and thus only one componet of the state vector $\mathbf{W}_\mathrm{O}$ is assigned, namely the density ρ_O, which, due to the barotropic state law is equivalent to assign the outlet pressure. Thus, the outflow convective flux is computed as:

$$\Phi_{i\mathrm{O}} = \Phi\left(\mathbf{W}_\mathrm{i}, \mathbf{W}_\mathrm{O}, \mathbf{n}_{i\mathrm{O}}\right) \tag{5.48}$$

where

$$\mathbf{W}_\mathrm{O}\left(\mathbf{W}_i\right) = \left(\rho_\mathrm{O},\ \rho u_i,\ \rho v_i,\ \rho w_i\right)^T \tag{5.49}$$

Concerning the free-slip conditions on the boundary Γ_s the standard impermeable condition:

$$\mathbf{u} \cdot \mathbf{n} = 0 \tag{5.50}$$

is adopted in this work. Thus, in this case the boundary condition can be computed as follows:

$$\Phi_{is} = (0, p_i \mathbf{n}_{is})^T \quad \text{and} \quad \Phi_{i\mathrm{b}} = (0, p_i \mathbf{n}_{ib})^T \tag{5.51}$$

Finally, in the laminar case the no-slip condition $\mathbf{u} = 0$ is strongly imposed on the solution. Instead, when turbulent simulations are considered approximate boundary conditions based on the Reichardt wall-law (see Sec. 2.5.1) are imposed. When the wall-law is applied, it is assumed that the wall nodes are not exactly on the body (Γ_b) but on an ideal surface Γ_{b^*} that is located at a small distance δ from Γ_b. The boundary conditions introduced on Γ_{b^*} are the following:

- $\mathbf{u} \cdot \mathbf{n} = 0$: the ideal surface Γ_{b^*} is assumed to be impermeable ;
- $\sigma_{ij} n_j = \tau_w n_j$: the viscous stresses on Γ_{b^*} are supposed to be assigned equal to τ_w;

where τ_w is defined by (2.41). The conditions on Γ_{b^*} leads to the following weak conditions:

$$\int_{\partial T_h \cap \Gamma_\mathrm{s}} \mu V_j \phi^{(i,T_h)} n_j^{T_h} \, \mathrm{d}\mathcal{S} \simeq |\partial T_h \cap \Gamma_\mathrm{s}| \, (0, \ \tau_w \mathbf{n})^T \tag{5.52}$$

where $|\partial T_h \cap \Gamma_\mathrm{s}|$ is the surface of the boundary. For the closure of the problem, it is necessary to assign the viscous stresses τ_w as a function of the unknown variables of the problem. This is done by applying the Reichardt wall-law which joins the velocity on Γ_i with a proper velocity profile that is imposed in the gap between Γ_{b^*} and Γ_b and satisfies the no-slip conditions on Γ_b. This procedure, which allows the shear stresses on Γ_{b^*} to be evaluated, is schematically sketched in Fig. 5.2.

5.8 3D time discretization

The time discretization for the 3D numerical method can be readily obtained from its 1D counterpart. Let consider a second-order accurate in space and time approach (the modifications for the first order case are straightforward). Then the implicit formulation of (5.4) is:

$$\begin{aligned} \mathcal{C}_i \frac{(1+2\tau)}{(1+\tau)\Delta^n t} \Delta^n \mathbf{W}_i + \sum_{j \in N(i)} \Delta^n \Phi_{ij} + \Delta^n \Upsilon_i - \Delta^n \mathbf{\Omega}_i = \\ \mathcal{C}_i \frac{\tau^2}{(1+\tau)\Delta^n t} \left(\mathbf{W}_i^{n-1} - \mathbf{W}_i^n \right) - \sum_{j \in N(i)} \Phi_{ij}^n - \Upsilon_i^n + \mathbf{\Omega}_i^n \end{aligned} \tag{5.53}$$

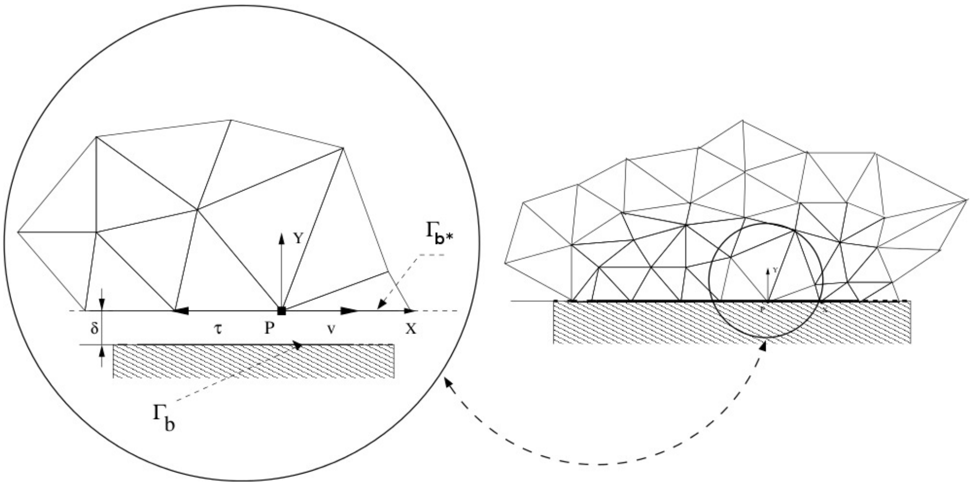


Figure 5.2: Schematic representation of the wall approach adopted.

In order to solve (5.53) it is necessary to define a suitable approximation for $\Delta^n\Phi_{ij}$, $\Delta^n\Upsilon_i$ and $\Delta^n\boldsymbol{\Omega}_i$.

Concerning the variation of the convective part, $\Delta^n\Phi_{ij}$, in view of the considerations of section 5.3, the time linearisations proposed in section 3.3 are applicable to the 3D case. In particular for the Rusanov and LD-HLL schemes both the simple (3.32) and the complete (3.52) time linearisation can be used in the 3D case.

Concerning the term $\Delta^n\Upsilon_i$, a simple linearisation technique is sufficient since, from (5.41), (5.39) and (5.35) it clearly appears that Υ_i is linearly dependent on the velocity vector. In particular $\Delta^n\sigma_{ij}^{T_h}$ can be written as:

$$\Delta^n\sigma_{ij}^{T_h} = \left(\sum_{L\in i(T_h)} -\frac{2}{3}\left(\Delta^n u_{L,k}\frac{\partial\phi^{(L,T_h)}}{\partial x_k}\right)\delta_{ij} \right.$$
$$\left. + \left(\Delta^n u_{L,i}\frac{\partial\phi^{(L,T_h)}}{\partial x_j} + \Delta^n u_{L,j}\frac{\partial\phi^{(L,T_h)}}{\partial x_i}\right)\right) \qquad (5.54)$$

Thus, the linearisation of $\Delta^n\sigma_{ij}^{T_h}$ is completed once the dependence of $\Delta^n u_{L,k}$ on $\Delta^n\mathbf{W}_L$ is explicitly taken into account. In particular, by approximating $\Delta^n u_{L,k}$ as:

$$\Delta^n u_{L,k} = -\frac{u_{L,k}^n}{\rho_L^n}\Delta^n\rho_L + \frac{1}{\rho_L^n}\Delta^n\rho u_{L,k} + O\left((\Delta^n t)^2\right) \qquad (5.55)$$

a second-order time linearisation for the viscous fluxes is achieved. Also, with a similar approach a second-order time linearisation for the source term is obtained. Thus, once $\Delta^n\Phi_{ij}$, $\Delta^n\Upsilon_i$ and $\Delta^n\mathbf{\Omega}_i$ are given, the second-order defect correction approach described in 3.4 can be applied also for the 3D case.

5.8.1 A linearised approach for turbulent flows

In this work a specific treatment has been considered for the linearisation of the turbulent case. In particular in the turbulent system (2.29) 6 components define the state vector $\mathbf{W}$, while in the laminar case only 4 components are considered. As a consequence, the matrix of the linear system which must be solved at each time step is a sparse matrix whose elements are 6×6 blocks instead of the 4×4 blocks of the laminar case. This clearly largely increases the memory and computational requirement of the numerical method.
In order to alleviate the memory and computational requirement, we introduced an additional approximation in the linearisation technique. In particular we set to zero the "mixed" derivatives, that is, in the mass and momentum equations the derivatives with respect to k and ε are nullified and similarly for the derivatives with respect ρ and $\rho\mathbf{u}$ of the equations for k and ε. This is equivalent to keep frozen at time t^n some quantities in each equation. As a consequence if B was a generic 6×6 block, then the modified block $\hat{B}$ reads:

$$B = \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} & b_{15} & b_{16} \\ b_{21} & b_{22} & b_{23} & b_{24} & b_{25} & b_{26} \\ b_{31} & b_{32} & b_{33} & b_{34} & b_{35} & b_{36} \\ b_{41} & b_{42} & b_{43} & b_{44} & b_{45} & b_{46} \\ b_{51} & b_{52} & b_{53} & b_{54} & b_{55} & b_{56} \\ b_{61} & b_{62} & b_{63} & b_{64} & b_{65} & b_{66} \end{pmatrix} \Longrightarrow$$

$$\hat{B} = \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} & 0 & 0 \\ b_{21} & b_{22} & b_{23} & b_{24} & 0 & 0 \\ b_{31} & b_{32} & b_{33} & b_{34} & 0 & 0 \\ b_{41} & b_{42} & b_{43} & b_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & b_{55} & b_{56} \\ 0 & 0 & 0 & 0 & b_{65} & b_{66} \end{pmatrix} \tag{5.56}$$

Considering the structure of the approximated block $\hat{B}$, the linear system which must be solved at each time step can be divided into two linear systems, one composed by 4×4 blocks, which updates the flow variables, ρ, ρu, ρv, ρw, and another one for the turbulent variables, ρk, $\rho\varepsilon$, composed by 2×2 blocks. Note that, even if (5.56) introduces an additional approximation to the discretization technique described in the previous sections, a linear system composed by modified block in the form (5.56) can be considered as particular instance of the DeC approach (3.11). As a consequence even if for this case we do not have theoretical results of a finite termination property, the second-order approximation should be recovered asymptotically when a full convergence of the DeC iteration is obtained.

Chapter 6

Numerical experiments for the validation of the 3D formulation

In this chapter, in order to validate and compare the previously defined numerical approaches, they have been applied to the simulation of different flow configurations. Two sets of test-cases have been considered: the first one, described in Sec. 6.1, is aimed at investigating the effects of the anti-diffusive term introduced in the LD-HLL scheme using as test problem the classical Blasius boundary layer problem.

Moreover, in Sec. 6.1.3 the methodology proposed in Chap. 5 for the discretization of the turbulent model (2.29) is validated by considering the turbulent flow over a flat plate. The second set of test-cases is aimed at investigating the accuracy, robustness and efficiency properties of the considered schemes in cavitating conditions by considering the flow around a NACA0015 hydrofoil mounted in a wind tunnel, for which experimental data are available [30]. For this last test-case the effects of viscosity have been neglected in the simulations: although the flow is obviously viscous, a sensible comparison against experimental data in terms of pressure distribution is possible also under the assumption of inviscid flow. On the other hand, this assumption clearly leads to an easier appraisal of the robustness and efficiency properties of the different numerical schemes in presence of cavitation in a realistic configuration.

Note that the aforementioned test-cases are essentially 2D. Nevertheless, the 3D formulation described in Chap. 5 has been used in these simulations by using computational domains with a little extension, i.e. containing only few nodes, in the spanwise direction.

The results shown in the following for the Rusanov and LD-HLL schemes, are obtained with the complete time linearisation (3.52), while the standard linearisation (3.32) is used only in Sec. 6.2.2 to compare with the more complete one. Indeed, only for these cases, in which cavitation is present, the problem of efficiency is substantial and thus significant differences are observed between the two approaches.

6.1 Flow Over a Flat Plate

In this section the flow over a flat plate is considered. More precisely, the laminar and the turbulent flow over a flat plate at zero angle of attack are analysed in order to validate the numerical ingredients introduced in Sec. 5. As previously mentioned, it is well-known that due to their excessive diffusivity in presence of contact discontinuities, the HLL schemes, as the standard Rusanov one, are not well-suited for viscous flow simulation. The Blasius boundary layer can thus be considered an appropriate test to validate the effectiveness of the anti-diffusive term introduced in the LD-HLL scheme. We consider hence the laminar viscous flow along a flat plate at zero angle of attack and water at $T = 293.16$ K is used as a working fluid. The considered free-stream conditions are described in Tab. 6.1 Note that the free-stream Mach number is approximately 10^{-5} and thus compressibility effects are negligible and comparison with the Blasius solution for incompressible flows can be made. The Reynolds number, based on the free-stream velocity and the flat plate length, is about 10000. The second test case is a turbulent flow. The set up of the turbulent problem is the same as for the laminar Blasius case except than for the free-stream velocity, which is increased, as shown in Tab. 6.1, in order to obtain a Reynolds number equal to 10^6. We remark that in this test-cases no cavitation phenomena occur in the flow. The computational domain is the following:

$$\begin{cases} -1 & \leq x \leq & 1 \\ 0 & \leq y \leq & 2 \end{cases}$$

where x is the streamwise direction and y the direction normal to the flat plate. The flat plate is located at $0 \leq x \leq 1$ and $y = 0$. The characteristic based boundary conditions described in Sec. 5.7 are used at the inflow ($x = -1$) and at the outflow ($x = 1$). Free-slip is imposed for $y = 0$ and $-1 < x < 0$ while no-slip conditions are considered at the flat-plate. Finally, free-slip conditions are used at $y = 2$. The independence of the results to the normal width of the domain has been checked.

6.1.1 Grid generation

The computational grids are generated as follows. Firstly, the square $0 \leq x \leq 1$ and $0 \leq y \leq 1$ is considered. The edge along the x direction is divided

Test case	ρ_∞	p_∞	u_∞	μ	L_{plate}	Re
Blasius	998	100000	0.01	0.001	1	10000
Turbulent	998	100000	1	0.001	1	1000000

Table 6.1: Main parameters of the flow along the flat plate; ρ_∞, p_∞ and u_∞ are the inflow density, pressure and velocity, μ is the fluid viscosity and L_{plate} is the flat plate length. All parameters are in SI units.

in a given number of segments, N_x, with length Δx varying in geometric progression with the smallest element (Δx_{min}) at the leading edge. Analogously the edge along the y direction is divided in N_y segments, with length Δy varying in geometric progression with the smallest element (Δy_{min}) at the wall. A structured quadrilateral grid is then generated and is reflected along the axis $x = 0$. For $y > 1$ the value of Δy is constant equal to the maximum value over the the square $0 \leq x \leq 1$ and $0 \leq y \leq 1$. Finally, each element of this grid is divided in 5 tetrahedrons by an English-flag type procedure. By using the previously described procedure, four grids whose main parameters are given in Tab. 6.2 were obtained.

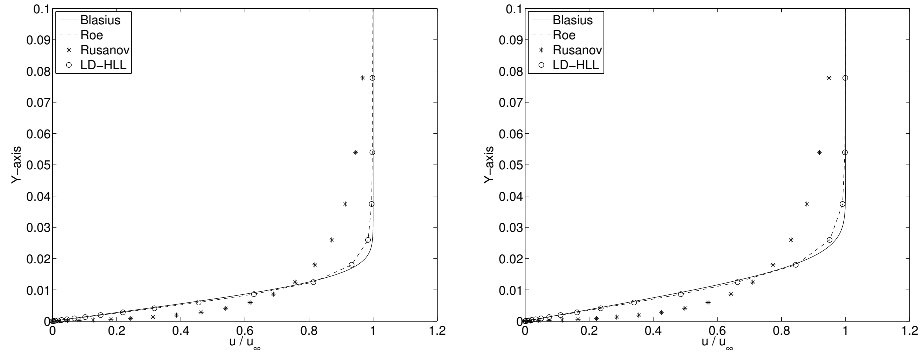
Grid	N_x	q_x	Δx_{min}	Δx_{max}	N_y	q_y	Δy_{min}	Δy_{max}
Grid1	25	1.21	0.0018	0.18	25	1.44	4.8×10^{-5}	0.33
Grid2	50	1.1	0.00086	0.1	50	1.2	2.2×10^{-5}	0.16
GR1	50	1	0.02	0.02	50	1.1	2.4×10^{-3}	0.068
GR2	100	1	0.01	0.01	100	1.04	1.2×10^{-3}	0.034

Table 6.2: Grids used for the flat plate simulations. N_x is the number of divisions of $0 \leq x \leq 1$, q_x is the ratio of the geometrical progression used for setting the length of the elements in the x direction, Δx_{min} and Δx_{max} are the minimum and maximum length of the elements in the x direction. N_y, q_y, Δy_{min}, Δy_{max} are the corresponding quantities in the y direction.

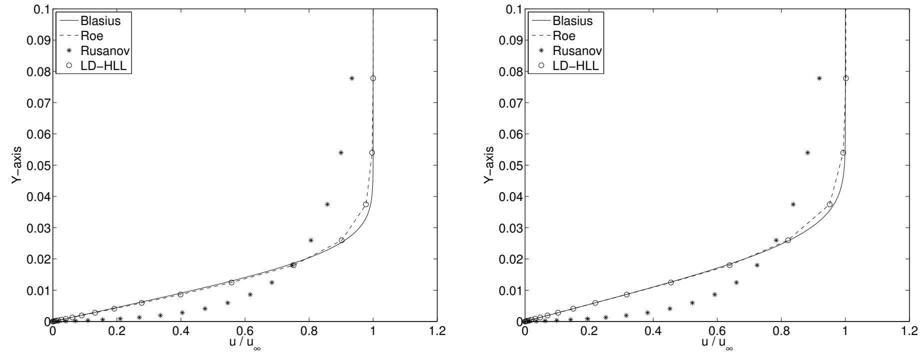
The grids Grid1 and Grid2 are used for the laminar simulations described in Sec. 6.1.2, while GR1 and GR2 are used for the turbulent computations. Note in particular that the considered grids for turbulent simulations are uniform in the x direction. This was done in order to avoid in these simulations a too large stretching of the grid elements and, as a consequence, to avoid numerical difficulties related to the discretization of the ε-equation. Also, the size of Δy for GR1 and GR2 has been chosen in order to define a near wall normal resolution corresponding to about 50-70 wall units for the first grid and to 25-40 for the second one.

6.1.2 Blasius Boundary Layer

The results of the simulations of the Blasius Boundary Layer with three different numerical schemes (Roe, Rusanov and LD-HLL) on grid Grid1 and Grid2 are compared and validated against the Blasius solution. Note that, for this test-case, all the simulations have been carried out by using the second-order accurate formulation. Figures 6.1 show the comparison of the velocity profiles at different sections over the flat plate, i.e. $x/c = 0.2$, $x/c = 0.4$, $x/c = 0.6$ and $x/c = 0.8$.



(a) $x/c = 0.2$ (b) $x/c = 0.4$



(c) $x/c = 0.6$ (d) $x/c = 0.8$

Figure 6.1: Velocity profiles at different sections over the flat plate obtained on grid Grid1.

Fig. 6.2 shows the comparison of the wall shear stress along the flat plate. While the results obtained with the standard Rusanov scheme lack in accuracy and confirm its over-diffusive character, the LD-HLL scheme gives very similar results to the Roe one and in good agreement with the Blasius solution. Small discrepancies are present between the Blasius solution and the numerical ones from LD-HLL and Roe approaches (see e.g. Fig. 6.2).

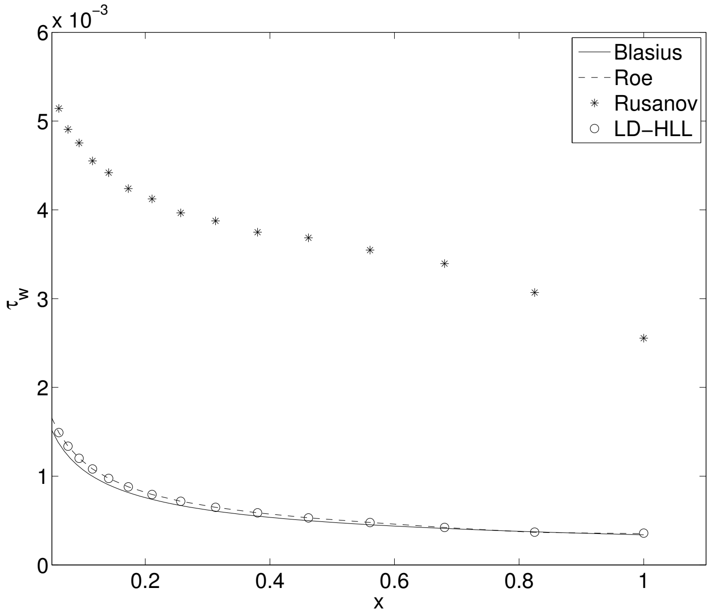


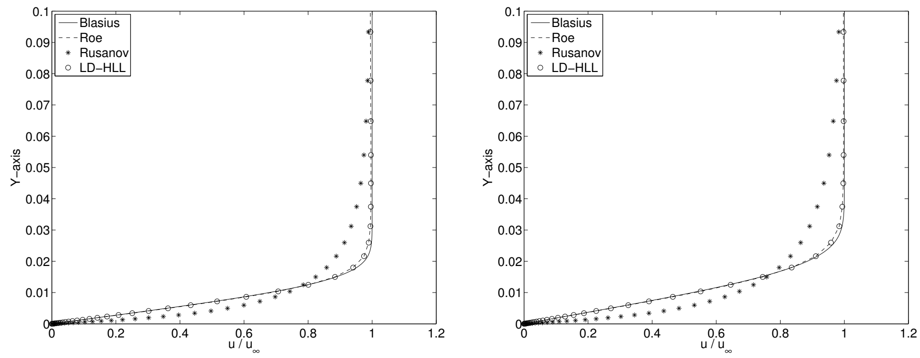
Figure 6.2: Comparison of the wall shear stress over the flat plate obtained on grid Grid1.

Nevertheless, they tend to vanish with grid refinement, as shown by the results obtained with the Roe and the LD-HLL schemes on grid Grid2 and reported in Figs. 6.3 and 6.4. Note how the standard Rusanov scheme gives inaccurate results also on this more refined grid. This test-case thus confirms that the anti-diffusive term introduced in the LD-HLL scheme actually counteracts the accuracy problems encountered by the standard Rusanov scheme in viscous flow simulations.

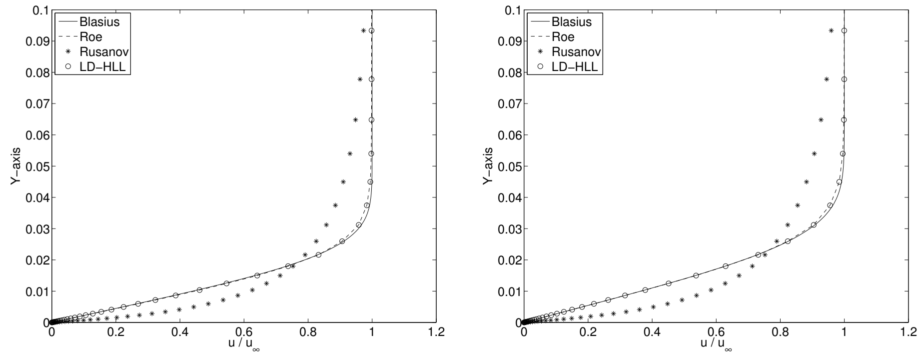
6.1.3 Validation for the turbulent flow around a flat plate at zero angle of attack

In this section, the results of the turbulent simulations performed using the LD-HLL scheme on the grids GR1 and GR2 are presented. Figure 6.5 shows the wall shear stress along the flat plate obtained in the present simulations compared to a well established empirical formula for the fully turbulent boundary layer over a flat plate, obtained by interpolating experimental data available in the literature:

$$c_f(x) = \frac{\tau_w(x)}{\frac{1}{2}\rho u_\infty^2} = \frac{0.0592}{Re_x^{0.2}} \qquad (6.1)$$



(a) $x/c = 0.2$ (b) $x/c = 0.4$



(c) $x/c = 0.6$ (d) $x/c = 0.8$

Figure 6.3: Velocity profiles at different sections over the flat plate obtained on grid Grid2.

where Re_x is the local Reynolds number based on the free-stream velocity, u_∞, and the distance from the plate leading edge, x. The agreement is reasonably good and improves with grid refinement. As it could be expected, the largest discrepancies are found near the leading-edge of the plate and they could be reduced by using a non uniform grid in the x direction, more refined near the leading edge, as the ones previously employed for the Blasius boundary layer. This is not done at this stage, since the objective of the present simulations was simply to validate the implementation of the turbulence model, which can be considered achieved already from the results reported in this section.

Fig. 6.6 shows the velocity profiles in wall units obtained at two different sections over the plate for the two considered grids. We recall that an approximate boundary condition based on the wall-law has been used; thus, the first computational point is located in the logarithmic region. The logarithmic behaviour is qualitatively well reproduced in the simulations, although

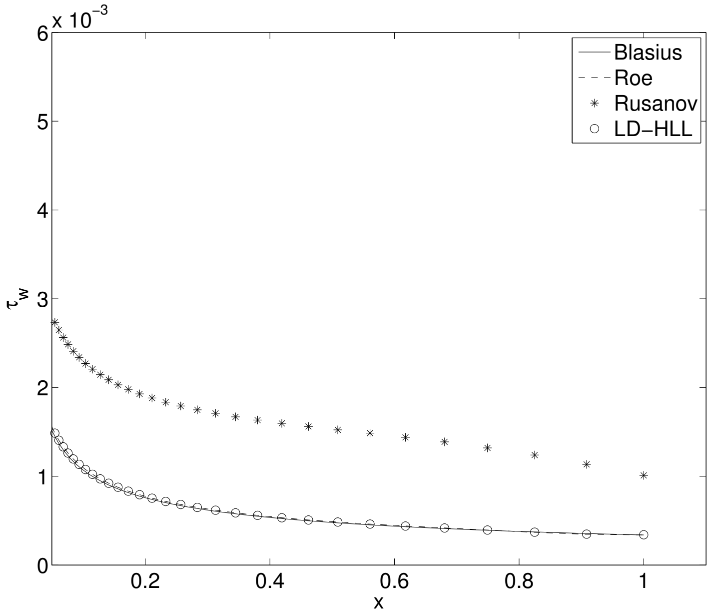


Figure 6.4: Comparison of the wall shear stress over the flat plate obtained on grid Grid2.

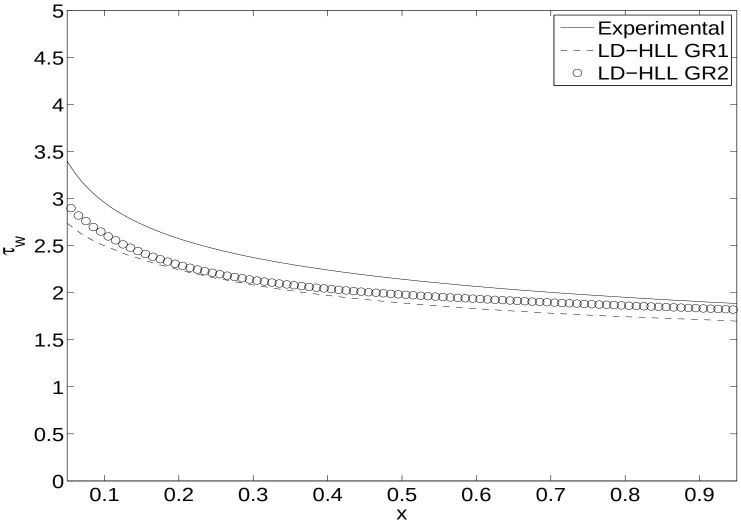
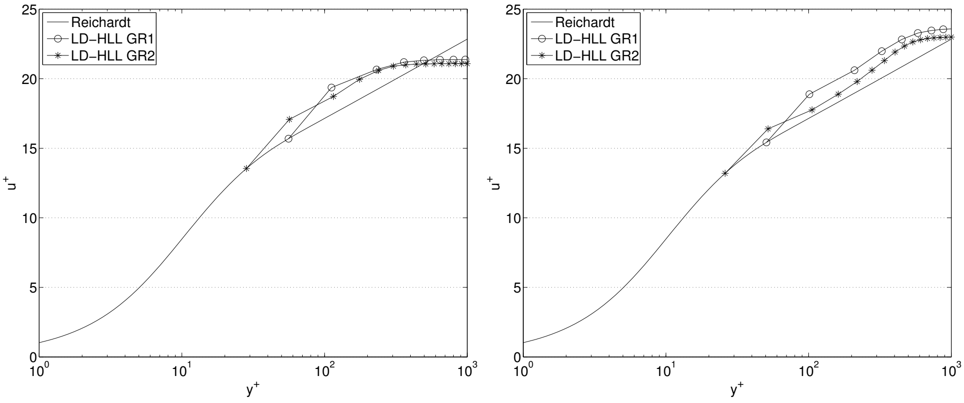


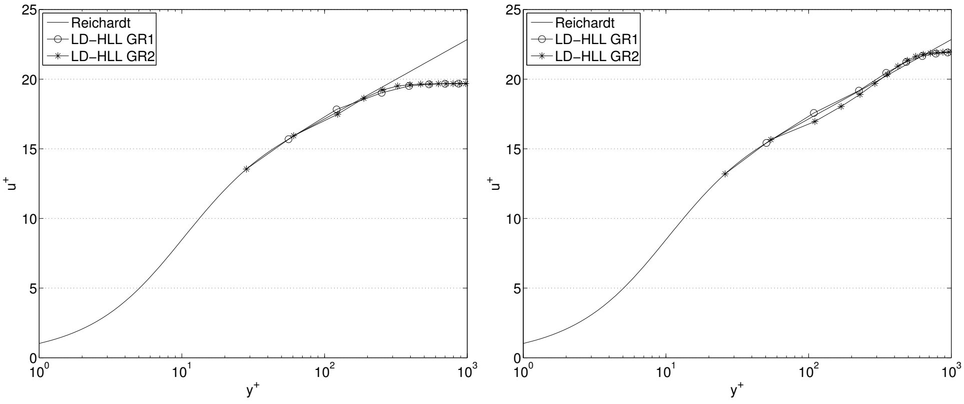
Figure 6.5: Comparison of the wall shear stress over the flat plate at Re=10^6.

there is a quantitative overestimation with respect to the wall law Fig. 6.6. However, this is due to the fact that u^+ is computed by using the local shear velocity obtained in the simulations, which is underestimated with respect to the reference one, as shown in Fig. 6.5. Indeed, if the profiles are rescaled by using the shear velocity given by the well established empirical formula in the computation of u^+ and y^+, a very good agreement is obtained as shown in Fig. 6.7.



(a) $x = 0.2$ **(b)** $x = 0.6$

Figure 6.6: Velocity profiles in wall units at different sections.



(a) $x = 0.2$ **(b)** $x = 0.6$

Figure 6.7: Rescaled velocity profiles in wall units at different sections.

Note that for the turbulent simulations a maximum CFL number of 10000 could be reached. This value is one order of magnitude smaller than the one found for laminar computations, 10^5. Indeed, this stability restriction could be a consequence of the specific linearisation proposed in Sec. 5.8. However, the allowed values of the CFL number found in this test-case are still two order of magnitude larger than the ones found in more complex applications and, as a consequence, no investigation of its cause has been carried on.

6.2 Flow Over a NACA 0015 Hydrofoil

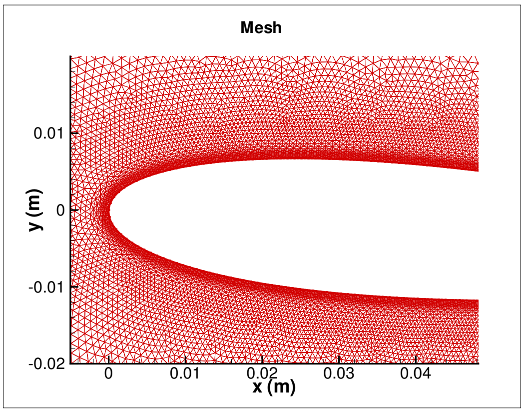
In this section the liquid flow over a NACA0015 hydrofoil in cavitating and non-cavitating conditions is considered. The hydrofoil of chord length c equal to 115 mm is mounted in a water tunnel at 4° incidence angle and spans the entire width of the rectangular test chamber section. The test section, which is obtained by cutting the chamber along its symmetry plane has a width of $1.28c$. This configuration has been considered in an experimental study [30] for which the pressure coefficient distribution on the symmetry plane of the hydrofoil is available.

Three different sets of inlet conditions are considered herein, as summarised in Tab. 6.3; the first conditions (TC1) correspond to a non-cavitating case while the second and the third ones (TC2 and TC3) generate a cavitating flow. The inlet conditions TC1 and TC2 correspond to those of the reference experiments. The equations of state (2.6)-(2.7) for cavitating flows are used by taking T_L equal the temperature T given by Tab. 6.3 and $\zeta = 0.01$. The dimensions of the computational domain in the lateral direction are the same of the experimental test section, while in the streamwise direction the inlet is at a distance of $3c$ from the leading edge and the outflow at a distance of $4c$ from the trailing edge. Finally, in the spanwise direction the domain thickness is $0.1c$, i.e. only a slice of the actual chamber width is considered. In [14] it was shown that this has not significant effects on the numerical pressure distribution over the hydrofoil, except for slight differences near the trailing edge (see also the brief discussion in the following). The characteristic based boundary conditions (see Sec. 5.7) are used at the inflow and outflow, while free slip is imposed at the remaining boundaries. Three different unstructured grids, shown in Figs. 6.8 and 6.9 are considered, the first

Test-case	U_∞ m/s	p_∞ Pa	T K	M_∞	σ_∞
TC1	3.115	59050	298.15	2.2×10^{-3}	11.71
TC2	3.46	12000	298.15	2.4×10^{-3}	1.49
TC3	3.46	11400	298.15	2.4×10^{-3}	1.38

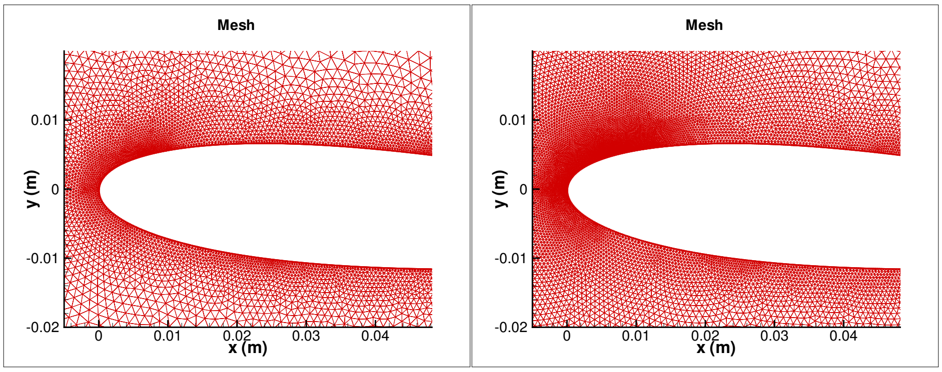
Table 6.3: Inlet conditions for the flow around a NACA0015 hydrofoil: σ_∞ is the inlet cavitation number defined as $\sigma_\infty = \dfrac{p_\infty - p_{sat}}{1/2\rho U_\infty^2}$.

one GR1, made of 115728 nodes, is used for the non-cavitating test TC1, the second (GR2) and third (GR3) grids are used for the cavitating cases and have 263832 and 502234 nodes respectively.



(a) GR1, 115728 nodes

Figure 6.8: Zoom of the grid used for the simulations of the non-cavitating water flow over a NACA0015 hydrofoil.



(a) GR2, 263832 nodes **(b)** GR3, 502234 nodes

Figure 6.9: Zoom of the different grids used for the simulations of the cavitating water flow over a NACA0015 hydrofoil.

6.2.1 Results of the non-cavitating simulations

For non-cavitating flows, large CFL values can be used for all the considered schemes. Here in particular a CFL value of 200 has been chosen for the different simulations. Since this already led to very reduced computational times on the considered grid GR1, a systematic analysis of the actual stability limit has not been carried out for these simulations.

The pressure coefficient distributions obtained on the hydrofoil in the simulations with the 1st- and 2nd-order versions of the Roe and LD-HLL schemes on GR1 are shown in Fig. 6.10, together with the experimental data. As expected, both schemes introduce a larger dissipation at 1st order than in their 2nd-order version, as can be seen in particular from the lower predicted values of the suction peak. Note how at 1st order the LD-HLL scheme appears to be more dissipative than the Roe one, as also shown by the overestimation of C_p at the stagnation point. However, at 2nd order the results obtained by the two different schemes are almost identical. As for accuracy, although there are no measurement points at the suction peak, a previously carried out potential flow simulation gave a suction peak value of approximately -1.7, in well agreement with simulations carried out with schemes of high-order of accuracy (see [14]). Thus, the predictions obtained with the 2nd-order versions of both considered schemes can definitely be considered as more accurate than those of the 1st-order ones. The small discrepancies observed near the trailing edge are due to the fact that the present simulations are inviscid and almost 2D; indeed, a similar behaviour was observed in a 2D potential flow solution and in other Euler calculations in [14].

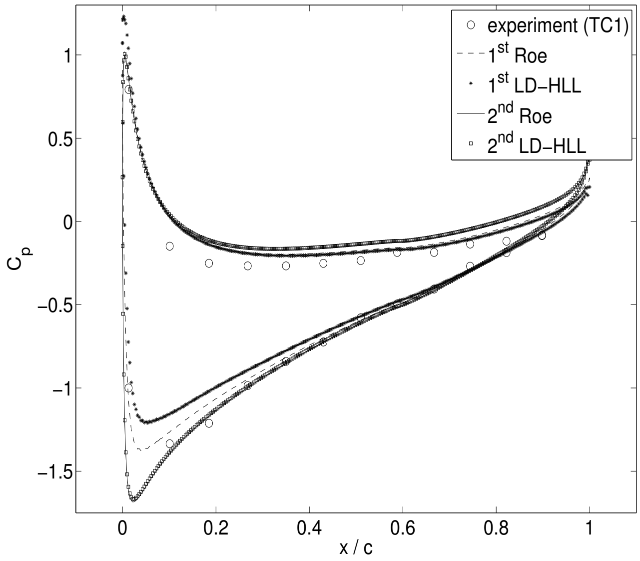


Figure 6.10: C_p distribution over the hydrofoil for the TC1 test-case (non-cavitating flow).

6.2.2 Results for cavitating conditions

For the cavitating test-cases, a CFL coefficient limitation of about 0.01 has been found for the Roe scheme and the LD-HLL one with the classical time linearisation, while the LD-HLL flux with the more complete linearisation permits to reach CFL 100, both at first order and 2nd order of accuracy. This confirms the previous observations made for the quasi-1D nozzle case (see Sec. 4.3.3) and, in particular, that the more complete time linearisation is a key point to significantly increase numerical efficiency in presence of cavitation.

Therefore, only the simulations carried out with the most efficient approach, i.e. the LD-HLL one with the more complete time linearisation, have been advanced in time sufficiently to obtain meaningful results. Note that an accurate prediction of the pressure near the leading edge and, in particular, of the suction peak is very important to correctly capture the cavitating region, which starts near the leading edge. The 1st-order version of the scheme, indeed, due to the underprediction of the suction peak, is not able to capture the cavitating region and gives no phase transition for the TC2 case. Consequently, also in accordance with the analysis carried in Sec. 6.2.1 for non-cavitating conditions, only 2nd-order accurate results are shown in the following (Figs. 6.11 to 6.14). Finally, two different grids are considered (GR2 and GR3), both more refined that the one used for the non-cavitating

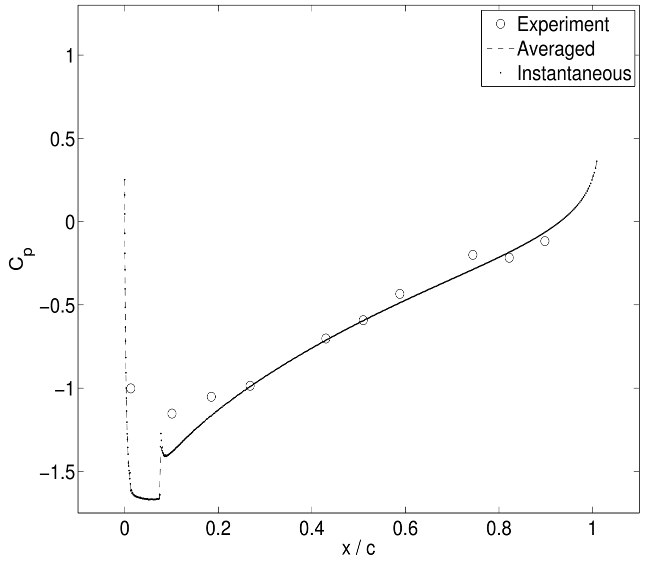


Figure 6.11: C_p distribution for the TC2 test-case on grid GR3 (cavitating flow): instantaneous and averaged values.

simulations. Fig. 6.11 shows the C_p distribution on the upper side of the hydrofoil obtained in the simulation on GR3, together with the relevant experimental data. Both instantaneous and averaged (over 5000 time steps) values are shown; practically no differences are found, thus, indicating that no significant unsteady effects are observed and that the solution is enough advanced in time. The cavitation region is clearly visible and coincides with the C_p plateau near the leading edge. This behaviour of pressure is typical of the adopted barotropic homogeneous-flow cavitation model, in which large density variations at almost constant pressure characterise the vapour regions (see, e.g., [102] or [37]). Note also the sharp pressure gradient at the vapour to liquid transition occurring at the end of the cavitation region, which also characterises this type of cavitation models. Conversely, the wiggle observed at that location in Fig. 6.11 seems to be due to numerics and is indeed largely reduced by grid refinement, as can be seen from the comparison between the solutions obtained on GR2 and GR3 in Fig. 6.12. Fig. 6.12 also shows that, except for this wiggle, grid independence of the results has been reached. To better highlight the behaviour of the solution in the cavitation region predicted by the adopted barotropic homogeneous-flow model, Fig. 6.13 shows the evolution of the C_p curve with the cavitation number value obtained on the most refined grid, i.e. GR3. As expected, a decrease of the cavitation number corresponds to a larger cavitation region. Note also that in all cases

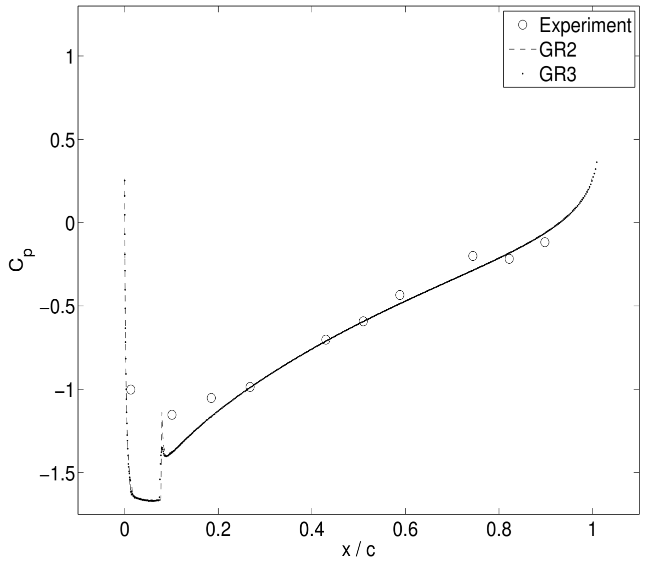


Figure 6.12: C_p distribution for the TC2 test-case (cavitating flow) for different grid resolutions.

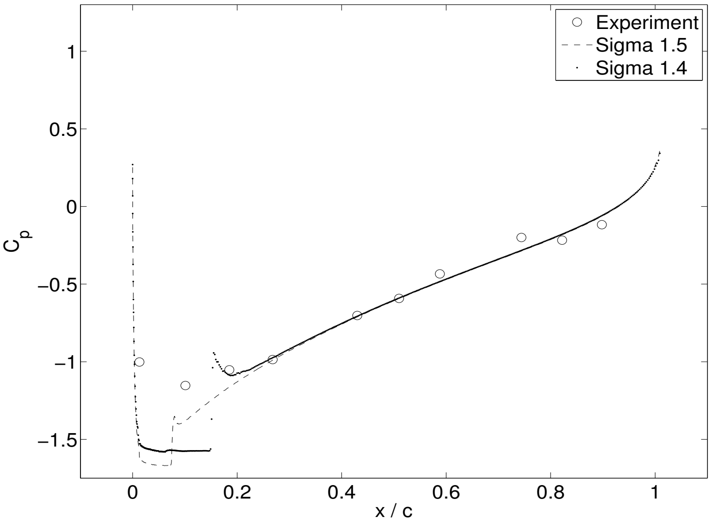


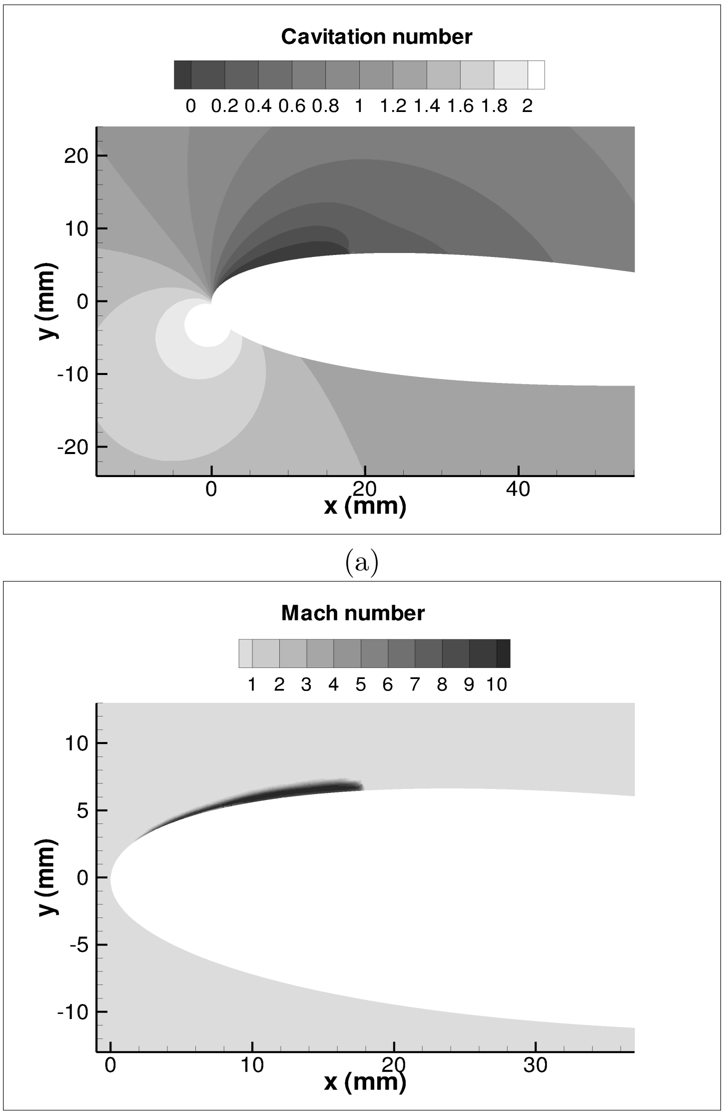
Figure 6.13: C_p distribution for cavitating flow on grid GR3: test-cases TC2 ($\sigma_\infty = 1.49$) and TC3 ($\sigma_\infty = 1.38$).

the pressure is characterised by a plateau in this region, with a value increasing as the cavitation number decreases, and that a sharp gradient is always present at the vapour to liquid transition. The agreement with the experiments for the TC2 solution (the same inlet conditions as in the experiments) appears to be reasonable, although there are no experimental points in the plateau region and some discrepancy is present at the end of the cavitating region. However, based on the previous analysis, this discrepancy seems to be due more to modelling than to numerics.

Finally, in order to give an idea of the numerical stiffness, the iso-contours of the local cavitation number and of the Mach number obtained on grid GR3 for the TC3 case are shown in Fig. 6.14. The local cavitation number is defined as

$$\sigma = \frac{p - p_{sat}}{1/2\rho U_\infty^2}$$

Hence, negative values of σ identify the cavitation region (dark grey in Fig. 6.14). Note how in the cavitation region the Mach number reaches the value of 10, while in the free-stream liquid flow it is equal to 2.4×10^{-3}. Therefore, there is an increase of about 4 orders of magnitude.



(a)

(b)

Figure 6.14: Test-case TC3: time-averaged isocontours of (a) cavitation number and (b) Mach number.

Chapter 7

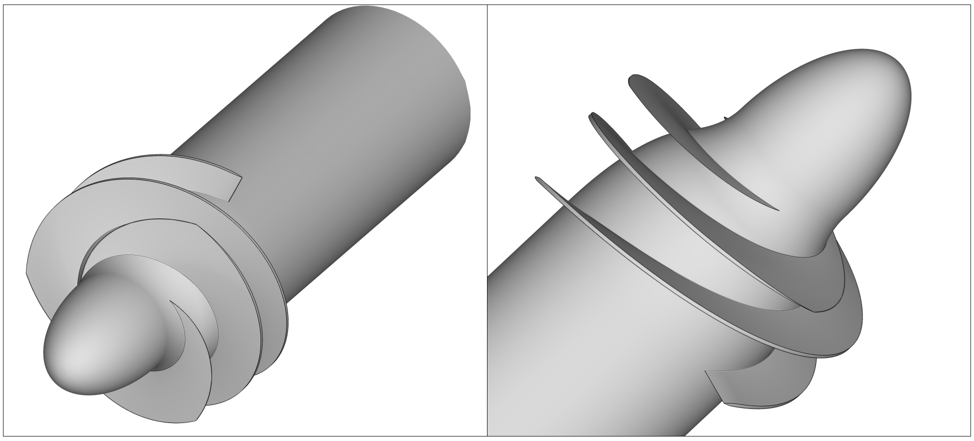
Numerical simulations in a real 3D inducer

In the previous chapter, a validation of the numerical tools introduced in Chaps. 3 and 5 has been carried out by using several test-cases for which theoretical as well as experimental data are available. Simulations of 3D flows for a real aerospace engineering application are proposed here. More specifically, the flow in a three blade axial inducer is considered. The aforementioned inducer has been developed and studied by Alta SpA in several works [31, 32, 38–40, 110–112]: it is a three blade inducer with a tip blade radius R_T of 81 mm and 2 mm radial clearance between the blade tip and the external case. Note that, as shown in Fig. 7.1, the considered geometry is really complex, and thus the simulations of the flow in a inducer is a challenging task. Experimental data are available for all the numerical simulations described in the following. In particular the pressure jump between two different stations has been measured for a wide range of working conditions: from small to large mass flow rates, non-cavitating and cavitating flows and different values of the rotational velocity ω_z of the inducer.

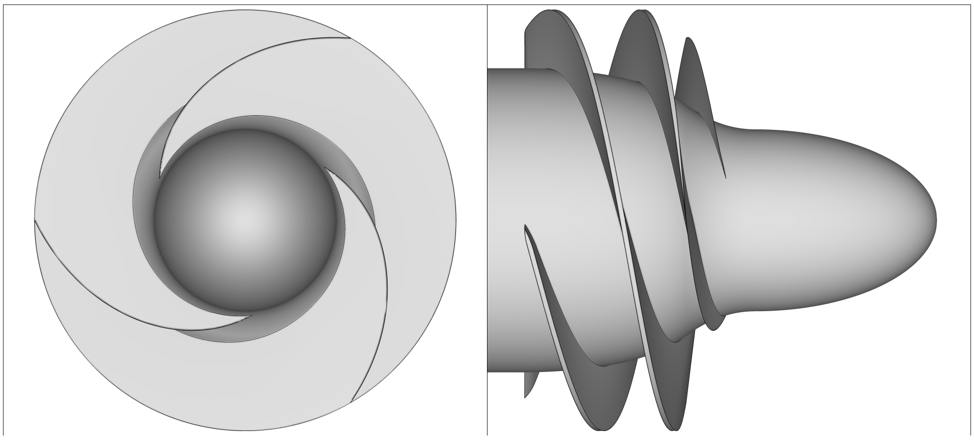
This chapter is organised as follows: first in Sec. 7.1 the main features of the computational grids developed for the numerical simulations are outlined, then in Sec. 7.2 the method of comparison between experimental and numerical data is described. In Secs. 7.3 and 7.4 the results obtained for, respectively, the non-cavitating and cavitating conditions are presented. Finally in Sec. 7.5 a preliminary analysis of the instabilities related to presence of cavitation is carried out.

7.1 Description of the computational grids

Two different computational domains and three different computational grids, shown in Tab. 7.1, have been considered. The external boundary is cylin-



(a) Global view **(b)** Detailed view



(c) Front view **(d)** Lateral view

Figure 7.1: Geometry of the inducer considered in this work.

Grid	Nodes	Tetrahedrons	z_{min}	z_{max}
G1	1926773	10572116	-319 mm	339 mm
G1L	2093770	11517393	-1190 mm	339 mm
G2	3431721	19133143	-319 mm	339 mm

Table 7.1: Grids used for the simulations of the inducer.

drical with radius R = 83 mm, while the internal boundary coincide with the inducer (see Fig. 7.1). The two domains differ by their extension in the axial direction z. For the first computational domain $z_{min} = -319$ mm and $z_{max} = 339$ mm: the inlet is placed 249 mm ahead of the inducer nose and the outlet is placed 409 mm behind. A second computational domain characterised by a larger streamwise length has also been considered, in particular $z_{min} = -1190$ mm (the inlet 1120 mm ahead the inducer nose). Two different grids have been generated to discretize the shorter domain: the basic one G1 (1926773 cells) and G2 (3431721 cells) obtained from G1 by refining the region between the blade tip and the external case. In particular, while in the grid G1 the clearance region is discretised by using 3-4 nodes, in the grid G2 9-10 nodes have been used. The larger domain has been discretized by grid G1L (2093770 cells), which coincides with G1 in the original domain.

As far as the computational costs are concerned, it is worth notice that the simulations on the grid G1 requires 32 processors and the simulation of one rotation of the inducer requires about 75 hours. Similarly the simulations on the grid G2 have been carried out on 64 processors and one non-cavitating rotation of the inducer requires about 200 hours of computational time while 1000 are required for a cavitating simulation. The difference between the computational time required by cavitating and non-cavitating simulations is the following: to increase the robustness of the numerical method, each time the numerical method detects a phase transition (i.e. a point whose density cross the saturation density) an additional DeC iteration is carried out. This improves the robustness of the numerical method, especially when near vacuum conditions are reached, at the cost of increasing the computational time.

However it is worth notice that this methodology has been proved to be robust and reliable. Indeed even the cavitating simulations in this complex geometry have been carried out using CFL numbers up to 500. Since 1000 hours are still required to simulate one rotation of the cavitating inducer, this clearly points out that less efficient numerical approaches would easily make these simulations not affordable from a practical point of view.

7.2 Comparison between numerical and experimental data

In this section the comparison between experimental and numerical data is addressed. In order to correctly perform a comparison with the experimental data, it is useful to define a time-averaged numerical solution, that is:

$$\widetilde{W}_i = \frac{\sum\limits_{k=1}^{n} W_i^k \Delta^k t}{\sum\limits_{k=1}^{n} \Delta^k t} \tag{7.1}$$

in which W_i^k is the instantaneous solution at time t^k. This averaging procedure is consistent with experimental data which is collected with measures that last 5 seconds, a time far greater than the period required by the inducer to make a complete revolution.
However, caution must be used when dealing with (7.1): the barotropic relation $p = p(\rho)$ is not linear and, as a consequence it is possible to have two different definitions for the averaged pressure, the time averaged pressure and the pressure corresponding to the time averaged density which are in general different, that is:

$$\tilde{p}_i \neq p(\tilde{\rho}_i)$$

As long as we deal with nearly incompressible flow the difference between these two averaged pressures is negligible but if we are interested in cavitating flows (i.e. highly compressible flows) the difference can be remarkable. Since the experimental data is collected using pressure probes we considered that to compare correctly the results we need to use the time-averaged pressure $\tilde{p}$ computed as in (7.1) and, in the following, unless explicitly stated, this should be the meaning of the term averaged pressure.

Another issue is the local character of the experimental data. Two pressure probes are positioned at a specific location, with a well defined axial coordinate and angular position. The first probe is located at 82.2 mm ahead the inducer nose, that is $z = -152.2$ mm, while the latter one 259.8 mm behind, that is $z = 189.8$ mm.

To compare properly our results we need to collect numerical data at the same locations of the experiments. The difficulty arises from the fact that the numerical simulation is carried out in a rotating frame. That is, a fixed point in space is a rotating point in the reference frame of the numerical simulation and, conversely, any fixed point in the computational domain is indeed a rotating one. However, since the experimental measures last 5 seconds, a time larger than the period of rotation of the inducer, it is reasonable to consider that any pressure variation along the circumference is averaged in the measurement process. Hence we used a double average: we computed the average over the outer circumference of radius R = 83 mm of the time-averaged pressure, that is we compare the experimental data with a time and circumferential averaged pressure field.

This process of double average of the pressure can introduce a discrepancy between numerical and experimental data, in particular if the flow field is not stationary and the pressure is not constant in the azimuthal direction. In order to make an estimation of the error that could be introduced by this technique we compute also the local maximum and minimum of the time-averaged pressure.

Finally we must take into account that, due to the use of unstructured grids, we have few nodes with the same axial coordinate of the probes. So, to avoid averaging only with few nodes, we select all the nodes within a specific tolerance from the geometric axial coordinate. That is, let be z_p the axial coordinate of the probe, we select all the nodes i with axial coordinate z_i which satisfies:

$$|z_i - z_p| < \delta \tag{7.2}$$

and in particular we choose:

$$\delta = 2.5 \text{ mm}$$

which corresponds to the local grid spacing in the z direction. To conclude, the averaged variable used in the comparison with the experimental data is the time-averaged pressure averaged over the outer circumference at a specific location, that is:

$$\widehat{\widetilde{P}}_z = \frac{\sum\limits_{i \in \delta_{z_p}} \widetilde{P}_i S_i}{\sum\limits_{i \in \delta_{z_p}} S_i} \tag{7.3}$$

while the estimation of the error is given by

$$\begin{aligned} \widetilde{P}_{z,min} &= \min_{i \in \delta_{z_p}} \widetilde{P}_i \\ \widetilde{P}_{z,max} &= \max_{i \in \delta_{z_p}} \widetilde{P}_i \end{aligned} \tag{7.4}$$

In (7.3) and (7.4) S_i is the surface of the i^{th} cell centred at node i and δ_{z_p} is the set containing all the cells which satisfy (7.2).

The results are presented in terms of the mean adimensionalised pressure jump $\varPsi$ as a function of the adimensionalised discharge $\varPhi$:

$$\varPsi = \frac{\Delta P}{\rho_L \omega_z^2 R_T^2} \qquad \varPhi = \frac{Q}{\pi R_T^2 \omega_z R_T} \tag{7.5}$$

where Q is the discharge, $R_T = 81$ mm is the radius of the tip of the blade, ρ_L the density of the liquid and ω_z is the angular velocity. In (7.5), ΔP is the difference between the pressure measured by the downstream experimental probe and the upstream one. This is an indicator of the inducer efficiency since the aim of inducers is to increase the inlet pressure of the pumps they are attached to, thus reducing or completely avoid cavitation. These adimensionalised variables are widely applied to characterise the performance of the

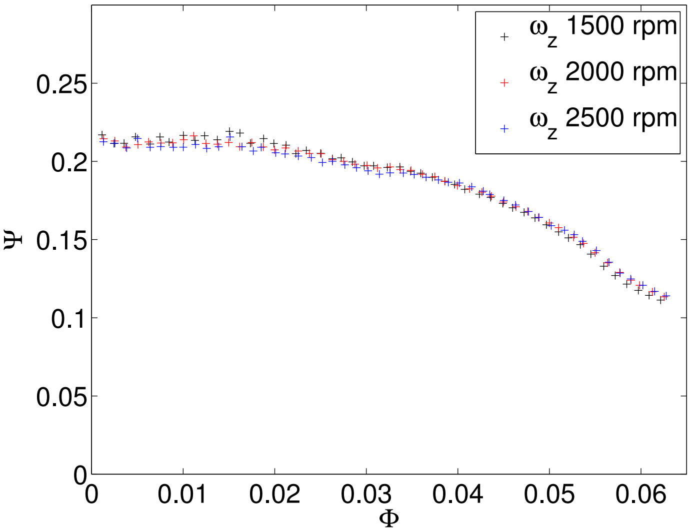


Figure 7.2: Experimental data given by Alta Spa for different values of the angular velocity.

inducer because in the plane $\Phi - \Psi$ the performance curves for different angular velocities collapse on a single curve as shown in Fig. 7.2. Finally, given the experimental value of the pressure jump Ψ_{exp}, it is possible to define the error of the simulation with respect the experimental data as:

$$\mathrm{Err}(\%) = \frac{\Psi_{ave} - \Psi_{exp}}{\Psi_{exp}} \cdot 100 \tag{7.6}$$

where Ψ_{ave} is the numerical pressure jump averaged .

7.3 Non-cavitating simulations

The working conditions considered for the non-cavitating simulations are shown in Tab. 7.2. All the considered cases use the same outlet pressure of the flow, p_{out}, and the same rotational velocity of 1500 rpm (revolution per minute). As previously pointed out, in the $\Phi - \Psi$ plane the experimental curves of the performances of the inducer are practically independent from the rotational velocity ω_z [110] (see also Fig. 7.2). As a consequence, validating the numerical tool for a specific rotational velocity and different flow rates should validate the proposed numerical tool for a generic rotational velocity. In particular, three different flow rates have been considered and the corresponding simulations, G1-Ind1, G1-Ind2, G1-Ind3 have been run on the grid G1. For the small and intermediate flow rates, Ind3 and Ind2, respectively,

Benchmark	Ind1	Ind2	Ind3
Φ	0.0584	0.0391	0.0185
ω_z (rpm)	1500	1500	1500
p_{out} (Pa)	125000	125000	125000
T (°C)	25°	25°	25°

Table 7.2: Conditions of the numerical simulations and of the experiments.

also the more refined grid G2 has been used for the simulations G2-Ind3 and G2-Ind2. Finally the influence of the upstream boundary conditions has been tested for the small flow rate Ind3 by using the extended computational domain with the grid G1L in the simulations G1L-Ind3 and G1L-Ind3-T. Note that all the aforementioned numerical simulations are laminar one: the only turbulent simulation carried out for this test-case is the last one, G1L-Ind3-T.

Fig. 7.3 and Tab. 7.3 show the numerical results obtained for the considered test-cases.

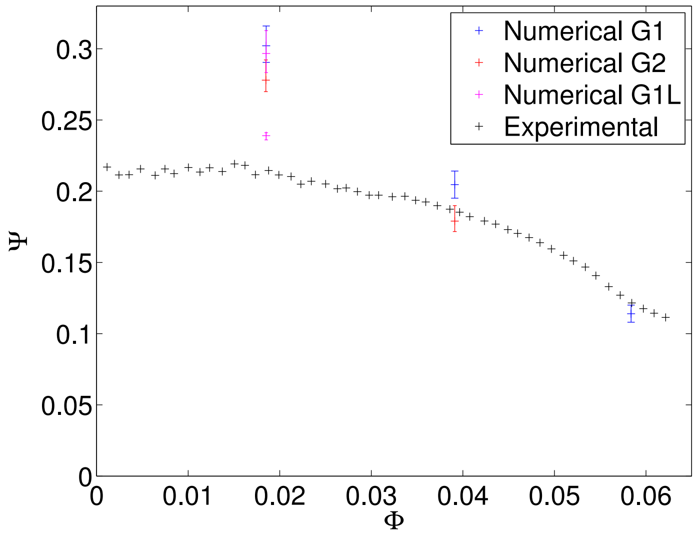


Figure 7.3: Comparison between laminar numerical simulations and experimental data.

From Fig. 7.3 it clearly appears that the lower is the discharge Φ the worse are the results. In particular the laminar simulations for the benchmark Ind3, G1-Ind3, G2-Ind3 and G1L-Ind3, give results that clearly overestimate the pressure jump while the turbulent simulation G1L-Ind3-T, the purple point closer to the experimental data in Fig. 7.3, greatly improves the results (see also Tab. 7.3). A detailed analysis of the results and, in particular of this

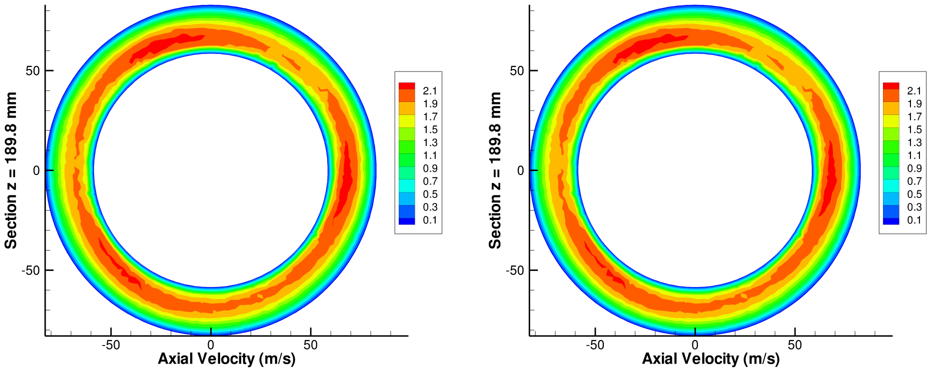
	Ψ_{exp}	Numerical Ψ_{ave}	Ψ_{min}	Ψ_{max}	Error%
G1-Ind1	0.122	0.114	0.108	0.120	−6.6%
G1-Ind2	0.186	0.204	0.193	0.215	+9.7%
G2-Ind2	0.186	0.179	0.171	0.192	−3.8%
G1-Ind3	0.214	0.302	0.290	0.316	+41%
G2-Ind3	0.214	0.278	0.270	0.292	+30%
G1L-Ind3	0.214	0.297	0.283	0.313	+39%
G1L-Ind3-T	0.214	0.239	0.236	0.241	+12%

Table 7.3: Pressure jump in non-cavitating conditions

point, is proposed in Secs. 7.3.1-7.3.3.

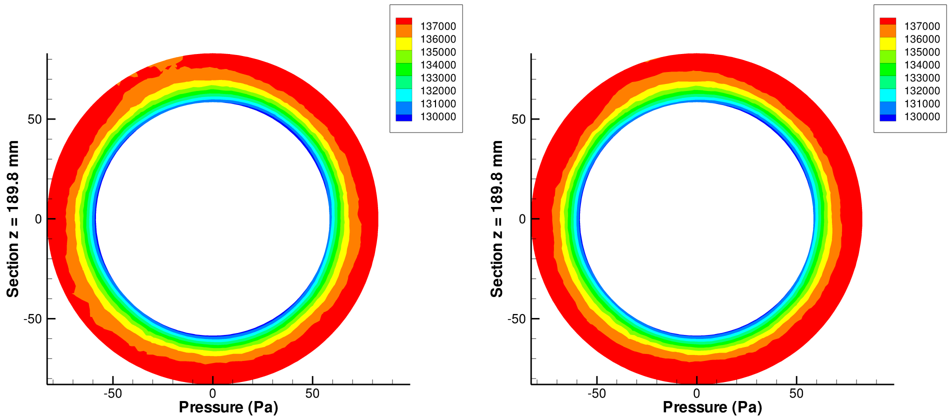
7.3.1 High-discharge conditions, Ind1

Considering the highest discharge value, corresponding to the case Ind1, already with the coarse grid G1 the pressure jump is underestimated by only 6.6%. The global features of the flow field at specific locations are illustrated in figures from 7.4 to 7.6. The aforementioned figures show a comparison between the instantaneous solution and the one averaged over one complete revolution of the inducer. The discrepancies between the two fields are hardly distinguishable, thus giving an additional proof that a steady condition has been reached. Furthermore, it is interesting to point out that Fig. 7.4 shows



(a) Instantaneous field **(b)** Averaged field

Figure 7.4: Simulation G1-Ind1, cross section of the axial velocity field at z = 189.8 mm, comparison between instantaneous and averaged solutions.



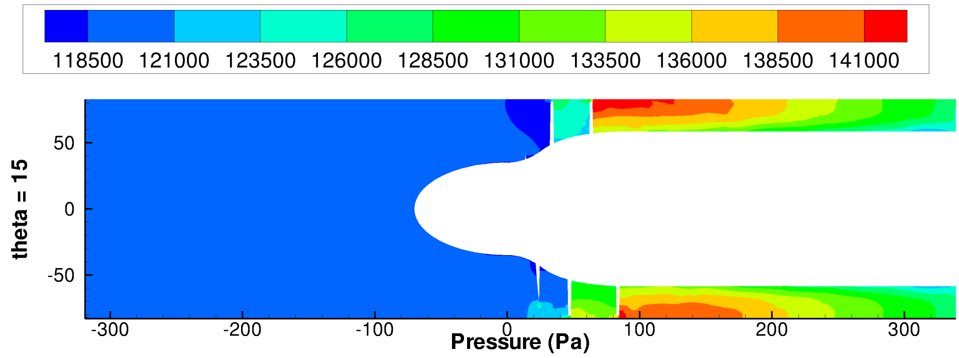
(a) Instantaneous field **(b)** Averaged field

Figure 7.5: Simulation G1-Ind1, cross section of the pressure field at z = 189.8 mm, comparison between instantaneous and averaged solutions.

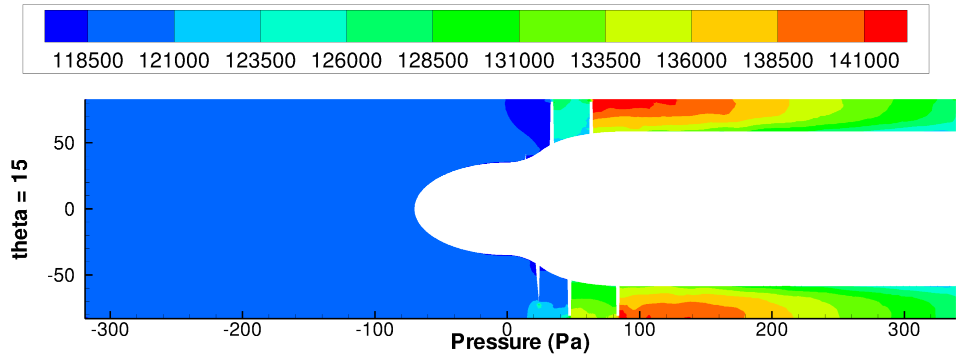
a nearly 120° symmetry of the flow field which is expectable since a 3 blade inducer is considered in these simulations. Finally, Fig. 7.6 shows an axial cross section of the pressure field: it is possible to observe how the maximum value of the pressure is located just downstream of the last inducer blade and, similarly the minimum pressure value is located upstream the first blade of the inducer.

7.3.2 Intermediate-discharge conditions, Ind2

For the intermediate value of the discharge Ind2, results qualitatively similar to the case Ind1 have been obtained. On the coarse grid G1 the pressure jump is over estimated by 9.7%: by considering the more refined grid G2, an absolute value of the error of 3.8% is obtained but in this case the pressure jump is underestimated. The numerical results for the second simulation with the more refined grid are shown in figures from 7.7 to 7.9. Figs. 7.7 and 7.8 show a comparison between the averaged solution computed on the grids G1 and G2. In particular, considering the axial velocity field shown in Fig. 7.7, the improved resolution of the refined grid clearly appears, even if, from a qualitative point of view the two grids computes similar results. The results for the pressure fields, shown in Fig. 7.8 confirm the benefits of grid refinement: while in all cases the pressure jump, as for the benchmark Ind1, takes place in the inducer region, its magnitude is different for the two grids. This is clearly understood by studying Fig. 7.9: the pressure upstream the inducer is the same for the two computational grids. Nevertheless the



(a) Instantaneous field



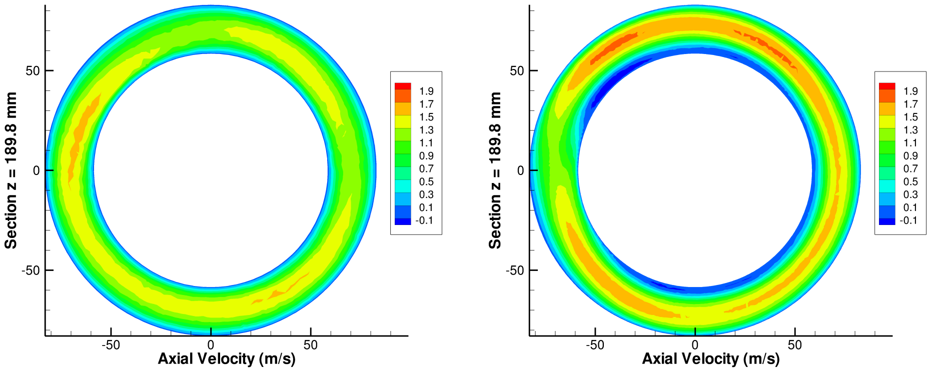
(b) Averaged field

Figure 7.6: Simulation G1-Ind1, cross section of the pressure field at $\theta = 15°$, comparison between instantaneous and averaged solutions.

pressure jump downstream the inducer is clearly higher for the grid G1, and this is confirmed by Fig. 7.8: this is consistent with the overestimation of the pressure jump computed with the first grid G1.

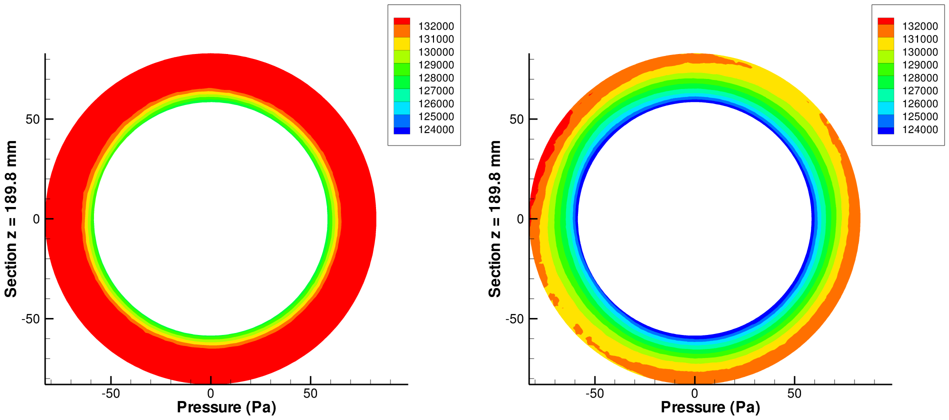
7.3.3 Small-discharge conditions Ind3

The lowest value of the discharge is qualitatively different from the previous ones. The results for both the high and intermediate flow rates, Ind1 and Ind2, can be considered in good agreement with the experimental data. Conversely, for the low discharge case, Ind3, the simulations with the grid G1 and G2 greatly overestimate the pressure jump by 41% and 30%, respectively. The Ind3 test-case is characterised by a strong backflow between the inducer blades and the external case which is the largest among the three considered cases, as it is clearly visible from Fig. 7.10.



(a) G1 **(b)** G2

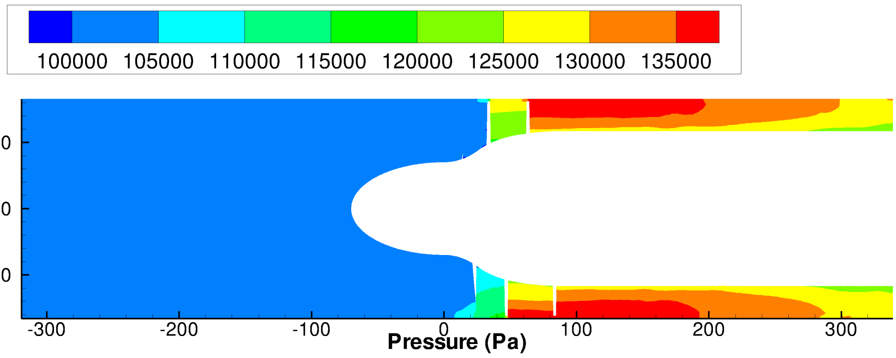
Figure 7.7: Benchmark Ind2, cross section of the axial velocity field at z = $189.8\ \mathrm{mm}$, comparison between grids G1 and G2.



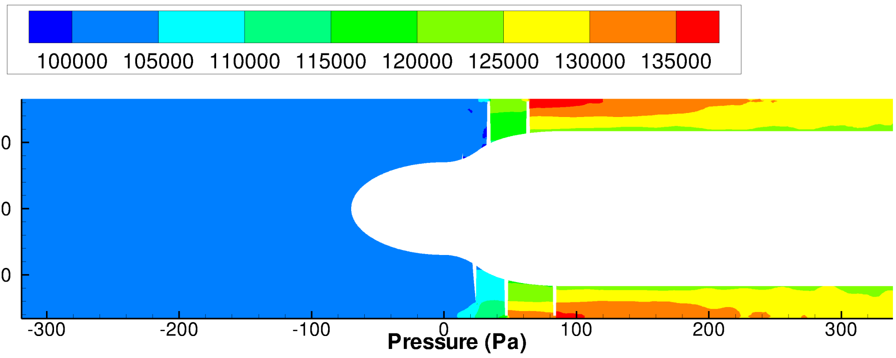
(a) G1 **(b)** G2

Figure 7.8: Benchmark Ind2, cross section of the pressure field at z = $189.8\ \mathrm{mm}$, comparison between grids G1 and G2.

From the experimental data it follows that the smaller is the mass flow rate, the larger is the pressure jump. It is natural to expect that the larger is the pressure gradient, the stronger will be the force pushing the flow backward through the tip clearance between the blades and the inducer case. This backward flow, with different pressure and velocity, interacts with the incoming flow. Thus, the correct resolution of the backward flow is of crucial importance for the determination of the performance of an inducer. A



(a) G1

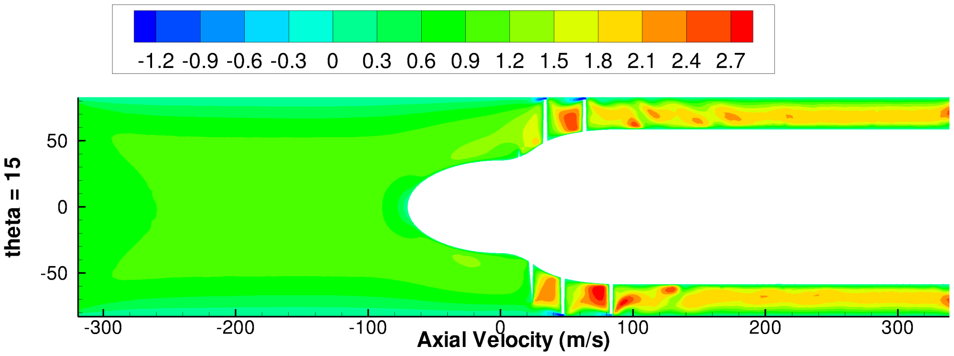


(b) G2

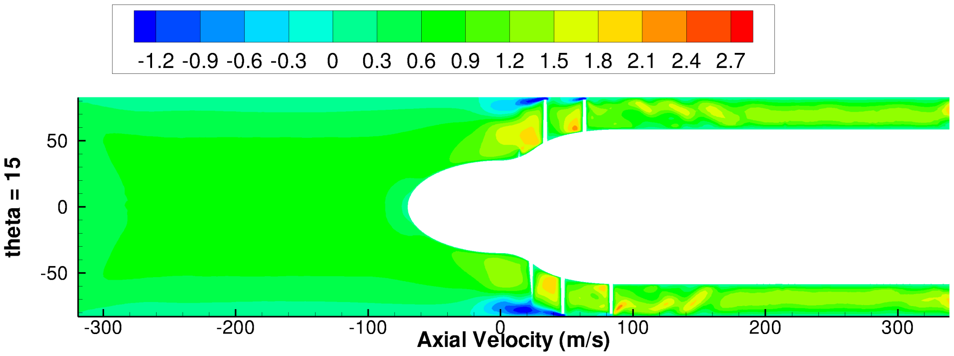
Figure 7.9: Benchmark Ind2, cross section of the pressure field at $\theta = 15°$, comparison between grids G1 and G2.

possible source of error comes from grid resolution: if the mesh size in the tip clearance is too coarse this could reduce the "effective" clearance and, as a consequence, increases the pressure jump, as shown in Fig. 7.11. However, grid refinement does not seem to be an effective solution. Indeed, by switching from grid G1 to G2 a remarkable reduction of the error (from 41% to 30%) has been achieved, but the error is still much larger than the one found in the previous cases. Thus, since grid G2 is already composed by more than 3 millions nodes, a further grid refinement does not seem a viable solution. Thus, for this test-case we investigated the influence on the numerical solution of two additional possible sources of error. The first one was a possibly too small distance of the inlet from the inducer nose and the second one was the neglection of the turbulence effects.

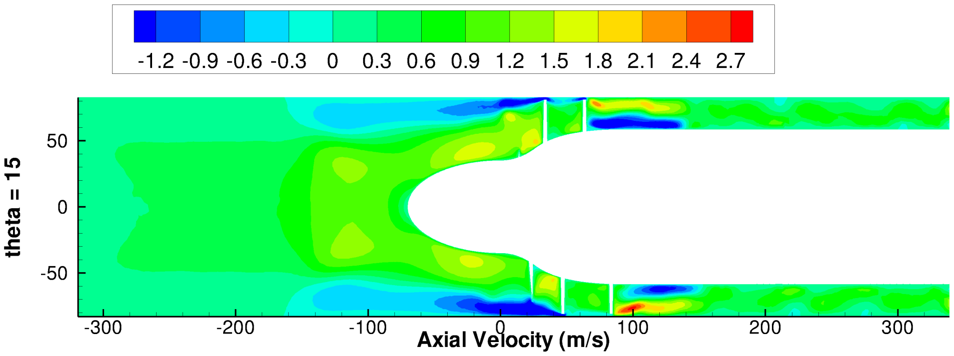
As shown in Fig. 7.10, for the benchmark Ind3 the backflow extension



(a) Benchmark Ind1

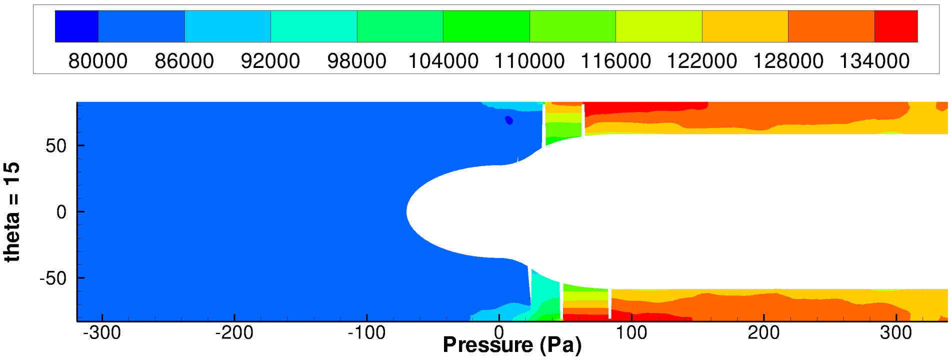


(b) Benchmark Ind2

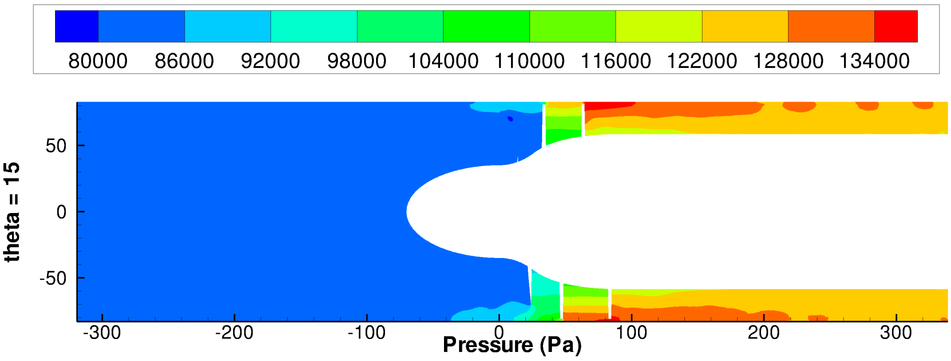


(c) Benchmark Ind3

Figure 7.10: Comparison of the averaged flow field for the three non cavitating benchmarks at $\theta = 15^{\circ}$, grid G1.



(a) G1

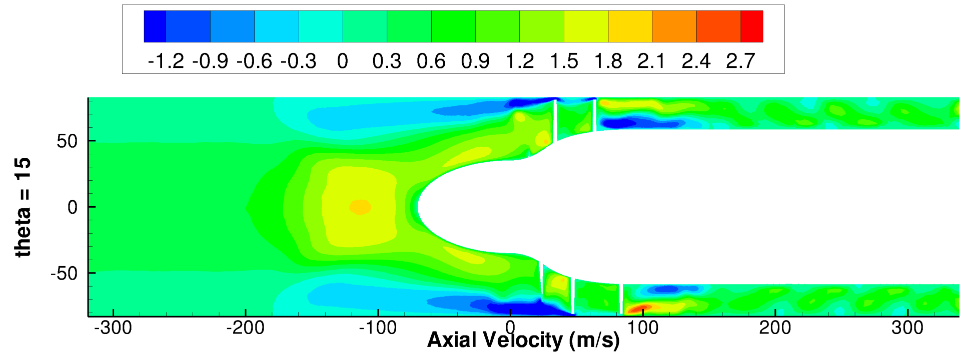


(b) G2

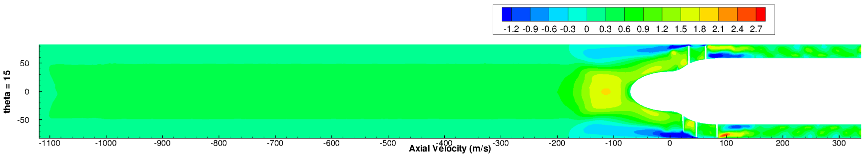
Figure 7.11: Benchmark Ind3, cross section of the pressure field at $\theta = 15°$, comparison between grids G1 and G2.

upstream the inducer blades is significant and thus, it is possible that the solution is influenced by the proximity of the inlet boundary. However, Tab. 7.3 shows that the increase of the streamwise dimension of the computational domain has a small effect in reducing the discrepancy with respect to the experimental data, from 41% to 39%, noticeably smaller than the reduction obtained using grid G2 instead of G1. Furthermore, Fig. 7.12 clearly shows that in the simulation with the larger domain G1L-Ind3 the backflow is not affected by the inlet conditions and, as consequence, no spurious effects on the solution are present (compare also Fig. 7.12a with 7.10c). Clearly, this error could be further reduced by carrying out a simulation on the refined grid GR2 with an extended width of the domain. However it seems unlike that this could lead to a satisfactory accuracy of the results.

Conversely, the results of the RANS simulation G1L-Ind3-T, which used



(a) Detailed view



(b) Global view

Figure 7.12: Cross section of the flow field at $\theta = 15°$, averaged axial velocity, simulation G1L-Ind3

the same numerical set-up and the same grid as for the G1L-Ind3, show that in this case turbulence is a key-issue. Indeed, the error falls down to 12%, less than the error obtained with the refined grid G2 in laminar conditions. It is therefore clear that the introduction of turbulence significantly improves the agreement with experiments. As expected the effects of turbulence are particularly important near the gap between the blades and the external case, as it is shown for instance by the averaged isocontours of k in Fig. 7.13. This strongly affects the backflow (compare Figs. 7.14 and 7.12) and, thus, the pressure jump. This also explains why for larger flow rates, for which the backflow is less important, the effects of turbulence are not so strong and a good agreement with experimental data can be obtained also in laminar simulations.

7.4 Cavitating simulations

In this section the results of the cavitating simulations are presented. The characteristics of the chosen benchmarks are described in Tab. 7.4, all the test-cases share the same discharge value $\varPhi = 0.0531$ and the same rotational velocity $\omega_z = 3000$ rpm. They differ in the outlet pressure or, equivalently,

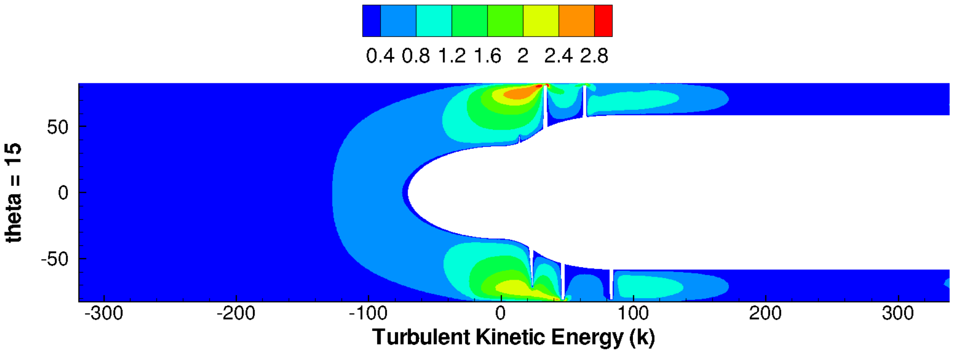


Figure 7.13: Cross section of the averaged k field at $\theta = 15°$, simulation G1L-Ind3-T

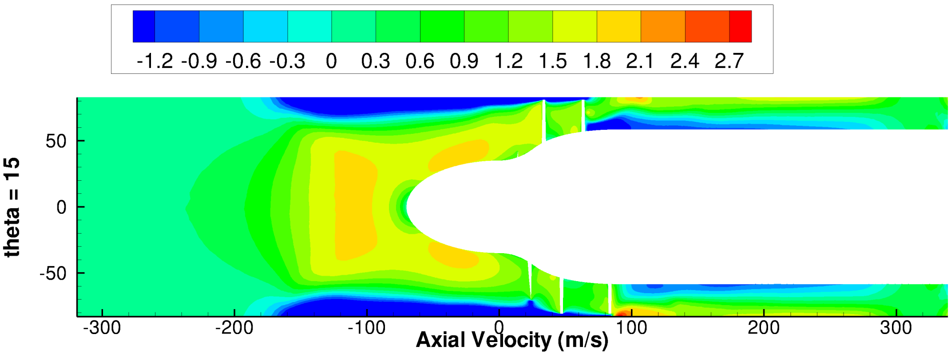


Figure 7.14: Cross section of the flow field at $\theta = 15°$, averaged axial velocity, simulation G1L-Ind3-T

Benchmark	Ind4	Ind5	Ind6
Φ	0.0531	0.0531	0.0531
ω_z (rpm)	3000	3000	3000
p_{out} (Pa)	60000	85000	82500
T (°C)	16.8°	16.8°	16.8°
σ	0.056	0.084	0.077

Table 7.4: Conditions of the numerical simulations and of the experiments.

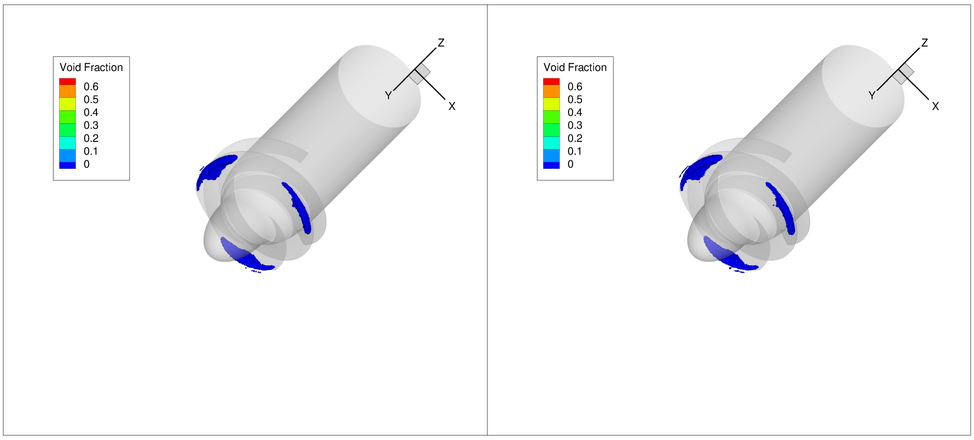
in the cavitation number σ of the test which, for the inducer simulations, is defined as:

$$\sigma = \frac{p - p_{Lsat}}{0.5 \rho \omega_z^2 R_T^2} \tag{7.7}$$

The choice of the benchmarks has been done by trying to reproduce different points on the experimental curves of the cavitating performance of the inducer. While in the non-cavitating case different flow rates at constant ω_z have been considered, in the cavitating case the performance at different σ and fixed ω_z and Φ have been analysed. Note that, since the mass flow rate for the cavitating cases is large enough to prevent the issues related to the backflow previously described, only laminar simulations have been considered. Tab. 7.5 reports the results obtained by the three cavitating simulations. It is clear that the first grid G1 is not enough refined to correctly describe cavitation for this case since the pressure jump is greatly overestimated by 36%. For the benchmark Ind4 the error is related to the underestimation of the cavitating region: the experimental data for $\sigma = 0.056$ show a large cavitating zone which completely floods the inducer and, consequently, the performance of the inducer is significantly deteriorated. Instead, in the simulation with grid G1 the cavitating region is small as shown in Fig. 7.15. Furthermore it is clearly noticeable from Fig. 7.15 that the cavitating region is nearly sta-

	Ψ_{exp}	Numerical Ψ_{ave}	Ψ_{min}	Ψ_{max}	Error%
G1-Ind4	0.105	0.143	0.135	0.151	+36%
G2-Ind5	0.143	0.130	0.123	0.139	−8.9%
G2-Ind6	0.137	0.130	0.123	0.139	−5.0%

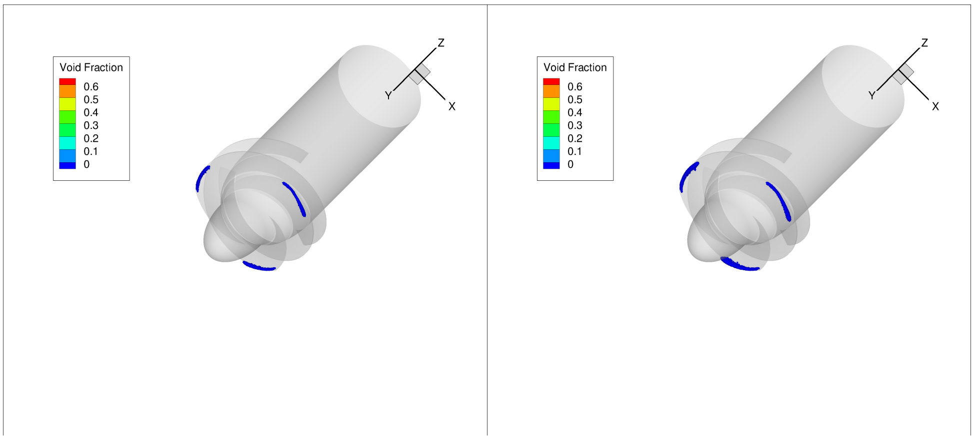
Table 7.5: Numerical results for the cavitating simulations



(a) Instantaneous field **(b)** Averaged field

Figure 7.15: Pressure isosurface for the cavitating simulation G1-Ind4: cavitating region.

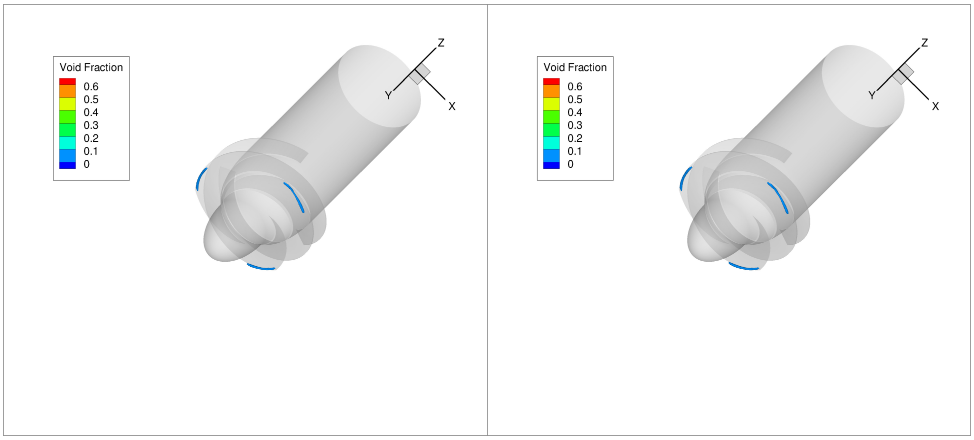
tionary, which is clearly in contrast with the experimental results. Thus, the "numerical" performance of the inducer is similar to the non cavitating case, as it can be also shown by comparing the numerical pressure jump obtained for the benchmark Ind4 and Ind1. Also for cavitating cases grid refinement has clearly a beneficial effect: in the simulations carried out on the more refined grid GR2 for, respectively, $\sigma = 0.084$ and $\sigma = 0.077$ (G2-Ind5 and G2-Ind6) cavitating zones can be observed. Conversely, for these values on grid G1 no cavitation has been found. The numerical prediction of the pressure jump is in this case in rather good agreement with the experiments (see Tab. 7.5), showing once again the positive effects of grid refinement. Fig. 7.16 finally shows that the extension of the cavitating region is slightly smaller than the one obtained with grid G1-Ind4. However, since the cavitation number is increased this clearly shows an improvement in the resolution of the cavitating region. Note in particular that the cavitating simulations on grid



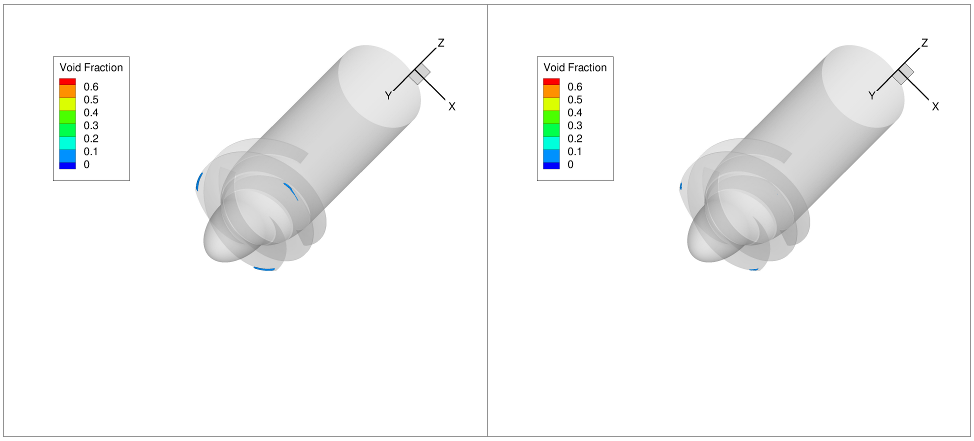
(a) Benchmark Ind5 **(b)** Benchmark Ind6

Figure 7.16: Pressure isosurface for the cavitating simulation on the grid G2: cavitating region.

G2 are not stationary. In order to show this characteristic of the flow, we visualise the cavitating regions on the blades, identified by an isosurface of the void fraction corresponding to the value of 0.05, at different instants during one inducer rotation. Fig. 7.17 and Fig. 7.18 shows these cavitating regions for 8 time instants uniformly distributed over one complete rotation of the inducer. It clearly appears that the cavitating regions are highly unsteady and progressively shrink in time. In the second half of the rotation the opposite occurs, i.e. the cavitating region sizes progressively increases to reach roughly the same dimensions as at the beginning of the rotation. Moreover, the cavitating regions over the three blades are slightly asymmetric and this is particularly evident when the regions reduce to their smallest dimensions; see for instance Fig. 7.17d in which the cavitating region has disappeared on one blade while this is not the case for the other two.

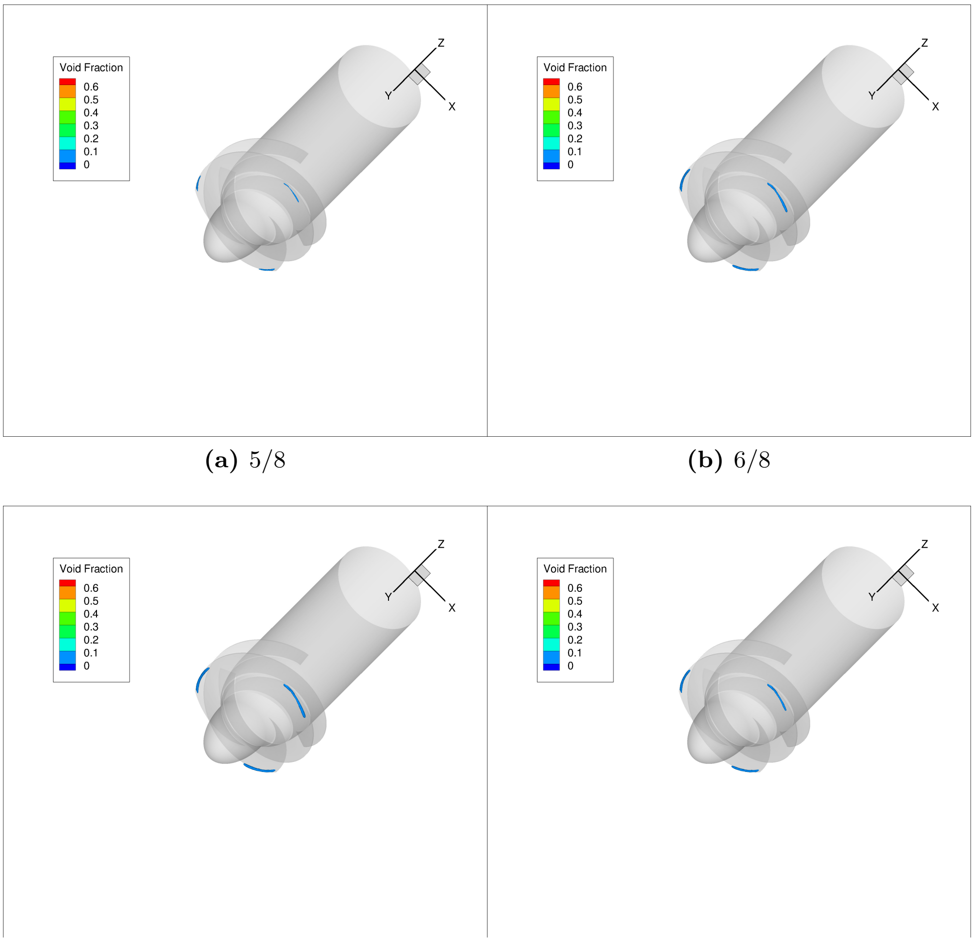


(a) 1/8 **(b)** 2/8



(c) 3/8 **(d)** 4/8

Figure 7.17: Isosurfaces of the void fraction for the cavitating simulation G2-Ind5 at different instants during the first half of the inducer rotation.



(a) 5/8 **(b)** 6/8

(c) 7/8 **(d)** 8/8

Figure 7.18: Isosurfaces of the void fraction for the cavitating simulation G2-Ind5 at different instants during the second half of the inducer rotation.

7.5 Preliminary analysis of instabilities related to cavitation

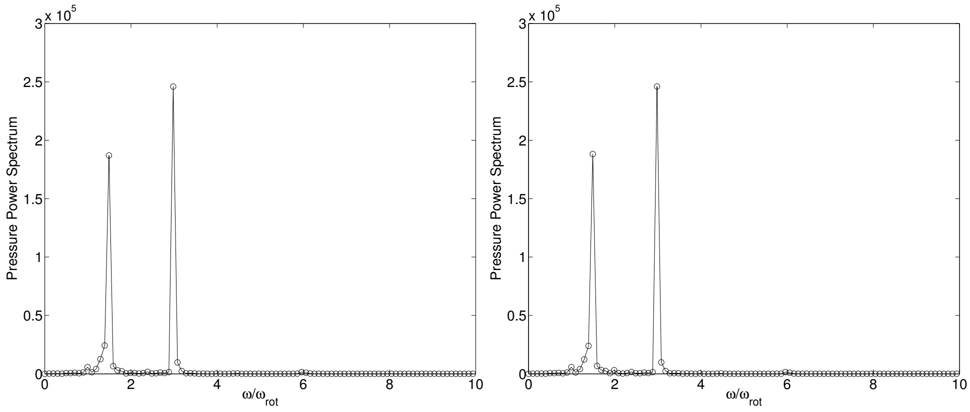
In this section a preliminary analysis aimed at investigating whether flow instabilities related to cavitation can be observed in the present numerical simulations, is presented. In particular the simulation G2-Ind6 carried out on the most refined grid is considered.

The aforementioned instabilities have been detected in the experiments carried out by Alta Spa in [110]. The asymmetric cavitation (also called *attached uneven cavitation* or *synchronous rotating cavitation*) is first considered. It is generated by uneven cavitation length on the inducer blades, with one blade cavitating more (or less) than the others and can therefore be considered as a 1-cell rotating instability with the same frequency as the inducer rotational speed (i.e. synchronous). In the experiments this behaviour may be caused by the fact that the inducer blades are not perfectly equal, due to uncontrolled small manufacturing defects, but also to uncontrolled asymmetric disturbances in the flow. In numerical simulations, the blades are exactly equal to the nominal ones, but uncontrolled small asymmetric disturbances might be introduced, for instance by the fact that the computational grid is not symmetric leading to an asymmetric response of the blades of the inducer, as described in Sec. 7.4 and shown in Fig. 7.17d.

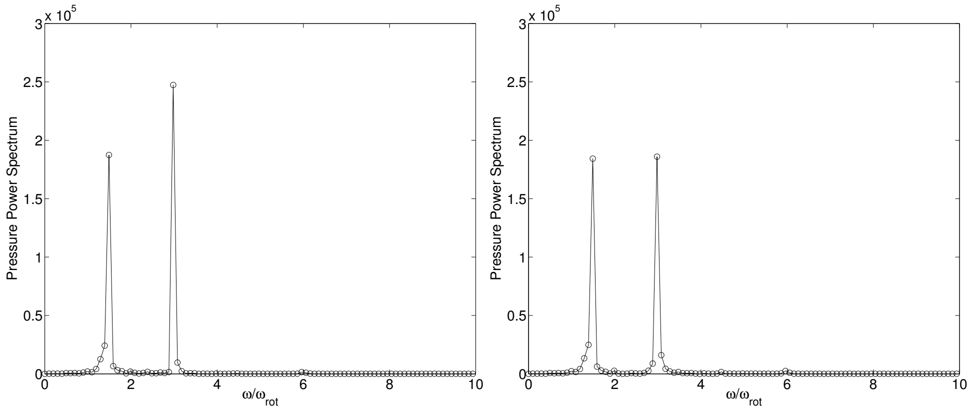
The detection of the instabilities from the experimental data was mainly based on the spectral analysis of the time pressure signals recorded by the inlet probe. On the basis of the experimental analysis (see [32, 110]), for the considered values of the discharge and of the cavitation number, the instability classified as surge may be expected, which is characterised by a global axial oscillation of the flow at a frequency of about 0.2 of the rotational frequency, ω_z. Additional instabilities, such as BVI1 (at approximately 0.7-0.8 ω_z), 3Ω-BVI1 (about 2.3 ω_z), BVI2 (about 3.3 ω_z), 2Ω+BVI2 (approximately 5.3 ω_z) have been found in experiments for the considered value of discharge but at lower values of the cavitation number [110]. They are related to different instabilities of backflow vortexes originating from the blades. Consistently with the experimental findings, we did not detect any significant instability of backflow vortexes.

In order to carry out a spectral analysis similar to the experimental one, time pressure signals have been recorded on the case at $z = -56.5$ mm, -28.8 mm, -7.2 mm, 189.8 mm. In particular $z = -7.2$ mm corresponds to the location of the experimental probe. For each z location we have extracted the signals at six different locations, namely $\theta = 0°$, $90°$, $120°$, $180°$, $240°$, $270°$.

The signals have been recorded over 13 inducer complete rotations, with 8000 time steps per rotation and a total computational cost of roughly 13000 hours of computational time. The signals have been then divided in four (intersecting) groups, each composed by ten rotations. As it is clear from figure 7.19, no significant differences have been observed in the spectra computed at different azimuthal locations. Two peaks are evident, the first at 1.5 ω_z and the second at 3 ω_z. The latter one was expected since it is related to the passage of the three blades. The first one might be simply a sub-harmonic or could be related to some phenomenon occurring at the passage of two blades



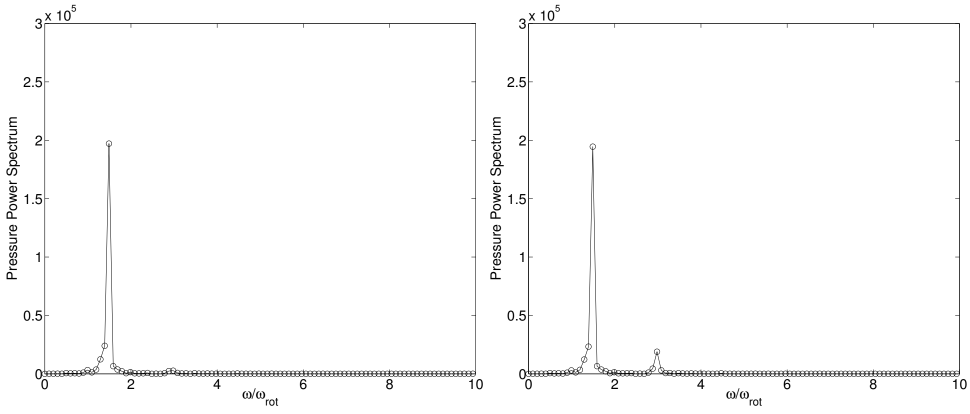
(a) $\theta = 0°$ **(b)** $\theta = 120°$



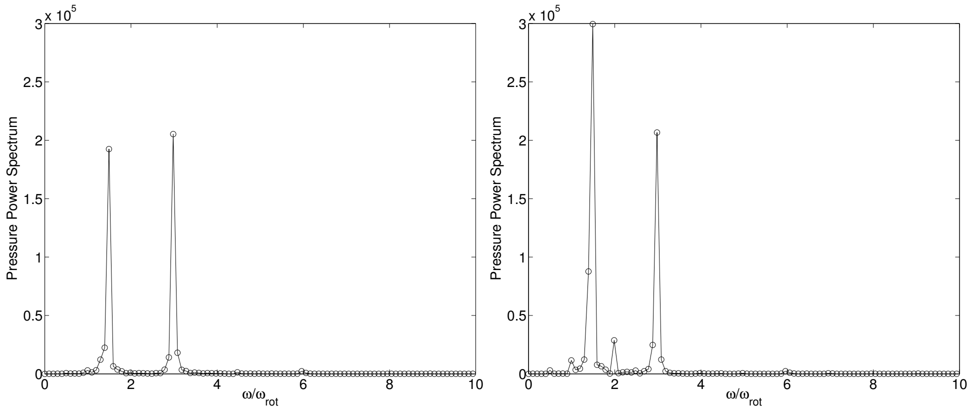
(c) $\theta = 240°$ **(d)** $\theta = 270°$

Figure 7.19: Comparison of the Pressure Power Spectrum at four different locations, z $= -7.2$ mm.

over three: it is interesting to notice that, as shown in Fig. 7.20, upstream the inducer nose the amplitude of the peak at 1.5 ω_z is roughly independent from the distance from the inducer blades. Conversely the magnitude of 3 ω_z peak increases approaching the inducer blades. This could suggest the presence of an axial instability which, for these working conditions, disagrees with the experimental results. It is also possible that the 1.5 ω_z peak is a consequence of a transient which has not ended as shown in 7.21. Indeed Fig. 7.21 shows that the power spectrum is still not time independent and, in particular, the magnitude of the 3 ω_z peak is increasing. Thus, it could be possible that at the very end of the transient only a greater 3 ω_z peak will be present at the expenses of the 1.5 ω_z. However, verifying this hypothe-



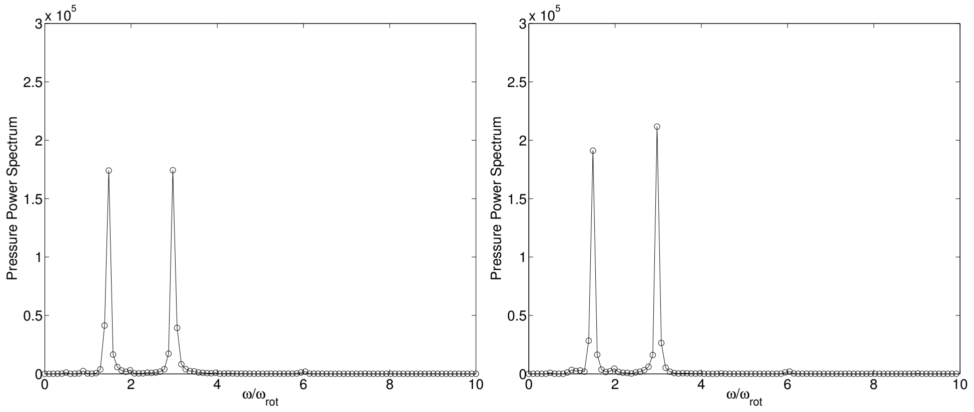
(a) z = -56.5 mm **(b)** z = -28.8 mm



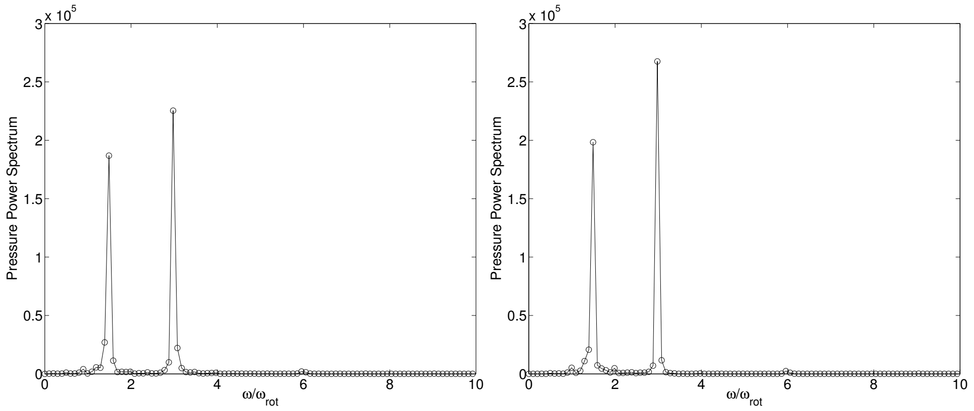
(c) z = -7.2 mm **(d)** z = 189.8 mm

Figure 7.20: Comparison of the Pressure Power Spectrum at different axial coordinates, $\theta = 180°$.

sis would require an huge amount of additional computational time which is beyond the scope of the present work. Finally, as done in the experiments, the frequencies which are entire multiples of ω_z have been filtered out; we use a notch filter having a band of $\pm 0.1\ \omega_z$. The filtered spectrum for the location $\theta = 180°$ is shown in Fig. 7.22: the range of the vertical axis has been reduced in order to also visualise local peaks characterised by a lower level than those at 1.5 ω_z.. Besides zones of significant power content, which seem residuals of peaks at entire multiples of ω_z, not eliminated by the notch filter, there is a peak in the lowest frequency region, which might be related to surge. This assumption is supported by the fact that this peak survives far upstream the blades, as it is expected for a frequency related to surge,



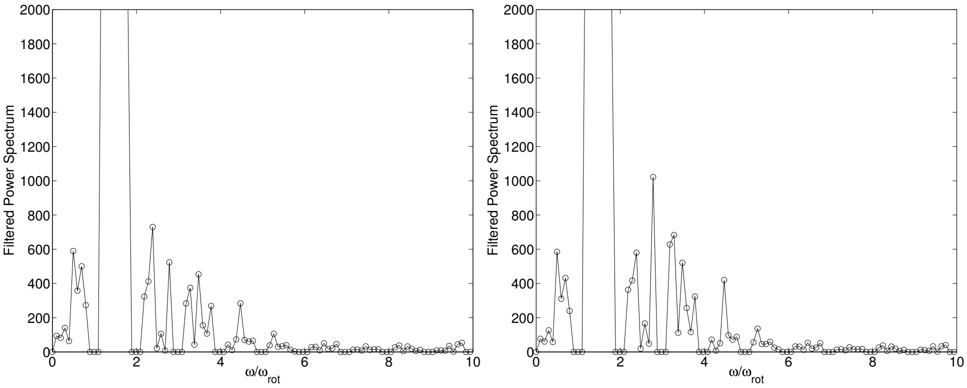
(a) Rotations 1-10 **(b)** Rotations 2-11



(c) Rotations 3-12 **(d)** Rotations 4-13

Figure 7.21: Evolution of the Pressure Power Spectrum, z $= -7.2$ mm.

which is a global axial flow oscillation. Note that, for a more accurate analysis of surge, pressure signals should be recorded over longer time intervals, but again this would imply additional computational costs.



(a) $z = -56.5$ mm **(b)** $z = -28.8$ mm

Figure 7.22: Filtered pressure power spectrum.

Part II

Numerical Simulation of Sediment Transport Problems

Chapter 8

Physical Model for Sediment Transport Problems

In this chapter different physical models for the study of sediment transport problems are presented. First, the standard Shallow Water equations, for both the 1D and 2D cases are shown in Sec. 8.1. After that, the 1D governing equations for sediment transport problems and their 2D extensions are described, respectively, in Sec. 8.2 and Sec. 8.3. The standard formulation of Shallow Water equations is derived starting from the incompressible Navier-Stokes equations [121] and by assuming that the velocity in one direction, let suppose it is x_3, and the corresponding acceleration are negligible. In addition, by depth averaging the the governing equations the standard Shallow Water formulation is recovered [121].

As for the mathematical notation, considering a fixed reference level L_{ref}, H is the bathymetry function with respect to L_{ref} and the height of the bottom is denoted by Z. Therefore, we have (see also Fig. 8.1):

$$H + Z = L_{ref} \tag{8.1}$$

The height of the flow above the bottom Z is denoted by h, while Q_i is the sediment transport flux in the x_i direction. Since only 1D and 2D physical models are considered in this part of the work, the axis x_1 and x_2 will be simply denoted by x and y. Finally the symbol reserved for the acceleration of gravity is g and the sediment porosity is denoted by p.

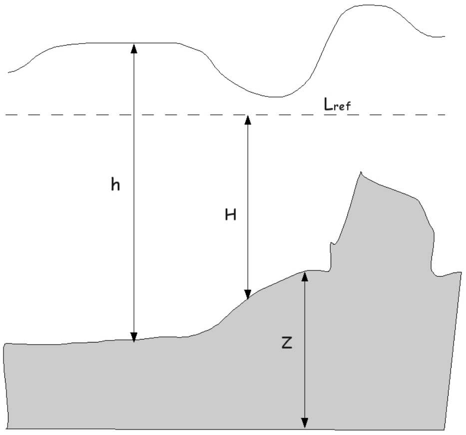


Figure 8.1: Physical variables considered for sediment transport problems.

8.1 Shallow Water equations

In this section the Shallow Water system of equations with variable bottom topography is presented: in the 1D case they can be formulated as follows:

$$\begin{cases} \dfrac{\partial h}{\partial t} + \dfrac{\partial hu}{\partial x} = 0 \\ \dfrac{\partial hu}{\partial t} + \dfrac{\partial}{\partial x}\left(hu^2 + \dfrac{1}{2}gh^2\right) = -gh\dfrac{dZ}{dx} \end{cases} \tag{8.2}$$

while in the 2D case they assume the following form:

$$\begin{cases} \dfrac{\partial h}{\partial t} + \dfrac{\partial hu}{\partial x} + \dfrac{\partial hv}{\partial y} = 0 \\ \dfrac{\partial hu}{\partial t} + \dfrac{\partial}{\partial x}\left(hu^2 + \dfrac{1}{2}gh^2\right) + \dfrac{\partial}{\partial y}(huv) = -gh\dfrac{\partial Z}{\partial x} \\ \dfrac{\partial hv}{\partial t} + \dfrac{\partial}{\partial x}(huv) + \dfrac{\partial}{\partial y}\left(hv^2 + \dfrac{1}{2}gh^2\right) = -gh\dfrac{\partial Z}{\partial y} \end{cases} \tag{8.3}$$

Note that, in standard Shallow Water formulation the bed level Z is fixed in time, thus $Z = Z(x)$ or $Z = Z(x, y)$ for (8.2) and (8.3),respectively. The variable bottom topography introduces an additional, non-homogeneous, term in the set of equations: the discretization of this term requires particular attention, as shown in Chap. 9. The mathematical structure of the Shallow Water equations is strictly related to the one of the Euler equations for barotropic flows described in Sec. 2.3. Indeed, by considering only the homogeneous part of the system or, equivalently, a constant bed level, the following 1D system is recovered:

$$\begin{cases} \dfrac{\partial h}{\partial t} + \dfrac{\partial hu}{\partial x} = 0 \\ \dfrac{\partial hu}{\partial t} + \dfrac{\partial}{\partial x}\left(hu^2 + \dfrac{1}{2}gh^2\right) = 0 \end{cases} \tag{8.4}$$

Except for the presence of the (decoupled) passive scalar equation, system (8.4) is a particular instance of the 1D system (2.11) when the following substitutions are operated:

$$\rho \Rightarrow h, \quad p \Rightarrow \frac{1}{2}gh^2 \tag{8.5}$$

Thus the mathematical properties of (8.4) can be readily derived from the ones of (2.11). In particular the eigenvalues of the homogeneous part of (8.2) are:

$$\lambda_1 = u - c, \quad \lambda_2 = u + c \tag{8.6}$$

where c is the equivalent of the speed of sound of system (2.11) defined as:

$$c \doteq \sqrt{gh} = \sqrt{\frac{\partial}{\partial h}\left(\frac{1}{2}gh^2\right)} \tag{8.7}$$

Similarly, for the 2D case (8.3), the eigenvalues are easily computed as:

$$\lambda_1 = u - c, \quad \lambda_2 = u, \quad \lambda_3 = u + c \tag{8.8}$$

8.2 1D Sediment Transport Model

In order to include the effects of sediment transport, an additional equation which describes the time evolution of the bed level must be added to system (8.2). In this work the Exner equation, a well-known approach to model the time evolution of the bottom topography, has been used:

$$(1-p)\frac{\partial Z}{\partial t} + \frac{\partial Q_1}{\partial x} = 0 \tag{8.9}$$

The Exner equation is a common choice since a wide range of formulations, deterministic and probabilistic as well as analytical and empirical ones, are available for the definition of the sediment transport flux Q_1. In this study the sediment transport fluxes are defined through the use of the classical Grass model [56]:

$$Q_1 = A_g u|u|^{m-1} \tag{8.10}$$

where $0 \leq A_g \leq 1$ takes into account the grain diameter and the cinematic viscosity [26] and $1 \leq m \leq 4$ is an experimental constant depending on the particular problem under consideration. More precisely, in the following only the standard value $m = 3$ is considered so that (8.10) reduces to:

$$Q_1 = A_g u^3 \tag{8.11}$$

Note that instead of the classical model (8.11), used here for the sake of simplicity, more complex formulations can also be considered for the formulation of Q_1 (see e.g. [26]). Using (8.3) and (8.9) the following coupled system is then obtained:

$$\begin{cases} \dfrac{\partial h}{\partial t} + \dfrac{\partial hu}{\partial x} + \dfrac{\partial hv}{\partial y} = 0 \\ \dfrac{\partial hu}{\partial t} + \dfrac{\partial}{\partial x}\left(hu^2 + \dfrac{1}{2}gh^2\right) = -gh\dfrac{\partial Z}{\partial x} \\ \dfrac{\partial Z}{\partial t} + \xi\dfrac{\partial Q_1}{\partial x} = 0 \end{cases} \tag{8.12}$$

where

$$\xi = \frac{1}{1-p} \tag{8.13}$$

It is possible to rewrite (8.12) as a system of conservation laws with a source term, that is:

$$\frac{\partial \mathbf{W}}{\partial t} + \frac{\partial \mathbf{F}_1(\mathbf{W})}{\partial x} = \mathbf{S}(\mathbf{W}) \tag{8.14}$$

where

$$\begin{cases} \mathbf{W} & = (\quad h, \quad hu, \quad Z \quad)^T \\ \mathbf{F}_1(\mathbf{W}) & = \left(\quad hu, \quad hu^2 + \dfrac{1}{2}gh^2, \quad \xi Q_1 \quad \right)^T \\ \mathbf{S}(\mathbf{W}) & = \left(\quad 0, \quad -gh\dfrac{\partial Z}{\partial x}, \quad 0 \quad \right)^T \end{cases} \tag{8.15}$$

The Jacobian matrix of $\mathbf{F}_1(\mathbf{W})$ in (8.15) is singular. Indeed, the eigenvalues and the matrix R, whose columns correspond to the right eigenvectors of the Jacobian of $\mathbf{F}_1(\mathbf{W})$ can be expressed as follows:

$$\lambda_1 = u - c, \quad \lambda_2 = 0, \quad \lambda_3 = u + c \tag{8.16}$$

$$R = \begin{pmatrix} 1 & 0 & 1 \\ u-c & 0 & u+c \\ \dfrac{-cd}{u-c} & 1 & \dfrac{cd}{u+c} \end{pmatrix} \tag{8.17}$$

where d is defined as:

$$d \doteq \frac{3\xi A_g u^2}{h} = \xi \frac{\partial Q_1}{\partial hu} \tag{8.18}$$

A singular Jacobian matrix may cause severe numerical difficulties in many finite volume schemes [9]. Furthermore, in order to obtain a well-balanced scheme [10], the presence of the source term requires a specific treatment, as it is shown is Chap. 9. Thus, in the literature several alternative formulations of system (8.14) have been derived to cope with these difficulties. Three of them, which are used in the numerical methods considered in this work, are shown in Secs. 8.2.1-8.2.3.

8.2.1 A formulation with non singular Jacobian

In order to avoid the singularity of the Jacobian of the flux function, it is possible to incorporate the variable Z into the flux function as shown in [64]. In particular, by using the product rule:

$$\frac{\partial hZ}{\partial x} = Z\frac{\partial h}{\partial x} + h\frac{\partial Z}{\partial x} \tag{8.19}$$

it is possible to recast system (8.12) as:

$$\frac{\partial \mathbf{W}}{\partial t} + \frac{\partial \tilde{\mathbf{F}}_1(\mathbf{W})}{\partial x} = \tilde{\mathbf{S}}(\mathbf{W}) \tag{8.20}$$

where

$$\begin{cases} \tilde{\mathbf{F}}_1(\mathbf{W}) & = \begin{pmatrix} hu, & hu^2 + \frac{1}{2}gh^2 + ghZ, & \xi Q_1 \end{pmatrix}^T \\ \tilde{\mathbf{S}}(\mathbf{W}) & = \begin{pmatrix} 0, & gZ\dfrac{\partial h}{\partial x}, & 0 \end{pmatrix}^T \end{cases} \tag{8.21}$$

The formulation (8.20) overcomes the issue of the singularity of the Jacobian matrix of (8.14). Indeed the right eigenvectors matrix $\tilde{R}$ of the Jabian of $\tilde{\mathbf{F}}_1(\mathbf{W})$ can be recasted in the following form:

$$\tilde{R} = \begin{pmatrix} 1 & 1 & 1 \\ \tilde{\lambda}_1 & \tilde{\lambda}_2 & \tilde{\lambda}_3 \\ \dfrac{\tilde{\beta}_1^2 - c^2 - gZ}{c^2} & \dfrac{\tilde{\beta}_2^2 - c^2 - gZ}{c^2} & \dfrac{\tilde{\beta}_3^2 - c^2 - gZ}{c^2} \end{pmatrix} \tag{8.22}$$

where

$$\tilde{\beta}_i \doteq \tilde{\lambda}_i - u \tag{8.23}$$

and $\tilde{\lambda}_i$ is the i^{th} eigenvalue of system (8.20).

The eigenvalues $\tilde{\lambda}_i$ cannot be easily obtained since they are the roots of the polynomial:

$$\tilde{\lambda}^3 - 2u\tilde{\lambda}^2 + \Big(u^2 - g\,(h + Z + hd)\Big)\,\tilde{\lambda} + ghud = 0 \tag{8.24}$$

where d is defined by (8.18). In [64] it is shown that the roots of the polynomial are always real and distinct if $h+Z > 0$ and, in this case, the eigenvalues can be expressed in the following form:

$$\begin{cases} \tilde{\lambda}_1 = 2\sqrt{-\tilde{L}}\cos\left(\dfrac{\tilde{\theta}}{3}\right) & +\dfrac{2}{3}u \\ \tilde{\lambda}_2 = 2\sqrt{-\tilde{L}}\cos\left(\dfrac{\tilde{\theta}+2\pi}{3}\right) & +\dfrac{2}{3}u \\ \tilde{\lambda}_3 = 2\sqrt{-\tilde{L}}\cos\left(\dfrac{\tilde{\theta}+4\pi}{3}\right) & +\dfrac{2}{3}u \end{cases} \text{ and } \begin{cases} \tilde{M} = \dfrac{u}{27}\left(9g\left(h - h\dfrac{d}{2} + Z\right) - u^2\right) \\ \tilde{L} = -\dfrac{1}{9}\Big(u^2 + 3g\,(h + hd + Z)\Big) \\ \cos\tilde{\theta} = \dfrac{\tilde{M}}{\sqrt{-\tilde{L}^3}} \end{cases} \tag{8.25}$$

8.2.2 A non conservative formulation

Following [26,28], it is possible to consider (8.12) as a non-conservative system in the form:

$$\frac{\partial \hat{\mathbf{W}}}{\partial t} + \frac{\partial \hat{\mathbf{F}}_1(\hat{\mathbf{W}})}{\partial x} = \hat{\mathbf{B}}_1(\hat{\mathbf{W}})\frac{\partial \hat{\mathbf{W}}}{\partial x} \tag{8.26}$$

where

$$\begin{cases} \hat{\mathbf{W}} & = \begin{pmatrix} h, & hu, & H \end{pmatrix}^T \\ \hat{\mathbf{F}}_1(\hat{\mathbf{W}}) & = \begin{pmatrix} hu, & hu^2 + \dfrac{1}{2}gh^2, & -\xi Q_1 \end{pmatrix}^T \end{cases} \tag{8.27}$$

and

$$\hat{\mathbf{B}}_1(\hat{\mathbf{W}}) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & gh \\ 0 & 0 & 0 \end{pmatrix} \tag{8.28}$$

Note that in (8.27) the variable H, defined by (8.1) is used instead of Z.

System (8.26)-(8.27) can be also rewritten as a homogeneous hyperbolic system in the form:

$$\frac{\partial \hat{\mathbf{W}}}{\partial t} + \hat{\mathbf{A}}_\mathbf{1}\frac{\partial \hat{\mathbf{W}}}{\partial x} = 0 \tag{8.29}$$

where the matrix $\hat{\mathbf{A}}_\mathbf{1}$ is defined as follows:

$$\hat{\mathbf{A}}_\mathbf{1} \doteq \frac{\partial \hat{\mathbf{F}}_1(\hat{\mathbf{W}})}{\partial \hat{\mathbf{W}}} - \hat{\mathbf{B}}_1(\hat{\mathbf{W}}) = \begin{pmatrix} 0 & 1 & 0 \\ c^2 - u^2 & 2u & -c^2 \\ -u\hat{d} & \hat{d} & 0 \end{pmatrix} \tag{8.30}$$

with:

$$\hat{d} \doteq \frac{-3\xi A_g u^2}{h} = -\xi \frac{\partial Q_1}{\partial hu} \tag{8.31}$$

The right eigenvectors matrix $\hat{\mathbf{R}}$ of $\hat{\mathbf{A}}_{\mathbf{1}}$ can be expressed as follows:

$$\hat{\mathbf{R}} = \begin{pmatrix} 1 & 1 & 1 \\ \hat{\lambda}_1 & \hat{\lambda}_2 & \hat{\lambda}_3 \\ 1 - \dfrac{\hat{\beta}_1^2}{c^2} & 1 - \dfrac{\hat{\beta}_2^2}{c^2} & 1 - \dfrac{\hat{\beta}_3^2}{c^2} \end{pmatrix} \tag{8.32}$$

where

$$\hat{\beta}_k = \hat{\lambda}_k - u \tag{8.33}$$

and $\hat{\lambda}_k$, the eigenvalues of $\hat{\mathbf{A}}_{\mathbf{1}}$, are the solutions of the following characteristic polynomial:

$$\hat{\lambda}^3 - 2u\hat{\lambda}^2 + \left(u^2 - gh\left(1 - \hat{d}\right)\right)\hat{\lambda} - ghu\hat{d} = 0 \tag{8.34}$$

whose solution can be expressed in the following form:

$$\begin{cases} \hat{\lambda}_1 = 2\sqrt{-\hat{L}}\cos\left(\dfrac{\hat{\theta}}{3}\right) + \dfrac{2}{3}u \\ \hat{\lambda}_2 = 2\sqrt{-\hat{L}}\cos\left(\dfrac{\hat{\theta} + 2\pi}{3}\right) + \dfrac{2}{3}u \\ \hat{\lambda}_3 = 2\sqrt{-\hat{L}}\cos\left(\dfrac{\hat{\theta} + 4\pi}{3}\right) + \dfrac{2}{3}u \end{cases} \quad \text{where} \quad \begin{cases} \hat{M} = \dfrac{u}{54}\left(9gh\left(2 + \hat{d}\right) - u^2\right) \\ \hat{L} = -\dfrac{1}{9}\left(u^2 + 3gh\left(1 - \hat{d}\right)\right) \\ \cos\hat{\theta} = \dfrac{\hat{M}}{\sqrt{-\hat{L}^3}} \end{cases} \tag{8.35}$$

8.2.3 Physical-variable based homogeneous formulation

Through the use of the primitive variables h, u and Z, system (8.14) can be also rewritten in order to eliminate the source term [64]. More precisely, the following homogeneous system is obtained:

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \check{\mathbf{F}}_1(\mathbf{U})}{\partial x} = 0 \tag{8.36}$$

where

$$\begin{cases} \mathbf{U} & = \begin{pmatrix} h, & u, & Z \end{pmatrix}^T \\ \check{\mathbf{F}}_1(\mathbf{U}) & = \begin{pmatrix} hu, & \dfrac{1}{2}u^2 + g\left(h + B\right), & \xi Q_1 \end{pmatrix}^T \end{cases} \tag{8.37}$$

Similarly to system (8.21) the right eigenvectors matrix $\check{R}$ of the Jacobian of $\check{\mathbf{F}}_1(\mathbf{U})$ can be formulated as a function of the eigenvalues as follows:

$$R = \begin{pmatrix} 1 & 1 & 1 \\ \dfrac{\check{\beta}_1}{h} & \dfrac{\check{\beta}_2}{h} & \dfrac{\check{\beta}_3}{h} \\ \dfrac{\check{\beta}_1^2 - c^2}{c^2} & \dfrac{\check{\beta}_1^2 - c^2}{c^2} & \dfrac{\check{\beta}_1^2 - c^2}{c^2} \end{pmatrix} \tag{8.38}$$

where

$$\beta_k \doteq \check{\lambda}_k - u \tag{8.39}$$

and $\check{\lambda}_k$ are the solutions of the third order polynomial:

$$\check{\lambda}^3 - 2u\check{\lambda}^2 + \left(u^2 - g\left(h + d\right)\right)\check{\lambda} + gud = 0 \tag{8.40}$$

where $\check{d}$ is defined as:

$$\check{d} \doteq 3\xi A_g u^2 = \xi\frac{\partial Q_1}{\partial u} \tag{8.41}$$

Similarly to the previous cases, the eigenvalues are always real and different, provided that $h > 0$, and can be computed by the following equation.

$$\begin{cases} \check{\lambda}_1 = 2\sqrt{-\check{L}}\cos\left(\dfrac{\check{\theta}}{3}\right) + \dfrac{2}{3}u \\ \check{\lambda}_2 = 2\sqrt{-\check{L}}\cos\left(\dfrac{\check{\theta} + 2\pi}{3}\right) + \dfrac{2}{3}u \\ \check{\lambda}_3 = 2\sqrt{-\check{L}}\cos\left(\dfrac{\check{\theta} + 4\pi}{3}\right) + \dfrac{2}{3}u \end{cases} \quad \text{where} \quad \begin{cases} \check{M} = \dfrac{u}{54}\left(9g\left(2h - d\right) - 2u^2\right) \\ \check{L} = -\dfrac{1}{9}\left(u^2 + 3g\left(h + d\right)\right) \\ \cos\check{\theta} = \dfrac{\check{M}}{\sqrt{-\check{L}^3}} \end{cases} \tag{8.42}$$

8.3 2D Physical Models

In this section the 2D extensions of the different formulations described in Sec. 8.2 are addressed. Concerning the fluid variables h, hu and hv, their evolution in time is governed by the 2D Shallow Water equations (8.3). As for the bottom topography Z, its evolution is governed by the 2D Exner equation:

$$\frac{\partial Z}{\partial t} + \xi\frac{\partial Q_1}{\partial x} + \xi\frac{\partial Q_2}{\partial y} = 0 \tag{8.43}$$

where Q_1 and Q_2 are, respectively, the sediment transport fluxes in the x and y directions. Similarly to the 1D case, several choices are possible for the definition of Q_1 and Q_2 and, in particular, in this study the sediment transport fluxes are defined through the use of the 2D Grass model [56]:

$$Q_1 = A_g u \left(u^2 + v^2\right)^{\frac{m-1}{2}}, \qquad Q_2 = A_g v \left(u^2 + v^2\right)^{\frac{m-1}{2}} \tag{8.44}$$

More precisely, as in the 1D case, the classical value $m = 3$ is considered so that (8.44) reduces to:

$$Q_1 = A_g u \left(u^2 + v^2\right), \qquad Q_2 = A_g v \left(u^2 + v^2\right) \tag{8.45}$$

Using (8.3) and (8.43) the following coupled system is then obtained:

$$\begin{cases} \dfrac{\partial h}{\partial t} + \dfrac{\partial hu}{\partial x} + \dfrac{\partial hv}{\partial y} = 0 \\ \dfrac{\partial hu}{\partial t} + \dfrac{\partial}{\partial x}\left(hu^2 + \dfrac{1}{2}gh^2\right) + \dfrac{\partial}{\partial y}(huv) \quad = - gh\dfrac{\partial Z}{\partial x} \\ \dfrac{\partial hv}{\partial t} + \dfrac{\partial}{\partial x}(huv) + \dfrac{\partial}{\partial y}\left(hv^2 + \dfrac{1}{2}gh^2\right) \quad = - gh\dfrac{\partial Z}{\partial y} \\ \dfrac{\partial Z}{\partial t} + \xi\dfrac{\partial Q_1}{\partial x} + \xi\dfrac{\partial Q_2}{\partial y} = 0 \end{cases} \tag{8.46}$$

which can be recasted in the following vector form:

$$\frac{\partial \mathbf{W}}{\partial t} + \frac{\partial \mathbf{F}_1(\mathbf{W})}{\partial x} + \frac{\partial \mathbf{F}_2(\mathbf{W})}{\partial y} = \mathbf{S}(\mathbf{W}) \tag{8.47}$$

where

$$\begin{cases} \mathbf{W} &= \left(h, \quad hu, \quad hv, \quad Z \right)^T \\ \mathbf{F}_1(\mathbf{W}) &= \left(hu, \quad hu^2 + \dfrac{1}{2}gh^2, \quad huv, \quad \xi Q_1 \right)^T \\ \mathbf{F}_2(\mathbf{W}) &= \left(hv, \quad hvu, \quad hv^2 + \dfrac{1}{2}gh^2, \quad \xi Q_2 \right)^T \\ \mathbf{S}(\mathbf{W}) &= \left(0, \quad -gh\dfrac{\partial Z}{\partial x}, \quad -gh\dfrac{\partial Z}{\partial y}, \quad 0 \right)^T \end{cases} \tag{8.48}$$

Let us consider the integral formulation of system (8.48): by integrating over a general control volume $\mathcal{V}$, the expression of the system is:

$$\int_{\mathcal{V}} \frac{\partial \mathbf{W}}{\partial t} \mathrm{d}\mathcal{V} + \int_{\mathcal{V}} \left(\frac{\partial}{\partial x}\mathbf{F_1}(\mathbf{W}) + \frac{\partial}{\partial y}\mathbf{F_2}(\mathbf{W}) \right) \mathrm{d}\mathcal{V} = \int_{\mathcal{V}} \mathbf{S}(\mathbf{W}) \mathrm{d}\mathcal{V} \tag{8.49}$$

Using the divergence theorem it is possible to reformulate (8.49) as follows:

$$\int_{\mathcal{V}} \frac{\partial \mathbf{W}}{\partial t} \mathrm{d}\mathcal{V} + \int_{\mathcal{S}} \left(n_x \mathbf{F}_1(\mathbf{W}) + n_y \mathbf{F}_2(\mathbf{W}) \right) \mathrm{d}\mathcal{S} = \int_{\mathcal{V}} \mathbf{S}(\mathbf{W}) \mathrm{d}\mathcal{V} \tag{8.50}$$

where, similarly as for the previously physical models for cavitating flows described in Chap.2, $\mathcal{S}$ is the surface of the control volume $\mathcal{V}$ and $\mathbf{n} = (n_x, n_y)^T$ is the unit vector normal to the surface pointing externally with respect to $\mathcal{V}$. Let us consider the Jacobian matrix of $n_x\mathbf{F}_1(\mathbf{W}) + n_y\mathbf{F}_2(\mathbf{W})$: the rotational invariance property introduced in Sec. 5.3 holds true also for this system of equations. Thus, in order to study the mathematical properties of system (8.50), it is possible to assume that $\mathbf{n}$ is coincident with the x axis and, as a consequence, the mathematical properties of the system can be derived by only considering for instance $\mathbf{F}_1(\mathbf{W})$. The Jacobian matrix of $\mathbf{F}_1(\mathbf{W})$ can be expressed as follows:

$$\frac{\partial \mathbf{F}_1(\mathbf{W})}{\partial \mathbf{W}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ c^2 - u^2 & 2u & 0 & 0 \\ -uv & v & u & 0 \\ -ud_2 - vd_3 & d_2 & d_3 & 0 \end{pmatrix} \tag{8.51}$$

where d_2 and d_3 are defined as:

$$d_2 \doteq \xi \frac{A_g\left(3u^2 + v^2\right)}{h} = \xi \frac{\partial Q_1}{\partial hu}, \quad d_3 \doteq \xi \frac{2A_g uv}{h} = \xi \frac{\partial Q_1}{\partial hv} \tag{8.52}$$

As for the 1D case, the Jacobian matrix of the sediment transport physical model is singular since its eigenvalues are:

$$\lambda_1 = u - c, \quad \lambda_2 = 0, \quad \lambda_3 = u + c \quad \lambda_4 = u \tag{8.53}$$

As a consequence, even for the 2D case different alternative formulations of system (8.46) are considered. More precisely, following the 1D, three different models are presented: an approach based on a non-singular matrix, a non-conservative one and a homogeneous formulation based on physical variables. These models will be used in the construction of the different numerical schemes presented in Chap. 9.

8.3.1 A 2D formulation with non singular Jacobian

Similarly to the 1D case, by using the product rule (8.19), the system (8.46) can be recast in the following form:

$$\frac{\partial \mathbf{W}}{\partial t} + \frac{\partial \tilde{\mathbf{F}}_1(\mathbf{W})}{\partial x} + \frac{\partial \tilde{\mathbf{F}}_2(\mathbf{W})}{\partial y} = \tilde{\mathbf{S}}(\mathbf{W}) \tag{8.54}$$

where $\mathbf{W}$ is unchanged with respect to (8.48) and

$$\begin{cases} \tilde{\mathbf{F}}_1(\mathbf{W}) & = \begin{pmatrix} hu, & hu^2 + \frac{1}{2}gh^2 + ghZ, & huv, & \xi Q_1 \end{pmatrix}^T \\ \tilde{\mathbf{F}}_2(\mathbf{W}) & = \begin{pmatrix} hv, & hvu, & hv^2 + \frac{1}{2}gh^2 + ghZ, & \xi Q_2 \end{pmatrix}^T \\ \tilde{\mathbf{S}}(\mathbf{W}) & = \begin{pmatrix} 0, & gZ\frac{\partial h}{\partial x}, & gZ\frac{\partial h}{\partial y}, & 0 \end{pmatrix}^T \end{cases} \tag{8.55}$$

Using this formulation, the Jacobian matrix of $\tilde{\mathbf{F}}_\mathbf{1}(\mathbf{W})$ is:

$$\frac{\partial \tilde{\mathbf{F}}_1(\mathbf{W})}{\partial \mathbf{W}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ c^2 - u^2 + gZ & 2u & 0 & c^2 \\ -uv & v & u & 0 \\ -ud_2 - vd_3 & d_2 & d_3 & 0 \end{pmatrix} \tag{8.56}$$

After some algebraic manipulation, the matrix of the right eigenvectors of the Jacobian (8.56) can be computed as:

$$\tilde{R} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ \tilde{\lambda}_1 & \tilde{\lambda}_2 & \tilde{\lambda}_3 & 0 \\ v & v & v & 1 \\ \frac{\tilde{\beta}_1^2 - c^2 - gZ}{c^2} & \frac{\tilde{\beta}_2^2 - c^2 - gZ}{c^2} & \frac{\tilde{\beta}_3^2 - c^2 - gZ}{c^2} & \frac{2A_g v}{h} \end{pmatrix} \tag{8.57}$$

where $\tilde{\beta}_i = \tilde{\lambda}_i - u$ and $\tilde{\lambda}_i$ are the eigenvalues of (8.56). The fourth eigenvalue $\tilde{\lambda}_4$ is:

$$\tilde{\lambda}_4 = u \tag{8.58}$$

while the other eigenvalues are the solutions of the third order polynomial:

$$\tilde{\lambda}^3 - 2u\tilde{\lambda}^2 + \Big(u^2 - g\left(h + Z + hd_2\right)\Big)\tilde{\lambda} + ghud_2 = 0 \tag{8.59}$$

which can be expressed as:

$$\begin{cases} \tilde{\lambda}_1 = 2\sqrt{-\tilde{L}}\cos\left(\frac{\tilde{\theta}}{3}\right) + \frac{2}{3}u \\ \tilde{\lambda}_2 = 2\sqrt{-\tilde{L}}\cos\left(\frac{\tilde{\theta}+2\pi}{3}\right) + \frac{2}{3}u \\ \tilde{\lambda}_3 = 2\sqrt{-\tilde{L}}\cos\left(\frac{\tilde{\theta}+4\pi}{3}\right) + \frac{2}{3}u \end{cases} \text{where} \begin{cases} \tilde{M} = \frac{u}{54}\Big(9g\left(2h - hd_2 + 2Z\right) \\ \qquad - 2u^2\Big) \\ \tilde{L} = -\frac{\left(u^2 + 3g\left(h + hd_2 + Z\right)\right)}{9} \\ \cos\tilde{\theta} = \frac{\tilde{M}}{\sqrt{-\tilde{L}^3}} \end{cases} \tag{8.60}$$

Equations (8.59) and (8.57) show an important feature of system (8.54)-(8.55). For the case of cavitating flows, multidimensional problems were equivalent to 1D problems with a suitable number of passive scalar equations. For sediment transport problems, even with the simple Grass model, this is not true in general. Due to the formulation of Q_1 in (8.45), the equation for hv is not decoupled from the other equations. This is clearly shown in (8.57) where the eigenvector associated to $\tilde{\lambda}_4 = u$ is not in general an element of the canonical basis (i.e. it is not $(0,0,1,0)^T$) as for the case of passive scalar equations. Note also that despite their formal analogy, even the polynomial (8.59) and (8.24) are different since $d_2 \neq d$. These differences between the 1D and 2D physical models clearly add to the difficulty of the problem.

8.3.2 A 2D non conservative formulation

Alternatively, following the work in [26, 28], it is possible to recast system (8.46) as a non-conservative system in the form:

$$\frac{\partial \hat{\mathbf{W}}}{\partial t} + \frac{\partial \hat{\mathbf{F}}_1(\hat{\mathbf{W}})}{\partial x} + \frac{\partial \hat{\mathbf{F}}_2(\hat{\mathbf{W}})}{\partial y} = \hat{\mathbf{B}}_1(\hat{\mathbf{W}})\frac{\partial \hat{\mathbf{W}}}{\partial x} + \hat{\mathbf{B}}_2(\hat{\mathbf{W}})\frac{\partial \hat{\mathbf{W}}}{\partial y} \tag{8.61}$$

where

$$\left\{\begin{array}{lllllll} \hat{\mathbf{W}} & = (& h, & hu, & hv, & H &)^T \\ \hat{\mathbf{F}}_1(\hat{\mathbf{W}}) & = (& hu, & hu^2 + \frac{1}{2}gh^2, & huv, & -\xi Q_1 &)^T \\ \hat{\mathbf{F}}_2(\hat{\mathbf{W}}) & = (& hv, & hvu, & hv^2 + \frac{1}{2}gh^2, & -\xi Q_2 &)^T \end{array}\right. \tag{8.62}$$

and

$$\hat{\mathbf{B}}_1(\hat{\mathbf{W}}) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & gh \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \hat{\mathbf{B}}_2(\hat{\mathbf{W}}) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & gh \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{8.63}$$

Similarly to the 1D case, system (8.61)-(8.62) can be recast in the following form:

$$\frac{\partial \hat{\mathbf{W}}}{\partial t} + \hat{\mathbf{A}}_1(\hat{\mathbf{W}})\frac{\partial \hat{\mathbf{W}}}{\partial x} + \hat{\mathbf{A}}_2(\hat{\mathbf{W}})\frac{\partial \hat{\mathbf{W}}}{\partial y} = 0 \tag{8.64}$$

where

$$\hat{\mathbf{A}}_k(\hat{\mathbf{W}}) = \frac{\partial \hat{\mathbf{F}}_k(\hat{\mathbf{W}})}{\partial \hat{\mathbf{W}}} - \hat{\mathbf{B}}_k(\hat{\mathbf{W}}) \quad k = 1, 2 \tag{8.65}$$

Note that the matrices $\hat{\mathbf{A}}_k$ are the sum of a conservative term, the Jacobian of $\hat{\mathbf{F}}_k(\hat{\mathbf{W}})$, and a non conservative one, $\hat{\mathbf{B}}_k(\hat{\mathbf{W}})$, which takes into account the spatial variation of the bed.

As for the previous cases, it is sufficient to restrict our attention to the matrix $\hat{\mathbf{A}}_1$ whose expression is as follows:

$$\hat{\mathbf{A}}_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ c^2 - u^2 & 2u & 0 & -c^2 \\ -uv & v & u & 0 \\ -u\hat{d}_2 - v\hat{d}_3 & \hat{d}_2 & \hat{d}_3 & 0 \end{pmatrix} \tag{8.66}$$

where

$$\hat{d}_2 \doteq -\xi \frac{A_g \left(3u^2 + v^2\right)}{h} = -\xi \frac{\partial Q_1}{\partial hu}, \quad \hat{d}_3 \doteq -\xi \frac{2A_g uv}{h} = -\xi \frac{\partial Q_1}{\partial hv} \tag{8.67}$$

The right eigenvector matrix $\hat{\mathbf{R}}$ of $\hat{\mathbf{A}}_1(\hat{\mathbf{W}})$ can be computed as:

$$\hat{\mathbf{R}} = \begin{pmatrix} 1 & 1 & 1 & -2\xi A_g v \\ \hat{\lambda}_1 & \hat{\lambda}_2 & \hat{\lambda}_3 & -2\xi A_g uv \\ v & v & v & 1 - 2\xi A_g v^2 \\ 1 - \dfrac{\hat{\beta}_1^2}{c^2} & 1 - \dfrac{\hat{\beta}_2^2}{c^2} & 1 - \dfrac{\hat{\beta}_3^2}{c^2} & -2\xi A_g v \end{pmatrix} \tag{8.68}$$

where $\hat{\beta}_i = \hat{\lambda}_i - u$ and $\hat{\lambda}_i$ are the eigenvalues of $\hat{\mathbf{A}}_1$. The fourth eigenvalue $\hat{\lambda}_4$ is:

$$\hat{\lambda}_4 = u \tag{8.69}$$

while the remaining eigenvalues are the solution of:

$$\hat{\lambda}^3 - 2u\hat{\lambda}^2 + \left(u^2 - gh\left(1 - \hat{d}_2\right)\right)\hat{\lambda} - ghu\hat{d}_2 = 0 \tag{8.70}$$

which can be expressed as:

$$\begin{cases} \hat{\lambda}_1 = 2\sqrt{-\hat{L}}\cos\left(\dfrac{\hat{\theta}}{3}\right) + \dfrac{2}{3}u \\ \hat{\lambda}_2 = 2\sqrt{-\hat{L}}\cos\left(\dfrac{\hat{\theta} + 2\pi}{3}\right) + \dfrac{2}{3}u \\ \hat{\lambda}_3 = 2\sqrt{-\hat{L}}\cos\left(\dfrac{\hat{\theta} + 4\pi}{3}\right) + \dfrac{2}{3}u \end{cases} \text{where} \begin{cases} \hat{M} = \dfrac{u}{54}\left(9gh\left(2 + \hat{d}_2\right) - u^2\right) \\ \hat{L} = -\dfrac{1}{9}\left(u^2 + 3gh\left(1 - \hat{d}_2\right)\right) \\ \cos\hat{\theta} = \dfrac{\hat{M}}{\sqrt{-\hat{L}^3}} \end{cases} \tag{8.71}$$

Again note that despite the formal analogy of (8.70) and (8.34), the two polynomials differs since $\hat{d}_2 \neq \hat{d}$.

8.3.3 2D physical-variable based formulation

Finally, a formulation based on the physical variables h, u, v, Z is presented. The aforementioned formulation is based on the approach described in [5]. Considering a generic control volume $\mathcal{V}$ and the normal vector $\mathbf{n}$ to the external surface $\mathcal{S}$ of $\mathcal{V}$, system (8.46) can be recast in the following differential form [5]:

$$\begin{cases} \dfrac{\partial h}{\partial t} + \dfrac{\partial h u_\eta}{\partial \eta} = 0 \\ \dfrac{\partial h u_\eta}{\partial t} + \dfrac{\partial}{\partial \eta}\left(h u_\eta^2 + \dfrac{1}{2} g h^2\right) = -gh \dfrac{\partial Z}{\partial \eta} \\ \dfrac{\partial h u_\tau}{\partial t} + \dfrac{\partial \left(h u_\eta u_\tau\right)}{\partial \eta} = 0 \\ \dfrac{\partial Z}{\partial t} + \dfrac{\partial}{\partial \eta}\left(\xi A_g u_\eta \left(u_\eta^2 + u_\tau^2\right)\right) = 0 \end{cases} \tag{8.72}$$

where u_η and u_τ are, respectively, the normal and tangential components of the velocity, i.e.:

$$\begin{cases} u_\eta = n_x u + n_y v \\ u_\tau = -n_y u + n_x v \end{cases} \tag{8.73}$$

Note in particular that when $\mathbf{n}$ is coincident with the x axis, the Jacobian of (8.72) is coincident with $\hat{\mathbf{A}}_1$ in (8.66) (taking into account the change of variables from Z to H). System (8.72) can be reformulated by using the physical variables as follows:

$$\begin{cases} \dfrac{\partial h}{\partial t} + u_\eta \dfrac{\partial h}{\partial \eta} + h \dfrac{\partial u_\eta}{\partial \eta} = 0 \\ \dfrac{\partial u_\eta}{\partial t} + u_\eta \dfrac{\partial u_\eta}{\partial \eta} + g \dfrac{\partial h}{\partial \eta} + g \dfrac{\partial Z}{\partial \eta} = 0 \\ \dfrac{\partial u_\tau}{\partial t} + u_\eta \dfrac{\partial u_\tau}{\partial \eta} = 0 \\ \dfrac{\partial Z}{\partial t} + d_\eta \dfrac{\partial u_\eta}{\partial \eta} + d_\tau \dfrac{\partial u_\tau}{\partial \eta} = 0 \end{cases} \quad \text{where} \quad \begin{cases} d_\eta \doteq \xi A_g \left(3 u_\eta^2 + u_\tau^2\right) \\ d_\tau \doteq \xi 2 A_g u_\eta u_\tau \end{cases} \tag{8.74}$$

or, in vector form, as:

$$\frac{\partial \mathbf{U}}{\partial t} + \mathbf{A}_\eta\left(\mathbf{U}\right) \frac{\partial \mathbf{U}}{\partial \eta} = 0 \tag{8.75}$$

with

$$\mathbf{U} = \begin{pmatrix} h \\ u_\eta \\ u_\tau \\ Z \end{pmatrix}, \quad \mathbf{A}_\eta = \begin{pmatrix} u_\eta & h & 0 & 0 \\ g & u_\eta & 0 & g \\ 0 & 0 & u_\eta & 0 \\ 0 & d_\eta & d_\tau & 0 \end{pmatrix} \tag{8.76}$$

Similarly to the previously described case, the eigenvalues and eigenvectors of matrix $\mathbf{A}_\eta$ can computed as follows:

$$\check{R} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ \check{\beta}_1 & \check{\beta}_2 & \check{\beta}_3 & 0 \\ 0 & 0 & 0 & 1 \\ \dfrac{\check{\beta}_1^2 - c^2}{c^2} & \dfrac{\check{\beta}_2^2 - c^2}{c^2} & \dfrac{\check{\beta}_3^2 - c^2}{c^2} & 2\xi A_g u_\tau \end{pmatrix} \tag{8.77}$$

where

$$\check{\beta}_i = \check{\lambda}_i - u_\eta \tag{8.78}$$

The fourth eigenvalue is found to be:

$$\check{\lambda}_4 = u_\eta \tag{8.79}$$

while the other ones are the solution of the third order polynomial:

$$\check{\lambda}^3 - 2u\check{\lambda}^2 + \left(u^2 - g\left(h + d_\eta\right)\right)\check{\lambda} + gud_\eta = 0 \tag{8.80}$$

which, provided $h > 0$, and can be expressed as:

$$\begin{cases} \check{\lambda}_1 = 2\sqrt{-\check{L}}\cos\left(\dfrac{\check{\theta}}{3}\right) + \dfrac{2}{3}u \\ \check{\lambda}_2 = 2\sqrt{-\check{L}}\cos\left(\dfrac{\check{\theta} + 2\pi}{3}\right) + \dfrac{2}{3}u \\ \check{\lambda}_3 = 2\sqrt{-\check{L}}\cos\left(\dfrac{\check{\theta} + 4\pi}{3}\right) + \dfrac{2}{3}u \end{cases} \text{where} \begin{cases} \check{M} = \dfrac{u}{54}\left(9g\left(2h - d\right) - 2u^2\right) \\ \check{L} = -\dfrac{1}{9}\left(u^2 + 3g\left(h + d\right)\right) \\ \cos\check{\theta} = \dfrac{\check{M}}{\sqrt{-\check{L}^3}} \end{cases} \tag{8.81}$$

Chapter 9

Numerical Methods for Sediment Transport Problems

In this chapter a finite-volume discretization of the physical models presented in Chap. 8 is described. Two different explicit numerical schemes, namely the SRNH scheme and the Modified Roe one (MR in the following) are considered. The explicit SRNH scheme is a predictor -corrector method developed by Benkhaldoun et al. in a series of paper [5–9, 96]. Conversely the MR explicit scheme, which has been developed by Castro et al. (see e.g. [26–29, 83]), is a Roe scheme modified in order to deal with non conservative systems of equations.

In this work the implicit counterpart of the aforementioned explicit numerical schemes is proposed. In particular the implicit numerical schemes are generated through the use of the automatic differentiation tool TAPENADE [59]. This chapter is organised as follows: first some general definitions for the considered finite-volume discretization are provided in Sec. 9.1. Then, the numerical formulation of the explicit numerical methods is presented in Sec. 9.2. Finally, at the end of this chapter the generation of the implicit schemes starting from their explicit counterpart is addressed in Sec. 9.3. As previously stated, for sediment transport problems the 2D case cannot be reformulated as a suitable 1D system including a passive scalar equation. Thus, in this chapter the 2D numerical discretization of the aforementioned numerical schemes is described, skipping the 1D case. However, once the 2D numerical methods are defined, the 1D formulations can easily be recovered by simply setting to zero the velocity component along the y-axis and all the derivatives in the y direction.

9.1 General definitions for the considered finite-volume formulation

The considered space discretization is based on a finite-volume approach. At a preliminary stage, the considered 2D computational domain $\mathcal{V} \in \mathbb{R}^2$ is approximated by means of a polygonal domain $\mathcal{V}^{pol}$ which, in turn, is divided into N_t triangles having vertexes $\mathbf{P}_i$, with $i \in \mathcal{I} := \{1, \dots, N_c\}$. Let T_h, with $h \in \mathcal{H} := \{1, \dots, N_t\}$, denotes the h^{th}-triangle: the i^{th} finite-volume cell V_i, associated with $\mathbf{P}_i$, is given by:

$$V_i = \bigcup_{h \in t(i)} V_i^{(h)}$$

where:

- $t(i)$ is the set of indexes marking those triangles which share $\mathbf{P}_i$ as a vertex;
- $V_r^{(h)}$ represents the subset of T_h which is defined by further dividing T_h into six sub-triangles by means of its medians and subsequently considering those two sub-triangles which share $\mathbf{P}_r$ as a vertex.

Clearly, this finite-volume discretisation is the 2D counterpart of the 3D case described in Chap. 5.

The following notation is considered: given a finite volume V_i, $|V_i|$ is its area and $\mathbf{G}_i$ is its centre of mass. $N(i)$ is the set of indexes corresponding to the neighbouring cells of the i^{th}-cell, $\mathcal{B}_i = N(i) \cup \{i\}$. Furthermore $\overline{\mathcal{B}}_i = \bigcup_{j \in \mathcal{B}_i} \mathcal{B}_j$ is the set of indexes marking the i^{th}-cell, its neighbours and the neighbours of the neighbours.

$\varGamma_{ij,1}$ and $\varGamma_{ij,2}$ are the two segments of the common interface between cell V_i and V_j while $|\varGamma_{ij,1}|$ and $|\varGamma_{ij,2}|$ are their length (see figure 9.1). The normal unit vector to $\varGamma_{ij,1}$ is $\mathbf{n}_{ij,1} = (n_{x,ij,1}, n_{y,ij,1})$ and a similar definition holds for $\mathbf{n}_{ij,2}$. The average normal vector for the interface between the i^{th} and j^{th}-cell is defined as follows:

$$|\varGamma_{ij}|\mathbf{n}_{ij} = |\varGamma_{ij,1}|\mathbf{n}_{ij,1} + |\varGamma_{ij,2}|\mathbf{n}_{ij,2}$$

where $\mathbf{n}_{ij}$ is the normal unit vector and $|\varGamma_{ij}|$ is the interface length. Finally $\mathbf{W}_i^n$ is the average value of the solution $\mathbf{W}$ in the i^{th} cell at time t_n:

$$\mathbf{W}_i^n \simeq \frac{1}{|V_i|} \int_{V_i} \mathbf{W}(\mathbf{x}, t_n) dV \tag{9.1}$$

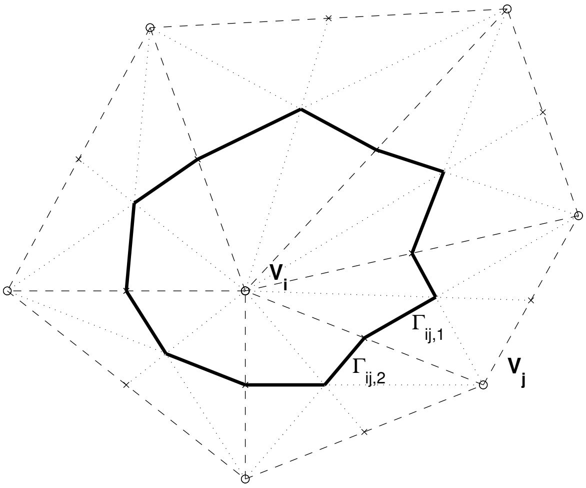


Figure 9.1: Generation of the dual Mesh.

9.2 Explicit time advancing

In this section the explicit formulations of the SNRH and MR numerical schemes are presented. For both numerical methods first the 1^{st}-order formulation is described then the 2^{nd}-order extensions are discussed. The main features of the schemes are briefly outlined without any proof. For additional details we refer to the references in the bibliography (see i.e. [5–8]).

9.2.1 SRNH Numerical Method

The SRNH numerical scheme is a numerical method developed and tested in several papers, see [6, 8] among others. Note that the 2D formulation presented in this work is slightly different from the original formulation in [8]: in the original paper the spatial discretization for the system of equations of the sediment transport problem is based on the formulation (8.47)-(8.48). In this study we considered the same numerical method based on a different form of the system of equations, namely (8.54)-(8.55), i.e. the formulation which avoids the presence of the singularity in the Jacobian matrix. This approach has been developed and tested in the 1D case in [9] but, at least in the author knowledge, the 2D extension has never been studied. Integrating (8.54) over a control volume V_i the resulting system of equations is:

$$\frac{\partial \mathbf{W}_i}{\partial t} = -\frac{1}{|V_i|} \sum_{j \in N(i)} \int_{\Gamma_{ij}} \tilde{\mathcal{F}}(\mathbf{W}, \mathbf{n}) d\sigma + \frac{1}{|V_i|} \int_{V_i} \tilde{\mathbf{S}}(\mathbf{W}) dV \tag{9.2}$$

where

$$\tilde{\mathcal{F}}(\mathbf{W}, \mathbf{n}) \doteq n_x \tilde{\mathbf{F}}_1(\mathbf{W}) + n_y \tilde{\mathbf{F}}_2(\mathbf{W}) \tag{9.3}$$

To discretize (9.2) it is necessary to define a numerical approximation for the flux function $\tilde{\mathcal{F}}$ as well as for the source term $\tilde{\mathbf{S}}$. The SRNH scheme [8] is composed by a predictor and a corrector stage: in the predictor stage an averaged state $\mathbf{W}_{ij}^n$ is computed, then this predicted state is used in the corrector stage to update the solution. In particular in the SRNH scheme the flux at the interface Γ_{ij} is approximated by the analytical flux function computed using the predicted state $\mathbf{W}_{ij}^n$:

$$\int_{\Gamma_{ij}} \tilde{\mathcal{F}}(\mathbf{W}, \mathbf{n}, t^n)\, d\sigma \simeq \tilde{\mathcal{F}}(\mathbf{W}_{ij}^n, \mathbf{n}_{ij})|\Gamma_{ij}| \tag{9.4}$$

The predictor stage

The first step is based on primitive variables and, in particular, on the homogeneous formulation (8.75)-(8.76). Let us introduce the function $\overline{\mathbf{U}}\left(\mathbf{U}_L, \mathbf{U}_R\right)$ which, given two generic states U_L and U_R, returns a Roe averaged state as follows:

$$\overline{\mathbf{U}}\left(\mathbf{U}_L, \mathbf{U}_R\right) = \begin{pmatrix} \overline{h}_{LR} \\ \overline{u}_{\eta,LR} \\ \overline{u}_{\tau,LR} \\ \overline{Z}_{LR} \end{pmatrix} \quad \text{where} \quad \begin{cases} \overline{h}_{LR} = \dfrac{h_L + h_R}{2} \\ \overline{u}_{\eta,LR} = \dfrac{u_{\eta,L}\sqrt{h_L} + u_{\eta,R}\sqrt{h_R}}{\sqrt{h_L} + \sqrt{h_R}} \\ \overline{u}_{\tau,LR} = \dfrac{u_{\tau,L}\sqrt{h_L} + u_{\tau,R}\sqrt{h_R}}{\sqrt{h_L} + \sqrt{h_R}} \\ \overline{Z}_{LR} = \dfrac{Z_L + Z_R}{2} \end{cases} \tag{9.5}$$

Furthermore, let us define the sign matrix $\mathrm{sgn}\left[\mathbf{A}_\eta(\mathbf{U})\right]$ as

$$\mathrm{sgn}\left[\mathbf{A}_\eta(\mathbf{U})\right] = \check{R}(\mathbf{U})\Lambda_{\mathrm{sgn}}(\mathbf{U})\check{R}^{-1}(\mathbf{U}) \tag{9.6}$$

where the elements of the diagonal matrix $\Lambda_{\mathrm{sgn}}(\mathbf{U})$ are the sign function of the eigenvalues of $\mathbf{A}_\eta(\mathbf{U})$ and $\mathcal{R}(\mathbf{U})$ is the corresponding right-eigenvector matrix defined in (8.77). With these definitions, the predictor step of the SRNH scheme is computed as follows:

$$\mathbf{U}_{ij}^n = \frac{1}{2}\left(\mathbf{U}_i^n + \mathbf{U}_j^n\right) - \frac{1}{2}\mathrm{sgn}\left[\mathbf{A}_\eta(\overline{\mathbf{U}}_{ij}^n)\right]\left(\mathbf{U}_j^n - \mathbf{U}_i^n\right) \tag{9.7}$$

where $\overline{\mathbf{U}}_{ij}^n = \overline{\mathbf{U}}\left(\mathbf{U}_i^n, \mathbf{U}_j^n\right)$. Once $\mathbf{U}_{ij}^n$ is available, the state $\mathbf{W}_{ij}^n$ appearing in (9.4) is simply $\mathbf{U}_{ij}^n$ expressed in conservative variables, that is:

$$\mathbf{W}_{ij}^n = \left(h_{ij}^n,\ h_{ij}^n\left(n_{x,ij}u_{\eta,ij}^n - n_{y,ij}u_{\tau,ij}^n\right),\ h_{ij}^n\left(n_{y,ij}u_{\eta,ij}^n + n_{x,ij}u_{\tau,ij}^n\right),\ Z_{ij}^n\right)^T \tag{9.8}$$

The source term

Finally, in order to complete the spatial discretization of (9.2), it is necessary to define an approximation of the source term such that:

$$\frac{1}{|V_i|}\int_{V_i}\tilde{\mathbf{S}}(\mathbf{W},t^n)dV \simeq \tilde{\mathbf{S}}_i^n \tag{9.9}$$

In [8], starting from the system (8.47)-(8.48) the discretization of the source term is defined so that the resulting numerical scheme is well-balanced and satisfy the C-property [10]. In the same spirit, let us define $\overline{Z}_{x,i}$ and $\overline{Z}_{y,i}$ as follows:

$$\overline{Z}_{x,i} = \frac{1}{2}\frac{\sum\limits_{j\in N(i)}(Z_{ij})^2\, n_{x,ij}|\mathit{\Gamma}_{ij}|}{\sum\limits_{j\in N(i)} Z_{ij}n_{x,ij}|\mathit{\Gamma}_{ij}|}, \quad \overline{Z}_{y,i} = \frac{1}{2}\frac{\sum\limits_{j\in N(i)}(Z_{ij})^2\, n_{y,ij}|\mathit{\Gamma}_{ij}|}{\sum\limits_{j\in N(i)} Z_{ij}n_{y,ij}|\mathit{\Gamma}_{ij}|} \tag{9.10}$$

Applying the same approach of [8] to the system (8.54)-(8.55) and, in particular by exchanging the roles of h and Z in the proof in the aforementioned work, the following discretization of the source term defines a well-balanced scheme:

$$\tilde{\mathbf{S}}_i^n = \tilde{\mathbf{S}}_i\left(\{\mathbf{W}_{ij}^n\}_{j\in N(i)}\right)$$

where

$$\tilde{\mathbf{S}}_i\left(\{\mathbf{W}_{ij}\}_{j\in N(i)}\right) = \Bigg(0, \quad g\overline{Z}_{x,i}\sum_{j\in N(i)} h_{ij}n_{x,ij}|\mathit{\Gamma}_{ij}|, \\ g\overline{Z}_{y,i}\sum_{j\in N(i)} h_{ij}n_{y,ij}|\mathit{\Gamma}_{ij}|, \quad 0\Bigg)^T \tag{9.11}$$

The complete discrete formulation

Once a numerical discretization for the flux function and the source term is defined, the following semi-discrete form of the system (9.2) is obtained by injecting (9.4) and (9.9) in (9.2):

$$\frac{\partial \mathbf{W}_i}{\partial t} = \mathrm{RHS}_1\left(\{\mathbf{W}_j^n\}_{j\in\mathcal{B}_i}\right) \tag{9.12}$$

where

$$\mathrm{RHS}_1\left(\{\mathbf{W}_j^n\}_{j\in\mathcal{B}_i}\right) = -\frac{1}{|V_i|}\sum_{j\in N(i)}\tilde{\mathcal{F}}(\mathbf{W}_{ij}^n,\mathbf{n}_{ij})|\mathit{\Gamma}_{ij}| + \tilde{\mathbf{S}}_i^n \tag{9.13}$$

and $\mathbf{W}_{ij}^n$ and $\tilde{\mathbf{S}}_i^n$ are defined by (9.8) and (9.11), respectively. Concerning the time discretization, an explicit first-order Euler scheme is considered as in [8]. Thus, the resulting numerical scheme can be expressed as:

$$\mathbf{W}_i^{n+1} = \mathbf{W}_i^n + \Delta t^n \mathrm{RHS}_1 \left(\{\mathbf{W}_j^n\}_{j \in \mathcal{B}_i} \right) \tag{9.14}$$

Second order extension

The extension to second-order accuracy in space can be achieved by using a classical MUSCL technique [115], in which the flux function is computed by using the extrapolated variable values at the cell interface. First, at each interface $\mathit{\Gamma}_{ij}$, the extrapolated values of the variables are computed:

$$\mathbf{W}_{ij}^{n,-} = \mathbf{W}_i^n + \frac{1}{2}\nabla \mathbf{W}_i \cdot \mathbf{d}_{ij}, \quad \mathbf{W}_{ij}^{n,+} = \mathbf{W}_j^n - \frac{1}{2}\nabla \mathbf{W}_j \cdot \mathbf{d}_{ij} \tag{9.15}$$

where $\mathbf{d}_{ij}$ is the vector joining the node i and the node j and $\nabla \mathbf{W}_i$ is an approximation of the cell gradient. More specifically, the 2D counterpart of the approach described in Sec. 5.4 is used.

First, a linear approximation of the gradient in each triangle T_j is considered. Defining $tn(j,k), k = 1,2,3$ the indexes of the nodes belonging to the triangle T_j, this linear approximation is given by:

$$\nabla \mathbf{W}_{|T_j} = \sum_{k=1}^{3} \mathbf{W}_{tn(j,k)} \nabla \mu_j^k \tag{9.16}$$

where μ_j^k is the barycentric coordinates associated with the k^{th}-vertex. Then a "centred" approximation of the gradient in the i^{th} cell is introduced as follows:

$$\nabla \mathbf{W}_i^c \simeq \frac{\displaystyle\sum_{j \in t(i)} |T_j| \nabla \mathbf{W}_{|T_j}}{\displaystyle\sum_{j \in t(i)} |T_j|} \tag{9.17}$$

where $|T_i|$ is the area of the i^{th} tetrahedron. Once $\nabla \mathbf{W}_i^c$ is available the terms $\nabla \mathbf{W}_i \cdot \mathbf{d}_{ij}$ and $\nabla \mathbf{W}_j \cdot \mathbf{d}_{ij}$ in (9.15) are computed as follows:

$$\left\{ \begin{array}{lcl} \nabla \mathbf{W}_i \cdot \mathbf{d}_{ij} & = & \mathrm{minmod}\{\mathbf{W}_j - \mathbf{W}_i, 2\nabla \mathbf{W}_i^c \cdot \mathbf{d}_{ij} - (\mathbf{W}_j - \mathbf{W}_i)\} \\ \nabla \mathbf{W}_j \cdot \mathbf{d}_{ij} & = & \mathrm{minmod}\{\mathbf{W}_j - \mathbf{W}_i, 2\nabla \mathbf{W}_j^c \cdot \mathbf{d}_{ij} - (\mathbf{W}_j - \mathbf{W}_i)\} \end{array} \right. \tag{9.18}$$

where $\mathrm{minmod}\{a,b\}$ is defined in (5.32).

The 2^{nd}-order accuracy in space is then achieved by computing the flux function and the source term using the extrapolated values (9.15), in analogy with the first-order scheme:

$$\left\{\begin{array}{lcl} \mathbf{U}_{ij}^{n,\pm} & = & \dfrac{1}{2}\left(\mathbf{U}_{ij}^{n,-} + \mathbf{U}_{ij}^{n,+}\right) \\ & & -\dfrac{1}{2}\mathrm{sgn}\left[\mathbf{A}_{\eta}(\overline{\mathbf{U}}_{ij}^{n,\pm})\right]\left(\mathbf{U}_{ij}^{n,+} - \mathbf{U}_{ij}^{n,-}\right) \\ \mathrm{RHS}_2\left(\{\mathbf{W}_j^n\}_{j\in\overline{\mathcal{B}}_i}\right) & = & -\dfrac{1}{|V_i|}\displaystyle\sum_{j\in N(i)} \tilde{\mathcal{F}}(\mathbf{W}_{ij}^{n,\pm}, \mathbf{n}_{ij})|\mathit{\Gamma}_{ij}| + \tilde{\mathbf{S}}_i^{n,\pm} \end{array}\right. \tag{9.19}$$

where $\tilde{\mathbf{S}}_i^{n,\pm} = \tilde{\mathbf{S}}_i\left(\{\mathbf{W}_{ij}^{n,\pm}\}_{j\in N(i)}\right)$ is the source term computed using the extrapolated states (9.15) and $\overline{\mathbf{U}}_{ij}^{n,\pm} = \overline{\mathbf{U}}\left(\mathbf{U}_{ij}^{n,-}, \mathbf{U}_{ij}^{n,+}\right)$ is the predicted state computed using the reconstructed states expressed in physical variables $\mathbf{U}_{ij}^{n,-}$ and $\mathbf{U}_{ij}^{n,+}$.

It is worth notice the difference between the equations in (9.19) and (9.13). The set of indexes j from which the solution W_i is dependent (the stencil) is larger for the 2^{nd}-order accurate flux function than for the first-order one. This is a well known feature of a second-order scheme with MUSCL reconstruction and in section 9.3 we discuss its consequences for the second-order implicit scheme. Using (9.19), the semi-discrete form of (9.2) becomes

$$\frac{\partial \mathbf{W}_i}{\partial t} = \mathrm{RHS}_2\left(\{\mathbf{W}_j^n\}_{j\in\overline{\mathcal{B}}_i}\right) \tag{9.20}$$

Finally, the second-order accuracy in time can be achieved through a two-step Runge-Kutta scheme for time advancing:

$$\left\{\begin{array}{lcl} \mathbf{W}_i^{n+1/2} & = & \mathbf{W}_i^n + \Delta t^n \mathrm{RHS}_2\left(\{\mathbf{W}_j^n\}_{j\in\overline{\mathcal{B}}_i}\right) \\ \mathbf{W}_i^{n+1} & = & \dfrac{\mathbf{W}_i^{n+1/2} + \mathbf{W}_i^n}{2} + \dfrac{1}{2}\Delta t^n \mathrm{RHS}_2\left(\{\mathbf{W}_j^{n+1/2}\}_{j\in\overline{\mathcal{B}}_i}\right) \end{array}\right. \tag{9.21}$$

9.2.2 Modified Roe Numerical Method

The following numerical method has been proposed by Castro et al. in a series of papers [26–28]. The Modified Roe scheme is based on the non conservative formulation (8.64) which, for the sake of clarity is recalled here:

$$\frac{\partial \hat{\mathbf{W}}}{\partial t} + \hat{\mathbf{A}}_1(\hat{\mathbf{W}})\frac{\partial \hat{\mathbf{W}}}{\partial x} + \hat{\mathbf{A}}_2(\hat{\mathbf{W}})\frac{\partial \hat{\mathbf{W}}}{\partial y} = 0 \tag{9.22}$$

where $\hat{\mathbf{A}}_k (k = 1, 2)$ are defined by (8.65).

A Roe-like numerical flux function

The MR numerical scheme is based on a Roe-like numerical flux function. First of all, let us introduce the Roe averages required for the definition of the MR scheme. Given two generic states $\hat{\mathbf{W}}_L$ and $\hat{\mathbf{W}}_R$, it is possible to define the following Roe averages:

$$\begin{cases} \hat{h}_{LR} = \dfrac{\hat{h}_L + \hat{h}_R}{2} \\ \hat{u}_{LR} = \dfrac{\hat{u}_L\sqrt{\hat{h}_L} + \hat{u}_R\sqrt{\hat{h}_R}}{\sqrt{\hat{h}_L} + \sqrt{\hat{h}_R}} \\ \hat{v}_{LR} = \dfrac{\hat{v}_L\sqrt{\hat{h}_L} + \hat{v}_R\sqrt{\hat{h}_R}}{\sqrt{\hat{h}_L} + \sqrt{\hat{h}_R}} \end{cases} \quad \text{and} \quad \begin{cases} \hat{c}_{LR} = \sqrt{g\hat{h}_{LR}} \\ \widehat{u^2_{LR}} = \dfrac{\hat{u}^2_L\sqrt{\hat{h}_R} + \hat{u}^2_R\sqrt{\hat{h}_L}}{\sqrt{\hat{h}_L} + \sqrt{\hat{h}_R}} \\ \widehat{v^2_{LR}} = \dfrac{\hat{v}^2_L\sqrt{\hat{h}_R} + \hat{v}^2_R\sqrt{\hat{h}_L}}{\sqrt{\hat{h}_L} + \sqrt{\hat{h}_R}} \end{cases} \tag{9.23}$$

Notice that $\widehat{u^2_{LR}}$ is not the square of $\hat{u}_{LR}$ and, thus, $\widehat{u^2_{LR}} \neq \hat{u}^2_{LR}$.

Using (9.23), in [28] a Roe-like matrix $\hat{\mathbf{A}}\left(\hat{\mathbf{W}}_L, \hat{\mathbf{W}}_R, n_{LR}\right)$ is defined as follows:

$$\hat{\mathbf{A}}\left(\hat{\mathbf{W}}_L, \hat{\mathbf{W}}_R, n_{LR}\right) = n_{x,LR}\hat{\mathbf{A}}_{1,LR} + n_{y,LR}\hat{\mathbf{A}}_{2,LR} \tag{9.24}$$

in which $\hat{\mathbf{A}}_{1,LR}$ is a Roe-like matrix derived from $\hat{\mathbf{A}}_1$ given by the following expression:

$$\hat{\mathbf{A}}_{1,LR} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \hat{c}^2_{LR} - \widehat{u^2_{LR}} & 2\hat{u}_{LR} & 0 & -\hat{c}^2_{LR} \\ -\hat{u}_{LR}\hat{v}_{LR} & \hat{v}_{LR} & \hat{u}_{LR} & 0 \\ -\hat{u}_{LR}\,\hat{f}_{1,LR} - \hat{v}_{LR}\,\hat{g}_{1,LR} & \hat{f}_{1,LR} & \hat{g}_{1,LR} & 0 \end{pmatrix} \tag{9.25}$$

with

$$\begin{cases} \hat{f}_{1,LR} \doteq -\xi A_g \dfrac{\hat{u}^2_L + \hat{u}_L\hat{u}_R + \hat{u}^2_R + \widehat{v^2_{LR}}}{\sqrt{\hat{h}_L}\sqrt{\hat{h}_R}} \\ \hat{g}_{1,LR} \doteq -\xi A_g \dfrac{\hat{u}_{LR}\left(\hat{v}_R + \hat{v}_L\right)}{\sqrt{\hat{h}_L}\sqrt{\hat{h}_R}} \end{cases} \tag{9.26}$$

Similarly, the definition of $\hat{\mathbf{A}}_{2,LR}$ is as follows:

$$\hat{\mathbf{A}}_{2,LR} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ -\hat{u}_{LR}\hat{v}_{LR} & \hat{v}_{LR} & \hat{u}_{LR} & 0 \\ \hat{c}^2_{LR} - \widehat{v^2_{LR}} & 0 & 2\hat{v}_{LR} & -\hat{c}^2_{LR} \\ -\hat{u}_{LR}\,\hat{f}_{2,LR} - \hat{v}_{LR}\,\hat{g}_{2,LR} & \hat{f}_{2,LR} & \hat{g}_{2,LR} & 0 \end{pmatrix} \tag{9.27}$$

where

$$\begin{cases} \hat{f}_{2,LR} \doteq -\xi A_g \dfrac{\hat{v}_{LR}\left(\hat{u}_R + \hat{u}_L\right)}{\sqrt{\hat{h}_L}\sqrt{\hat{h}_R}} \\[2ex] \hat{g}_{2,LR} \doteq -\xi A_g \dfrac{\hat{v}_L^2 + \hat{v}_L\hat{v}_R + \hat{v}_R^2 + \widehat{u^2_{LR}}}{\sqrt{\hat{h}_L}\sqrt{\hat{h}_R}} \end{cases} \tag{9.28}$$

Finally, let us introduce $\hat{\mathbf{F}}_{\mathbf{n}}$ and $\left|\hat{\mathbf{A}}\right|$ as follows:

$$\hat{\mathbf{F}}_{\mathbf{n}}(\hat{\mathbf{W}}) \doteq n_x\hat{\mathbf{F}}_1(\hat{\mathbf{W}}) + n_y\hat{\mathbf{F}}_2(\hat{\mathbf{W}}) \tag{9.29}$$

and

$$\left|\hat{\mathbf{A}}\right| = \hat{R}\,|\Lambda|\,\hat{R}^{-1} \tag{9.30}$$

where $|\Lambda|$ is the diagonal matrix whose elements are the absolute values of the eigenvalues of $\hat{\mathbf{A}}$.

By using the definitions (9.23)-(9.30), in [26, 28] the following Roe-like numerical flux function $\hat{\mathcal{F}}(\hat{\mathbf{W}}_j^n, \hat{\mathbf{W}}_i^n, \mathbf{n}_{ij})$ is introduced:

$$\begin{aligned} \hat{\mathcal{F}}(\hat{\mathbf{W}}_i^n, \hat{\mathbf{W}}_j^n, \mathbf{n}_{ij}) = \frac{1}{2}\left(\hat{\mathbf{F}}_{\mathbf{n}_{ij}}(\hat{\mathbf{W}}_i) + \hat{\mathbf{F}}_{\mathbf{n}_{ij}}(\hat{\mathbf{W}}_j)\right)& \\ - \frac{1}{2}|\hat{\mathbf{A}}|(\hat{\mathbf{W}}_j^n, \hat{\mathbf{W}}_i^n, \mathbf{n}_{ij})\left(\hat{\mathbf{W}}_j^n - \hat{\mathbf{W}}_i^n\right)& \end{aligned} \tag{9.31}$$

The complete discrete formulation

Once the Roe-like numerical flux function $\hat{\mathcal{F}}$ is introduced, in [28] a Roe-type scheme for the non conservative system (9.22) is derived. In the particular the semi-discrete formulation of the aforementioned system is:

$$\frac{\partial\hat{\mathbf{W}}_i^n}{\partial t} = \widehat{\mathrm{RHS}}_1\left(\{\hat{\mathbf{W}}_j^n\}_{j\in\mathcal{B}_i}\right) \tag{9.32}$$

where

$$\begin{aligned} \widehat{\mathrm{RHS}}_1\left(\{\hat{\mathbf{W}}_j^n\}_{j\in\mathcal{B}_i}\right) = -\frac{1}{|V_i|}\sum_{j\in N(i)}|\mathit{\Gamma}_{ij}|\Big(\hat{\mathcal{F}}(\hat{\mathbf{W}}_i^n, \hat{\mathbf{W}}_j^n, \mathbf{n}_{ij})& \\ -\frac{1}{2}\hat{\mathbf{B}}_{ij}\left(\hat{\mathbf{W}}_j^n - \hat{\mathbf{W}}_i^n\right)\Big)& \end{aligned} \tag{9.33}$$

and $\hat{\mathbf{B}}_{ij} = \hat{\mathbf{B}}\left(\hat{\mathbf{W}}_i, \hat{\mathbf{W}}_j, n_{ij}\right)$ is defined as follows:

$$\hat{\mathbf{B}}_{ij} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & g\hat{h}_{ij}n_{x,ij} \\ 0 & 0 & 0 & g\hat{h}_{ij}n_{y,ij} \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{9.34}$$

Finally, to obtain an explicit scheme first order accurate in time, the time-discretization of the scheme can be carried out using an explicit Euler scheme so that the resulting numerical method is:

$$\mathbf{W}_i^{n+1} = \mathbf{W}_i^n + \Delta t^n \widehat{\mathrm{RHS}}_1 \left(\{\mathbf{W}_j^n\}_{j \in \mathcal{B}_i} \right) \tag{9.35}$$

Second-order extension

The formulation of the second-order Modified Roe scheme follows a different approach from the one of the SRNH scheme. First, a reconstruction operator $P_i(\mathbf{x})$ is defined at each cell. This reconstruction operator depends on the values of the solution in the cell V_i and in its neighbouring cells, that is:

$$P_i(\mathbf{x}) = P_i(\mathbf{x}, \{\hat{\mathbf{W}}_j\}_{j \in \mathcal{B}_i}) \tag{9.36}$$

Note that, for the sake of simplicity, the specific formulation of the reconstruction operator used in this work will be defined in the next (see (9.43)-(9.48)). Nevertheless, once (9.36) is defined, it is possible to extend the Modified Roe scheme to the second-order accuracy. The semi-discrete formulation of (9.22) is then defined as:

$$\begin{aligned} \frac{\partial \hat{\mathbf{W}}}{\partial t} = & -\frac{1}{|V_i|} \sum_{j \in N(i)} \int_{\Gamma_{ij}} \left(\hat{\mathcal{F}} \left(\hat{\mathbf{W}}_{ij}^-(\sigma), \hat{\mathbf{W}}_{ij}^+(\sigma), \mathbf{n}_{ij} \right) \right. \\ & \left. - \frac{1}{2} \hat{\mathbf{B}}_{ij}(\sigma) \left(\hat{\mathbf{W}}_{ij}^+(\sigma) - \hat{\mathbf{W}}_{ij}^-(\sigma) \right) \right) d\sigma \\ & - \frac{1}{|V_i|} \int_{V_i} \left(\hat{\mathbf{B}}_1(P_i(\mathbf{x})) \frac{\partial P_i}{\partial x}(\mathbf{x}) + \hat{\mathbf{B}}_2(P_i(\mathbf{x})) \frac{\partial P_i}{\partial y}(\mathbf{x}) \right) dV \end{aligned} \tag{9.37}$$

where σ is a point of the interface Γ_{ij}, $\hat{\mathbf{W}}_{ij}^-(\sigma) = P_i(\sigma)$ and $\hat{\mathbf{W}}_{ij}^+(\sigma) = P_j(\sigma)$. Even if it does not clearly appear, the right-hand-side of (9.37) is time-dependent. Indeed, due to (9.36), $P_i(\mathbf{x})$ is a function of the space but also of the solution $\mathbf{W}(\mathbf{x}, t)$. Let us remark a key difference between the SRNH scheme and the Modified Roe one. In the SRNH scheme the second-order extension is dependent only from the extrapolated values of the solution at the cell interfaces. Conversely, in the Modified Roe scheme, due to the non-conservative formulation, the second-order scheme is also function of the extrapolated values of the solution in the interior of the cells.

The integrals in (9.37) are numerically approximated and, in order to preserve the second-order spatial accuracy of the scheme the order of the quadrature formula must be higher than that of the reconstruction operator. In [28] it has been shown that the third-order formula for the line integrals and the barycentre quadrature formula for the surface integrals satisfy both criteria. As a consequence it is possible to approximate the line integrals over Γ_{ij} as:

$$\int_{\Gamma_{ij}} \left(\hat{\mathcal{F}} \left(\hat{\mathbf{W}}_{ij}^{-}(\sigma), \hat{\mathbf{W}}_{ij}^{+}(\sigma), \mathbf{n}_{ij} \right) - \frac{1}{2} \hat{\mathbf{B}}_{ij}(\sigma) \left(\hat{\mathbf{W}}_{ij}^{+}(\sigma) - \hat{\mathbf{W}}_{ij}^{-}(\sigma) \right) \right) d\sigma =$$
$$\sum_{l=1}^{2} \int_{\Gamma_{ij,l}} \left(\hat{\mathcal{F}} \left(\hat{\mathbf{W}}_{ij}^{-}(\sigma), \hat{\mathbf{W}}_{ij}^{+}(\sigma), \mathbf{n}_{ij} \right) - \frac{1}{2} \hat{\mathbf{B}}_{ij}(\sigma) \left(\hat{\mathbf{W}}_{ij}^{+}(\sigma) - \hat{\mathbf{W}}_{ij}^{-}(\sigma) \right) \right) d\sigma \simeq$$
$$\sum_{l=1}^{2} |\Gamma_{ij,l}| \sum_{m=1}^{2} w_{lm} \left(\hat{\mathcal{F}} \left(\hat{\mathbf{W}}_{ij,lm}^{-}, \hat{\mathbf{W}}_{ij,lm}^{+}, \mathbf{n}_{ij,l} \right) - \frac{1}{2} \hat{\mathbf{B}}_{ij,lm} \left(\hat{\mathbf{W}}_{ij,lm}^{+} - \hat{\mathbf{W}}_{ij,lm}^{-} \right) \right) \quad (9.38)$$

where $\hat{\mathbf{W}}_{ij,lm}^{-} = P_i(\sigma_{lm})$, $\hat{\mathbf{W}}_{ij,lm}^{+} = P_j(\sigma_{lm})$ and w_{lm} and σ_{lm} are respectively, the weights and the points of the quadrature rule. For the normalised domain of integration $[-1, 1]$, the corresponding weights and points of the quadrature rule, $\bar{\sigma}_m$ and $\bar{w}_m$, respectively, are reported in Tab. 9.1. Thus, once $\bar{\sigma}_m$ and $\bar{w}_m$ are available, σ_{lm} and w_{lm} are simply obtained by applying the linear transformation which takes $[-1, 1]$ into $\Gamma_{ij,l}$.

m	1	2
$\bar{\sigma}_m$	$-1/\sqrt{3}$	$1/\sqrt{3}$
$\bar{w}_m$	1	1

Table 9.1: Weights and nodes of the quadrature rule for the normalised domain of integration $[-1, 1]$.

Similarly, for the volume integral the barycentre quadrature formula reads as follows:

$$\int_{V_i} \left(\hat{\mathbf{B}}_1(P_i(\mathbf{x})) \frac{\partial P_i}{\partial x}(\mathbf{x}) + \hat{\mathbf{B}}_2(P_i(\mathbf{x})) \frac{\partial P_i}{\partial y}(\mathbf{x}) \right) dV \simeq$$
$$|V_i| \left(\hat{\mathbf{B}}_1(P_i(\mathbf{G}_i)) \frac{\partial P_i}{\partial x}(\mathbf{G}_i) + \hat{\mathbf{B}}_2(P_i(\mathbf{G}_i)) \frac{\partial P_i}{\partial y}(\mathbf{G}_i) \right) \quad (9.39)$$

where we recall that $\mathbf{G}_i$ is the centroid of the i^{th} cell.

By injecting (9.38) and (9.39) in (9.37), the semi-discrete expression of the second-order Modified Roe scheme is:

$$\frac{\partial \hat{\mathbf{W}}_i}{\partial t} = \widehat{\mathrm{RHS}}_2 \left(\{\hat{\mathbf{W}}_j\}_{j \in \mathcal{B}_i} \right) \quad (9.40)$$

where $\widehat{\mathrm{RHS}}_2\left(\{\hat{\mathbf{W}}_j\}_{j\in\overline{\mathcal{B}}_i}\right)$ is defined as follows:

$$\begin{aligned}\widehat{\mathrm{RHS}}_2\left(\{\hat{\mathbf{W}}_j\}_{j\in\overline{\mathcal{B}}_i}\right) &= \hat{\mathbf{B}}_1(P_i(\mathbf{G}_i))\frac{\partial P_i}{\partial x}(\mathbf{G}_i) + \hat{\mathbf{B}}_2(P_i(\mathbf{G}_i))\frac{\partial P_i}{\partial y}(\mathbf{G}_i)\\ &- \frac{1}{|V_i|}\sum_{j\in N(i)}\sum_{l=1}^{2}|\Gamma_{ij,l}|\sum_{m=1}^{2} w_{lm}\left(\hat{\mathcal{F}}\left(\hat{\mathbf{W}}^-_{ij,lm},\hat{\mathbf{W}}^+_{ij,lm},\mathbf{n}_{ij,l}\right)\right.\\ &\left.-\frac{1}{2}\hat{\mathbf{B}}_{ij,lm}\left(\hat{\mathbf{W}}^+_{ij,lm}-\hat{\mathbf{W}}^-_{ij,lm}\right)\right)\end{aligned} \tag{9.41}$$

Note that the dependency from $j\in\overline{\mathcal{B}}_i$ in (9.40) results from the reconstruction operator and, more specifically, from $\hat{\mathbf{W}}^-_{ij,lm}$ and $\hat{\mathbf{W}}^+_{ij,lm}$. To conclude, in order to obtain an explicit scheme second-order accurate in time, the time discretization can be carried out using a second-order TVD Runge-Kutta method [55]. The final expression of the second order numerical scheme is:

$$\begin{cases}\hat{\mathbf{W}}^{n+1/2}_i &= \hat{\mathbf{W}}^n_i + \Delta t^n\widehat{\mathrm{RHS}}_2\left(\{\hat{\mathbf{W}}^n_j\}_{j\in\overline{\mathcal{B}}_i}\right)\\ \hat{\mathbf{W}}^{n+1}_i &= \dfrac{\hat{\mathbf{W}}^{n+1/2}_i+\hat{\mathbf{W}}^n_i}{2} + \dfrac{1}{2}\Delta t^n\widehat{\mathrm{RHS}}_2\left(\{\hat{\mathbf{W}}^{n+1/2}_j\}_{j\in\overline{\mathcal{B}}_i}\right)\end{cases} \tag{9.42}$$

Gradient approximation The numerical scheme (9.42) is completely defined once a suitable formulation for P_i is given. In particular in this work we consider a MUSCL-like reconstruction operator which at time t^n can be expressed as:

$$P_i(\mathbf{x},\{\hat{\mathbf{W}}^n_j\}_{j\in\mathcal{B}_i}) = \hat{\mathbf{W}}^n_i + \widehat{\nabla\hat{\mathbf{W}}^n_i}\left(\mathbf{x}-\mathbf{G_i}\right) \tag{9.43}$$

where $\widehat{\nabla\hat{\mathbf{W}}^n_i}$ is an approximation of the gradient in the i^{th}-cell, possibly taking into account flux-limiters. The same technique described in [28] is considered for the computation of $\widehat{\nabla\hat{\mathbf{W}}^n_i}$. First, a linear approximation of the gradient in each triangle T_j is considered: defining $tn(j,k), k=1,2,3$ the indexes of the nodes belonging to the triangle T_j, this linear approximation is given by:

$$\widehat{\nabla\hat{\mathbf{W}}}_{|T_j} = \sum_{k=1}^{3}\hat{\mathbf{W}}_{tn(j,k)}\nabla\lambda^k_j \tag{9.44}$$

where λ^k_j is the barycentric coordinates associated with the k^{th}-vertex. Once $\widehat{\nabla\hat{\mathbf{W}}}_{|T_j}$ is available, a first-order approximation of $\widehat{\nabla\hat{\mathbf{W}}}_i$ is the following:

$$\widehat{\nabla\hat{\mathbf{W}}^c_i} \simeq \frac{\sum\limits_{j\in t(i)}|T_j|\widehat{\nabla\hat{\mathbf{W}}}_{|T_j}}{\sum\limits_{j\in t(i)}|T_j|} \tag{9.45}$$

The proof that (9.45) is a first-order approximation of the gradient of $\hat{\mathbf{W}}$ for regular solutions can be found in [28].

A common characteristic of high-order methods for hyperbolic systems is to compute unphysical oscillations near discontinuities. A possible way to prevent the generation of those oscillations can be to introduce slope limiters in the reconstructor operator (9.43). More precisely, as in [28], the following approximate gradient is defined through its l^{th}-component:

$$\widehat{\nabla \hat{\mathbf{W}}_{i,l}} = \hat{\varphi}_{i,l} \widehat{\nabla \hat{\mathbf{W}}^c_{i,l}} \tag{9.46}$$

where $\hat{\varphi}_{i,l}$ is the slope limiter function associated with $\widehat{\nabla \hat{\mathbf{W}}^c_{i,l}}$, defined as:

$$\left\{ \begin{array}{rcl} \hat{\varphi}_{i,l} & = & \min\limits_{j \in N(i)} \{ \max\{0, \min\{r_{ij,l}, 1\}\}\} \\ r_{ij,l} & = & \left\{ \begin{array}{ll} \dfrac{\hat{W}^{max}_{i,l} - \hat{W}_{i,l}}{\hat{W}^*_{ij,l} - \hat{W}_{i,l}} & \text{if} \quad \hat{W}^*_{ij,l} - \hat{W}_{i,l} > 0 \\ \dfrac{\hat{W}^{min}_{i,l} - \hat{W}_{i,l}}{\hat{W}^*_{ij,l} - \hat{W}_{i,l}} & \text{if} \quad \hat{W}^*_{ij,l} - \hat{W}_{i,l} < 0 \\ 1 & \text{if} \quad \hat{W}^*_{ij,l} \end{array} \right. \end{array} \right. \tag{9.47}$$

where

$$\hat{W}^{min}_{i,l} = \min_{j \in \mathcal{B}_i} \{\hat{W}_{i,l}\}, \quad \hat{W}^{max}_{i,l} = \max_{j \in \mathcal{B}_i} \{\hat{W}_{i,l}\}$$

and

$$\hat{W}^*_{ij} = \hat{W}_{ij} + \widehat{\nabla \hat{\mathbf{W}}^c_i} (\mathbf{c}_{ij} - \mathbf{G}_i) \tag{9.48}$$

$\mathbf{c}_{ij}$ being the middle point of the segment connecting the i^{th} and j^{th} nodes.

9.3 Implicit time advancing

In this section the issue of generating an implicit scheme, starting from its explicit counterpart, is addressed. Initially only first-order numerical schemes are considered, then the second-order extensions are discussed.

9.3.1 First-order schemes

Generally speaking, the implicit counterpart of a first order explicit Euler method is obtained by considering the right-hand-side term as a function of the solution at time $n+1$ instead of n. Hence, a fully implicit first-order version of the schemes in equations (9.14) and (9.35) can be obtained by using the following backward Euler method:

$$\left\{ \begin{array}{rcl} \mathbf{W}^{n+1}_i - \Delta t^n \text{RHS}_1 \left(\{\mathbf{W}^{n+1}_j\}_{j \in \mathcal{B}_i} \right) & = & \mathbf{W}^n_i \\ \hat{\mathbf{W}}^{n+1}_i - \Delta t^n \widehat{\text{RHS}_1} \left(\{\hat{\mathbf{W}}^{n+1}_j\}_{j \in \mathcal{B}_i} \right) & = & \hat{\mathbf{W}}^n_i \end{array} \right. \tag{9.49}$$

However, from a practical point of view this would require the solution of a large non-linear system of equations at each time step. As already pointed out in Part. I, the computational cost for this operation is in general not affordable in practical applications and, in general, significantly overcomes any advantage that an implicit scheme could have with respect to its explicit counterpart. A common technique to overcome this difficulty is to linearise the numerical scheme, that is to find an approximation of $\mathrm{RHS}_1\left(\{\mathbf{W}_j^{n+1}\}_{j\in\mathcal{B}_i}\right)$ or $\widehat{\mathrm{RHS}}_1\left(\{\hat{\mathbf{W}}_j^{n+1}\}_{j\in\mathcal{B}_i}\right)$ in the form:

$$\begin{cases} \mathrm{RHS}_1\left(\{\mathbf{W}_j^{n+1}\}_{j\in\mathcal{B}_i}\right) & \simeq & \mathrm{RHS}_1\left(\{\mathbf{W}_j^{n}\}_{j\in\mathcal{B}_i}\right) + \sum\limits_{j\in\mathcal{B}_i} \mathcal{D}_{ij}\Delta^n\mathbf{W}_j \\ \widehat{\mathrm{RHS}}_1\left(\{\hat{\mathbf{W}}_j^{n+1}\}_{j\in\mathcal{B}_i}\right) & \simeq & \widehat{\mathrm{RHS}}_1\left(\{\hat{\mathbf{W}}_j^{n}\}_{j\in\mathcal{B}_i}\right) + \sum\limits_{j\in\mathcal{B}_i} \hat{\mathcal{D}}_{ij}\Delta^n\hat{\mathbf{W}}_j \end{cases} \tag{9.50}$$

where $\Delta^n(\,\cdot\,) = (\,\cdot\,)^{n+1} - (\,\cdot\,)^n$ and $\mathcal{D}_{ij}$ and $\hat{\mathcal{D}}_{ij}$ are matrices depending on the solution in the i^{th} and j^{th} cells. Using this approximation, the following linear system must be solved at each time step for the SRNH scheme:

$$\frac{\Delta^n\mathbf{W}_i}{\Delta t^n} - \sum_{j\in\mathcal{B}_i} \mathcal{D}_{ij}\Delta^n\mathbf{W}_j = \mathrm{RHS}_1\left(\{\mathbf{W}_j^{n}\}_{j\in\mathcal{B}_i}\right) \tag{9.51}$$

and similarly for the Modified Roe scheme:

$$\frac{\Delta^n\hat{\mathbf{W}}_i}{\Delta t^n} - \sum_{j\in\mathcal{B}_i} \hat{\mathcal{D}}_{ij}\Delta^n\hat{\mathbf{W}}_j = \widehat{\mathrm{RHS}}_1\left(\{\hat{\mathbf{W}}_j^{n}\}_{j\in\mathcal{B}_i}\right) \tag{9.52}$$

The implicit linearised scheme is completely defined once a suitable definition for the matrices $\mathcal{D}_{ij}$ is given. If the right hand side is differentiable, a common choice is to use the Jacobian matrices, hence:

$$\begin{cases} \mathcal{D}_{ij} & \simeq & \dfrac{\partial \mathrm{RHS}_1\left(\{\mathbf{W}_l^{n}\}_{l\in\mathcal{B}_i}\right)}{\partial \mathbf{W}_j^n} \\ \\ \hat{\mathcal{D}}_{ij} & \simeq & \dfrac{\partial \widehat{\mathrm{RHS}}_1\left(\{\hat{\mathbf{W}}_l^{n}\}_{l\in\mathcal{B}_i}\right)}{\partial \hat{\mathbf{W}}_j^n} \end{cases} \tag{9.53}$$

Nevertheless, it is not always possible nor convenient to exactly compute the Jacobian matrices. In fact, it is not unusual to have some lack of differentiability of the numerical flux functions or in the source term. Furthermore in the particular case of the SRNH scheme (9.14) the difficulty in using linearization (9.53) is increased by the fact that the scheme is composed by a predictor and a corrector stage. This problem has been solved herein through the use of the automatic differentiation software Tapenade [59]. The operational principle of an automatic differentiation software is as follows: given the source code of a routine which computes the function $y = F(x)$, the automatic differentiation software generates a new source code which compute the analytical derivative of the original program. In practise, each time the original program performs some operation, the differentiated program performs additional operations dealing with the differential values. For example, if the original program, at some time executes the following instruction on variables a, b, c:

$$a = b \cdot c \tag{9.54}$$

then the differentiated program computes also the differentials da, db, dc of these variable [59]:

$$da = db \cdot c + b \cdot dc \tag{9.55}$$

Through an automatic differentiation software it is possible to quickly implement an implicit linearised scheme of the form (9.51) or (9.52), once a routine which computes the explicit flux function is available. As a consequence using an automatic differentiation tool, starting from a first-order explicit method, it is possible to automatically compute the matrices $\mathcal{D}_{ij}$ (or $\hat{\mathcal{D}}_{ij}$) and then implement the linearised implicit methods (9.51) and (9.52) without additional modifications.

9.3.2 Second-order schemes

A second-order implicit scheme can be obtained from its explicit counterpart by using the same approach as described for the first-order schemes in section 9.3.1. Therefore the second-order implicit SRNH scheme would be:

$$\left\{ \begin{array}{lcl} \mathbf{W}_i^{n+1/2} - \Delta t^n \mathrm{RHS}_2 \left(\{\mathbf{W}_j^{n+1/2}\}_{j \in \overline{\mathcal{B}}_i} \right) & = & \mathbf{W}_i^n \\ \mathbf{W}_i^{n+1} - \dfrac{1}{2} \Delta t^n \mathrm{RHS}_2 \left(\{\mathbf{W}_j^{n+1}\}_{j \in \overline{\mathcal{B}}_i} \right) & = & \dfrac{\mathbf{W}_i^{n+1/2} + \mathbf{W}_i^n}{2} \end{array} \right. \tag{9.56}$$

and similarly for the Modified Roe scheme:

$$\left\{ \begin{array}{lcl} \hat{\mathbf{W}}_i^{n+1/2} - \Delta t^n \widehat{\mathrm{RHS}}_2 \left(\{\hat{\mathbf{W}}_j^{n+1/2}\}_{j \in \overline{\mathcal{B}}_i} \right) & = & \hat{\mathbf{W}}_i^n \\ \hat{\mathbf{W}}_i^{n+1} - \dfrac{1}{2} \Delta t^n \widehat{\mathrm{RHS}}_2 \left(\{\hat{\mathbf{W}}_j^{n+1}\}_{j \in \overline{\mathcal{B}}_i} \right) & = & \dfrac{\hat{\mathbf{W}}_i^{n+1/2} + \hat{\mathbf{W}}_i^n}{2} \end{array} \right. \tag{9.57}$$

The application of this method would require the solution of two non linear systems of equations at each time step, thus dramatically increasing the computational costs with respect to the explicit version. An alternative approach, generally more efficient in terms of computational costs, is to use a second-order backward differentiation formula in time:

$$\begin{cases} \dfrac{(1+2\tau)\mathbf{W}_i^{n+1} - (1+\tau)^2\mathbf{W}_i^n + \tau^2\mathbf{W}_i^{n-1}}{\Delta^n t\,(1+\tau)} - \mathrm{RHS}_2\left(\{\mathbf{W}_j^{n+1}\}_{j\in\overline{\mathcal{B}}_i}\right) = 0 \\ \dfrac{(1+2\tau)\hat{\mathbf{W}}_i^{n+1} - (1+\tau)^2\hat{\mathbf{W}}_i^n + \tau^2\hat{\mathbf{W}}_i^{n-1}}{\Delta^n t\,(1+\tau)} - \widehat{\mathrm{RHS}}_2\left(\{\hat{\mathbf{W}}_j^{n+1}\}_{j\in\overline{\mathcal{B}}_i}\right) = 0 \end{cases} \tag{9.58}$$

where $\tau = \dfrac{\Delta^n t}{\Delta^{n-1} t}$.

Similarly to the first-order case, a linearization of $\mathrm{RHS}_2\left(\{\mathbf{W}_j^{n+1}\}_{j\in\overline{\mathcal{B}}_i}\right)$ and $\widehat{\mathrm{RHS}}_2\left(\{\hat{\mathbf{W}}_j^{n+1}\}_{j\in\overline{\mathcal{B}}_i}\right)$ must be carried out in order to avoid the solution of a nonlinear system at each time step. Clearly, the same approach as for the first-order scheme could be considered, that is to find an approximation of $\mathrm{RHS}_2\left(\{\mathbf{W}_j^{n+1}\}_{j\in\overline{\mathcal{B}}_i}\right)$ in the form:

$$\begin{cases} \mathrm{RHS}_2\left(\{\mathbf{W}_j^{n+1}\}_{j\in\overline{\mathcal{B}}_i}\right) \simeq \mathrm{RHS}_2\left(\{\mathbf{W}_j^{n}\}_{j\in\overline{\mathcal{B}}_i}\right) + \sum\limits_{j\in\overline{\mathcal{B}}_i} \mathcal{D}_{2,ij}\Delta^n\mathbf{W}_j \\ \widehat{\mathrm{RHS}}_2\left(\{\hat{\mathbf{W}}_j^{n+1}\}_{j\in\overline{\mathcal{B}}_i}\right) \simeq \widehat{\mathrm{RHS}}_2\left(\{\hat{\mathbf{W}}_j^{n}\}_{j\in\overline{\mathcal{B}}_i}\right) + \sum\limits_{j\in\overline{\mathcal{B}}_i} \hat{\mathcal{D}}_{2,ij}\Delta^n\hat{\mathbf{W}}_j \end{cases} \tag{9.59}$$

However, as discussed in Part. I, the linearization for the second-order accurate fluxes and the solution of the resulting linear system implies significant computational costs and memory requirements. This is a consequence of the more complex expression of second-order schemes with respect to their first-order counterparts (compare (9.13) with (9.19) and (9.33) with (9.42) and, in particular, of the larger stencil of the second order flux function (i. e. considering an uniform triangular grid the set $\mathcal{B}_i$ contains 7 nodes while 19 nodes are in $\overline{\mathcal{B}}_i$). In order to reduce the computational costs, an alternative approach is to use a defect-correction technique [81], as done in the first part of the thesis (see Chap. 3). The defect-correction iterations write as:

$$\begin{cases} \mathcal{W}^0 = \mathbf{W}^n \\ \mathcal{L}_i^s\Delta^s\mathcal{W}_i - \sum\limits_{j\in N(i)} \mathcal{D}_{ij}^s\Delta^s\mathcal{W}_j = \mathcal{C}_i^s \quad s = 0,\cdots,r-1 \\ \mathbf{W}^{n+1} = \mathcal{W}^r \end{cases} \tag{9.60}$$

in which:

$$\begin{cases} \mathcal{L}_i^s = \dfrac{(1+2\tau)}{\Delta^n t\,(1+\tau)}\, I - \mathcal{D}_{ii}^s \\[2ex] \mathcal{D}_{ij}^s = \mathcal{D}_{ij}\left(\{\mathcal{W}_l^s\}_{l\in\mathcal{B}_i}\right) \\[2ex] \mathcal{C}_i^s = -\left(\dfrac{(1+2\tau)\mathcal{W}_i^s - (1+\tau)^2\mathbf{W}_i^n + \tau^2\mathbf{W}_i^{n-1}}{\Delta^n t\,(1+\tau)}\right) + \mathrm{RHS}_2\left(\{\mathcal{W}_j^s\}_{j\in\overline{\mathcal{B}}_i}\right) \end{cases} \tag{9.61}$$

$\mathcal{D}_{ij}$ being the generic matrices of the approximation (9.50); r is typically chosen equal to 2. Indeed, it can be shown [81, 103] that only one defect-correction iteration is theoretically needed to reach a second-order accuracy, while few additional iterations (one or two) can improve the robustness. An expression similar to (9.60)-(9.61), omitted for the sake of brevity, can be derived for the Modified Roe scheme by simply considering $\hat{\mathbf{W}}$ and $\widehat{\mathrm{RHS}}_2\left(\{\hat{\mathbf{W}}_j\}_{j\in\overline{\mathcal{B}}_i}\right)$ instead of $\mathbf{W}$ and $\mathrm{RHS}_2\left(\{\mathbf{W}_j\}_{j\in\overline{\mathcal{B}}_i}\right)$.

Chapter 10

Numerical Simulations for 1D Sediment Transport Problems

In this chapter a 1D numerical test-case is considered in order to compare the different numerical approaches explicit or implicit, as well as 1^{st} or 2^{nd}-order accurate, described in Chap. 9. The structure of this chapter is as follows: first the set-up of the considered benchmark is described in Sec. 10.1, then the numerical experiments corresponding to different speeds of interaction between the flow and the bed are shown in Secs. 10.2-10.5.

10.1 1D numerical experiments

The proposed test-case corresponds to a standard problem already considered in several papers (see e.g. [8, 26]). It is a sediment transport problem in a channel of length $l = 1000$ m with a non constant bottom profile. The initial bottom topography is given by a hump shape function, that is:

$$\left\{ \begin{array}{rcl} Z(0,x) & = & \left\{ \begin{array}{ll} 0.1 + \sin^2\left(\dfrac{(x-300)\pi}{200}\right) & \text{if } 300 \leq x \leq 500 \\ 0.1 & \text{elsewhere} \end{array} \right. \\ h(0,x) & = & 10 - Z(0,x) \\ u(0,x) & = & \dfrac{10}{h(0,x)} \end{array} \right. \tag{10.1}$$

in which all the variables are in SI units.

Two different uniform grids are considered for the discretization of the computational domain: a coarse grid, GR1, which is composed by 100 cells and a refined one, GR2, composed by 250 cells. The results computed by first and second-order schemes, both explicit and implicit, are compared in terms of accuracy and computational costs. Four sets of simulations have been carried out, characterised by A_g equal to $0.001, 0.01, 0.1$ and 1 respectively, A_g being the free parameter in the Grass model (8.11). Each value of the parameter A_g corresponds to a specific speed of interaction between the flow and the bedload and, as a consequence, to a specific time scale for the evolution of the bottom topography. The first value corresponds to a weak interaction between the flow and the bedload, the last to a strong one, while the other values to intermediate conditions. Therefore, in order to observe significant variations of the bed profile, the simulations corresponding to small values of A_g are advanced in time for longer periods, as shown in table 10.1. Note, in particular, that in the case of weak interaction between the bed and the flow ($A_g = 0.001$) a stiff problem is obtained, since the characteristics time scales for the evolution of the bed and of the flow are different. The CFL number used in the following simulations is defined as follows:

A_g	1	0.1	0.01	0.001
Simulation time	700	7000	50000	500000

Table 10.1: Final simulation time (seconds) for the considered values of A_g.

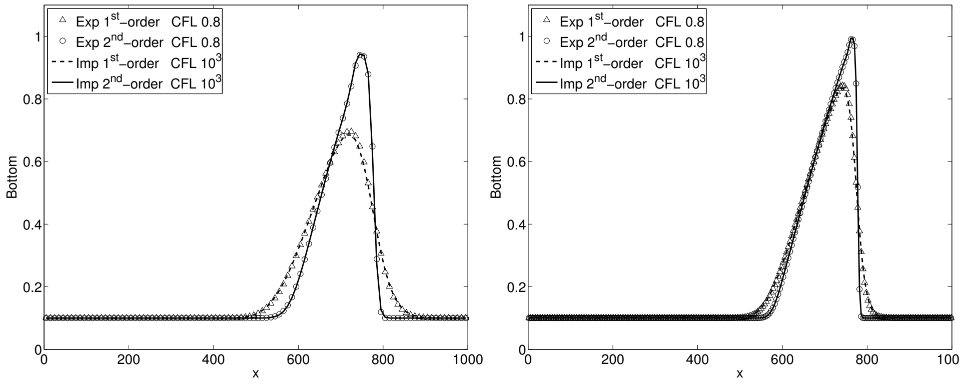
$$\text{CFL} = \frac{\lambda_{max}\Delta t}{\Delta x} \quad \text{where} \quad \lambda_{max} = \max_{i,j,l} \lambda_{ij}^{l} \tag{10.2}$$

and λ_{ij}^{l} is the l^{th}-eigenvalue of the matrix $\hat{\mathbf{A}}\left(\hat{\mathbf{W}}_i, \hat{\mathbf{W}}_j, n_{ij}\right)$ in (9.24) or of the matrix $\mathbf{A}_\eta(\overline{\mathbf{U}}_{ij}^{n})$ in (9.7) for, respectively, the MR scheme and the SRNH scheme. All the results and CPU times shown in the following are at the final instant of each simulation. All the simulations have been carried out on a 3 GHz Intel Pentium 4 processor with 2Gb RAM.

10.2 Slow speed of interaction between bedload and water flow

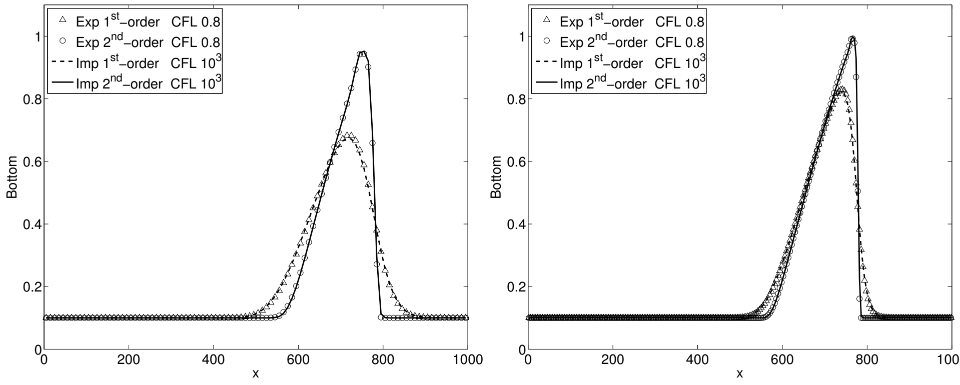
The first test-case is characterised by a slow speed of interaction between the bedload and the water flow, corresponding to the case $A_g = 0.001$ in Tab. 10.1. Figure 10.1 shows a comparison of the results obtained by means of the explicit version of the SRNH scheme at CFL$= 0.8$ with those of the implicit one at CFL$= 1000$, both for 1^{st} and 2^{nd}-order accuracy. Figure 10.2 shows the same comparison for the MR scheme. In both cases, there is practically no difference between the solutions obtained with the implicit and explicit version of the schemes, while the results obtained at 1^{st}-order of accuracy significantly differ from the 2^{nd}-order ones. Also note that the results shown in figure 10.1 and 10.2 for the second-order implicit scheme are obtained by using only one iteration of the Defect Correction method.

Furthermore, note the similarity between the solutions computed by the SRNH and the MR schemes. As for the efficiency, both the SRNH and the MR implicit scheme seem to be unconditionally stable: the CFL has been increased up to 10^5 while obtaining stable solutions. However, the accuracy of the results obviously decreases if the time step is too large. As it is shown



(a) 100 Cells **(b)** 250 Cells

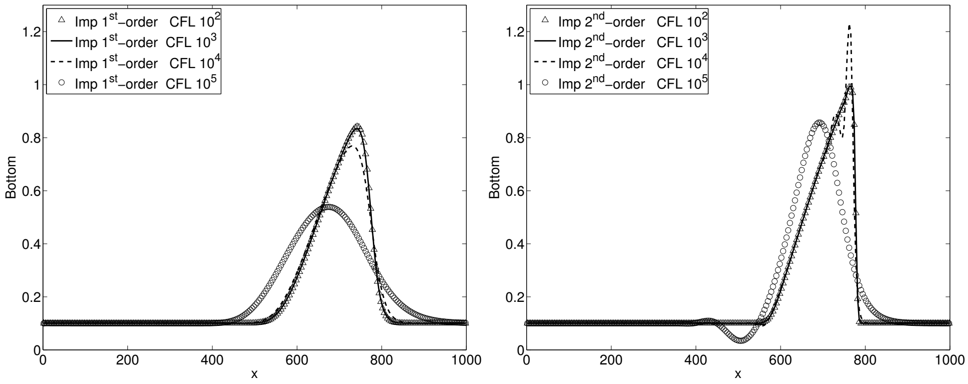
Figure 10.1: Bottom height results computed by the SRNH scheme and $A_g = 10^{-3}$: comparison between explicit and implicit, 1^{st} and 2^{nd}-order formulations.



(a) 100 Cells **(b)** 250 Cells

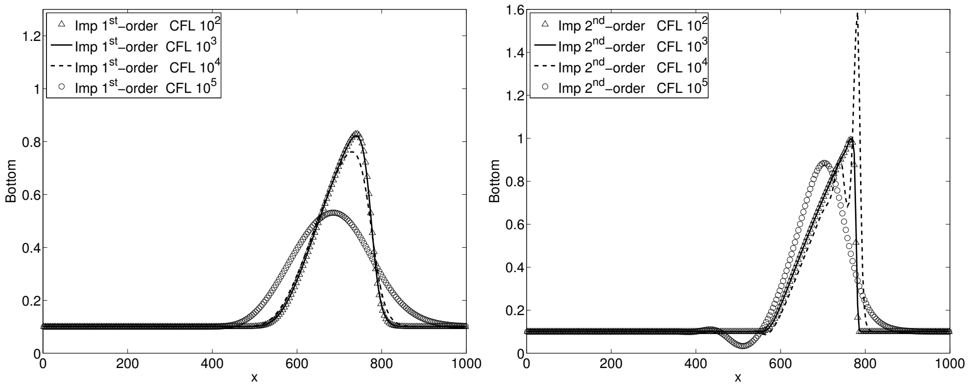
Figure 10.2: Bottom height results computed by the MR scheme and $A_g = 10^{-3}$: comparison between explicit and implicit, 1^{st} and 2^{nd}-order formulations.

in figures 10.3 and 10.4, in this case the quality of the results computed by the implicit schemes is not significantly deteriorated up to a CFL number of 1000. Increasing the CFL above 1000 significantly reduces the accuracy of the implicit schemes and, in the specific case of the second-order scheme, can introduce unphysical oscillations, as shown in figures 10.3 and 10.4. These oscillations can be eliminated or at least largely reduced by increasing the number of Defect Correction iterations, as shown in figure 10.5. Two Defect Correction iterations not only eliminate the oscillations but they also increase



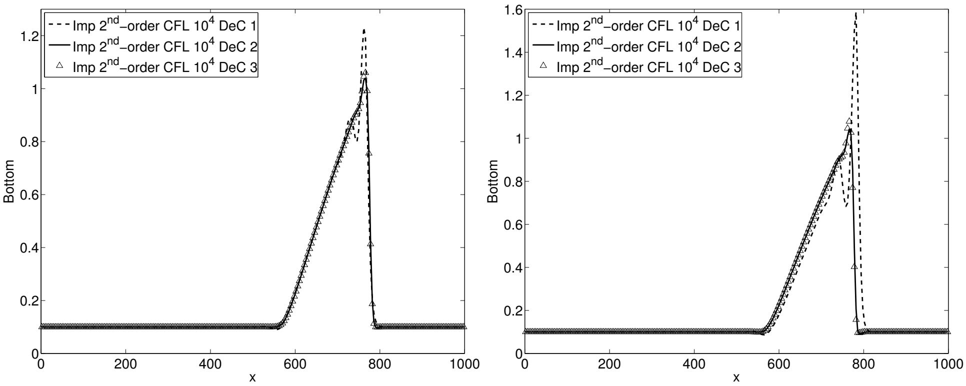
(a) First Order **(b)** Second Order

Figure 10.3: Bottom height obtained with the SRNH implicit scheme at different values of CFL, $A_g = 10^{-3}$, 250 Cells.



(a) First Order **(b)** Second Order

Figure 10.4: Bottom height obtained with the MR implicit scheme at different values of CFL, $A_g = 10^{-3}$, 250 Cells.



(a) SRNH scheme **(b)** MR scheme

Figure 10.5: Effect of the number of Defect Correction iterations on the refined grid using a CFL value of 10^4: $A_g = 10^{-3}$.

the accuracy of the numerical method and an additional iteration leads to a further slight improvement. For all the simulations considered in this work, no additional accuracy improvement was observed by considering a number of Defect Correction larger than 3. For both the SRNH scheme and the MR one, the results computed using 3 Defect Correction iterations and CFL=10^4 are equivalent, in terms of accuracy, to those computed using 1 DeC iteration at CFL= 10^3.

	MR scheme		SRNH scheme	
Method	GR1	GR2	GR1	GR2
Explicit 1^{st} order, CFL= 0.8	$57.0s$	$355.6s$	$165.6s$	$1025s$
Explicit 2^{nd} order, CFL= 0.8	$136.7s$	$852.2s$	$247.2s$	$1537s$
Implicit 1^{st} order, CFL= 10^3	$0.3s$	$2.0s$	$0.4s$	$2.6s$
Implicit 2^{nd} order, CFL= 10^3, 1 DeC	$0.3s$	$2.1s$	$0.5s$	$3.0s$
Implicit 2^{nd} order, CFL= 10^4, 3 DeC	$0.1s$	$0.6s$	$0.1s$	$0.8s$

Table 10.2: CPU time required (seconds), case $A_g = 10^{-3}$.

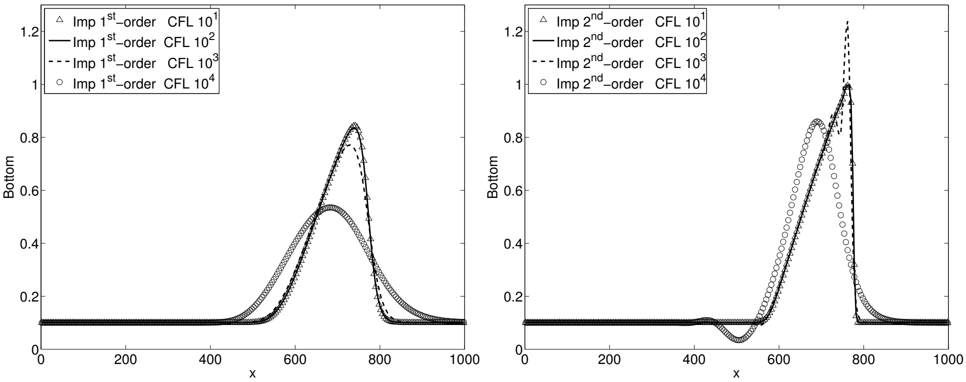
As for the computational costs, table 10.2 shows that already at CFL= 1000 the gain in CPU time obtained with the implicit scheme is large, for both the SRNH and MR schemes at both 1^{st} and 2^{nd}-order of accuracy. The CPU gain obtained with the implicit scheme is significantly larger for 2^{nd}-order accuracy. Indeed, when the implicit formulation is used, there are not significant differences, in terms of CPU time, between the 1^{st} and 2^{nd}-order simulations. Instead in the explicit case an important computational cost increase is observed to reach 2^{nd}-order accuracy: the second-order approach is about 1.5 times more expensive than the first-order approach for the SRNH scheme and about 2.4 times for the MR one. As a consequence, already at CFL= 1000 using 1 DeC iteration the 2^{nd}-order implicit approach is more than 400 times faster than the explicit one, for both the MR and SRNH schemes. The CPU gain of the 2^{nd}-order implicit approach can be further increased considering 3 DeC iterations and CFL= 10^4. However using more than one DeC iteration increases the memory requirement of the numerical method and the complexity of the algorithm. In the case of slow speed of interaction the gain in CPU time using 1 DeC iteration is remarkable so the use of 3 DeC iterations is probably not necessary. However in the following it will be shown that there are conditions in which using 3 DeC iterations can be useful.

Furthermore, comparing the two different numerical schemes, the numerical formulation based on the modified Roe scheme seems to be more robust than the SRNH one (see in particular the second-order results in Figs. 10.3-10.5). Moreover the MR formulation is also less demanding from a computational point of view even if, as shown in Tab. 10.2, the differences between the implicit formulations are reduced.

10.3 Slow/Intermediate speed of interaction between bedload and water flow

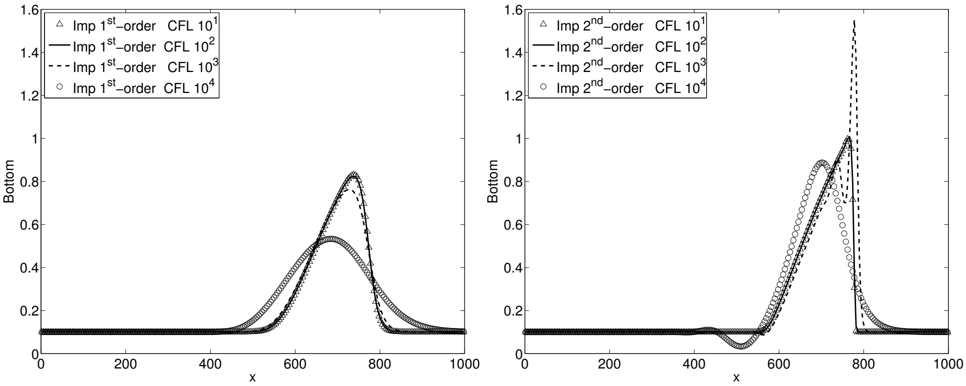
The increase of the A_g value allows problems with faster speed of interaction between the bedload and the water flow to be considered. A first intermediate case has been performed by taking $A_g = 10^{-2}$. Also in this case, the implicit schemes seem to be unconditionally stable: the CFL has been increased up to 10^4 and the simulations remained stable (with the aforementioned CFL number only 6 iterations are required on the coarse grid to complete the tests). Nevertheless, as shown in figures 10.6 and 10.7, the upper bound for the CFL number of the 1^{st} and 2^{nd}-order implicit approaches is reduced of one order of magnitude, with respect to the previous test-case, for both the SRNH scheme and the MR one.

Figures 10.8 and 10.9 show a comparison of the results obtained by means of the explicit versions of the SRNH and MR schemes at CFL= 0.8 with those of the respective implicit versions at CFL= 100, both for 1^{st} and 2^{nd}-order accuracy for the case $A_g = 10^{-2}$. Considerations similar to the slow speed of interaction case are still valid for this intermediate speed of interaction: there is practically no difference between the solutions obtained with the implicit and explicit schemes, while the results obtained at 1^{st}-order of accuracy significantly differ from the 2^{nd}-order ones. Similarly to the case $A_g = 10^{-3}$, an



(a) First Order **(b)** Second Order

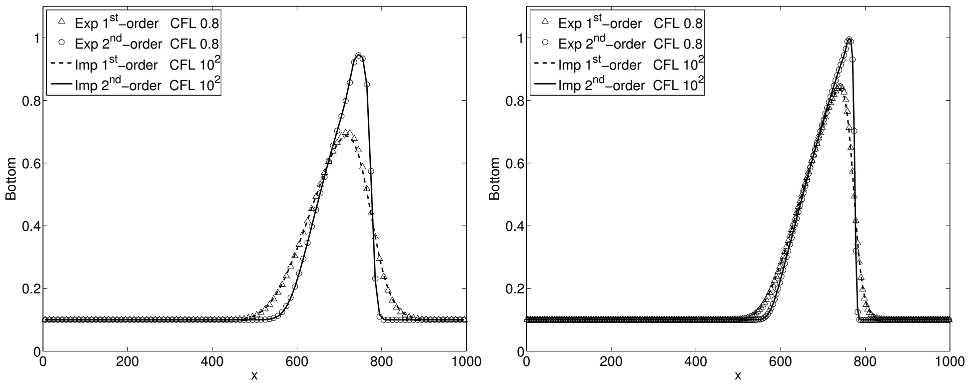
Figure 10.6: Bottom height obtained with the SRNH implicit scheme at different values of CFL, $A_g = 10^{-2}$, 250 Cells.



(a) First Order **(b)** Second Order

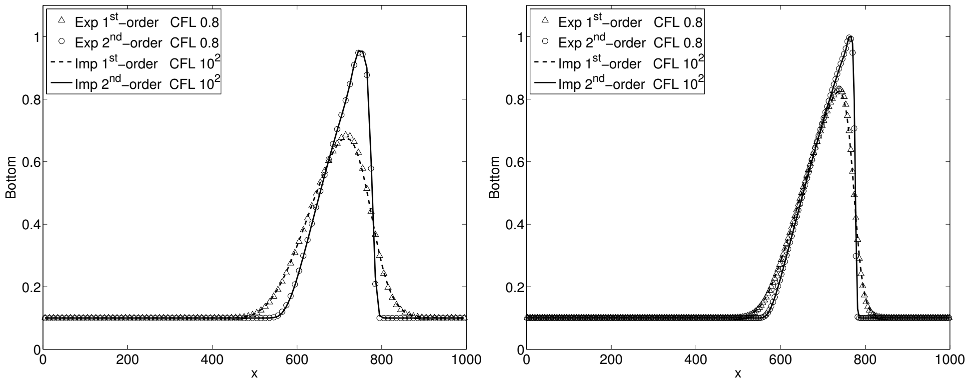
Figure 10.7: Bottom height obtained with the MR implicit scheme at different values of CFL, $A_g = 10^{-2}$, 250 Cells.

increase of the CFL number above those values can induce unphysical oscillations in the solutions computed by the second-order implicit scheme: these oscillations can be reduced with additional Defect Correction iterations, thus increasing the accuracy. As in the weak interaction case, the results computed using 1 Defect Correction iteration and CFL= 100 are equivalent to the result using CFL= 1000 and 3 DeC iterations. As for computational costs, table 10.3 shows that also when considering an intermediate speed of interaction between the bedload and the water flow, in spite of the lower upper bound on the CFL number, there is a gain in CPU time obtained



(a) 100 Cells **(b)** 250 Cells

Figure 10.8: Bottom height results computed by the SRNH scheme and $A_g = 10^{-2}$: comparison between explicit and implicit, 1^{st} and 2^{nd}-order formulations.



(a) 100 Cells **(b)** 250 Cells

Figure 10.9: Bottom height results computed by the MR scheme and $A_g = 10^{-2}$: comparison between explicit and implicit, 1^{st} and 2^{nd}-order formulations.

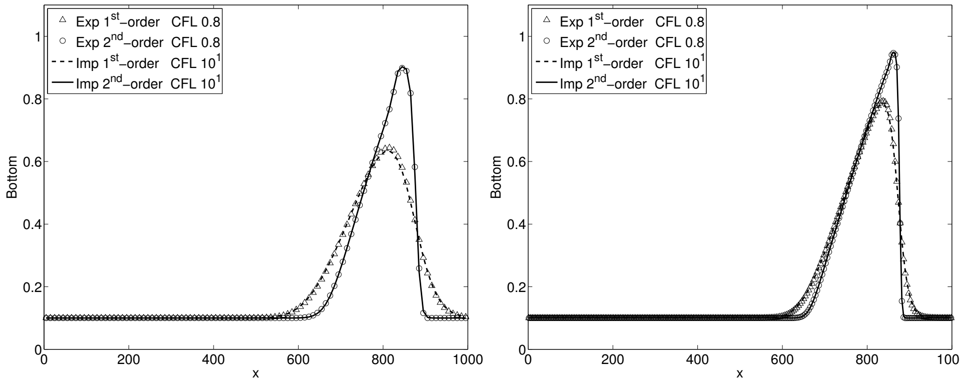
with the implicit scheme, both at 1^{st} and 2^{nd}-order of accuracy. Even if the CPU time gain is not as large as the one in the previous case, the 2^{nd}-order implicit approach is still more than 40 times faster than the explicit one and more than 100 hundred time faster when 3 DeC iterations and CFL$= 1000$ are considered.

	MR scheme		SRNH scheme	
Method	GR1	GR2	GR1	GR2
Explicit 1^{st} order, CFL= 0.8	$5.8s$	$35.7s$	$16.7s$	$103.3s$
Explicit 2^{nd} order, CFL= 0.8	$13.6s$	$85.6s$	$24.8s$	$153.7s$
Implicit 1^{st} order, CFL= 10^2	$0.3s$	$2.0s$	$0.4s$	$2.6s$
Implicit 2^{nd} order, CFL= 10^2, 1 DeC	$0.3s$	$2.1s$	$0.5s$	$3.0s$
Implicit 2^{nd} order, CFL= 10^3, 3 DeC	$0.1s$	$0.6s$	$0.1s$	$0.8s$

Table 10.3: CPU time required (seconds), case $A_g = 10^{-2}$.

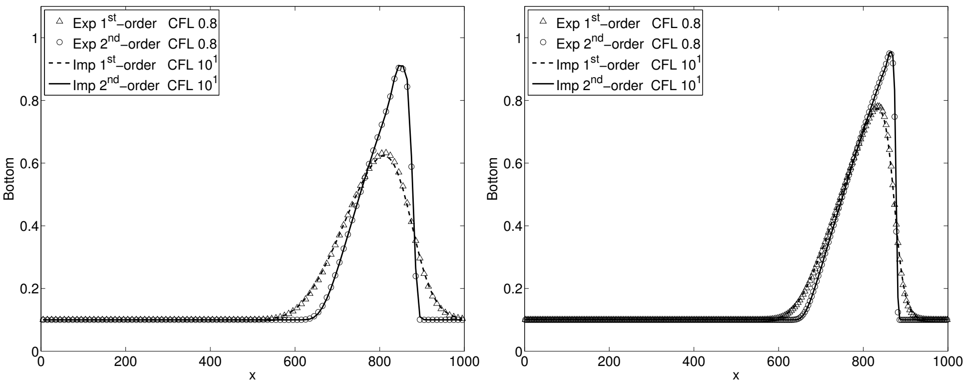
10.4 Intermediate/Fast speed of interaction between bedload and water flow

This case corresponds to $A_g = 10^{-1}$. The conclusions drawn for the previous case hold also for this case: the results obtained through the SRNH and the MR schemes are practically the same. The implicit schemes at CFL = 10 and the explicit scheme at CFL=0.8 give similar results, as shown in figures 10.10 and 10.11, and the implicit schemes seem to be unconditionally stable. However, in order to prevent loss in accuracy, in this case a further reduction of the maximum time step must be considered. Fig. 10.12 shows that the upper limit for the CFL number must be set to 10 in this case. Notwithstanding this lower limitation, the implicit scheme, both for SRNH and the MR scheme, is still preferable in terms of CPU time, even if the gain is reduced with respect to the previous cases, as shown in table 10.4. In this



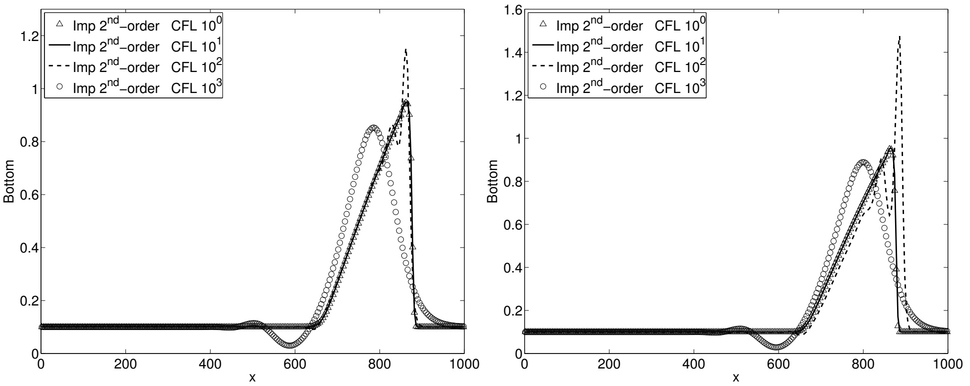
(a) 100 Cells **(b)** 250 Cells

Figure 10.10: Bottom height results computed by the SRNH scheme and $A_g = 10^{-1}$: comparison between explicit and implicit, 1^{st} and 2^{nd}-order formulations.



(a) 100 Cells **(b)** 250 Cells

Figure 10.11: Bottom height results computed by the MR scheme and $A_g = 10^{-1}$: comparison between explicit and implicit, 1^{st} and 2^{nd}-order formulations.



(a) SRNH scheme **(b)** MR scheme

Figure 10.12: Bottom height obtained with the implicit scheme at different values of CFL, $A_g = 10^{-1}$, 250 Cells.

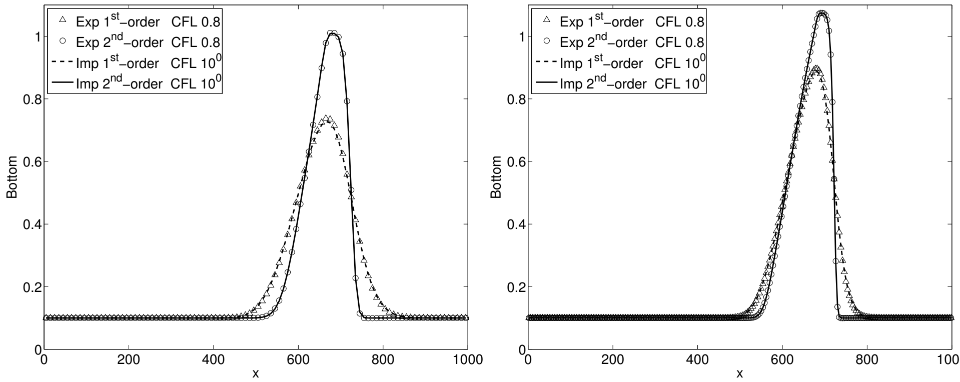
case the 2^{nd}-order implicit approach with 1 DeC iteration is roughly 4 times faster than its explicit counterpart. As a consequence, for this value of A_g it could be interesting to use 3 DeC iterations with a consequent CPU time gain of about 12 times for both the MR and SRNH schemes.

	MR scheme		SRNH scheme	
Method	GR1	GR2	GR1	GR2
Explicit 1^{st} order, CFL$= 0.8$	$0.8s$	$5.1s$	$2.4s$	$14.6s$
Explicit 2^{nd} order, CFL$= 0.8$	$2.0s$	$12.1s$	$3.6s$	$21.8s$
Implicit 1^{st} order, CFL$= 10^1$	$0.5s$	$2.9s$	$0.6s$	$3.7s$
Implicit 2^{nd} order, CFL$= 10^1$, 1 DeC	$0.5s$	$3.0s$	$0.7s$	$4.2s$
Implicit 2^{nd} order, CFL$= 10^2$, 3 DeC	$0.1s$	$0.9s$	$0.2s$	$1.1s$

Table 10.4: CPU time required (seconds), case $A_g = 10^{-1}$.

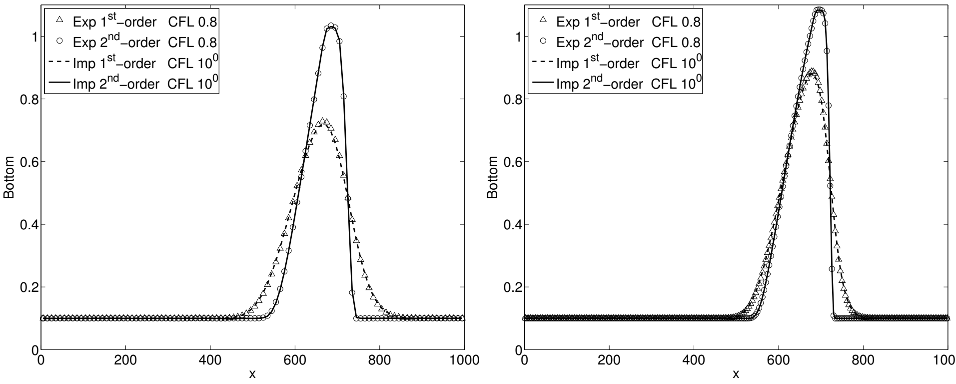
10.5 Fast speed of interaction between bedload and water flow

The fast speed of interaction case, $A_g = 1$ is considered in this section. Figures 10.13 and 10.14 show a comparison of the results obtained by means of the explicit versions of the SRNH and MR schemes at CFL$= 0.8$ with those of the implicit versions at CFL$= 1$, both for 1^{st} and 2^{nd}-order accuracy. As for the previous case there is no difference between the solutions obtained with the implicit and explicit schemes. In this case of fast interaction between bedload and water flow, the quality of the results for the implicit schemes imposes a maximum CFL number equal to 1, although the implicit schemes seem again to be unconditionally stable. As a consequence, as it is shown in table 10.5, in this test case the computational cost of the first-order implicit approach is larger than for the explicit one. Also the 2^{nd}-order implicit



(a) 100 Cells **(b)** 250 Cells

Figure 10.13: Bottom height results computed by the SRNH scheme and $A_g = 10^{-0}$: comparison between explicit and implicit, 1^{st} and 2^{nd}-order formulations.



(a) 100 Cells **(b)** 250 Cells

Figure 10.14: Bottom height results computed by the MR scheme and $A_g = 10^{-0}$: comparison between explicit and implicit, 1^{st} and 2^{nd}-order formulations.

method with 1 DeC iteration is more expensive than its explicit counterpart. Nevertheless, as previously pointed out, if 3 DeC iterations are considered,

	MR scheme		SRNH scheme	
Method	GR1	GR2	GR1	GR2
Explicit 1^{st} order, CFL= 0.8	$0.1s$	$0.6s$	$0.3s$	$1.8s$
Explicit 2^{nd} order, CFL= 0.8	$0.2s$	$1.4s$	$0.4s$	$2.6s$
Implicit 1^{st} order, CFL= 10^0	$0.6s$	$3.6s$	$0.7s$	$4.4s$
Implicit 2^{nd} order, CFL= 10^0, 1 DeC	$0.6s$	$3.6s$	$0.8s$	$5.2s$
Implicit 2^{nd} order, CFL= 10^1, 3 DeC	$0.2s$	$1.0s$	$0.2s$	$1.3s$

Table 10.5: CPU time required (seconds), case $A_g = 1$.

the second-order implicit schemes compute accurate results also for larger CFL values, namely CFL=10 for this case. Thus, by using 3 DeC iterations, it is possible to obtain a second-order implicit approach more efficient than the explicit one also for the case $A_g = 1$. The only case in which the implicit approach seems to be more efficient than the explicit scheme seems to be the second-order method with 3 DeC iterations. For this speed of interaction the use of multiple DeC iterations seems to be mandatory. However, since the CPU time needed for the entire simulation is very small, to really quantify the gain for this case, more demanding cases are required.

Summarising, for the considered 1D case, the implicit time advancing seems to be unconditionally stable in all the considered cases, the CFL limitation to avoid loss of accuracy is roughly inversely proportional to A_g and, in presence of unphysical oscillations, increasing the number of Defect Correction iterations can significantly increase the accuracy. The implicit time advancing is computationally efficient even using only one DeC iteration for slow and intermediate speeds of interaction. For fast speed of interaction 3 DeC iterations are required, as shown in Figs. 10.15 and 10.16.

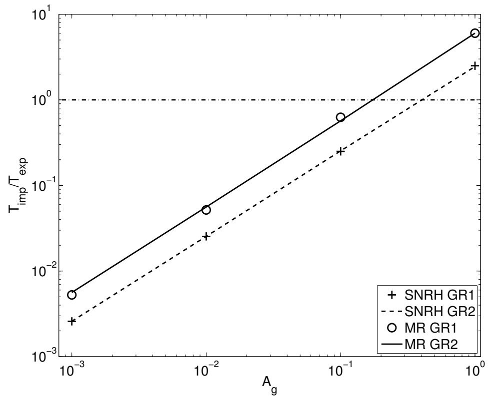
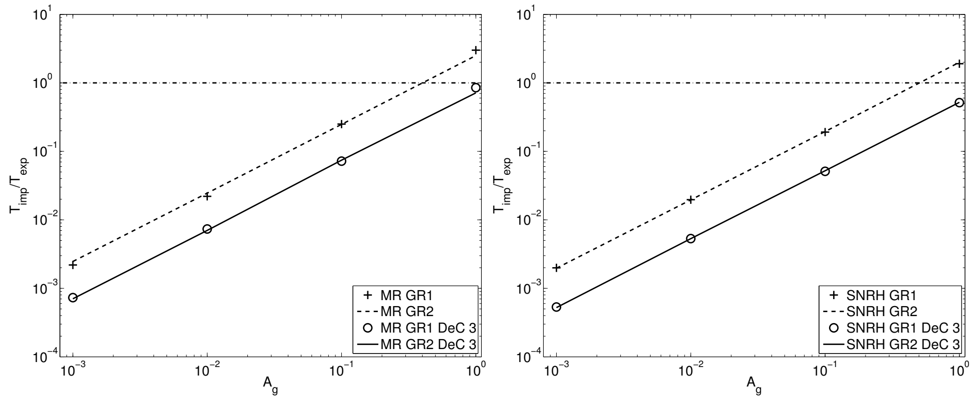


Figure 10.15: Ratio between the implicit computational time and the explicit one as a function of A_g for the first-order approaches.



(a) MR scheme **(b)** SRNH scheme

Figure 10.16: Ratio between the implicit computational time and the explicit one as a function of A_g for the second-order approaches.

Chapter 11

Numerical Simulations for 2D Sediment Transport Problems

In this chapter a 2D numerical benchmark is considered in order to validate the methodology proposed in section 9.3 to develop an implicit scheme starting from its explicit counterpart. Considering the results obtained for the 1D simulations in Chap. 10, in the 2D case only two working conditions are considered, namely the fast and the slow speeds of interaction, which are the most favourable working conditions for the explicit and implicit time-advancing approaches, respectively. The organisation of this chapter is as follows: first the set-up of the considered benchmark is described in Sec. 11.1, then the numerical experiments corresponding to a fast interaction between the bed and the flow are shown in Sec. 11.2 and finally in Sec. 11.3 the results obtained for the slow interaction case are presented.

11.1 2D Numerical Experiments

The 2D test-case considered herein is a well-known benchmark test, adopted in several papers (e.g. [8, 28]) and it is the 2D generalisation of the 1D test-case described in section 10.1. It is a sediment transport problem in a square domain Ω of dimensions 1000×1000 m^2 with a non constant bottom relief.

The initial bottom topography is defined as follows:

$$\begin{cases} Z(0,x) & = & \begin{cases} \sin^2\left(\dfrac{(x-300)\pi}{200}\right)\sin^2\left(\dfrac{(y-400)\pi}{200}\right) & \text{if } (x,y)\in Q_h \\ 0 & \text{elsewhere} \end{cases} \\ h(0,x,y) & = & 10 - Z(0,x,y) \\ u(0,x,y) & = & \dfrac{10}{h(0,x,y)} \\ v(0,x,y) & = & 0 \end{cases} \tag{11.1}$$

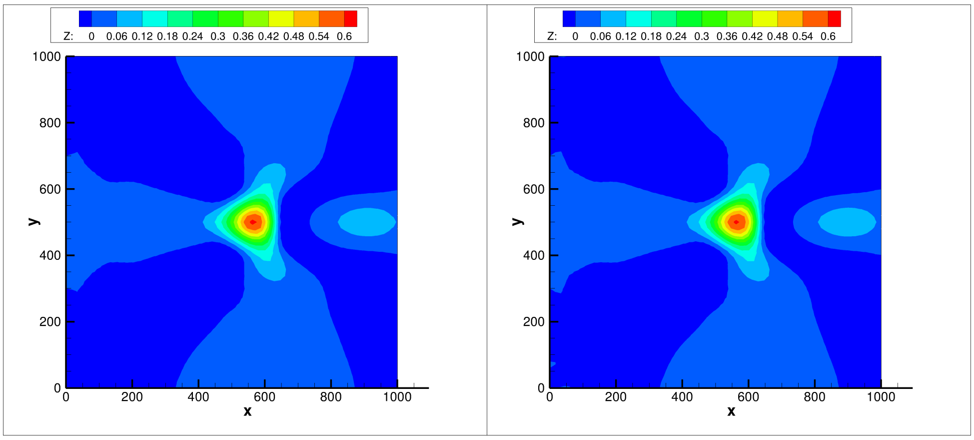
in which all the variables are in SI units and $Q_h = [300,500] \times [400,600]$. Dirichlet boundary conditions are imposed at the inlet, while at the outlet characteristic based conditions are used. Finally, free-slip is imposed on the lateral boundaries. The spatial discretization of the computational domain has been carried out by using two different grids. Both of them are symmetric with respect to the axis $y = 500$ m and are composed by uniform-size triangular elements. The main characteristics of the coarse (GR1) and of the refined (GR2) grid are reported in Tab. 11.1 where l_m denotes the characteristic length of the elements. As previously pointed out, for the 2D case two different values of the parameter A_g are considered, namely the slow interaction case, $A_g = 0.001$ and the fast one, $A_g = 1$, corresponding to the most favourable condition for the implicit time-advancing and the explicit one, respectively. Since different values of the parameter A_g corresponds to different time scales for the evolution of the bottom topography, different time intervals have been simulated for the two considered cases. In particular for the case $A_g = 1$ the total simulation time is 500 seconds while for the case $A_g = 0.001$ is 100 hours (360000 seconds). All the results and CPU times shown in the following are at the final instant of each simulation and they have been obtained on a 3 GHz Intel Pentium 4 processor with 2Gb RAM.

	Nodes	Elements	l_m
GR1	2901	5600	20
GR2	11425	22448	10

Table 11.1: Main characteristics of the grids used in the simulations.

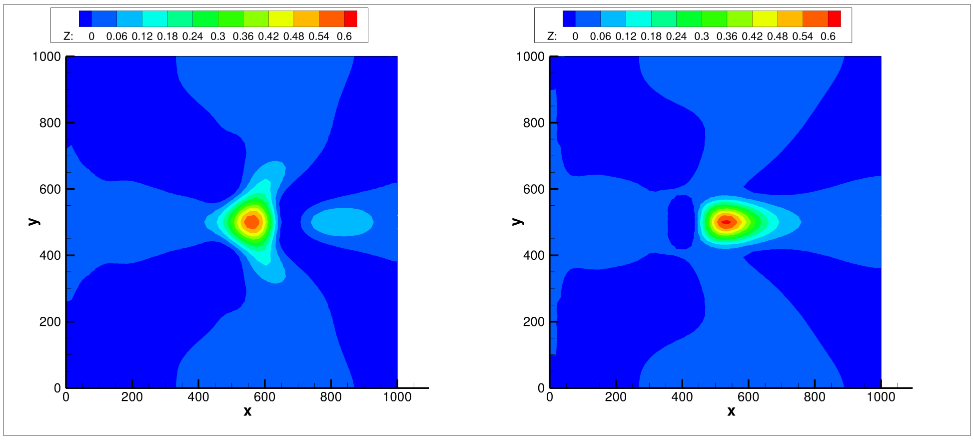
11.2 Fast speed of interaction between the bedload and water flow

For the fast speed of interaction, $A_g = 1$ in terms of accuracy the results essentially confirm the analysis of the 1D case. For this value of A_g, to avoid loss of accuracy the CFL number of the implicit scheme must be lowered down to 1. As an example figure 11.1 shows a comparison between the explicit and implicit approaches at different CFL values for the MR scheme. On the other hand, as for the 1D case, by increasing the number of DeC iterations, it is possible to increase the maximum CFL value by a factor 10 for



(a) Explicit CFL= 0.8

(b) Implicit, CFL= 1, DeC 1

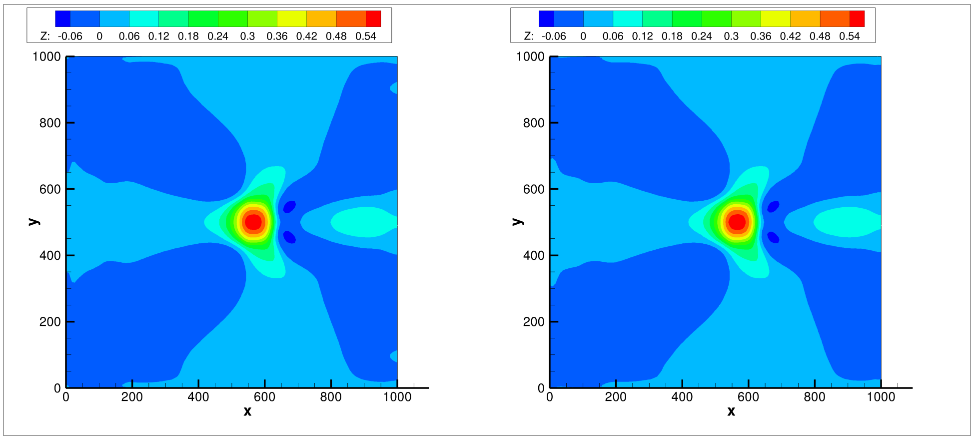


(c) Implicit, CFL= 10, DeC 1

(d) Implicit, CFL= 100, DeC 1

Figure 11.1: Grid GR1, $A_g = 1$: comparison of the results for the bed profile of the 2^{nd}-order MR scheme.

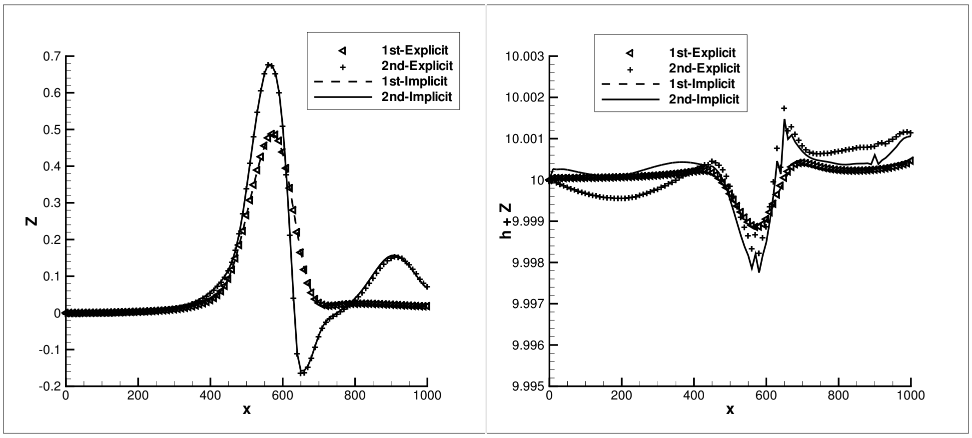
the second-order implicit approach without loosing in accuracy, as shown in figure 11.2. To be more quantitative, in Fig. 11.3 a comparison of the results



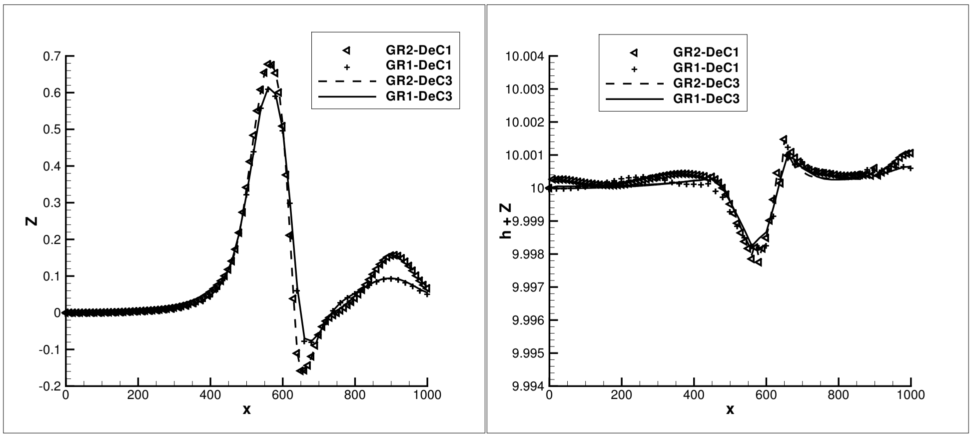
(a) Explicit CFL= 0.8 **(b)** Implicit, CFL= 10, 3 DeC

Figure 11.2: Grid GR2, $A_g = 1$: comparison of the results for the bed profile of the 2^{nd}-order SRNH scheme.

obtained with several versions of the MR scheme along the symmetry line $y = 500$ m is shown. In particular, Fig. 11.3a shows a comparison between the explicit and implicit time advancing, both at first and second-order of accuracy for the bottom profile on the grid GR2. It is possible to notice that, as in the 1D case, there is a significant difference between the first and second-order of accuracy while the results of the explicit and implicit approaches are hardly distinguishable. Fig. 11.3b shows the same comparison for the variable $h + Z$. The correct reproduction of this variable is rather critical and, while in the first-order solutions no oscillations have been found, in the second-order computations they are present. The magnitude of these oscillations is small and compared to the absolute value of the variation of Z or h is less than 10^{-3}. Furthermore, these oscillations are not introduced by the proposed implicit time advancing technique since they are present also in the second-order explicit method. It is also interesting to notice that in the regions of nearly constant Z the second-order implicit approach presents smaller variations in $h+Z$. Figs. 11.3c and 11.3d show instead the comparison between the second-order implicit scheme at CFL $=1$ with one DeC iteration and the same scheme at CFL $=10$ with three DeC iterations. As previously pointed out, both for GR1 and GR2 the same results have been obtained by using a value of the CFL number ten times larger and three DeC iterations. Interestingly enough this is true both for the variable Z and for $h + Z$. The



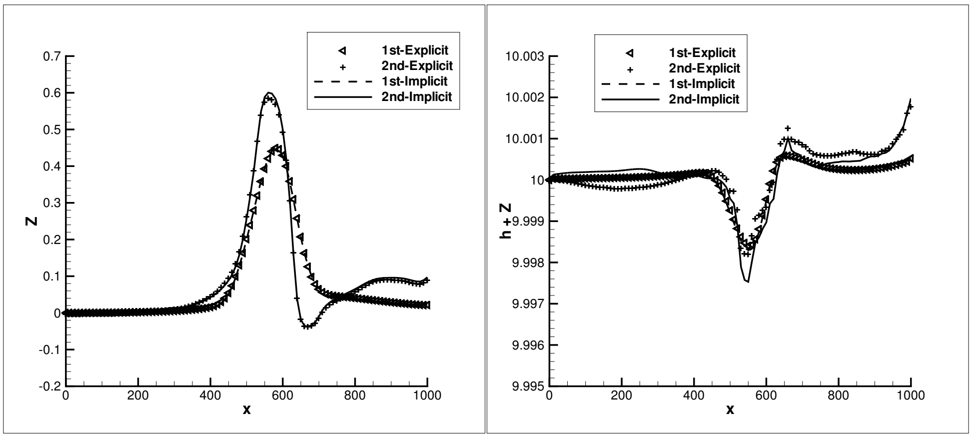
(a) GR2, Z **(b)** GR2, h+Z



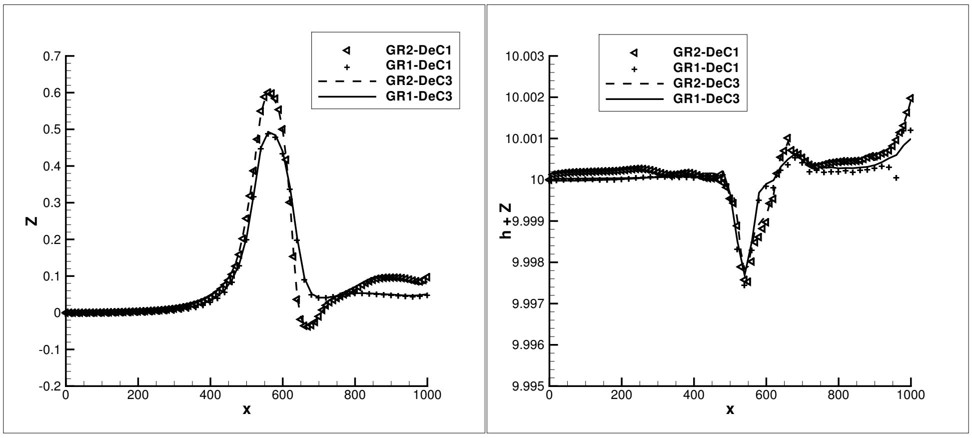
(c) Z **(d)** h+Z

Figure 11.3: $A_g = 1$: Comparison of the results for the bed profile for the MR scheme at $y = 500$ m.

same comparisons are visualised in Fig. 11.4 for the SNRH scheme: the same conclusions hold true also for this scheme with the only exception of the variable $h+Z$ for which the second-order implicit scheme is more oscillating than the explicit one, see Fig. 11.4b. Considering the stability of the numerical method, we found that using CFL= 10^2 the SRNH scheme is unstable on the refined grid. It is possible that this stability limitation is the consequence of a non-optimal choice for the solver of the linear system associated with the implicit approach. However, since for such a high CFL number, the quality of the results is anyway highly decreased this possibility has not been investigated. A new feature deserving attention is indeed the effect on the results



(a) Z **(b)** h+Z

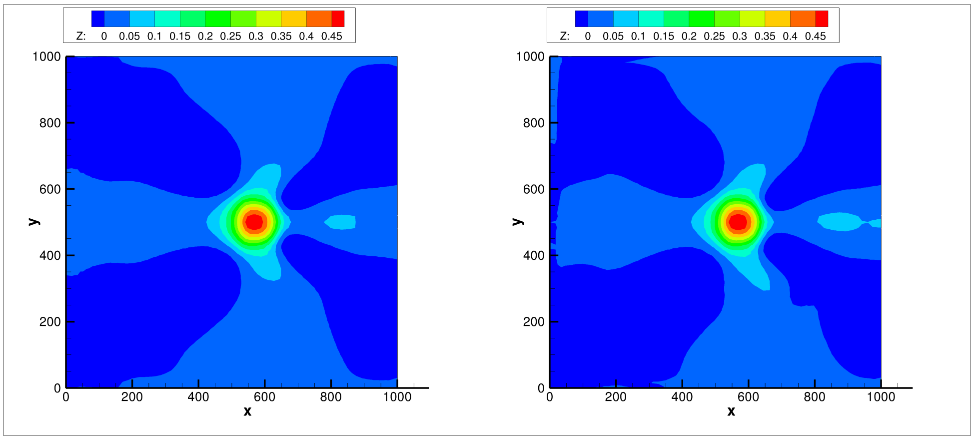


(c) Z **(d)** h+Z

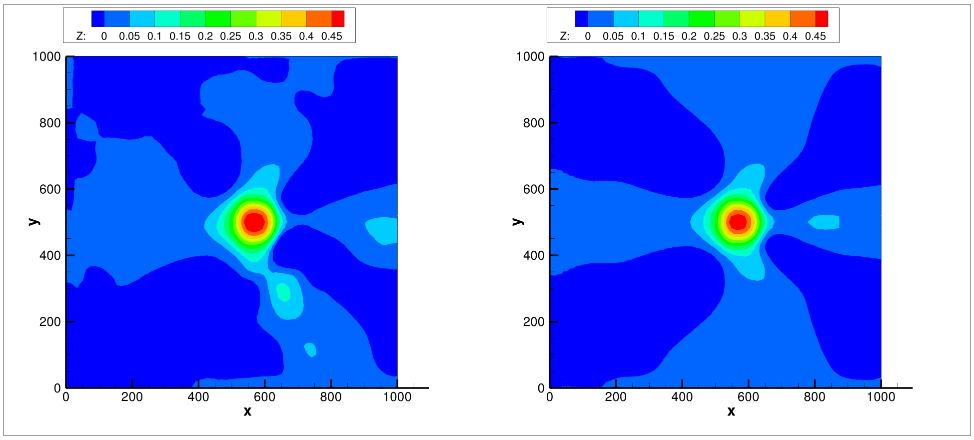
Figure 11.4: $A_g = 1$: Comparison of the results for the bed profile for the SRNH scheme at $y = 500$ m.

of the solution of the linear system associated with the implicit scheme. In the 1D case the linear system was solved using a direct method thus, except for roundoff errors, the exact solution was computed. Instead, for the 2D case, the linear system associated with the implicit scheme is solved using an iterative GMRES method. This iterative solver naturally introduces an additional approximation, since it is necessary to define a convergence criterion to terminate the iteration loop. In this work the iterative solver is stopped when the residual of the linear system is below a fixed fraction of the initial residual, that is:

$$\text{res} \leq \text{res}_0 * Tol$$



(a) Explicit **(b)** Implicit $Tol = 10^{-6}$, DeC 1



(c) Implicit $Tol = 10^{-8}$, DeC 1 **(d)** Implicit $Tol = 10^{-4}$, DeC 3

Figure 11.5: Grid GR1, CFL$= 1$: comparison of the effect of the value of Tol and of the number of DeC iterations for the SRNH scheme on the bed profile.

Clearly, the smaller is Tol, the smaller is the influence of the error in the solution of the linear system on the results. However, exceedingly decreasing Tol can abruptly increase the computational costs. Whatever is the CFL number, for the Modified Roe scheme we found that a reasonable value is $Tol = 10^{-4}$. The behaviour of the SRNH scheme is different: for the first order SRNH setting $Tol = 10^{-6}$ is enough to preserve the accuracy of the method. Instead the second-order scheme requires a different approach: using $Tol = 10^{-6}$ the solution computed by the implicit scheme is not completely similar to the explicit one. Lowering Tol below 10^{-6} could introduce roundoff errors, contaminating the solution. Under these circumstances, it is interesting to compare the effect of the value Tol against the one of the number of Defect Correction iterations. Figure 11.5 shows that the increase of the number of DeC iterations has a more pronounced effect than decreasing Tol.

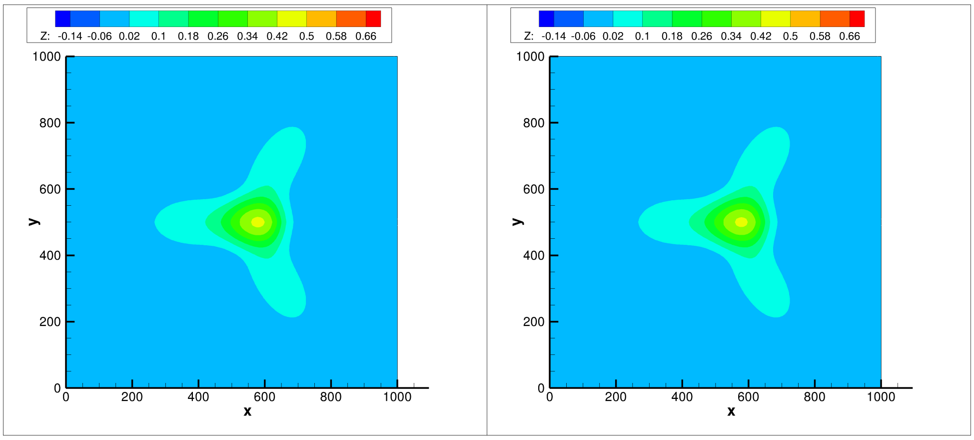
The CPU times for all the considered schemes and grids are reported in Tab. 11.2. An interesting feature of the 2D case is that, also for the fast interaction speed, the MR scheme at CFL= 10 with 3 DeC iterations is computationally more efficient than its explicit counterpart. This is not true for the SRNH scheme, as shown in table 11.2. The origin of this different behaviour is related to the different formulations of the two numerical methods: the second-order extension of the MR scheme is more computationally demanding that the SRNH one (see equation (9.38)) and, as a consequence, a Defect Correction approach is more efficient in this case. Also, the solution of the linear linear system associated to the implicit SRNH scheme seems to be more demanding than the MR one and this increases the computational time in the implicit case.

	MR scheme		SRNH scheme	
Method	GR1	GR2	GR1	GR2
Explicit 1^{st} order, CFL= 0.8	$16.1s$	$130.4s$	$21.0s$	$169.7s$
Explicit 2^{nd} order, CFL= 0.8	$140.4s$	$1137s$	$52.4s$	$409.9s$
Implicit 1^{st} order, CFL= 1	$197.0s$	$1560s$	$191.5s$	$1541s$
Implicit 2^{nd} order, CFL= 1, 1 DeC	$239.0s$	$1919s$	$198.7s$	$1582s$
Implicit 2^{nd} order, CFL= 10, 3 DeC	$84.3s$	$689.0s$	$74.5s$	$606.8s$

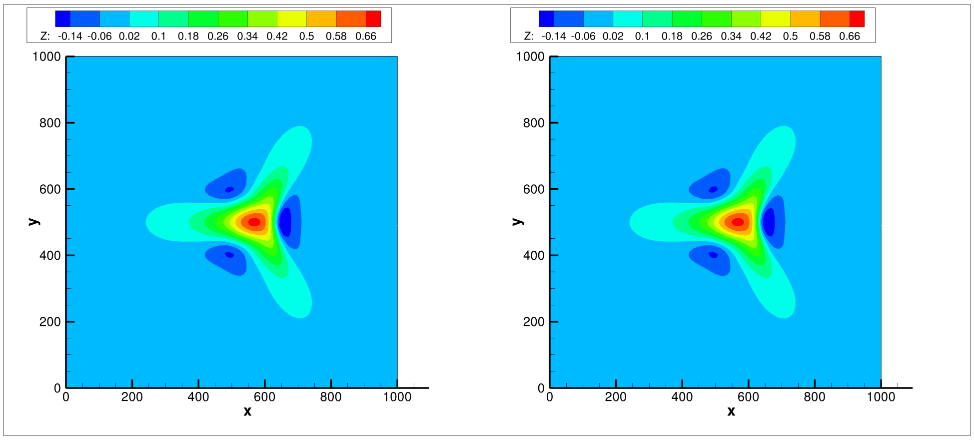
Table 11.2: CPU time required (seconds), case $A_g = 1$.

11.3 Slow speed of interaction between the bedload and water flow

For the slow speed of interaction, the results of the 2D computations essentially confirm the analysis of the 1D case. Both for the SRNH and for the MR scheme there is practically no difference between the solutions obtained with the implicit schemes at CFL= 1000 and explicit ones at CFL= 0.8, while the solutions computed using 1^{st}-order of accuracy significantly differ from the 2^{nd}-order ones as shown in figures 11.6 and 11.7. Again, by increasing the number of DeC iterations for the second-order scheme, an increase of the CFL number limit without losing accuracy can be obtained, as shown in figure 11.8. Also for this case a more quantitative analysis of the results has been carried out. Fig. 11.9 shows a comparison for the MR scheme: the results of the explicit and implicit formulation are practically superimposed



(a) 1^{st}-order Explicit CFL=0.8 **(b)** 1^{st}-order Implicit CFL=1000

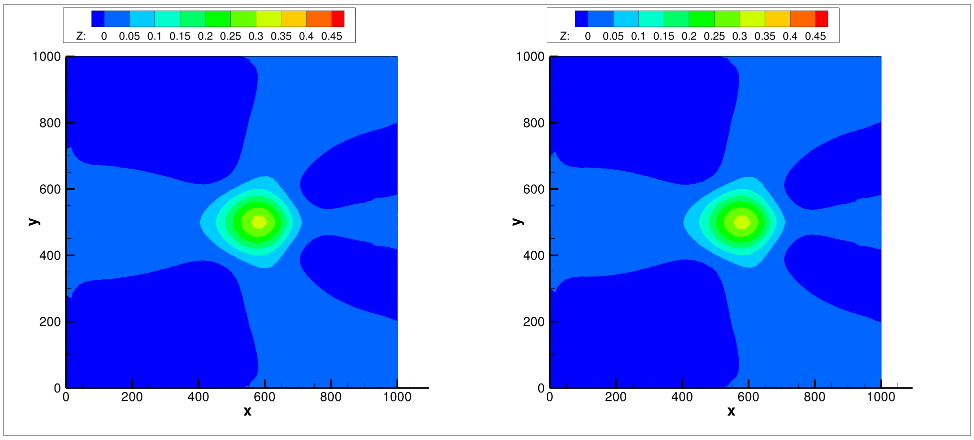


(c) 2^{nd}-order Explicit CFL=0.8 **(d)** 2^{nd}-order Implicit CFL=1000

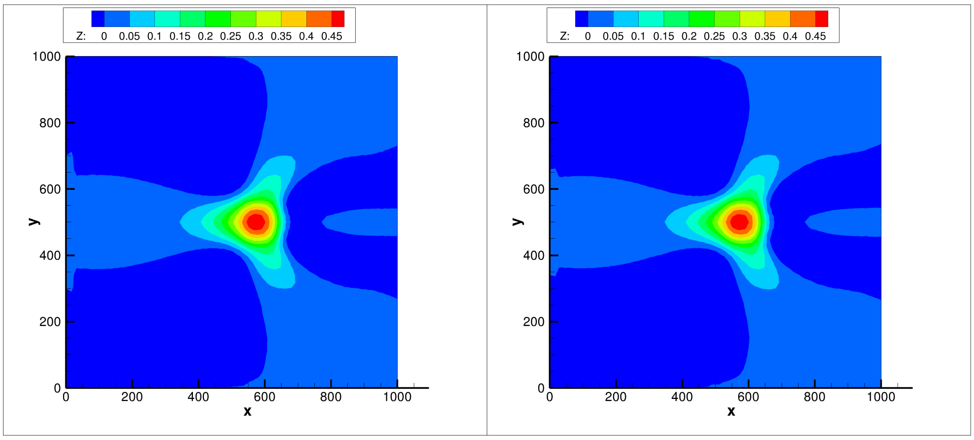
Figure 11.6: Grid GR2, $A_g = 10^{-3}$: comparison of the results of the bed profile of the explicit and implicit MR scheme.

for both Z and $h + Z$. Furthermore, also for the 2D case with $A_g = 10^{-3}$ it is possible to increase the CFL number of one order of magnitude, namely from 10^3 to 10^4, without loosing in accuracy, by using 3 DeC iterations in place of one, as shown in Figs.11.9c and 11.9d. The results obtained by using the SRNH implicit scheme, shown in Fig.11.10 globally confirm this analysis. The only noticeable difference is the fact that the second-order implicit scheme computes slightly larger oscillations than the explicit one in the solution for the variable $h + Z$, as shown in Fig.11.10b.

As for the fast speed of interaction case, the MR implicit scheme seems to be unconditionally stable while the SRNH implicit scheme is unstable using



(a) 1^{st}-order Explicit CFL=0.8 **(b)** 1^{st}-order Implicit CFL=1000

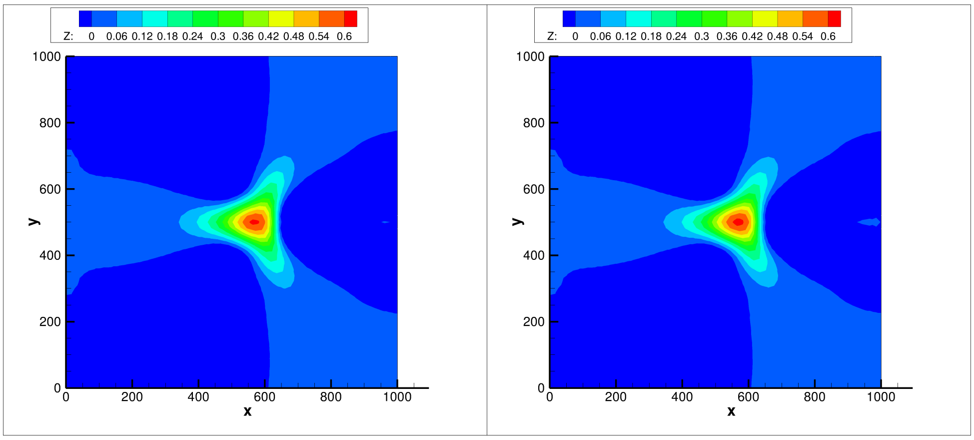


(c) 2^{nd}-order Explicit CFL=0.8 **(d)** 2^{nd}-order Implicit CFL=1000

Figure 11.7: Grid GR1, $A_g = 10^{-3}$: comparison of the results of the bed profile of the explicit and implicit SRNH scheme.

a CFL number of 10^5 on the grid GR2.

Finally, as for the computational costs, for both the SRNH and MR schemes, already at CFL= 1000 the gain in CPU time obtained with the implicit scheme is large, both at 1^{st} and 2^{nd}-order of accuracy as shown in table 11.3. Increasing the number of DeC iterations together with the CFL number can further decrease the computational cost of the second-order implicit scheme. As previously pointed out, in the 2D case the explicit second-order extension of the MR scheme is more computationally demanding than the SRNH one (see Tabs. 11.2 and 11.3). On the contrary, when the implicit formulation is considered, this analysis is more complex. Indeed, while the



(a) CFL= 10^3 with 1 DeC iteration

(b) CFL= 10^4 with 3 DeC iteration

Figure 11.8: Grid GR1, $A_g = 10^{-3}$: comparison of the results for the bed profile of the 2^{nd}-order implicit MR scheme.

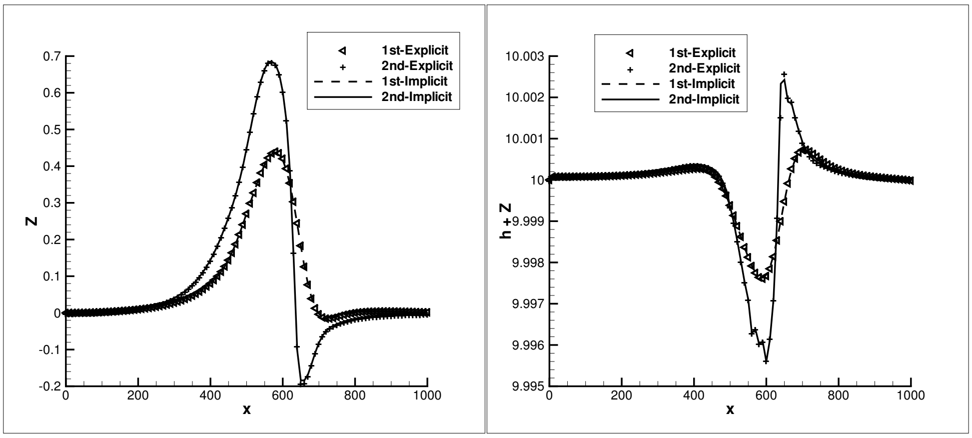
SRNH scheme is still more efficient for the case $A_g = 1$, in the case of a slow speed of interaction the MR scheme is roughly two times faster than the SRNH one, or even three time faster if the grid G2 and 3 DeC iterations are considered.

	MR scheme		SRNH scheme	
Method	GR1	GR2	GR1	GR2
1^{st}-Explicit CFL= 0.8	$9497s$	$78073s$	$12824s$	$103238s$
2^{nd}-Explicit CFL= 0.8	$82993s$	$670770s$	$30996s$	$247215s$
1^{st}-Implicit CFL= 10^3	$199.8s$	$2134s$	$323.6s$	$4336s$
2^{nd}-Implicit CFL= 10^3, 1 DeC	$293.7s$	$2776s$	$481.5s$	$8537s$
2^{nd}-Implicit CFL= 10^4, 3 DeC	$136.4s$	$1625s$	$265.9s$	$4866s$

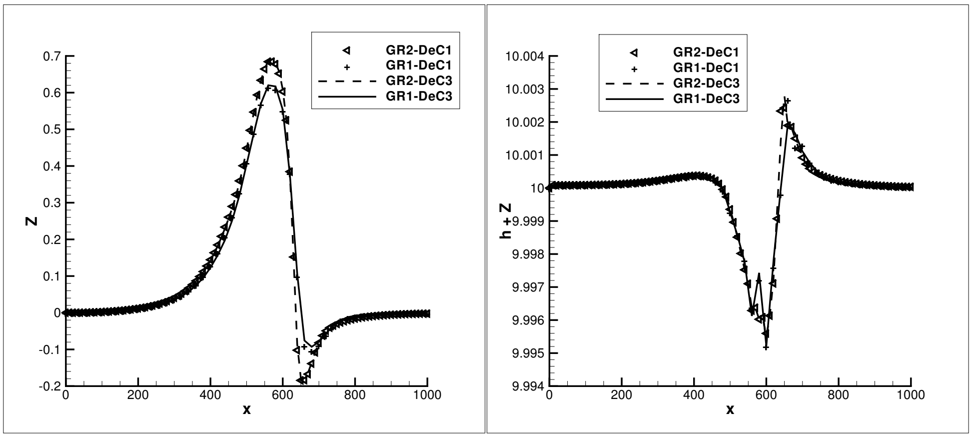
Table 11.3: CPU time required (seconds), case $A_g = 10^{-3}$.

This behaviour is the consequence of the interaction of at least three main factors: the time required to compute the explicit fluxes which constitutes the right-hand side of the linear system which must be solved at each time step, the time required to assembly the matrix associated with the linear system and the time required to solve the linear system. Unfortunately in the output of our simulations only the total simulations time has been recorded thus no exact figures are available for the aforementioned quantities. That notwithstanding, it still possible to give some reasonable estimation.

First, let us consider the time required to compute the explicit fluxes: in



(a) Z **(b)** h+Z

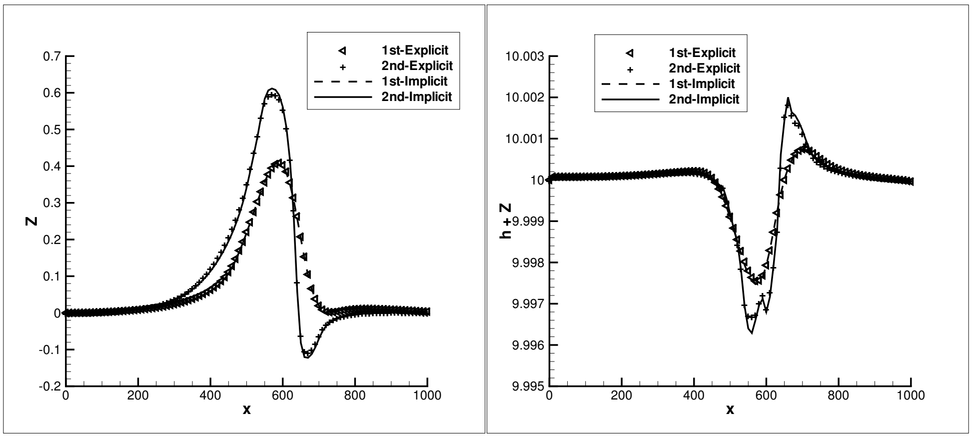


(c) Z **(d)** h+Z

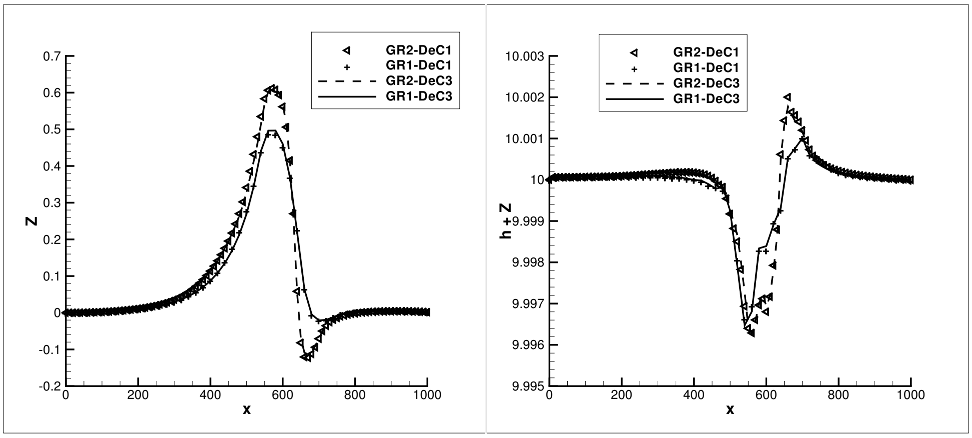
Figure 11.9: $A_g = 10^{-3}$: Comparison of the results for the bed profile for the MR scheme at $y = 500$ m.

the explicit numerical approaches, once the fluxes have been computed the time required to update the solution is without doubt negligible. Thereby the time required to run the explicit simulations can be safely ascribed to the computation of the fluxes and of the source term. Thus, by considering Tabs.11.2 and 11.3 it clearly appears that the time required to compute the first-order explicit fluxes is roughly the same for the SRNH and the MR schemes, the latter being 1.3 times cheaper. On the contrary the computation of the second-order explicit MR fluxes is roughly 2.7 times more demanding than the SRNH ones.

Estimating the time needed to assembly the matrix of the linear system



(a) Z **(b)** h+Z



(c) Z **(d)** h+Z

Figure 11.10: $A_g = 1$: Comparison of the results for the bed profile for the SRNH scheme at $y = 500$ m.

is more difficult. However, due to the considered DeC technique associated to the automatic differentiation procedure, this time should be roughly proportional to the one needed for the computation of first-order fluxes. Thus, in view of the previous considerations, it should be similar between the MR and the SRNH schemes, with the former one slightly cheaper.

Finally, the time required for the solution of the linear system must be defined. Even in this case no exact figures are available. Nevertheless, in Sec. 11.3 it has been pointed out that the linear system associated with the SRNH scheme seems to stiffer than the one associated with the MR scheme, thus requiring stronger convergence criterion and, as a consequence, more iteration to compute the solution. Therefore it is reasonable to consider that the stiffer is the problem, the longer will take to solve the linear system. Since both decreasing A_g and refining the grid increase the stiffness of the linear system to be solved, it follows that for the SRNH scheme the case $A_g = 1$ on the coarse grid is the most favourable among those considered in this work and, conversely, $A_g = 10^{-3}$ on the refined grid is the worst.

These considerations give further insight for the determination of which approach is the most efficient. The DeC implicit approach is preferable when computationally expensive explicit methods are considered, in particular if the computational cost greatly varies between first and second-order of accuracy. On the contrary, the comparison between different numerical schemes should take into account others factors and, in particular, when stiff problem are considered the time required to solve the linear system is of the greatest importance.

Chapter 12

Concluding Remarks

In this work different numerical methods for the simulation of barotropic flows in complex geometries have been presented. Two main applications, which are characterised by possible problem stiffness and for which barotropic models can be used, have been considered, namely cavitating flows and sediment transport problems.

12.1 Cavitating Flows

In the present study, a numerical formulation is proposed for the simulation of 3D viscous cavitating flows, which combines good properties of accuracy, robustness and efficiency.

For the particular application to cavitating flows, in the non-cavitating regions, a weakly-compressible liquid at constant temperature is considered, while for the cavitating regime, the homogeneous-flow model explicitly accounting for thermal cavitation effects and for the concentration of the active cavitation nuclei in the pure liquid is adopted. This chosen homogeneous-flow cavitation model leads to a complex barotropic state equation and, as a consequence, the numerical schemes have been developed for generic barotropic flows. However, the simulation of cavitating flows, in particular for the chosen cavitation model, in which huge variations of flow quantities are present and nearly-incompressible zones coexist with highly-supersonic regions, leads to specific numerical difficulties from a view point of both accuracy and efficiency.

Our starting point was a numerical solver for 3D inviscid barotropic flows introduced in previous works [16, 103, 104]. The numerical scheme was first-order accurate both in space and time.

As for physical modelling, viscous terms and the standard $k-\epsilon$ RANS turbulence model have been added in the present work. The present numerical approach is based on finite-volume and finite-element discretizations for the convective and the viscous terms respectively. Compressible flows are considered and approximate Riemann solvers are used to compute the numerical fluxes involved in the finite-volume discretisation of the convective fluxes. Suitable preconditioning is used to avoid accuracy problems in the low-Mach regime, while maintaining time consistency. Time advancing is carried out through a linearised implicit approach. However, in previous works it was found that the use of an implicit time advancing is not a sufficient guarantee when cavitation occurs; indeed, severe CFL limitations were observed for a linearised implicit formulation based on the Roe scheme [14, 102, 104]. Two new ingredients have been introduced here. Starting from the Rusanov scheme, which is a simple averaged-state approximate Riemann solver known for its reliability and robustness, a low-diffusive HLL numerical flux is proposed. The latter one is obtained by introducing an anti-diffusive term in the Rusanov flux in order to deal with its excessive diffusivity in presence of contact discontinuities. Thanks to the particular features of the Riemann problem for cavitating flows, the simple structure of the Rusanov scheme could be maintained in the LD-HLL one. A suitable preconditioning for the low-Mach regime has also been defined. Second, two different time linearisations are considered for the LD-HLL flux; a classical one in which the upwind part of the flux function is partially frozen in the time differentiation, and a more complete one which takes into account, at least in an approximate way, the neglected term. Finally, second-order accuracy is obtained through a MUSCL reconstruction for space accuracy and a second-order backward differentiation formula for time advancing associated with defect-correction iterations.

The proposed numerical ingredients have been validated through various test-cases, both 1D and 3D, which cover a large range of flow features and for which analytical solution or experimental results are available. Three different numerical flux functions adapted to barotropic flows, the Roe, Rusanov and LD-HLL ones, are compared, by including first and second-order accurate formulations and the two kinds of linearisation for the HLL-like schemes. The capability to obtain accurate solutions for contact discontinuities and viscous flows has been investigated by considering a Riemann problem as well as the Blasius boundary layer over a flat plate. These test-cases show that the anti-diffusive term introduced in the LD-HLL scheme is able to counteract the accuracy problem in presence of contact discontinuity with results very similar to the ones of the Roe scheme and in good accordance with theoretical solutions. Therefore, this scheme appears to be suitable for the simulation of flows characterised by the presence of contact discontinuities, and, in particular, of viscous flows.

A second set of test-cases has also been considered in order to assess the accuracy, robustness and efficiency of the different numerical formulations in cavitating conditions. It includes a quasi-1D cavitating flow in a convergent-divergent nozzle as well as the flow around a hydrofoil mounted in a tunnel, both in cavitating and non-cavitating conditions. It clearly follows from these numerical experiments that the use of the more complete linearisation is essential for efficiency in presence of cavitation. Indeed, very strong stability limitations appear when cavitation occurs, this for the Roe linearised implicit formulation as well as for the LD-HLL implicit approach using the classical linearisation. An increase of five and four orders of magnitude (for 1D and 2D flows respectively) is obtained, in terms of CFL number, when the more complete linearisation for the LD-HLL scheme is used. Significant improvements have been observed by considering the second-order accurate version of the LD-HLL with a behaviour much closer to experimental data and physical expectations. Thus, it emerges that the second-order preconditioned implicit LD-HLL approach associated with the more complete linearisation appears an adequate formulation for viscous flows in cavitating conditions.

Furthermore, the implementation of the standard $k - \varepsilon$ model has been validated for the turbulent flow over a flat plate at zero angle of attack.

Finally, applications to real 3D configurations have been carried out by considering the viscous flow in a three blade axial inducer, for which experimental data are available. Different working conditions have been reproduced by varying the flow rates and the cavitation number. A first set of simulations have been carried out in non cavitating conditions, without turbulence model. In non-cavitating conditions the simulations associated with high flow rates give a satisfactory agreement with the experiments while for lower flow rates the accuracy of the solution decreases. These discrepancies are related with the presence of a large backflow flow. In particular, a very fine grid resolution is needed near the blade tip to capture the steep variation of the flow velocity in this zone. Furthermore, for low flow rates, the backflow extends to the inlet of the computational domain, and thus, spurious effects could be present on the solution. Based on these considerations, two additional simulations have been carried out; in the first one the same grid resolution is used but the inlet of the computational domain is moved upstream, while in the second one the computational domain is the shorter one but a more refined grid is used. Both the increase of the computational domain axial length and the grid refinement lead to a reduction of the discrepancy in the pressure-jump prediction with respect to the experimental value, with grid refinement having a larger impact on the accuracy of the numerical results. Nonetheless, the error made in the numerical predictions remains significant also on the more refined grid. A further possible explanation of this disagreement is that the effects of turbulence are probably more important at low discharge rates when a strong backflow is present. In order to assess this point a RANS simulation was also carried for the lowest flow rate. The introduction of turbulence has been found to significantly improve the agreement with the experiments. As expected the effects of turbulence are particularly important near the gap between the blades and the external case and this strongly affects the backflow and, thus, the pressure jump. This also explains why for larger flow rates, for which the backflow is less important, the effects of turbulence are not so strong and a good agreement with experimental data can be obtained also in laminar simulations.

In addition, cavitating cases have also been considered. Note that only the most robust and efficient numerical formulation considered herein, i.e. the preconditioned linearised LD-HLL approach associated with the complete linearisation, is able to deal with 3D cavitating flows in the inducer. Indeed, considering the CFL number, the results found in the preliminary test-cases, i.e. cavitating flow over a NACA 0015 hydrofoil, have been confirmed, since the scheme is stable with values of the CFL number up 500.

The first simulation carried out in cavitating conditions showed some discrepancies with experimental data and, in particular, an underestimation of the cavitation zone extension. In order to investigate this issue, two additional simulations on the more refined grid have been considered, for two different cavitation numbers. The predictions of the pressure jump obtained in these simulations are in good accordance with the experimental values, showing the positive effect of grid refinement also in cavitating conditions.

Finally, a preliminary analysis aimed at investigating whether flow instabilities, related to cavitation and detected in the experiments, has been carried out for the cavitating-flow simulation on the most refined grid by visualising the time behaviour of the cavitating regions over the three blade and through spectral analysis of pressure times signals recorded at different axial locations over the external case. Traces of asymmetric and surge instabilities, also detected in experiments for the considered values of discharge and cavitation number, have been found in the simulations, showing that the developed numerical approach is able to at least qualitatively capture these instabilities. Nonetheless, we would like to stress that this type of analysis is difficult to be carried out numerically since it requires the simulation of large time intervals, especially to characterise low-frequency instabilities as surge; since the discretisation of complex 3D domains also implies a large number of nodes, all this leads to highly demanding simulations both as regards CPU times and memory requirements. On the other hand, numerical simulations can furnish details on the flow dynamics that are difficult or even impossible to be obtained in experiments and, thus, can be a complementary tool to experimental analysis in the characterisation and even in the design of turbopump inducers.

12.2 Sediment transport problems

The main focus of the present work was on the comparison between implicit and explicit schemes for the simulation of sediment transport problems, in terms of accuracy and computational requirements. The problem is modeled through the shallow-water equations coupled with the Exner equation to describe the time evolution of the bed profile. The Grass model is used for the sediment transport.

We started from two finite-volume discretisation schemes previously proposed in the literature, combined with explicit time advancing, namely the SRNH solver [5, 6, 8] and a modified Roe scheme [26–28]. The implicit linearised versions of these schemes have been generated starting from their explicit counterpart by using an automatic differentiation tool, Tapenade [59], to compute the flux Jacobians. In the context of the Shallow-Water model 1D and 2D test cases have been considered, characterised by different rates of interaction between the bed and the water flow. In the 1D simulations, for both the SRNH and MR schemes, the implicit method was found to run with a virtually unlimited CFL number without stability problems. However, to avoid loss of accuracy, the CFL number of the implicit scheme must be reduced to a value roughly inversely proportional to the constant determining the velocity of the interaction between the flow and the bed-load. Only one DeC iteration seems to be enough to preserve the accuracy of the second-order implicit scheme even if increasing the number of DeC iterations can increases the maximum CFL number allowable. In the 1D case the implicit code has been found to be computationally more efficient than the explicit one for all the considered rates of interactions between the bed and the flow. The 2D tests globally confirm these results, but the influence of the number of Defect Correction iterations seems to be more pronounced, in particular for the SRNH scheme. Furthermore, it was found that for fast bed load/water interaction only the MR implicit scheme can be competitive with its explicit counterpart. As a consequence, the proposed methodology, implicit time advancing and Defect Correction technique, seems to be particularly suitable for slow and intermediate speeds of interaction and for computationally expensive explicit methods. Also, since the CFL number is limited by the speed of interaction, another field deserving attention is the application of the considered methodology to more complex models for the evolution of the bed.

As a concluding remark it may be worth emphasising that the proposed numerical methodologies are rather general and, thus, there is room for further improvements. As far as cavitating flows are concerned, it could be interesting to further study the influence of the turbulence model on the numerical results and to take into consideration the possibility of implementing different turbulence models which explicitly take into account the physical effects of rotation and cavitation. Furthermore, a systematic study of the instabilities related to the cavitation, also considering lower values of the cavitation number, could be interesting.

Switching to sediment transport problems, it turns out that the implicit time advancing is particularly interesting when the problem is stiff, that is when the difference in the characteristic time scales of the problems is more than one order of magnitude. Thus, it could be interesting to consider more complex models for sediment transport than the Grass one or problems involving less regular solutions (i.e. problems with discontinuities). Furthermore, as a natural development, the proposed methodology could be applied in the future to realistic problems, i.e. to the simulation of coastal zones or river flows.

Part III

Appendices

Appendix A

Solution of the Riemann problem for a convex barotropic state law

Let us consider the following Riemann problem:

$$\left\{\begin{array}{rcll} \dfrac{\partial W}{\partial t}(x,t) + \dfrac{\partial F}{\partial x}(x,t) & = & 0 & \text{in} \quad \mathbb{R}\times]0,\infty[\\ \\ W(x,0) & = & \left\{\begin{array}{lll} W_L & if & x<0 \\ \\ W_R & if & x>0 \end{array}\right. & \end{array}\right. \tag{A.1}$$

with $W = (\rho, \rho u, \rho\xi)^T$ and $F = (\rho u, \rho u^2 + p, \rho u \xi)^T$.

This system of partial differential equations is supposed to be closed by a convex barotropic equation of state. In this case, the convexity condition implies that the second derivative of the pressure with respect to the inverse of the density be strictly positive and therefore can be expressed as follows:

$$\frac{a}{\rho} + \frac{da}{d\rho} > 0. \tag{A.2}$$

It has been shown in [102] that, in this case, the involved waves are either genuinely non-linear or linearly-degenerate and, consequently, a local Riemann solution can be constructed simply through the juxtaposition of basic wave solutions (i.e. rarefactions, shocks and contact discontinuities). A constructive procedure to solve the Riemann problem (A.1), associated with a generic but convex barotropic equation of state, is then proposed. The construction of the resulting solution is recalled below; we refer to [102] for a detailed derivation.

The solution of the Riemann problem (A.1) consists of four states, W_L, $W_L^\star$, $W_R^\star$ and W_R separated by three waves: the 1-wave and 3-wave can be

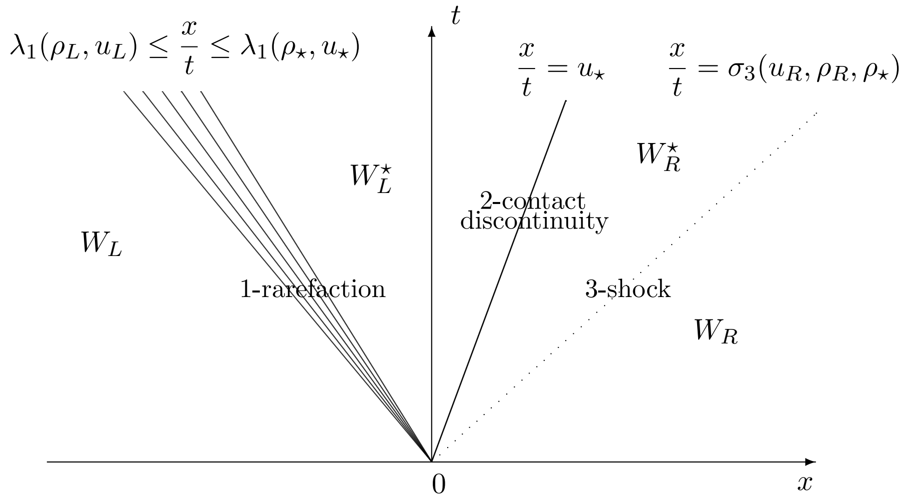


Figure A.1: Schematic representation of the solution of the Riemann problem (A.1) for the particular case in which the 1-wave is a rarefaction and the 3-wave is a shock.

either a shock or a rarefaction while the 2-wave is a contact discontinuity. Fig. A.1 shows a sketch of the solution structure for a case in which all types of waves are involved. The states W_L and W_R are given by the initial condition of the Riemann problem while the intermediate states, $W_L^\star$ and $W_R^\star$, have the following expression:

$$W_S^\star = (\rho_\star, \rho_\star u_\star, \rho_\star \xi_S)^T \qquad \text{for } S \in \{L, R\}.$$

The solution is completely determined by $\rho_\star$ and $u_\star$, i.e. the values of the intermediate density and velocity (which determine the nature and the position of the different waves); the corresponding solution region is commonly called “star” region. In [103], it is proposed to firstly define $\rho_\star$ by solving the following non linear equation:

$$g_L(\rho) + g_R(\rho) + u_R - u_L = 0 \tag{A.3}$$

the functions g_L and g_R being expressed as

$$g_S(\rho) = I_{[\rho \leq \rho_S]}(\rho)\, \psi(\rho_S, \rho) - I_{[\rho > \rho_S]}(\rho)\, \varphi(\rho_S, \rho) \qquad \text{for } S \in \{L, R\}$$

in which I represents the characteristic function while ψ and φ are defined as follows for any $\rho_1, \rho_2 > 0$:

$$\varphi(\rho_1, \rho_2) = -\left(\frac{(p(\rho_2) - p(\rho_1))\,(\rho_2 - \rho_1)}{\rho_1\, \rho_2}\right)^{\frac{1}{2}}, \quad \psi(\rho_1, \rho_2) = \int_{\rho_1}^{\rho_2} \frac{a(r)}{r}\, dr.$$

Then, $u_\star$ can be obtained directly from $\rho_\star$ using either the left or the right initial state through the following relation:

$$u_\star = u_S - g_S(\rho_\star) \qquad \text{for } S \in \{L, R\}. \tag{A.4}$$

The knowledge of $u_\star$ allows to completely define the 2-contact discontinuity since its speed σ_2 is simply equal to $u_\star$ (i.e. the contact discontinuity between $W_L^\star$ and $W_R^\star$ is located on the line $x/t = u_\star$).

Note that the exact Riemann solution cannot be achieved, in general, in a completely analytical way (indeed it also depends of the particular state law) and therefore a numerical procedure is needed to solve the non-linear equation (A.3). Let us consider the interval $[\rho_{min}, \rho_{sup}[$ in which the barotropic state law can be adequately used ($\rho_{min} > 0$ while ρ_{sup} can be $+\infty$). The existence and the uniqueness of $\rho_\star$ in $[\rho_{min}, \rho_{sup}[$ is ensured if and only if the initial velocity variation verifies:

$$\sum_{S\in\{L,R\}} \varphi\left(\rho_{sup}, \rho_S\right) < u_R - u_L \leq \sum_{S\in\{L,R\}} \psi(\rho_{min}, \rho_S). \tag{A.5}$$

While verifying the condition above, the wave structure of the Riemann problem solution changes as $u_R - u_L$ varies from the two extreme values. More precisely, the following three different cases can be identified, in which ρ_m and ρ_M are defined by $\rho_M = \max(\rho_L, \rho_R)$ and $\rho_m = \min(\rho_L, \rho_R)$, respectively:

- $u_R - u_L < \varphi(\rho_m, \rho_M)$: in this case $\rho_M < \rho_\star < \rho_{sup}$ and both the $1-$wave and the $3-$wave are shocks;
- $\varphi(\rho_m, \rho_M) \leq u_R - u_L < \psi(\rho_m, \rho_M)$: in this case, $\rho_m < \rho_\star \leq \rho_M$; there is a shock between $\rho_\star$ and ρ_m, and a rarefaction between $\rho_\star$ and ρ_M;
- $u_R - u_L \geq \psi(\rho_m, \rho_M)$: then $\rho_{min} \leq \rho_\star \leq \rho_m$ and both the $1-$wave and the $3-$wave are rarefactions.

The location of the rarefaction waves are defined as follows:

$$1 - \text{rarefaction} : \lambda_1\left(\rho_\mathrm{L}, \mathrm{u_L}\right) \leq \frac{\mathrm{x}}{\mathrm{t}} \leq \lambda_1\left(\rho_\star, \mathrm{u}_\star\right)$$

$$3 - \text{rarefaction} : \lambda_3\left(\rho_\star, \mathrm{u}_\star\right) \leq \frac{\mathrm{x}}{\mathrm{t}} \leq \lambda_3\left(\rho_\mathrm{R}, \mathrm{u_R}\right)$$

with $\lambda_k\left(\rho, u\right) = u + (k-2)\, a(\rho) \quad \text{for } k \in \{1, 3\}$. Moreover, the speed σ_k of the k-shock (which completely determines its position) can be expressed as follows:

$$\text{for } k \in \{1, 3\} \qquad \sigma_k = u_{S_k} + (k-2)\left(\frac{\rho_\star}{\rho_{S_k}} \frac{p(\rho_{S_k}) - p(\rho_\star)}{\rho_{S_k} - \rho_\star}\right)^{\frac{1}{2}}$$

where the notation $S_1 = L$ and $S_3 = R$ is understood. The expression at hand can easily be prolonged by continuity at $\rho_{S_k} = \rho_\star$ [103]; however, only a simplified representation is reported here for ease of discussion.

The solution for ξ depends on the location of the contact discontinuity and thus on $\rho_\star$ and $u_\star$ but, in turns, it does not affect the solution for ρ and u. This point, which is due to the decoupling between the passive scalar and the underlying 1D flow field permits to straightforwardly extend the structure of the considered solution to the case of an arbitrary number $m > 1$ of passive scalars.

Appendix B

The alternative formulation of the Roe scheme for barotropic flows

In this appendix, the proof that the Roe scheme for barotropic state law, as defined in [102, 103] can be also rewritten in the context of the unified Godunov-type schemes as presented in [84], is proposed. More precisely, it is derived the proof that the numerical flux function defined by (3.17)-(3.59) coincides with the one obtained from (3.65) and (3.69).

Let us explicitly write the matrix $|\tilde{J}|$, $\tilde{J}$ being the Roe matrix defined in Sec. 3.5.1:

$$|\tilde{J}| = \begin{pmatrix} \dfrac{|\tilde{\lambda}_1|\tilde{\lambda}_2 - \tilde{\lambda}_1|\tilde{\lambda}_2|}{\tilde{\lambda}_2 - \tilde{\lambda}_1} & \dfrac{|\tilde{\lambda}_2| - |\tilde{\lambda}_1|}{\tilde{\lambda}_2 - \tilde{\lambda}_1} & 0 \\[2ex] \dfrac{\tilde{\lambda}_2\tilde{\lambda}_1\left(|\tilde{\lambda}_1| - |\tilde{\lambda}_2|\right)}{\tilde{\lambda}_2 - \tilde{\lambda}_1} & \dfrac{|\tilde{\lambda}_2|\tilde{\lambda}_2 - |\tilde{\lambda}_1|\tilde{\lambda}_1}{\tilde{\lambda}_2 - \tilde{\lambda}_1} & 0 \\[2ex] \dfrac{\tilde{\xi}_{lr}\left(|\tilde{\lambda}_1|\tilde{\lambda}_2 - \tilde{\lambda}_1|\tilde{\lambda}_2| - |\tilde{\lambda}_3|\left(\tilde{\lambda}_2 - \tilde{\lambda}_1\right)\right)}{\tilde{\lambda}_2 - \tilde{\lambda}_1} & \dfrac{\tilde{\xi}_{lr}\left(|\tilde{\lambda}_2| - |\tilde{\lambda}_1|\right)}{\tilde{\lambda}_2 - \tilde{\lambda}_1} & |\tilde{\lambda}_3| \end{pmatrix} \tag{B.1}$$

where we recall that $\tilde{\lambda}_h$ are the eigenvalues of $\tilde{J}$:

$$\tilde{\lambda}_1 = \tilde{u}_{lr} + \tilde{a}_{lr}\ ,\ \tilde{\lambda}_2 = \tilde{u}_{lr} - \tilde{a}_{lr} \ \text{ and } \ \tilde{\lambda}_3 = \tilde{u}_{lr} \tag{B.2}$$

and r_h are the corresponding right eigenvectors:

$$r_1 = \begin{pmatrix} 1 \\ \tilde{\lambda}_1 \\ \tilde{\xi}_{lr} \end{pmatrix}, \quad r_2 = \begin{pmatrix} 1 \\ \tilde{\lambda}_2 \\ \tilde{\xi}_{lr} \end{pmatrix}, \quad r_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{B.3}$$

Besides, the Roe average for the speed of sound $\tilde{a}_{lr}$ is defined by (3.60), while for the velocity and the passive scalar, the Roe average is obtained substituting ϕ by respectively u or ξ in the following relation:

$$\tilde{\phi}_{lr} = \frac{\sqrt{\rho_l}\phi_l + \sqrt{\rho_r}\phi_r}{\sqrt{\rho_l} + \sqrt{\rho_r}} \tag{B.4}$$

Let us recall, here, two useful properties of the Roe averages which have been intensively used for the proof. Let ψ and φ be any variable such that their Roe average can be expressed as in (B.4), i.e. here, whichever between u and ξ, not necessarily different:

$$\rho_r\psi_r - \rho_l\psi_l = \tilde{\rho}_{lr}\Delta^{lr}\psi + \tilde{\psi}_{lr}\Delta^{lr}\rho \tag{B.5}$$

$$\rho_r\psi_r\varphi_r - \rho_l\psi_l\varphi_l = \tilde{\rho}_{lr}\tilde{\psi}_{lr}\Delta^{lr}\varphi + \tilde{\rho}_{lr}\tilde{\varphi}_{lr}\Delta^{lr}\psi + \tilde{\varphi}_{lr}\tilde{\psi}_{lr}\Delta^{lr}\rho \tag{B.6}$$

where $\Delta^{lr}(\,\cdot\,) = (\,\cdot\,)_r - (\,\cdot\,)_l$ and $\tilde{\rho}_{lr}$ is the Roe average for the density defined as:

$$\tilde{\rho}_{lr} = \sqrt{\rho_l\rho_r} \tag{B.7}$$

Finally, let us recall that from the knowledge of the right eigenvectors r_p the coefficients α_p of $\mathbf{W}_r - \mathbf{W}_l$ in the basis (r_1, r_2, r_3) can be also calculated. More precisely:

$$\begin{cases} \alpha_1 &= \dfrac{1}{2}\left(\Delta^{lr}\rho + \dfrac{\tilde{\rho}_{lr}\Delta^{lr}u}{\tilde{a}_{lr}}\right) \\ \alpha_2 &= \dfrac{1}{2}\left(\Delta^{lr}\rho - \dfrac{\tilde{\rho}_{lr}\Delta^{lr}u}{\tilde{a}_{lr}}\right) \\ \alpha_3 &= \tilde{\rho}_{lr}\Delta^{lr}\xi \end{cases} \tag{B.8}$$

The proof is split into two different cases, i.e. considering either supersonic or subsonic interfaces.

B.1 Supersonic interface

Since the speed of sound is always positive, the eigenvalues of the Roe scheme always satisfy :

$$\tilde{\lambda}_1 > \tilde{\lambda}_3 > \tilde{\lambda}_2$$

It is possible to refer to the interface as supersonic when all the eigenvalues have the same sign, that is:

$$\tilde{\lambda}_1 > \tilde{\lambda}_3 > \tilde{\lambda}_2 > 0 \quad \text{or} \quad 0 > \tilde{\lambda}_1 > \tilde{\lambda}_3 > \tilde{\lambda}_2 \tag{B.9}$$

If all the eigenvalues are positive then $|\tilde{J}|$ and $\tilde{J}$ coincide, and thus, using the third property of the Roe matrix the classical expression (3.17) of the Roe scheme reduces to $\Phi(W_l, W_r) = F(W_l)$. On the other hand, if all the Roe matrix eigenvalues are positive, from (3.69) it follows that $b^+ = \tilde{\lambda}_1$ and $b^- = 0$, and, as a consequence (3.65) associated with (3.69) also reduces to $\Phi(W_l, W_r) = F(W_l)$. The equivalence of the two formulation in the case of all negative eigenvalues follows similarly and this concludes the proof for a supersonic interface.

B.2 Subsonic interface

The flow is subsonic at the interface when

$$\tilde{\lambda}_1 > 0 \quad \text{and} \quad 0 > \tilde{\lambda}_2 \tag{B.10}$$

while $\tilde{\lambda}_3$ can assume whichever sign. In this case, the matrix $|\tilde{J}|$ in (B.1) can be rewritten as:

$$|\tilde{J}_{lr}| = \begin{pmatrix} \dfrac{\tilde{a}_{lr}^2 - \tilde{u}_{lr}^2}{\tilde{a}_{lr}} & \dfrac{\tilde{u}_{lr}}{\tilde{a}_{lr}} & 0 \\[2ex] \left(\tilde{a}_{lr}^2 - \tilde{u}_{lr}^2\right) \dfrac{\tilde{u}_{lr}}{\tilde{a}_{lr}} & \dfrac{\tilde{a}_{lr}^2 + \tilde{u}_{lr}^2}{\tilde{a}_{lr}} & 0 \\[2ex] \tilde{\xi}_{lr} \left(\dfrac{\tilde{a}_{lr}^2 - \tilde{u}_{lr}^2}{\tilde{a}_{lr}} - |\tilde{u}_{lr}| \right) & \tilde{\xi}_{lr} \dfrac{\tilde{u}_{lr}}{\tilde{a}_{lr}} & |\tilde{u}_{lr}| \end{pmatrix} \tag{B.11}$$

Then, substituting (B.11) in (3.17) it is possible to write the three components of the Roe numerical flux function as follows:

$$\Phi_{lr,1} = \frac{\rho_l u_l + \rho_r u_r}{2} + \frac{\tilde{a}_{lr}^2 - \tilde{u}_{lr}^2}{2\tilde{a}_{lr}} (\rho_l - \rho_r) + \frac{\tilde{u}_{lr}}{2\tilde{a}_{lr}} (\rho_l u_l - \rho_r u_r) \tag{B.12a}$$

$$\begin{aligned} \Phi_{lr,2} = \frac{\rho_l u_l^2 + p_l + \rho_r u_r^2 + p_r}{2} &+ \left(\frac{\tilde{a}_{lr}^2 - \tilde{u}_{lr}^2}{2} \right) \frac{\tilde{u}_{lr}}{\tilde{a}_{lr}} (\rho_l - \rho_r) \\ &+ \frac{\tilde{a}_{lr}^2 + \tilde{u}_{lr}^2}{2\tilde{a}_{lr}} (\rho_l u_l - \rho_r u_r) \end{aligned} \tag{B.12b}$$

$$\begin{aligned} \Phi_{lr,3} = \frac{\rho_l u_l \xi_l + \rho_r u_r \xi_r}{2} &+ \frac{\tilde{\xi}_{lr}}{2} \left(\frac{\tilde{a}_{lr}^2 - \tilde{u}_{lr}^2}{\tilde{a}_{lr}} - |\tilde{u}_{lr}| \right) (\rho_l - \rho_r) \\ &+ \frac{\tilde{\xi}_{lr} \tilde{u}_{lr}}{2\tilde{a}_{lr}} (\rho_l u_l - \rho_r u_r) + \frac{|\tilde{u}_{lr}|}{2} (\rho_l \xi_l - \rho_r \xi_r) \end{aligned} \tag{B.12c}$$

Let us, now, consider φ the numerical flux function defined by (3.65) with b^+, b^- and $\tilde{\delta}$ obtained from (3.69). In the subsonic case, i.e. when (B.10) is satisfied, φ can be expressed as:

$$\begin{aligned}\varphi_{l,r} = & \frac{(\tilde{u}_{lr} + \tilde{a}_{lr})\, F(W_l) - (\tilde{u}_{lr} - \tilde{a}_{lr})\, F(W_r)}{2\tilde{a}_{lr}} \\ & + \frac{\tilde{u}_{lr}^2 - \tilde{a}_{lr}^2}{2\tilde{a}_{lr}} (W_r - W_l) - \frac{\tilde{\rho}_{lr}}{2} \frac{\tilde{u}_{lr}^2 - \tilde{a}_{lr}^2}{|\tilde{u}_{lr}| + \tilde{a}_{lr}} (\xi_r - \xi_l)\, r_3 \end{aligned} \tag{B.13}$$

The first component of (B.13) can be written as:

$$\varphi_{lr,1} = \frac{\tilde{a}_{lr}(\rho_l u_l + \rho_r u_r) + \tilde{u}_{lr}(\rho_l u_l - \rho_r u_r)}{2\tilde{a}_{lr}} + \frac{\tilde{a}_{lr}^2 - \tilde{u}_{lr}^2}{2\tilde{a}_{lr}} (\rho_l - \rho_r) \tag{B.14}$$

and, after simplification, this equation is the same as (B.12a).
To show that (B.12b) is equivalent to the second component of (B.13) it is necessary to recall that from the definition of $\tilde{a}_{lr}$:

$$p_l - p_r = \tilde{a}_{lr}^2 (\rho_l - \rho_r) \tag{B.15}$$

and to use the equality:

$$\rho_l u_l^2 - \rho_r u_r^2 = \tilde{u}_{lr} \left(2 (\rho_l u_l - \rho_r u_r) - \tilde{u}_{lr} (\rho_l - \rho_r)\right) \tag{B.16}$$

which follows from (B.5) and (B.6). Then, the second component of (B.13) can be recasted as

$$\begin{aligned}\varphi_{lr,2} = & \frac{(\rho_l u_l^2 + p_l + \rho_r u_r^2 + p_r)}{2} + \frac{\tilde{u}_{lr}(\rho_l u_l^2 - \rho_r u_r^2)}{2\tilde{a}_{lr}} \\ & + \frac{\tilde{u}_{lr}(p_l - p_r)}{2\tilde{a}_{lr}} + \frac{\tilde{a}_{lr}^2 - \tilde{u}_{lr}^2}{2\tilde{a}_{lr}} (\rho_l u_l - \rho_r u_r) \end{aligned} \tag{B.17}$$

and substituting (B.15) and (B.16) into (B.17), equation (B.12b) is recovered.
For the last component of (B.13) an additional equation is needed:

$$\tilde{u}_{lr}^2 - \tilde{a}_{lr}^2 = (|\tilde{u}_{lr}| + \tilde{a}_{lr})(|\tilde{u}_{lr}| - \tilde{a}_{lr}) \tag{B.18}$$

Using (B.5) and (B.18), the third component of (B.13) can be recasted as

$$\begin{aligned}\varphi_{lr,3} = & \frac{(\rho_l u_l \xi_l + \rho_r u_r \xi_r)}{2} + \frac{\tilde{u}_{lr}(\rho_l u_l \xi_l - \rho_r u_r \xi_r)}{2\tilde{a}_{lr}} + \frac{\tilde{a}_{lr}^2 - \tilde{u}_{lr}^2}{2\tilde{a}_{lr}} (\rho_l \xi_l - \rho_r \xi_r) \\ & - \frac{|\tilde{u}_{lr}| - \tilde{a}_{lr}}{2} \Big(\rho_r \xi_r - \rho_l \xi_l - \tilde{\xi}_{lr} (\rho_r - \rho_l)\Big) \end{aligned} \tag{B.19}$$

Substituting (B.6) and collecting the terms with common factors in (B.19) the result is:

$$\begin{aligned}\varphi_{lr,3} = {} & \frac{(\rho_l u_l \xi_l + \rho_r u_r \xi_r)}{2} + \frac{\tilde{u}_{lr}\left(\tilde{u}_{lr}\left(\rho_l \xi_l - \rho_r \xi_r\right) + \tilde{\rho}_{lr}\tilde{\xi}_{lr}\left(u_l - u_r\right)\right)}{2\tilde{a}_{lr}} \\ & - \frac{|\tilde{u}_{lr}|}{\tilde{a}_{lr}}\frac{|\tilde{u}_{lr}| - \tilde{a}_{lr}}{2}\left(\rho_l \xi_l - \rho_r \xi_r\right) + \tilde{\xi}_{lr}\frac{|\tilde{u}_{lr}| - \tilde{a}_{lr}}{2}\left(\rho_r - \rho_l\right) \quad \text{(B.20)}\end{aligned}$$

Using (B.5) for the term $\tilde{\rho}_{lr}\left(u_l - u_r\right)$ and collecting the terms with common factors:

$$\begin{aligned}\varphi_{lr,3} = {} & \frac{(\rho_l u_l \xi_l + \rho_r u_r \xi_r)}{2} + \frac{|\tilde{u}_{lr}|}{2}\left(\rho_l \xi_l - \rho_r \xi_r\right) + \frac{\tilde{\xi}_{lr}}{2}\frac{\tilde{u}_{lr}}{\tilde{a}_{lr}}\left(\rho_l u_l - \rho_r u_r\right) \\ & - \frac{\tilde{\xi}_{lr}}{2}\frac{\tilde{u}_{lr}}{\tilde{a}_{lr}}\left(\tilde{u}_{lr}\left(\rho_l - \rho_r\right)\right) + \tilde{\xi}_{lr}\frac{|\tilde{u}_{lr}| - \tilde{a}_{lr}}{2}\left(\rho_r - \rho_l\right) \quad \text{(B.21)}\end{aligned}$$

Finally collecting the terms with common factors in (B.21), (B.12c) is recovered and the proof is completed.

Appendix C

Low Mach Number asymptotic study

C.1 The continuous case

Let us, first, recall the principal result concerning the asymptotic behaviour of the continuous equations for a barotropic case (see [102] for details).

Considering an asymptotic expansion in power of $M_\star$ inside the non-dimensional 1-D Euler equations associated by a barotropic state law, the following expression can be obtained for the momentum equation:

$$\frac{1}{M_\star^2}\frac{\partial p_0}{\partial x} + \frac{1}{M_\star}\frac{\partial p_1}{\partial x} + \left(\frac{\partial \rho_0 u_0}{\partial t} + \frac{\partial p_2}{\partial x} + \rho_0 u_0^2\right) + M_\star(\cdots) = 0$$

Thus, in order to solve this equation in the limit $M_\star \to 0$, the following conditions should be satisfied:

$$\frac{\partial p_0}{\partial x} = \frac{\partial p_1}{\partial x} = 0 \tag{C.1}$$

Consequently, when $M_\star \to 0$ the pressure field solution is of the form:

$$p(x,t) = p_0(t) + M_\star p_1(t) + M_\star^2\, \hat{p}_2(x,t) \tag{C.2}$$

Note that a similar behaviour for the pressure has been also observed in the case of a perfect-gas state law (see [57]).

C.2 The semi-discrete case

We carry out, now, an analysis of the semi-discrete governing equation in a general form when the Mach number tends to zero, when the following family of flux functions is considered:

$$\Phi_{ij} = \frac{F(W_i) + F(W_j)}{2} - \frac{1}{2}\, H_{ij}\, \Delta^{ij} W \tag{C.3}$$

where

$$\Delta^{ij} W = \begin{pmatrix} \rho_j - \rho_i \\ \rho_j u_j - \rho_i u_i \end{pmatrix} \quad \text{and} \quad H_{ij} = \begin{pmatrix} h_{11,ij} & h_{12,ij} \\ h_{21,ij} & h_{22,ij} \end{pmatrix} \tag{C.4}$$

H_{ij} is a general upwind matrix.

Note that only the mass and momentum balances are considered here; anyway this study is also valid if a passive scalar equation is present, since this equation is decoupled from the previous ones.

Since it is possible to operate on the upwind term without affecting the consistency in time (see [57]) we want to identify a family of H_{ij} matrix that shows the correct asymptotic behaviour in the low Mach number limit. Defining

$$\bar{\rho}_{ij} \doteq \frac{\rho_j + \rho_i}{2} \quad \bar{u}_{ij} \doteq \frac{u_j + u_i}{2}$$

then the following equality holds true:

$$\rho_j u_j - \rho_i u_i = \bar{\rho}_{ij} \left(u_j - u_i\right) + \bar{u}_{ij} \left(\rho_j - \rho_i\right) \tag{C.5}$$

Introducing the average numerical sound speed:

$$\bar{a}_{ij}^2 \doteq \frac{p_j - p_i}{\rho_j - \rho_i}$$

it is possible to recast (C.3) as

$$\Phi_{ij} = \frac{F(W_i) + F(W_j)}{2} - \frac{1}{2} H_{ij} \begin{pmatrix} \dfrac{p_j - p_i}{\bar{a}_{ij}^2} \\ \\ \bar{\rho}_{ij} \left(u_j - u_i\right) + \dfrac{\bar{u}_{ij} \left(p_j - p_i\right)}{\bar{a}_{ij}^2} \end{pmatrix} \tag{C.6}$$

First of all we adimensionalise the semi-discrete equations:

$$\delta_i \frac{\mathrm{d}W_i}{\mathrm{d}t} + \Phi_{i(i+1)} - \Phi_{(i-1)i} = 0 \tag{C.7}$$

To the purpose, we introduce the following reference quantities: $\delta_\star$ (an arbitrary length scale), $\rho_\star = \max_x \rho(x, 0)$, $u_\star = \max_x u(x, 0)$ and $a_\star^2 = \max_x a^2(x, 0)$. Then, a reference Mach number may be defined as: $M_\star = u_\star / a_\star$ where $a_\star = \sqrt{a_\star^2}$. By exploiting these quantities, we define the non-dimensional variables as: $x' = x/\delta_\star$ $\rho' = \rho/\rho_\star$, $u' = u/u_\star$, $t' = t u_\star/\delta_\star$ and $p' = p/(\rho_\star a_\star^2)$. Substituting the non dimensional variables in (C.7) the following non-dimensional equations are obtained (the prime is dropped for the sake of clarity):

$$\begin{aligned}
2\delta_i\frac{\mathrm{d}\rho_i}{\mathrm{d}t} = {} & (\rho_{i-1}u_{i-1}-\rho_{i+1}u_{i+1}) \\
- {} & \frac{1}{u_\star}\left(\frac{h_{11,i(i+1)}\Delta^{i(i+1)}p}{\bar{a}^2_{i(i+1)}}-\frac{h_{11,(i-1)i}\Delta^{(i-1)i}p}{\bar{a}^2_{(i-1)i}}\right) \\
- {} & h_{12,i(i+1)}\left(\bar{\rho}_{i(i+1)}\left(u_{i+1}-u_i\right)+\frac{\bar{u}_{i(i+1)}\left(p_{i+1}-p_i\right)}{\bar{a}^2_{i(i+1)}}\right) \\
+ {} & h_{12,(i-1)i}\left(\bar{\rho}_{(i-1)i}\left(u_i-u_{i-1}\right)+\frac{\bar{u}_{(i-1)i}\left(p_i-p_{i-1}\right)}{\bar{a}^2_{(i-1)i}}\right)
\end{aligned}\tag{C.8}$$

$$\begin{aligned}
2\delta_i\frac{\mathrm{d}\rho_i u_i}{\mathrm{d}t} = {} & (\rho_{i-1}u^2_{i-1}-\rho_{i+1}u^2_{i+1})+\frac{1}{M^2_\star}\left(p_{i-1}-p_{i+1}\right) \\
- {} & \frac{1}{u^2_\star}\left(\frac{h_{21,i(i+1)}\Delta^{i(i+1)}p}{\bar{a}^2_{i(i+1)}}-\frac{h_{21,(i-1)i}\Delta^{(i-1)i}p}{\bar{a}^2_{(i-1)i}}\right) \\
- {} & \frac{h_{22,i(i+1)}}{u_\star}\left(\bar{\rho}_{i(i+1)}\left(u_{i+1}-u_i\right)+\frac{\bar{u}_{i(i+1)}\left(p_{i+1}-p_i\right)}{\bar{a}^2_{i(i+1)}}\right) \\
+ {} & \frac{h_{22,(i-1)i}}{u_\star}\left(\bar{\rho}_{(i-1)i}\left(u_i-u_{i-1}\right)+\frac{\bar{u}_{(i-1)i}\left(p_i-p_{i-1}\right)}{\bar{a}^2_{(i-1)i}}\right)
\end{aligned}\tag{C.9}$$

We look for semi-discrete solutions to equations (C.9) and (C.8) as an asymptotic expansion in power of $M_\star$:

$$\left\{\begin{array}{l}
\rho_i(t)=\rho_{0i}(t)+M_\star\,\rho_{1i}(t)+M^2_\star\,\rho_{2i}(t)+\cdots \\
u_i(t)=u_{0i}(t)+M_\star\,u_{1i}(t)+M^2_\star\,u_{2i}(t)+\cdots \\
p_i(t)=p_{0i}(t)+M_\star\,p_{1i}(t)+M^2_\star\,p_{2i}(t)+\cdots
\end{array}\right.\tag{C.10}$$

Note 1 *The power expansion (C.10) is applied to* $\bar{u}_{ij}$*,* $\bar{\rho}_{ij}$ *and* $\bar{a}_{ij}$ *as well. Let us consider, as an example, the expansion of* $\bar{a}_{ij}$*.* $\bar{a}_{ij}$ *is always strictly positive. Hence,*

$$\bar{a}_{0ij}=\bar{a}_{ij}|_{M_\star=0}>0$$

and $\bar{a}^{-1}_{ij}$ *may be expanded as follows:*

$$\frac{1}{\bar{a}_{ij}}=\frac{1}{\bar{a}_{0ij}}\left(1+M_\star\frac{\bar{a}_{1ij}}{\bar{a}_{0ij}}+\cdots\right)^{-1}=\frac{1}{\bar{a}_{0ij}}\left(1-M_\star\frac{\bar{a}_{1ij}}{\bar{a}_{0ij}}+\cdots\right)$$

More precisely, $\bar{a}_{0ij}$ *can be explicitly written as follows:*

- *if* $\Delta^{ij}\rho = 0$ *then* $\rho_i = \rho_j = \bar{\rho}$ *and* $\bar{a}_{ij} = a(\bar{\rho})$*:*

  $$\bar{a}_{ij} = a(\bar{\rho}_0) + \frac{da}{d\rho}(\bar{\rho}_0)\bar{\rho}_1 M_\star + \cdots$$

  Clearly, $\bar{a}_{0ij} = a(\bar{\rho}_0)$*;*

- *if* $\Delta^{ij}\rho \neq 0$ *then, by definition:* $\Delta^{ij}\rho\bar{a}^2_{ij} = \Delta^{ij}p$*; if we consider only the terms of order zero in the expansion in power of* M_**, we obtain:* $\Delta^{ij}\rho_0\bar{a}^2_{0ij} = \Delta^{ij}p_0$*. Since, as previously pointed out,* $\bar{a}^2_{0ij} > 0$*, the following equivalence holds:*

  $$\forall i, j \quad \Delta^{ij}\rho_0 = 0 \;\Leftrightarrow\; \Delta^{ij}p_0 = 0$$

  - *if* $\Delta^{ij}\rho_0 \neq 0$*, and, thus,* $\Delta^{ij}p_0 \neq 0$*, then:*

    $$\begin{aligned} \bar{a}_{ij} = & \left(\frac{\Delta^{ij}p}{\Delta^{ij}\rho}\right)^{\frac{1}{2}} \\ = & \left(\frac{\Delta^{ij}p_0}{\Delta^{ij}\rho_0}\right)^{\frac{1}{2}} \frac{\left(1 + \dfrac{\Delta^{ij}p_1}{\Delta^{ij}p_0}M_\star + \cdots\right)^{\frac{1}{2}}}{\left(1 + \dfrac{\Delta^{ij}\rho_1}{\Delta^{ij}\rho_0}M_\star + \cdots\right)^{\frac{1}{2}}} \\ = & \left(\frac{\Delta^{ij}p_0}{\Delta^{ij}\rho_0}\right)^{\frac{1}{2}} + \frac{1}{2}\left(\frac{\Delta^{ij}p_0}{\Delta^{ij}\rho_0}\right)^{\frac{1}{2}} \left(\frac{\Delta^{ij}p_1}{\Delta^{ij}p_0} - \frac{\Delta^{ij}\rho_1}{\Delta^{ij}\rho_0}\right) M_\star + \cdots \end{aligned}$$

    In this case $\bar{a}_{0ij} = \left(\dfrac{\Delta^{ij}p_0}{\Delta^{ij}\rho_0}\right)^{\frac{1}{2}}$*;*

  - *if* $\Delta^{ij}\rho_0 = 0$ *and, thus* $\Delta^{ij}p_0 = 0$*, then, by exploiting the same kind of linearization as above:*

    $$\bar{a}_{0ij} = \left(\frac{\Delta^{ij}p_s}{\Delta^{ij}\rho_s}\right)^{\frac{1}{2}}$$

    where s *is the first integer such that* $\Delta^{ij}\rho_s \neq 0$*.*

As for the continuous case, once the expansions have been introduced into equations (C.9) and (C.8), all the terms associated with $M_\star^{-k}$ $(k > 0)$ must be set to zero in order to find the asymptotic solution. At this point it is necessary to identify one or more families of matrix H_{ij} that allows to recover the correct asymptotic solution, i.e. a pressure solution having a asymptotic behaviour compatible with the one obtained in the continuous case, i.e. as in (C.2).

C.3 A first family of matrix

A first class of suitable matrices can be defined by imposing:

$$h_{11} = c_{11}\,\frac{a_\star}{M_\star^k} \quad \text{with } \ k \geq 1 \tag{C.11}$$

$$h_{12} = c_{12}\,M_\star^k \quad \text{with } \ k \geq 0 \tag{C.12}$$

$$h_{21} = c_{21}\,{a_\star}^2 M_\star^k \quad \text{with } \ k \geq 0 \tag{C.13}$$

$$h_{22} = c_{22}\,a_\star M_\star^k \quad \text{with } \ k \geq 1 \tag{C.14}$$

where c_{11}, c_{12}, c_{21} and c_{22} do not depend on $M_\star$. Note that (C.11) imposes a lower bound while the other three conditions are relative to upper bounds.

Let us first consider the mass equation, if (C.12) is verified for h_{12} then only the second term in the right hand side of (C.8) can have a wrong asymptotic behaviour. Then, if now (C.11) is also verified the following conditions have to be imposed to set to zero the term of order $M_\star^{-k}$ ($k \geq 2$):

$$c_{11,i}\,\frac{\Delta^{i(i+1)}p_0}{\bar{a}^2_{0,i(i+1)}} - c_{11,i-1}\,\frac{\Delta^{(i-1)i}p_0}{\bar{a}^2_{0,(i-1)i}} = 0 \tag{C.15}$$

Moreover, if (C.15) holds, then to set to zero the term of order $M_\star^{-k+1}$ it is sufficient to have:

$$c_{11,i}\,\frac{\Delta^{i(i+1)}p_1}{\bar{a}^2_{0,i(i+1)}} - c_{11,i-1}\,\frac{\Delta^{(i-1)i}p_1}{\bar{a}^2_{0,(i-1)i}} = 0 \tag{C.16}$$

Let us now consider the semi-discrete momentum equation. If h_{22} satisfies (C.14) the last two terms of the right hand side of (C.9) do not have wrong asymptotic behaviour. Let us look at the third term of the right hand side of (C.9) when (C.13) holds; in the worst case, i.e. when $h_{21} = c_{21}\,{a_\star}^2$, the following terms of order $M_\star^{-2}$ and $M_\star^{-1}$ should be taken into account:

1. for order $M_\star^{-2}$:

$$c_{21,i}\,\frac{\Delta^{i(i+1)}p_0}{\bar{a}^2_{0,i(i+1)}} - c_{21,i-1}\,\frac{\Delta^{(i-1)i}p_0}{\bar{a}^2_{0,(i-1)i}} \tag{C.17}$$

2. for order $M_\star^{-1}$:

$$c_{21,i}\,\frac{\Delta^{i(i+1)}p_1}{\bar{a}^2_{0,i(i+1)}} - c_{21,i-1}\,\frac{\Delta^{(i-1)i}p_1}{\bar{a}^2_{0,(i-1)i}} \tag{C.18}$$

The terms defined in (C.17) and (C.18) are similar to the ones which appear in (C.15)-(C.16). If for some θ independent of the index i, we have:

$$c_{21,i} = \theta\, c_{11,i} \quad \forall i \tag{C.19}$$

then from the previous conditions (C.15)-(C.16), (C.17) and (C.18) are equal to zero. Consequently, the only conditions from the momentum equation are coming from the second term in the right hand side of (C.9), which has terms of order $M_\star^{-2}$ and in $M_\star^{-1}$. Nevertheless, these terms can be set to zero by imposing:

1. for order $M_\star^{-2}$:
$$\Delta^{i(i+1)} p_0 + \Delta^{(i-1)i} p_0 = 0 \tag{C.20}$$
2. for order $M_\star^{-1}$:
$$\Delta^{i(i+1)} p_1 + \Delta^{(i-1)i} p_1 = 0 \tag{C.21}$$

Equations (C.15) and (C.20) constitute a homogeneous system of two equations for the two unknowns $\Delta^{(i-1)i}p_0$ and $\Delta^{i(i+1)}p_0$, which is also linearly independent if

$$\frac{c_{11,i}}{\bar{a}^2_{0,i(i+1)}} + \frac{c_{11,i-1}}{\bar{a}^2_{0,(i-1)i}} \neq 0 \quad \forall i \tag{C.22}$$

Note that this condition is not hard to satisfy: for instance, it is sufficient that the $c_{11,i}$ are all different from zero and of the same sign. If (C.22) holds then

$$\Delta^{ij} p_0 = 0 \quad \forall (i,j) \tag{C.23}$$

Therefore p_{0i} is independent of the index i: $p_{0i}(t) = p_0(t)$.

Furthermore, if (C.22) holds then we also have that (C.16) and (C.21) constitute a system of two homogeneous linearly independent equations for the two unknowns $\Delta^{(i-1)i}p_1$ and $\Delta^{i(i+1)}p_1$. As a consequence the only solution is:

$$\Delta^{ij} p_1 = 0 \quad \forall (i,j) \tag{C.24}$$

Therefore p_{1i} is independent of the index i: $p_{1i}(t) = p_1(t)$.

As a conclusion, if (C.11) to (C.14) hold true (with also the conditions on the coefficients (C.19)-(C.22)), when $M_\star \to 0$ the pressure solution to equations (C.8) and (C.9) is of the form:

$$p_i(t) = p_0(t) + M_\star\, p_1(t) + M_\star^2\, \hat{p}_{2i}(t) \tag{C.25}$$

thus recovering, from a qualitative point of view, a behaviour which is similar to that of the continuous case (see (C.2). (C.16) holds true and h_{21} satisfies the third equation of system (C.11)-(C.14). To prove this statement it is sufficient to use an argument similar to that used to show that h_{21} has no effect on the terms associated with $M_\star^{-2}$. valid: if h_{12} is larger than $M_\star^0$ or h_{12} is larger than $a_\star M_\star$ then in (C.15), (C.16) and (C.21) terms associated with $\Delta^{i(i+1)}u$ and $\Delta^{(i-1)i}u$ would appear and (C.25) would be false.

C.3.1 The Rusanov case

The Rusanov upwind matrix is a diagonal matrix with diagonal elements of the form:

$$\lambda = u + a$$

as a consequence we have:

$$\left\{ \begin{array}{lcl} h_{11} & \propto & a_\star \\ h_{12} & = & 0 \\ h_{21} & = & 0 \\ h_{22} & \propto & a_\star \end{array} \right. \tag{C.26}$$

(C.12)-(C.13) are verified but h_{11} is too small and h_{22} [1] is too large. It can be easily checked by direct substitution that with this matrix the asymptotic solution permits pressure oscillation larger than those in the continuous case. It is necessary to modify the Rusanov upwind matrix, in particular it is sufficient to consider:

$$h_{11,p} = h_{11} \cdot f(M)$$

$$h_{22,p} = h_{22} \cdot g(M)$$

with

$$\lim_{M \to 0} f(M) \propto M^{-1}$$

and

$$\lim_{M \to 0} g(M) \propto M^{1}$$

Using $h_{11,p}$ and $h_{22,p}$ instead of h_{11} and h_{22} makes the Rusanov matrix to satisfy (C.11) to (C.14) and then the correct behaviour of the asymptotic solution is recovered.

C.4 A wider class of matrix

It is possible to define a wider class of matrix, relaxing the first equation of system (C.11). If we consider

$$h_{11} = c_{11}\, a_\star M_\star^k \quad \text{with} \quad k = 0 \tag{C.27}$$

---

[1] This term is the main responsible for the generation of spurious pressure oscillations

then, substituting in (C.8) and (C.9) and following the steps of section (C.3) we obtain:

$$\Delta^{ij} p_0 = 0 \quad \forall (i,j) \tag{C.28}$$

$$p_{1,(i-1)} = p_{1,(i+1)} \quad \forall i \tag{C.29}$$

Equation (C.29) is an approximation of $\frac{\partial p_1(x,t)}{\partial x} = 0$. This condition is sometimes considered sufficient especially when the numerical scheme is used with unstructured meshes [79].

Appendix D

Algorithm to compute the steady solution for the nozzle

In this chapter we study the problem of a barotropic flow in a nozzle in the one dimensional case. In this study we are not concerned with the assumptions used to transform the Navier-Stokes equations in the 1D nozzle problem. Thus, for the sake of simplicity, we state that we want to solve the following set of equations:

$$\begin{cases} \dfrac{\partial}{\partial t}\displaystyle\int_\alpha^\beta A\rho \,\mathrm{d}x + \rho u A|_\beta - \rho u A|_\alpha = 0 \\[2ex] \dfrac{\partial}{\partial t}\displaystyle\int_\alpha^\beta A\rho u \,\mathrm{d}x + \rho u^2 A|_\beta - \rho u^2 A|_\alpha = -pA|_\beta + pA|_\alpha + \overline{p}(A(\beta) - A(\alpha)) \\[2ex] p = p(\rho) \end{cases} \tag{D.1}$$

In this system of equation α and β are two generic points on the x-axis, p, ρ and u are unknown functions of space and time, while A is a known function depending only on the space variable[1]. Hence, we can write:

$$p = p(x,t) \quad \rho = \rho(x,t) \quad u = u(x,t) \quad A = A(x)$$

Finally $\overline{p}$ is the space average of p over the interval $[\alpha, \beta]$ and the system is closed with a barotropic state law $p = p(\rho)$. In a general context, it is not possible to substitute p with ρ in the first two equations of system (D.1) because the state law is seldom in explicit form.

At this point of the study we are interested in a steady solution when $A \in C^1$,

---

[1] The area of the duct for our problem

then system (D.1) can be reduced to:

$$\begin{cases} \rho u A|_\beta - \rho u A|_\alpha = 0 \\ \\ \rho u^2 A|_\beta - \rho u^2 A|_\alpha = -pA|_\beta + pA|_\alpha + \overline{p}(A(\beta) - A(\alpha)) \\ \\ p = p(\rho) \end{cases} \tag{D.2}$$

Since α and β are arbitrary, the first equation of system (D.2) can be recasted as

$$\rho u A = \text{constant} \doteq \dot{\mathcal{M}}$$

and consequently (D.2) can be rewritten as:

$$\begin{cases} \rho u A = \dot{\mathcal{M}} \\ \\ \dot{\mathcal{M}}(u|_\beta - u|_\alpha) = -pA|_\beta + pA|_\alpha + \overline{p}(A(\beta) - A(\alpha)) \\ \\ p = p(\rho) \end{cases} \tag{D.3}$$

Then an explicit algorithm to compute the analytical solution

$$\rho = \rho(A(x)) \quad u = u(A(x))$$

is achievable under the hypothesis that ρ and u are piecewise differentiable functions, with the possibility to have a finite number of discontinuous points [2].

D.1 Equations for a discontinuity

Let suppose that $\bar{x}$ is a point of discontinuity, then system (D.3) can be recasted as

$$\begin{cases} \rho u|_{\bar{x}^-} = \rho u|_{\bar{x}^+} \\ \\ (p + \rho u^2)|_{\bar{x}^-} = (p + \rho u^2)|_{\bar{x}^+} \\ \\ p = p(\rho) \end{cases} \tag{D.4}$$

Thus, in presence of a discontinuity, we obtain the well known Rankine-Hugoniot equations for a normal shock.

---

[2] Many situations of practical interest satisfy this requirement so it is not a strong restriction

D.2 Analytical solutions in absence of discontinuity

If we look for solutions that are piecewise differentiable, it is possible to write system (D.3) as:

$$\begin{cases} \displaystyle\int_\alpha^\beta \frac{\partial \rho u A}{\partial x}\,\mathrm{d}x = 0 \\ \displaystyle\int_\alpha^\beta \frac{\partial \rho u^2 A}{\partial x}\,\mathrm{d}x = -\int_\alpha^\beta \frac{\partial p A}{\partial x}\,\mathrm{d}x + \int_\alpha^\beta p\frac{\partial A}{\partial x}\,\mathrm{d}x \\ p = p(\rho) \end{cases} \tag{D.5}$$

These relations are independent from the particular choice of the interval $[\alpha, \beta]$ so the result is:

$$\begin{cases} \displaystyle\frac{\partial \rho u A}{\partial x} = 0 \\ \displaystyle\frac{\partial \rho u^2 A}{\partial x} = -A\frac{\partial p}{\partial x} \\ p = p(\rho) \end{cases} \tag{D.6}$$

Substituting the first equation in the second:

$$\frac{\dot{\mathcal{M}}}{A}\frac{\partial u}{\partial x} = -\frac{\partial p}{\partial x} \tag{D.7}$$

and using the equality

$$u = \frac{\dot{\mathcal{M}}}{\rho A}$$

it is possible to derive the following ordinary differential equation for density:

$$\frac{\dot{\mathcal{M}}^2}{\rho^3 A^3}\frac{\partial \rho A}{\partial x} = \frac{1}{\rho}\frac{\partial p}{\partial x} \tag{D.8}$$

If we integrate over interval $[\alpha, \beta]$, we obtain:

$$\frac{\dot{\mathcal{M}}^2}{2}\left(\frac{1}{\rho^2(\alpha)A^2(\alpha)} - \frac{1}{\rho^2(\beta)A^2(\beta)}\right) = \int_{\rho(\alpha)}^{\rho(\beta)} \frac{a^2}{\rho}\mathrm{d}\rho \tag{D.9}$$

where, due to the barotropic closure of the system a^2, defined as $\frac{\mathrm{d}p}{\mathrm{d}\rho}$, is a function of ρ only. Finally, thanks to the additive property of integrals, it is possible to rewrite (D.9) as follows:

$$\frac{\dot{\mathcal{M}}^2}{2}\frac{1}{\rho^2(\alpha)A^2(\alpha)} + \int_{\rho(\gamma)}^{\rho(\alpha)} \frac{a^2}{\rho}\mathrm{d}\rho = \frac{\dot{\mathcal{M}}^2}{2}\frac{1}{\rho^2(\beta)A^2(\beta)} + \int_{\rho(\gamma)}^{\rho(\beta)} \frac{a^2}{\rho}\mathrm{d}\rho \tag{D.10}$$

where γ is an arbitrary point. Now we can define the function:

$$\mathcal{K}(\rho, A) \doteq \frac{\dot{\mathcal{M}}^2}{2}\frac{1}{\rho^2 A^2} + \int_{\bar{\rho}}^{\rho}\frac{a^2(\chi)}{\chi}\mathrm{d}\chi$$

With this notation the solution of (D.5) must satisfy:

$$\mathcal{K}(\rho, A) = \mathcal{K}_0 \tag{D.11}$$

where $\mathcal{K}_0$ is a constant that can be defined using the boundary conditions.

D.2.1 Properties of the solution

The definition of $\mathcal{K}$ is a compact notation to define the relation that a solution to problem (D.6) must satisfy. However, yet we have no information on the existence and uniqueness of the solution. Before answering these questions we have to show the main features of the function $\mathcal{K}$. The first step is to take the derivative of $\mathcal{K}$ respect to ρ:

$$\frac{\partial \mathcal{K}}{\partial \rho} = \frac{a^2}{\rho} - \frac{u^2}{\rho} = \frac{a^2}{\rho}(1 - M^2) \tag{D.12}$$

Since $\frac{a^2}{\rho}$ is always positive the sign of $\frac{\partial \mathcal{K}}{\partial \rho}$ depends only on M, the Mach number.

It can be easily shown that $\frac{\partial \mathcal{K}}{\partial A}$ is always negative, thus, we have proved that, as in the case of perfect gas, area variations have opposite effects on subsonic and supersonic flows. Indeed, let us assume that ρ_1 and ρ_2 solve equation (D.12) for A_1 and A_2. If $A_1 > A_2$ then $\mathcal{K}_0 = \mathcal{K}(\rho_1, A_1) < \mathcal{K}(\rho_1, A_2)$: due to equation (D.12) if $M < 1$ then ρ_2 must be lower than ρ_1, while if $M > 1$ then ρ_1 must be lower than ρ_2.

D.3 Solution in the presence of shocks

It is worth pointing out that only the second equation in system (D.4) is important. Indeed through a discontinuity we have $A^+ = A^- = A(x)$ thus, if ρ^+, u^+ and ρ^-, u^- are the C^1 solutions on the two sides of the discontinuity, due to the first equation in (D.3), they satisfy also the first in (D.4). The second of (D.4) can be recasted as follow

$$(p + \frac{\dot{\mathcal{M}}^2}{\rho A^2})|_{\bar{x}^-} = (p + \frac{\dot{\mathcal{M}}^2}{\rho A^2})|_{\bar{x}^+} \tag{D.13}$$

This equation is function only of ρ^+ and ρ^-.

D.4 Algorithm for the computation of the solution

It is important to analyse the derivative of the Mach number with respect to density:

$$\frac{\partial M}{\partial \rho} = -\frac{\dot{\mathcal{M}}}{A\rho^2 a^2} \cdot \frac{\partial \rho a}{\partial \rho} \tag{D.14}$$

If:

$$\frac{\partial \rho a}{\partial \rho} > 0 \quad \forall \rho \tag{D.15}$$

then it is possible to write an algorithm to compute the solution of problem (D.3) when a finite number of discontinuities is present. Indeed from (D.12) we have that:

$$\frac{\partial \mathcal{K}}{\partial \rho} = 0 \Longleftrightarrow M = 1 \tag{D.16}$$

On another hand, assuming that (D.15) is satisfied and considering $\dot{\mathcal{M}} > 0$, i.e. $u > 0$ (note that this assumption can be made without loosing generality), from (D.14) M is a strictly decreasing function of ρ. Consequently, for each possible value $M^\star$ of M there is only one value $\rho^\star$ of ρ such that $M(\rho^\star) = M^\star$. In particular, from (D.16), there is only one density value $\hat{\rho}$ which makes $\frac{\partial \mathcal{K}}{\partial \rho}$ vanishing. Note that $\hat{\rho}$ should depend on the area A. Furthermore for $M < 1$, i.e. $\rho > \hat{\rho}$, we have from (D.12) that $\frac{\partial \mathcal{K}}{\partial \rho} > 0$ while for $M > 1$, i.e. $\rho < \hat{\rho}$, we have $\frac{\partial \mathcal{K}}{\partial \rho} < 0$. Thus $\hat{\rho}$ is a minimum for $\mathcal{K}$ that is

$$\mathcal{K}(\hat{\rho}, A) < \mathcal{K}(\rho, A) \quad \forall \rho \neq \hat{\rho}$$

As a consequence we have three different situations:

$$\begin{cases} \mathcal{K}(\hat{\rho}(A), A) > \mathcal{K}_0 & 0 \; \textit{solutions} \\ \mathcal{K}(\hat{\rho}(A), A) = \mathcal{K}_0 & 1 \; \textit{solutions} \\ \mathcal{K}(\hat{\rho}(A), A) < \mathcal{K}_0 & 1 \; \textit{or} \, 2 \; \textit{solutions} \end{cases} \tag{D.17}$$

Even if there are two solutions of equation (D.11), one being always subsonic and the other supersonic, only one of them satisfies the hypothesis that the solution is C^1 from which (D.11) is derived. Thus, equation (D.17) answers the questions of existence and unicity of the solution when there are no discontinuities. For the sake of simplicity, assume that there is only one discontinuity. Then the solution consists of two smooth regions, with constants $\mathcal{K}_0$ and $\mathcal{K}_1$[3], divided by a shock. In the two smooth regions (D.17) holds true, while at interface (D.13) must be considered. Defining:

$$\mathcal{S}(\rho, A) \doteq p(\rho) + \frac{\dot{\mathcal{M}}^2}{\rho A^2}$$

---

[3] It can be easily checked that $\mathcal{K}_0 \neq \mathcal{K}_1$ must be true to have a discontinuity

it can be easily obtained that:

$$\frac{\partial \mathcal{S}(\rho, A)}{\partial \rho} = a^2 \left(1 - M^2\right) \tag{D.18}$$

It is possible to recast equation (D.13) as:

$$\mathcal{S}(\rho^+, A) = \mathcal{S}(\rho^-, A)$$

and, due to (D.18), this equation can have one or two solutions. The solution

$$\rho^+ = \rho^-$$

is always possible and is not a discontinuity: if this is the unique solution, then it is impossible to have discontinuities. When another solution is possible, the couple ρ^+ and ρ^- are the solution in presence of a discontinuity for a given area of the nozzle. It is worth noting that one between the + and − state is always supersonic and the other subsonic. Even if we have not introduced any entropy inequality, due to the physical meaning of the problem that we are studying, only compression shock wave (transition from supersonic to subsonic) will be considered acceptable. In order to define the position of the shock we have to find $\hat{A}$ that solve the system:

$$\begin{cases} \mathcal{K}(\rho^+, \hat{A}) & = & \mathcal{K}_0 \\ \mathcal{K}(\rho^-, \hat{A}) & = & \mathcal{K}_1 \\ \mathcal{S}(\rho^+, \hat{A}) & = & \mathcal{S}(\rho^-, \hat{A}) \end{cases} \tag{D.19}$$

D.4.1 Algorithm to compute the solution, with one discontinuity

The solution described in the previous sections can be computed using the following steps:

1. Choose a particular value of A, $A^\star$
2. Compute $\hat{\rho}(A^\star)$
3. Compute $\mathcal{K}(\hat{\rho}(A^\star), A^\star)$ and check the existence of the solution both for $\mathcal{K}_0$ and $\mathcal{K}_1$
4. If the solution exist, solve $\mathcal{K}(\rho_0, A^\star) = \mathcal{K}_0$ and $\mathcal{K}(\rho_1, A^\star) = \mathcal{K}_1$
5. Finally check if system (D.19) holds true with ρ_0, ρ_1 and $A^\star$ to find the position of the discontinuity.

This algorithm works if a discontinuity is present, however it is possible to take into account the presence of more discontinuities.

List of Figures

List of Tables

Bibliography

[1] I. Abalakin, A. Dervieux, and T. Kozubskaya, *A vertex centered high order MUSCL scheme applying to linearised Euler acoustics*, INRIA Rapport de Recherche 4459, INRIA, 2002.

[2] I. Abalakin and B. Koobus, *The Behavior of Two Near-Wall Models for $k - \varepsilon$ Prediction of Stall*, INRIA Rapport de Recherche 4075, INRIA, 2000.

[3] M. R. Baer and J. W. Nunziato, *A two-phase mixture theory for the deflagration-to-detonation transition (ddt) in reactive granular materials*, International Journal of Multiphase Flow, 12 (1986), pp. 861–889.

[4] P. Batten, M. A. Leschziner, and U. C. Goldberg, *Average-state* Jacobians *and implicit methods for compressible viscous and turbulent flows*, Journal of Computational Physics, 137 (1997), pp. 38–78.

[5] F. Benkhaldoun, S. Daoudi, I. Elmahi, and M. Seaïd, *Numerical modelling of sediment transport in the Nador lagoon (Morocco)*, (preprint).

[6] F. Benkhaldoun, I. Elmahi, and M. Seaïd, *Well-balanced finite volume schemes for pollutant transport by shallow water equations on unstructured meshes*, Journal of Computational Physics, 226 (2007), pp. 180–203.

[7] F. Benkhaldoun, S. Sahmim, and M. Seaïd, *Solution of the Sediment Transport Equations Using a Finite Volume Method Based on Sign Matrix*, SIAM Journal on Scientific Computing, 31 (2009), pp. 2866–2889.

[8] F. Benkhaldoun, S. Sahmim, and M. Seaïd, *A two-dimensional finite volume morphodynamic model on unstructured triangular grids*,

International Journal for Numerical Methods in Fluids, 63 (2010), pp. 1296–1327.

[9] F. Benkhaldoun, S. Sahmim, and M. Seaïd, *Mathematical development and verification of a finite volume model for morphodynamic flow applications*, (preprint).

[10] A. Bermudez and M. Vazquez, *Upwind methods for hyperbolic conservation laws with source terms*, Computers & Fluids, 23 (1994), pp. 1049–1071.

[11] F. Beux, M. Bilanceri, G. Pagano, and M. V. Salvetti, *TN6: Work Package WP3300 - 2D Cavitation Model Development*, technical Report for the ESA Project: *Scaling of thermal cavitation effects on cavitation-induced instabilities.*

[12] F. Beux, M. Bilanceri, and M. V. Salvetti, *TN5: Work Package WP3200 - 1D Cavitation Model Development*, technical Report for the ESA Project: *Scaling of thermal cavitation effects on cavitation-induced instabilities.*

[13] F. Beux, M. Bilanceri, and M. V. Salvetti, *TN7: Work Package WP3400 - 3D Cavitation Model Development*, technical Report for the ESA Project: *Scaling of thermal cavitation effects on cavitation-induced instabilities.*

[14] F. Beux, M. Salvetti, A. Ignatyev, D. Li, C. Merkle, and E. Sinibaldi, *A numerical study of non-cavitating and cavitating liquid flow around a hydrofoil*, Mathematical Modelling and Numerical Analysis, 39 (2005), pp. 577–590.

[15] M. Bilanceri, F. Beux, I. Elmahi, H. Guillard, and M. V. Salvetti, *Linearised implicit time-advancing applied to sediment transport simulations*, INRIA Rapport de Recherche 7492, 2010.

[16] M. Bilanceri, F. Beux, and M. V. Salvetti, *Investigation on Numerical Schemes, Preconditioning and Time Advancing in the Simulation of 1D Cavitating Flows*, ADIA 2008-7, University of Pisa, October 2008.

[17] M. Bilanceri, F. Beux, and M. V. Salvetti, *Investigation on Numerical Schemes in the Simulation of Barotropic Cavitating Flows*, in Proceedings of the 7^{th} International Symposium on Cavitation, Ann Arbor, Michigan, USA, 17–22 August 2009.

[18] M. Bilanceri, F. Beux, and M. V. Salvetti, *An Implicit Low-Diffusive HLL Scheme with Complete Time Linearization: Application to Cavitating Barotropic Flows*, Computer & Fluids, 39 (2010), pp. 1990–2006.

[19] M. Bilanceri, F. Beux, and M. V. Salvetti, *An Implicit Low-Diffusive HLL Scheme for Cavitating Flow Simulation*, in V European Conference on Computational Fluid Dynamics, Lisbon, Portugal, 14–17 June 2010.

[20] M. Bilanceri, I. Elmahi, H. Guillard, M. V. Salvetti, and F. Beux, *Implicit Time-Advancing Applied to Shallow-Water Problems with Mobile Bed.* Presented at the Workshop on Numerical Method for Interaction between Sediment & Water, Paris, France, 20–24 September 2010.

[21] M. Bilanceri, H. Guillard, M. V. Salvetti, and F. Beux, *Implicit Simulations of Shallow-Water Equations with Mobile Bed*, in V European Conference on Computational Fluid Dynamics, Lisbon, Portugal, 14–17 June 2010.

[22] M. Bilanceri, E. Sinibaldi, F. Beux, and M. V. Salvetti, *A Preconditioned Second-Order Linearized Implicit Formulation For Barotropic Cavitating Flows.* Presented at the 8^{th} World Congress on Computational Mechanics (WCCM8), 5^{th} European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS 2008), Venice, Italy, 30 June– 5 July 2008.

[23] J. Bredberg, *Turbulence Modelling for Internal Cooling of Gas-Turbine Blades*, PhD thesis, Chalmers University of Technology, 2002.

[24] S. Camarri, M. Salvetti, B. Koobus, and D. A., *A low diffusion* MUSCL *scheme for* LES *on unstructured grids*, Computer & Fluids, 33 (2004), pp. 1101–1129.

[25] R. Campos-Amezcua, S. Kehelladi, F. Bakir, Z. Mazur-Czerwiec, C. Sarraf, and R. Rey, *Numerical analysis of unsteady cavitating flow in an axial inducer*, Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy, 224 (2010), pp. 223–238.

[26] M. Castro Díaz, E. Fernández-Nieto, and A. Ferreiro, *Sediment transport models in Shallow Water equations and numerical ap-*

proach by high order finite volume methods, Computers & Fluids, 37 (2008), pp. 299–316.

[27] M. Castro Díaz, E. Fernández-Nieto, A. Ferreiro, J. García-Rodríguez, and C. Parés, *High Order Extensions of Roe Schemes for Two-Dimensional Nonconservative Hyperbolic Systems*, Journal of Scientific Computing, 39 (2009), pp. 67–114.

[28] M. Castro Díaz, E. Fernández-Nieto, A. Ferreiro, and C. Parés, *Two-dimensional sediment transport models in shallow water equations. A second order finite volume approach on unstructured meshes*, Computer Methods in Applied Mechanics and Engineering, 198 (2009), pp. 2520–2538.

[29] M. Castro Díaz, J. Galardo, and C. Parés, *High order finite volume schemes based on reconstruction of states for solving hyperbolic systems with nonconservative products. Applications to shallow-water systems*, Mathematics of Computation, 75 (2006), pp. 1103–1134.

[30] A. Cervone, C. Bramanti, E. Rapposelli, and L. d'Agostino, *Thermal cavitation experiments on a* NACA0015 *hydrofoil*, ASME Journal of Fluids Engineering, 128 (2006), pp. 326–331.

[31] A. Cervone, L. Torre, A. Pasini, and L. d'Agostino, *Cavitation and turbopump hydrodynamics research at alta s.p.a. and pisa university*, in 4^{th} International Symposium on Fluid Machinery and Fluid Engineering, Beijing, China, 24–27 November 2008.

[32] A. Cervone, L. Torre, A. Pasini, and L. d'Agostino, *Cavitation and flow instabilities in a 3-bladed axial inducer designed by means of a reduced order analytical model*, in Proceedings of the 7^{th} International Symposium on Cavitation, Ann Arbor, Michigan, USA, 17–22 August 2009.

[33] O. Coutier-Delgosha, R. Fortes-Patella, J. Reboud, N. Hakimi, and C. Hirsch, *Numerical simulation of cavitating flow in 2*D *and 3*D *inducer geometries*, International Journal for Numerical Methods in Fluids, 48 (2005), pp. 135–167.

[34] O. Coutier-Delgosha, R. Fortes-Patella, J.-L. Reboud, N. Hakimi, and C. Hirsch, *Stability of preconditioned* N*avier-*S*tokes equations associated with a cavitation model*, Computers & Fluids, 34 (2005), pp. 319–349.

[35] O. COUTIER-DELGOSHA, P. MOREL, R. FORTES-PATELLA, AND J. L. REBOUD, *Numerical simulation of turbopump inducer cavitating behavior*, International Journal of Rotating Machinery, 2 (2005), pp. 135–142.

[36] B. CUI, Z. ZHU, AND Y. LIN, *Numerical simulation of inner flow in equal-pitch inducer*, Jixie Gongcheng Xuebao/Journal of Mechanical Engineering, 46 (2010), pp. 158–163.

[37] L. D'AGOSTINO, E. RAPPOSELLI, C. PASCARELLA, AND A. CIUCCI, *A modified bubbly isenthalpic model for numerical simulation of cavitating flows*, in 37^{th} AIAA/ASME/SAE/ASEE Joint Propulsion Conference, Salt Lake City, UT, USA., 2001.

[38] L. D'AGOSTINO, L. TORRE, A. PASINI, D. BACCARELLA, A. CERVONE, AND A. MILANI, *A reduced order model for preliminary design and performance prediction of tapered inducers: Comparison with numerical simulations*, in 44^{th} AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, Hartford, Connecticut, USA, 21–23 July 2008.

[39] L. D'AGOSTINO, L. TORRE, A. PASINI, AND A. CERVONE, *On the preliminary design and noncavitating performance of tapered axial inducers*, ASME Journal of Fluids Engineering, 130 (2008), pp. 111303–1–111303–8.

[40] L. D'AGOSTINO, L. TORRE, A. PASINI, AND A. CERVONE, *A reduced order model for preliminary design and performance prediction of tapered inducers*, in 12^{th} International Symposium on Transport Phenomena and Dynamics of Rotating Machinery, Honolulu, Hawaii, USA, 12–17 February 2008.

[41] S. DAVIS, *Simplified second-order* G*odunov-type methods*, SIAM Journal on Scientific and Statistical Computing, 9 (1988), pp. 445–473.

[42] Y. DELANNOY AND J. KUENY, *Cavity flow predictions based on the* E*uler equations*, ASME Cavitation and Multiphase Flow Forum, 1990, pp. 153–158.

[43] J. A. DÉSIDÉRI AND P. W. HEMKER, *Convergence analysis of the defect-correction iteration for hyperbolic problems*, SIAM Journal on Scientific Computing, 16 (1995), pp. 88–118.

[44] B. EINFELDT, C. MUNZ, P. L. ROE, AND B. SJÖGREEN, *On* G*odunov-type methods near low densities*, Journal of Computational Physics, 92 (1991), pp. 273–295.

[45] C. Farhat, L. Fezoui, and S. Lanteri, *Computational fluid dynamics with irregular grids on the connection machine*, INRIA Rapport de Recherche 1411, INRIA, 1991.

[46] E. Faucher, J. Herardc, M. Barret, and C. Toulemonde, *Computation of Flashing Flows In Variable Cross-Section Ducts*, International Journal of Computational Fluid Dynamics, 13 (2000), pp. 365–391.

[47] G. Fernandez, *Simulation numérique d'écoulements réactifs à petit nombre de Mach*, PhD thesis, Université de Nice, 1989.

[48] L. Fézoui and B. Stoufflet, *A class of implicit upwind schemes for Euler simulations with unstructured meshes*, INRIA Rapport de Recherche 517, INRIA, 1986.

[49] L. Fézoui and B. Stoufflet, *A class of implicit upwind schemes for* E*uler simulations with unstructured meshes*, Journal of Computational Physics, 84 (1989), pp. 174–206.

[50] R. Fortes-Patella, O. Coutier-Delgosha, J. Perrin, and J. L. Reboud, *Numerical model to predict unsteady cavitating flow behavior in inducer blade cascades*, Journal of Fluids Engineering, Transactions of the ASME, 129 (2007), pp. 128–135.

[51] T. Gallouët, J. Hérard, and N. Seguin, *Some approximate Godunov schemes to compute Shallow-Water equations with topography*, Computers & Fluids, 32 (2003), pp. 479–513.

[52] E. Godlewski and P. Raviart, *Numerical approximation of hyperbolic systems of conservation laws*, no. 118 in Applied Mathematical Sciences, Springer, 1996.

[53] S. K. Godunov, *A difference scheme for numerical solution of discontinuous solution of hydrodynamic equations*, Math. Sbornik, 47 (1959), pp. 271–306. Translated US Joint Publ. Res. Service, JPRS 7226, 1969.

[54] E. Goncalves and R. F. Patella, *Numerical simulation of cavitating flows with homogeneous models*, Computers & Fluids, 38 (2009), pp. 1682–1696.

[55] S. Gottlieb and C. W. Shu, *Total variation diminishing runge-kutta schemes*, Mathematics of Computation, 67 (1998), pp. 73–85.

[56] A. Grass, *Sediments transport by waves and currents*, Report No. FL29, SERC London Cent. Mar. Technol., 1981.

[57] H. Guillard and C. Viozat, *On the behaviour of upwind schemes in the low* M*ach number limit*, Computers & Fluids, 28 (1999), pp. 63–86.

[58] A. Harten, P. Lax, and B. van Leer, *On upstreaming differencing and* G*odunov-type schemes for hyperbolic conservation laws*, SIAM Review, 25 (1983), pp. 35–61.

[59] L. Hascoët and V. Pascual, *TAPENADE 2.1 User's Guide*, Technical Report n 300, INRIA, 2004.

[60] K. Hejranfar, E. Ezzatneshan, and K. F. Hesary, *A dual-time implicit preconditioned Navier-Stokes method for solving 2D steady/unsteady laminar cavitating/noncavitating flows using a Barotropic model* , in Proceedings of the 7^{th} International Symposium on Cavitation, Ann Arbor, Michigan, USA, 17–22 August 2009.

[61] J. Hinze, *Turbulence*, McGraw-Hill, 1959.

[62] H. Hoeijmakers, M. Janssens, and W. Kwan, *Numerical simulation of sheet cavitation*, in Proc. Third International Symposium on Cavitation, Grenoble (France), April 1998.

[63] A. Hosangadi, V. Ahuja, R. J. Ungewitter, and J. Busby, *Analysis of thermal effects in cavitating liquid hydrogen inducers*, Journal of Propulsion and Power, 23 (2007), pp. 1225–1234.

[64] J. Hudson and P. K. Sweby, *Formulations for Numerically Approximating Hyperbolic Systems Governing Sediment Transport*, Journal of Scientific Computing, 19 (2003), pp. 225–252.

[65] M. Ishii, *Thermo-fluid dynamic theory of two-phase flow*, Paris, Eyrolles, 1975.

[66] D. Kang, K. Yonezawa, T. Ueda, N. Yamanishi, C. Kato, and Y. Tsujimoto, *Large eddy simulation of the dynamic response of an inducer to flow rate fluctuations*, in AIP Conference Proceedings, vol. 1225, 2010, pp. 769–779.

[67] A. K. Kapila, R. Menikoff, J. B. Bdzil, S. F. Son, and D. S. Stewart, *Two-phase modeling of deflagration-to-detonation transition in granular materials: Reduced equations*, Physics of Fluids, 13 (2001), pp. 3002–3024.

[68] S.-E. Kim, *A Numerical Study of Unsteady Cavitation on A Hydrofoil*, in Proceedings of the 7^{th} International Symposium on Cavitation, Ann Arbor, Michigan, USA, 17–22 August 2009.

[69] T. Kimura, Y. Yoshida, T. Hashimoto, and M. Shimagaki, *Numerical simulation for vortex structure in a turbopump inducer: Close relationship with appearance of cavitation instabilities*, Journal of Fluids Engineering, Transactions of the ASME, 130 (2008), pp. 0511041–0511049.

[70] D. Kleine, *Numerical simulation of unsteady flow in hydraulic turbomachines*, PhD thesis, Technische Universiteit Eindhoven, 2009.

[71] F. Kong, H. Zhang, X. Zhang, and Z. Wang, *Design on variable-pitch inducer based on numerical simulation for cavitation flow*, Paiguan Jixie Gongcheng Xuebao/Journal of Drainage and Irrigation Machinery Engineering, 28 (2010), pp. 12–17.

[72] A. Koop, *Numerical Simulation of Unsteady Three-Dimensional Sheet Cavitation*, PhD thesis, University of Twente, 2008.

[73] A. Koop and H. Hoeijmakers, *Numerical Simulation of Unsteady Three-Dimensional Sheet Cavitation*, in Proceedings of the 7^{th} International Symposium on Cavitation, Ann Arbor, Michigan, USA, 17–22 August 2009.

[74] R. Kunz, D. Boger, D. Stinebring, T. Chyczewski, J. Lindau, H. Gibeling, S. Venkateswaran, and T. Govindan, *A preconditioned* N*avier-*S*tokes method for two-phase flows application to cavitation prediction*, Computers & Fluids, 29 (2000), pp. 849–875.

[75] S. Lanteri, *Simulation d'écoulements aérodynamiques Instationnaires sur une Architecture S.I.M.D. Massivement Parallèle*, PhD thesis, Université de Nice, 1991.

[76] B. Launder and D. Spalding, *The numerical computation of turbulent flows*, Computer Methods in Applied Mechanics and Engineering, 3 (1974), pp. 269–289.

[77] O. Le Métayer, J. Massoni, and R. Saurel, *Modelling evaporation fronts with reactive riemann solvers*, Journal of Computational Physics, 205 (2005), pp. 567–610.

[78] R. J. Leveque, *Numerical Methods for Conservation Laws*, Birkhäuser-Verlag, Basel, 1994.

[79] X.-s. Li and C.-w. Gu, *An all-speed* Roe*-type scheme and its asymptotic analysis of low* Mach *number behaviour*, Journal of Computational Physics, 227 (2008), pp. 5144–5159.

[80] T. Liu, B. Khoo, and W. Xie, *Isentropic one-fluid modelling of unsteady cavitating flow*, Journal of Computational Physics, 201 (2004), pp. 80–108.

[81] R. Martin and H. Guillard, *A second order defect correction scheme for unsteady problems*, Computers & Fluids, 25 (1996), pp. 9–27.

[82] I. Mejri, F. Bakir, R. Rey, and T. Belamri, *Comparison of computational results obtained from a homogeneous cavitation model with experimental investigations of three inducers*, Journal of Fluids Engineering, Transactions of the ASME, 128 (2007), pp. 1308–1323.

[83] C. Parés and M. Castro Díaz, *On the well-balance property of Roe's method for nonconservative hyperbolic systems. applications to shallow-water systems*, ESAIM: Mathematical Modelling and Numerical Analysis, 38 (2004), pp. 821–852.

[84] S. Park and J. Kwon, *On the dissipation mechanism of* G*odunov-type schemes*, Journal of Computational Physics, 188 (2003), pp. 524–542.

[85] B. Pouffary, J. L. Patella, R. F.and Reboud, and P. A. Lambert, *Numerical analysis of cavitation instabilities in inducer blade cascade*, Journal of Fluids Engineering, Transactions of the ASME, 130 (2008), pp. 0413021–0413028.

[86] B. Pouffary, R. F. Patella, J. L. Reboud, and P. A. Lambert, *Numerical simulation of 3d cavitating flows: Analysis of cavitation head drop in turbomachinery*, Journal of Fluids Engineering, Transactions of the ASME, 130 (2008), pp. 061301–061310.

[87] Q. Qin, C. Song, and R. Arndt, *A virtual single-phase natural cavitation model and its application to* CAV*2003 hydrofoil*, in Proc. CAV2003 - Fifth International Symposium on Cavitation, Osaka (Japan), November 2003.

[88] T. Rebollo, A. Delgado, and E. Fernández-Nieto, *A family of stable numerical solvers for the shallow water equations with source terms*, Computer Methods in Applied Mechanics and Engineering, 192 (2003), pp. 203–225.

[89] T. Rebollo, E. Fernández-Nieto, and M. Mármol, *A flux-splitting solver for Shallow Water equations with source terms*, International Journal of Numerical Methods in Fluids, 42 (2003), pp. 23–55.

[90] M. Ricchiuto, R. Abgrall, and H. Deconinck, *Application of conservative residual distribution schemes to the solution of the shallow water equations on unstructured meshes*, Journal of Computational Physics, 222 (2007), pp. 287–331.

[91] P. Roe, *Approximate* Riemann *solvers, parameter vectors, and difference schemes*, Journal of Computational Physics, 43 (1981), pp. 357–372.

[92] G. Rosatti, L. Bonaventura, A. Deponti, and G. Garegnani, *An accurate and efficient semi-implicit method for section-averaged free-surface flow modelling*, International Journal for Numerical Methods in Fluids, 65 (2011), pp. 448–473.

[93] G. Rosatti and L. Fraccarollo, *A well-balanced approach for flows over mobile-bed with high sediment-transport*, Journal of Computational Physics, 220 (2006), pp. 312–338.

[94] G. Rostand, *Sur une méthode de Volumes Finis en maillage non structuré pour le calcul d'écoulements visqueux compressibles*, PhD thesis, Université de Nice, 1989.

[95] V. Rusanov, *The calculation of the interaction of non-stationary shock waves with barriers*, J. Comput. Math. Phys. USSR, 1 (1961), pp. 267–279.

[96] S. Sahmim, F. Benkhaldoun, and F. Alcrudo, *A sign matrix based scheme for non-homogeneous PDE's with an analysis of the convergence stagnation phenomenon*, Journal of Computational Physics, 226 (2007), pp. 1753–1783.

[97] M. Salvetti, E. Sinibaldi, and F. Beux, *Towards the simulation of cavitating flows in inducers through a homogeneous barotropic flow model*, in Fluid Dynamics of Cavitation and Cavitating Turbopumps, no. 496 in CISM Courses and Lectures, SpringerWienNewYork, 2007, pp. 317–351.

[98] M. V. Salvetti, F. Beux, M. Bilanceri, and E. Sinibaldi, *A Numerical Method for Barotropic Flow Simulation with Applications*

to Cavitation. Presented by invitation at the Third Workshop "Micro-Macro Modelling and Simulation of Liquid-Vapour Flows", Strasbourg, France, 22–24 Jannuary 2008.

[99] R. Saurel, J. P. Cocchi, and P. Butler, *Numerical study of cavitation in the wake of a hypervelocity underwater projectile*, Journal of Propulsion and Power, 15 (1999), pp. 513–522.

[100] R. Saurel, F. Petitpas, and R. Abgrall, *Modelling phase transition in metastable liquids: Application to cavitating and flashing flows*, Journal of Fluid Mechanics, 607 (2008), pp. 313–350.

[101] I. Senocak and W. Shyy, *A pressure-based method for turbolent cavitating flow computations*, Journal of Computational Physics, 176 (2002), pp. 363–383.

[102] E. Sinibaldi, *Implicit preconditioned numerical schemes for the simulation of three-dimensional barotropic flows*, PhD thesis, Scuola Normale Superiore di Pisa, 2006.

[103] E. Sinibaldi, F. Beux, M. Bilanceri, and M. V. Salvetti, *A Second-Order Linearised Implicit Formulation for Hyperbolic Conservation Laws, with Application to Barotropic Flows*, ADIA 2008-4, University of Pisa, June 2008.

[104] E. Sinibaldi, F. Beux, and M. V. Salvetti, *A numerical method for* 3D *barotropic flows in turbomachinery*, Flow, Turbulence and Combustion, 76 (2006), pp. 371–381.

[105] C. Song and J. He, *Numerical simulation of cavitating flows by single-phase flow approach*, in Proc. Third International Symposium on Cavitation, Grenoble (France), April 1998.

[106] K. W. Thompson, *Time dependent boundary conditions for hyperbolic systems ii*, Journal of Computational Physics, 89 (1990), pp. 439–461.

[107] E. F. Toro, *Riemann solvers and numerical methods for fluid dynamics*, Springer, 1997.

[108] E. F. Toro, *Shock-Capturing Methods for Free-Surface Shallow Flows*, Wiley, 2001.

[109] E. F. Toro and P. Garcia-Navarro, *Godunov-type methods for free-surface shallow flows: A review*, Journal of Hydraulic Research, 45 (2007), pp. 736–751.

[110] L. Torre, G. Pace, P. Miloro, A. Pasini, A. Cervone, and L. d'Agostino, *Flow instabilities on a three bladed axial inducer at variable tip clearance*, in 13th International Symposium on Transport Phenomena and Dynamics of Rotating Machinery, Honolulu, Hawaii, USA., 2010.

[111] L. Torre, A. Pasini, A. Cervone, and L. d'Agostino, *Experimental performance of a tapered axial inducer: Comparison with analytical predictions*, in 45th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, Denver, Colorado, USA, 2–5 August 2009.

[112] L. Torre, A. Pasini, A. Cervone, L. Pecorari, A. Milani, and L. d'Agostino, *Rotordynamic forces on three bladed inducer*, in Space Propulsion Conference, San Sebastian, Spain, 3–6 May 2010.

[113] G. van Albada, B. van Leer, and W. Roberts, Jr., *A comparative study of computational methods in cosmic gas dynamics*, Astronomy and Astrophysics, 108 (1982), pp. 76–84.

[114] D. van der Heul, C. Vuik, and P. Wesseling, *Stability analysis of segregated solution methods for compressible flows*, Applied Numerical Mathematics, 38 (2001), pp. 257–274.

[115] B. van Leer, *Towards the ultimate conservative difference scheme* V*: a second-order sequel to* G*odunov's method*, Journal of Computational Physics, 32 (1979), pp. 101–136.

[116] M. E. Vásquez-Cendón, *Improved treatment of source terms in upwind schemes for the shallow water equations in channels with irregular geometry*, Journal of Computational Physics, 148 (1999), pp. 497–526.

[117] D. C. Wilcox, *Turbulence modeling for CFD*, D C W Industrie, 1998.

[118] N. Yamanishi, S. Fukao, X. Qiao, C. Kato, and Y. Tsujimoto, *Les simulation of backflow vortex structure at the inlet of an inducer*, Journal of Fluids Engineering, Transactions of the ASME, 129 (2007), pp. 587–594.

[119] H. C. Yee, *Construction of explicit and implicit symmetric* TVD *schemes and their applications*, Journal of Computational Physics, 68 (1987), pp. 151–179.

[120] J. G. Zhou, D. M. Causon, C. G. Mingham, and D. M. Ingram, *The surface gradient method for the treatment of source terms*

in the Shallow-Water equations, Journal of Computational Physics, 168 (2001), pp. 1–25.

[121] O. C. Zienkiewicz and R. L. Taylor, *Finite Element Method, Volume 3 - Fluid Dynamics, 5th edition*, Elsevier, 2000.