Scuola Normale Superiore di Pisa

Classe di Scienze

Ph.D. Thesis

Implicit preconditioned numerical schemes for the simulation of three-dimensional barotropic flows

Candidate	Supervisors
E. Sinibaldi	Dott. F. Beux Prof. M. V. Salvetti

Pisa, Italy - 2006

*...to **Roberta**,*

who patiently supported me in this work.
With never-ending Love.

Abstract

A numerical method for simulating three-dimensional, generic barotropic flows on unstructured grids is developed. Space and time discretizations are separately considered. A finite volume compressible approach, based on a suitable Roe numerical flux function, is proposed and the accuracy of the resulting semi-discrete formulation for nearly-incompressible flows is ensured by *ad hoc* preconditioning. Moreover, a linearized implicit time-advancing technique is proposed, only relying on the algebraic properties of the Roe flux function and therefore applicable to a variety of problems. This implicit strategy is extended so as to incorporate the aforementioned preconditioning. The considered numerical ingredients are firstly defined in a one-dimensional context; after validation, they are extended to three-dimensional non-rotating as well as rotating frames. Finally, the resulting numerical method is validated by considering complex industrial flows, namely the water flow around a hydrofoil (for which specific experimental data are available) and the water flow around a rotating turbo-pump inducer.

By starting from a particular industrial problem (namely the numerical simulation of propellant flows around an axial inducer belonging to the feed turbo-pump system of a liquid propellant rocket engine), a numerical method which can be applied to generic barotropic flows is defined. Along the way, a constructive procedure for solving the 1D Riemann problem associated with a generic convex barotropic state law is proposed. This solution, also exploited for defining a Godunov numerical flux suitable for incorporation into finite volume schemes, is systematically used in order to define exact benchmarks for the quantitative validation of the proposed one-dimensional numerical methods.

Contents

Introduction

Motivation of the study

The present thesis documents the starting efforts made for constructing a numerical frame aimed at simulating propellant flows occurring in the feed turbo-pumps of modern liquid propellant rocket engines. More precisely, the numerical simulation of the three-dimensional (hereafter 3D as well) unsteady propellant flows around the axial inducers which are part of the aforementioned turbo-machines (see sec. 1.1) is the long-term goal of the present research.

The suction performance and, consequently, the global performance of the rocket engine significantly depend on the flow pattern within the turbo-pumps [9]; hence, it is of interest to understand and control their hydrodynamics, in order to conceive a correct design. Besides the experimental investigations, undoubtedly expensive as well as quite dangerous when dealing with many liquid propellants [97], the numerical simulation may provide a deep insight into the flows under consideration at a generally affordable cost, thus motivating the development of suitable numerical tools. Nevertheless, several factors make it challenging to accurately and efficiently simulate the considered liquid flows. Firstly, the severe weight and size constraints to which the considered engines are subjected impose high rotor speeds which, in turn, systematically entail the occurrence of cavitation phenomena (see sec. 1.2). The simultaneous presence of the pure liquid (which is almost incompressible) and the liquid-vapour mixture (which behaves like a highly compressible fluid) dramatically changes the local flow properties at the unknown interface, thus rendering common numerical methods hardly applicable. Furthermore, the very complex geometry of the axial inducers adds to the difficulty of the problem. Indeed, it imposes to adopt huge computational grids which, if unstructured, generally do no permit to straightforwardly extend some well-known numerical techniques and, in any case, require efficient, hopefully parallel, algorithms to be defined.

In view of the aforementioned considerations, the subject of the present study seems to possess aspects of interest from both an academic and an industrial point of view. As far as the latter point is concerned, a significant part of the documented research activity has been funded by the Italian Space Agency (ASI) under a 16 month industrial program, namely the FAST2 (Future Advanced Space Transportation Technologies) project [1].

[1] The author joined the project as a collaborator of the Aerospace Engineering Dept. of the University of Pisa which, in turn, was involved in the program as subcontractor of

Overview and choice of the numerical methodology

In consideration of the fact that, under typical operational conditions, cavitation phenomena can take place within the aforementioned flows, it is necessary to select a suitable cavitation model as a first step in the development of the numerical tool under consideration. Indeed, the cavitation model specifically adopted directly affects the mathematical formulation of the problem (through the closure of the governing equations) and therefore its numerical discretization. A concise overview of the current cavitation models is reported in sec. 1.3. For the present purposes it suffices to mention that almost all the formulations used for computations of industrial interest are based on *equivalent fluid models*, namely:

(BH) barotropic *homogeneous flow models* (see sec. 1.3), according to which the pressure and the density are linked to each other by an invertible relation within both the pure liquid and the cavitating mixture. Examples of this approach may be found, for instance, in [16], [21], [22], [23], [24], [25], [45], [54], [76], [78], [82], [96], [104] and [105];

(DS) "dual species" models (a particular class of *non-homogeneous flow models*, see sec. 1.3) in which a convection equation for the volume or mass void fraction is introduced, explicitly accounting for the mass transfer at phase transition. Examples of this approach may be found, for instance, in [3], [46], [47], [58], [70], [86], [87], [89] and [116].

With the exception of [46] and [47], the aforementioned works deal with structured grids (possibly involving generalized curvilinear coordinates as for [58] and [89]) and adopt a finite volume spatial discretization (see secs. 3.1.1 and 5.1). Some of them, belonging either to the (BH) or to the (DS) class, implement a pressure-based approach typically extending well-known pressure-correction algorithms originally conceived for incompressible flows (e.g. the SIMPLE algorithm [74] or some related variants, like PISO) in order to suitably cope with the compressible, cavitating flow sub-domains. Other works, instead, adopt a density-based approach which modifies (by preconditioning techniques, see e.g. sec. 3.4) common algorithms originally conceived for compressible flows so as to account for the very weak liquid compressibility. In most cases, the convective component of the numerical flux function (see the relevant paragraph in sec. 3.1.1) is discretized by upwinding: typically a TVD scheme (see e.g. [98]) or, for the (BH) class, an artificial dissipation approach (see e.g. [53]) is exploited. Conversely, the diffusive component of

the Italian Aerospace Research Center (CIRA).

the numerical flux (if any) is discretized by central differencing for almost all the considered works. As far as the time-advancing is concerned, both explicit and implicit techniques (see sec. 3.1.2) are considered but only the latter allow for the construction of efficient schemes. A dual time-stepping is adopted in certain works (e.g. [22], [23], [24] and [58]), in which a suitable preconditioning technique as well as an under-relaxation of the density are introduced in order to speed-up the convergence of the internal iterations. It is worth noticing that only Coutier-Delgosha and coworkers currently manage to compute cavitating flows in inducers with a certain degree of accuracy [2] while other researchers only succeed in dealing with less ambitious (even if still challenging) applications like nozzle and hydrofoil flows.

On the basis of the literature reviewed at the beginning of the research project here documented, both the pressure-based and the density-based approach seemed to possess points of strength as well as weaknesses, so that there was no clear advantage in *a priori* preferring one to the other. Because of this point, a resource-driven choice was made. In particular, a density-based approach was selected, suitable for incorporation within a numerical framework for the simulation of compressible flows, which was available to the research group [3]. The numerical tool under consideration is the AERO code (see e.g. [32], [35] and [71]), derived from a collaboration between the French national institute for research in computer science and control (INRIA, “Institut National de Recherche en Informatique et en Automatique”) and the University of Boulder (Colorado, USA).

The AERO code discretizes both the laminar and the turbulent Navier-Stokes equations (written in conservation form) for ideal gases; in the latter case either a Reynolds-averaged formulation (closed by several turbulence models) or a LES (Large-Eddy Simulation) formulation can be adopted. Moreover, it permits to simulate one-way fluid-structure interactions. Space and time discretizations are kept separate (“method of lines”). The space discretization is carried out by a mixed finite volume-finite element formulation (see e.g. [85]) based on tetrahedral unstructured grids. The first-order approximation of the convective fluxes is obtained by means of the Roe flux function [84]; higher order extensions are based on a MUSCL-like reconstruction [108] conceived for unstructured grids (see e.g. [28]). As for the diffusive fluxes, P1 finite elements are exploited. A Roe-Turkel preconditioning technique is adopted for low Mach number (i.e. nearly incompressible) flows, which can be exploited for unsteady simulations as well (see e.g. [42]

[2] The most advanced results, reported in e.g. [22] and [24], were not published at the time the research project here documented started.

[3] Namely, the CFD group of the Aerospace Engineering Dept. of the University of Pisa.

and [112]). As far as the time discretization is concerned, either explicit or implicit time-advancing strategies are available. In the former case, a low-storage 4−th order Runge-Kutta scheme is adopted while, in the latter one, a linearized implicit scheme designed for the ideal gas state law (see e.g. [36]) is implemented and the extension to the second order in time is achieved by a "defect-correction" strategy [67]. An efficient implementation of the whole numerical framework is achieved by means of a message-passing (MPI-1 standard) parallelization strategy.

In order to construct a numerical solver for cavitating flows in complex geometries starting from AERO, only the inviscid portion of the laminar governing equations (i.e. the Euler equations) has been considered (see sec. 2.2 for the rationale). A barotropic homogeneous flow cavitation model (able to take into account thermal cavitation effects and, possibly, the concentration of the active cavitation nuclei) as well as a barotropic state law for the pure liquid have been chosen, thus providing a unified barotropic state law for the working fluid (see sec. 4.1).

Note 1 *For the sake of generality, only a very few, physically-based, constraints have been imposed on the considered barotropic state law (see sec. 1.5). Hence, all the proposed numerical ingredients (with the only exception of the Godunov numerical flux discussed in sec. 3.2, which additionally requires the state law to be convex) can be applied to generic barotropic laws.*

The adopted barotropic formulation permits to decouple the energy balance from the rest of the governing equations and therefore a "reduced" system only comprising the mass and the momentum balance has been considered. The chosen state law directly affects the space discretization of the AERO solver through the definition of the Roe numerical flux. Moreover, it also affects the time discretization, since the linearized implicit time-advancing strategy originally appearing in AERO directly exploits the first-order homogeneity of the analytical flux, which holds for the ideal gas case but not for the barotropic one (see sec. 3.5). In addition, the adopted state law indirectly appears within the Roe-Turkel preconditioning strategy as well (see sec. 3.4.3). In view of the above considerations, a Roe numerical flux function suitable for generic barotropic flows has been proposed as a basic numerical ingredient. The spatial accuracy of the numerical solution obtained by applying the resulting finite volume scheme to nearly-incompressible flows has then been addressed, by performing the asymptotic analysis originally proposed in [42]. It has been shown that the preconditioning technique proposed in [42] for the ideal gas case can be extended to the barotropic one. Carried out one-dimensional (hereafter 1D as well) numerical experiments have confirmed the predicted accuracy problems occurring at low Mach numbers as well as the effectiveness of the considered preconditioning strategy [91]. However, it has been also observed that the preconditioning at hand restricts the stability region of common explicit time-advancing schemes (see sec. 3.4.4), thus decreasing their efficiency [91]. In order to counteract this problem, a linearized implicit time-advancing strategy has been proposed, which only relies on the algebraic properties of the Roe flux function (and therefore it is well suited to a variety of problems) and which can be applied to the preconditioned formulation as well. As shown by the aforementioned 1D numerical experiments, the implicit scheme allows for a very efficient time-advancing to be performed when non-cavitating flows are considered; however, when cavitation occurs, the efficiency of the implicit scheme is noticeably reduced. A 3D numerical method has been subsequently derived from the considered 1D techniques by exploiting the tensorial character of the governing equations. In view of the time-schedule imposed by the supporting industrial program, the proposed 3D numerical method has been directly implemented within the AERO mainframe and the resulting numerical tool has been validated by considering the water flow around a NACA0015 hydrofoil. The proposed 3D numerical method has finally been extended so as to deal with rotating frames and the corresponding implementation has been validated by considering the water flow around an axial turbo-pump inducer. The efficiency issues originally noticed in a 1D context have been systematically observed in the 3D case as well. A more systematic investigation of the aforementioned 1D numerical ingredients has been consequently started. In this context, the exact solution of a 1D Riemann problem involving a generic convex barotropic state law has been constructed (based on classical elements of the theory of hyperbolic partial differential equations), to be exploited for defining exact benchmarks for the analysis and the validation of the considered 1D numerical schemes, also when considering cavitating test-cases. A Godunov numerical flux function based on the aforementioned exact solution has been defined as well.

Thesis outline

- In sec. 1 the considered industrial problem is presented. More in detail, once chosen a suitable (barotropic homogeneous flow) cavitation model, a concise statement of the industrial problem under consideration is reported, highlighting the generality of the chosen barotropic state law;

- in sec. 2 a hierarchy of governing equation is presented, each of which is considered at a subsequent stage of the discussion. Once underlined the hyperbolic character of the systems at hand, the attention is focused on the 1D Riemann problem (hereafter RP as well); a constructive procedure for determining its exact solution when considering a generic convex barotropic state law is proposed;

- in sec. 3 all the proposed 1D numerical ingredients are presented. After introducing some basic material on the numerical discretization, a Godunov numerical flux for generic convex barotropic state laws as well as a Roe numerical flux for generic barotropic state laws are proposed. The behaviour of the considered semi-discrete formulation (based on a finite volume approach involving the Roe numerical flux) dealing with nearly-incompressible flows is addressed and a suitable preconditioning strategy is presented following [42]. Moreover, a linearized implicit time-advancing strategy is proposed, only relying on the algebraic properties of the Roe numerical flux function and therefore applicable to a variety of problems. Finally, the linearized implicit strategy is extended so as to deal with the preconditioned numerical flux function. All the proposed ingredients are validated against exact (namely, solutions to 1D RPs) or nearly-exact benchmarks. The issue of the efficiency of the considered scheme when dealing with discontinuous flow fields (mimicking cavitating conditions) is put forward;

- in sec. 4 the barotropic state law specifically adopted for the subsequent simulation of the industrial test-cases is introduced. Moreover, an illustrative 1D numerical experiment involving cavitation phenomena is considered, in order to highlight some difficulties that are systematically encountered when dealing with the chosen cavitation model (or similar ones);

- in sec. 5 the proposed (preconditioned) Roe numerical flux is extended to the 3D case. Moreover, the discretization of the domain as well as the numerical treatment of the convective fluxes are discussed. The considered 3D numerical method is then extended to rotating frames. Finally, the linearized implicit time-advancing originally proposed in a 1D context is extended to the 3D rotating case;

- in sec. 6 the applications of the proposed 3D numerical method, namely the simulation of water flows around a hydrofoil and an axial turbopump inducer, are presented. In the former case, a quantitative appraisal is given for both non-cavitating and cavitating conditions, based on available experimental data. In the latter one, a qualitative appraisal is given for a non-cavitating flow;

- in sec. 7 the main achievements of the present study, as well as its open questions, are summarized, together with some research perspectives.

Auxiliary material (e.g. some mathematical derivations and proofs) is finally reported in the appendices A and B, for ease of presentation.

Note 2 *No details are given in the present document concerning the implementation of the proposed numerical schemes within the aforementioned parallel numerical frame, because they are behind the scope of the discussion.*

Related scientific documentation

Some numerical ingredients discussed in secs. 3.3, 3.4, 3.5 and 5, as well as the applications reported in sec. 6 have been documented through:

- the international publications [6] and [93];

- the INRIA research report [91];

- the proceedings of the international conferences [92] and [94];

- the proceedings of the national conference [90].

Other issues (e.g. those presented in secs. 2.5 and 3.2) originally appear in the present document.

1 Industrial problem

The present study is aimed at developing a numerical method suitable for the numerical simulation of propellant flows occurring in the feed turbo-pumps of modern liquid propellant rocket engines. More in detail, the numerical simulation of 3D unsteady liquid flows through axial inducers is the long-term goal of the present research.

In secs. 1.1 to 1.4 several issues are concisely presented, concerning the physical modelling of the industrial problem of interest. Among the wide variety of technical and conceptual aspects potentially arising during the discussion, only those required by the subsequent treatment are introduced, for ease of presentation. In sec. 1.5 the general form of the state law adopted for the working fluid is defined. Finally, in sec. 1.6, a statement of the industrial problem under consideration is reported, based on the material presented through the previous sections.

1.1 Axial inducers

Axial flow inducers are hydraulic devices suitably designed to improve the performance of the (usually centrifugal) pumps they are attached to, by increasing the inlet pressure to the pump to a level at which it can operate without excessive loss of performance due to cavitation (see sec. 1.2). Typically they consist of an axial flow stage, like that one shown in Figs. 1 and 2, placed just upstream of the inlet to the main impeller. They are designed to operate at small incidence angles and to have thin blades so that the perturbation to the flow is small in order to minimize the production of cavitation and its deleterious effect upon the flow: the objective is to raise the pressure very gradually to the desired level [4] [9].

Axial inducers can be “shrouded” or “unshrouded”. In the former case, there is no gap between the tip of the blades and the external case while in the latter one such a gap is present. An example of an unshrouded axial inducer is shown in Figs. 1 and 2. The shrouded geometry makes the inducer more robust with respect to cavitation instabilities [5]. In addition, the absence of the gap prevents the creation of very complex secondary flows

[4] The reason why the design incidence angle is not zero is that, under these conditions, cavitation could form on either the pressure or the suction surfaces of the blades or it could oscillate between the two. It is preferable to use a few degrees of incidence to eliminate this uncertainty and ensure suction surface cavitation [9].

[5] For an exhaustive treatment of cavitation instabilities in inducers the interested reader can refer to the work of Y. Tsujimoto, not reported in the bibliography because beyond the scope of the present work.

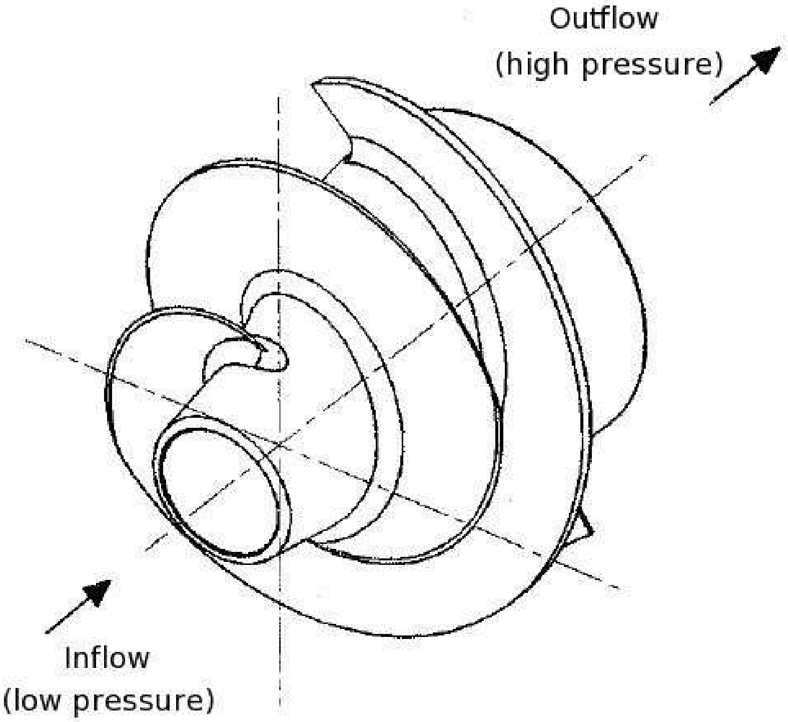

Figure 1: Schematic of a two-bladed unshrouded helical inducer. The far-field inflow is aligned with the rotation axis; the (swirled) outflow is directed towards the main pump impeller (not shown).

(synthetically referred to as "tip leakage") which affect the inducer fluid dynamics [59], generally weakening the inducer pumping performance. In spite of their attractive features, the manufacturing process required by shrouded inducers is very complex (and expensive); hence, most inducers are nowadays unshrouded [9].

As for the vast majority of modern turbo-machines typical of space propulsion applications, also for axial inducers the very strict weight and size constraints impose, for a given power, a high rotational speed. This, in turn, entails high tip speeds and paves the way for cavitation to take place. A certain understanding of cavitation phenomena as well as suitable modelling techniques are therefore needed in order to describe the liquid flows within this kind of machines, even when they are not expressly designed to operate in cavitating conditions.

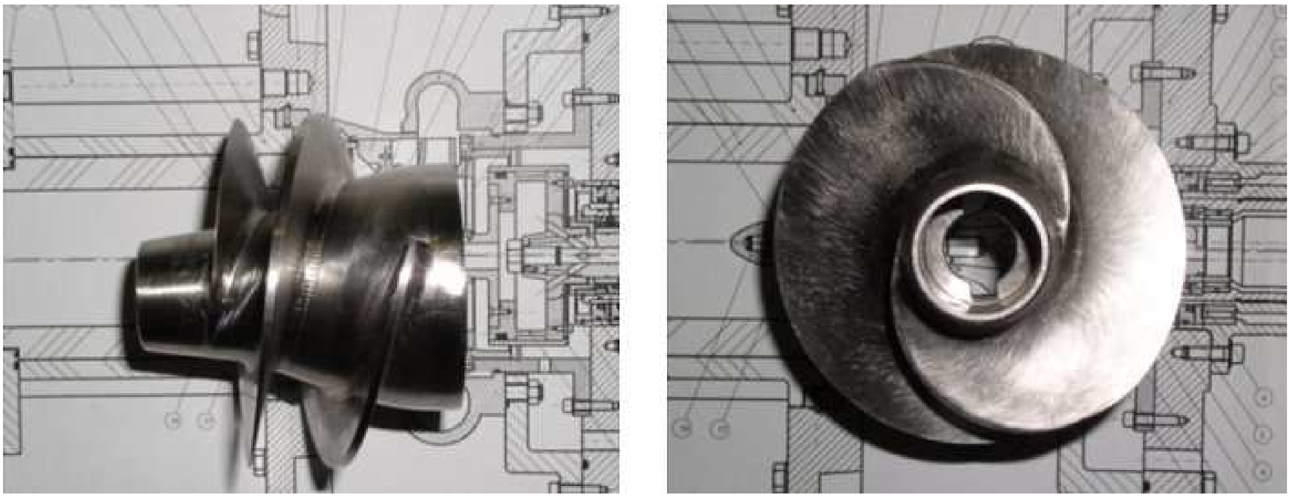

Figure 2: Side and front views (left to right) of the two-bladed unshrouded helical inducer sketched in Fig. 1.

1.2 Cavitation

Cavitation is a complex fluid dynamic phenomenon, involving the extremely rapid growth and subsequent collapse of liquid cavities originating from weak spots (cavitation nuclei) when the pressure falls below the saturation value for a sufficiently long time for the nuclei to become active [10]. The relative abundance and susceptibility of nuclei in the bulk of the liquid and on its low pressure boundaries determines the dispersed or attached form of cavitation. A major difficulty in the analysis of cavitating flows is the presence of free surfaces, whose shape, location and evolution are not known *a priori* and must in principle be obtained as part of the solution of the flow field. Cavitating flows are therefore intrinsically unsteady on a length scale comparable to the cavity size and often also on the global (macroscopic) scale, especially in internal reverberating flows. The dynamic nature of cavitation, with the occurrence of appreciable inertial effects in the liquid and rate-controlled evaporation/condensation at the interface, adds to the complexity of the phenomenon, since thermodynamic equilibrium is not satisfied and the usual barotropic behaviour of common fluids [6] should in principle be replaced by a differential relation between the local density and pressure. Therefore cavitation poses formidable obstacles in terms of both physical and numerical modelling.

[6] I.e. the possibility of expressing the thermal state law of a certain fluid by means of a one-to-one correspondence between density and pressure.

1.3 Cavitation modelling

Current models for the description of cavitating flows can be classified as follows:

- *free streamline models*, where the cavity region is separated by a sharp interface from the region occupied by the pure liquid (e.g. [117]);
- *equivalent fluid models*, where volume, time or ensemble averaging is used to account for the presence of two phases (e.g. [50]);
- *direct simulation models*, where the coupled Navier-Stokes equations of the two phases are solved simultaneously.

In turn, equivalent fluid models can be divided in:

- *homogeneous flow models*, where the macroscopic features of cavitation are represented in terms of a single-phase fluid whose properties are obtained by introducing suitable simplifying assumptions (e.g. [30]);
- *non-homogeneous flow models*, based on the separate characterization of the two phases with the relevant interaction terms (e.g. [2] and [70]);
- *non-homogeneous flow models with cavity dynamics*, similar to the previous ones except for the inclusion of the evolutionary effects connected to the transfer of mass, momentum and energy between the two phases (e.g. [15], [57] and [77]).

None of these models is free from inherent limitations. Free streamline models, where a well-defined interface separates the pure liquid from the cavity region occupied by the non-condensed phase, introduce prohibitive complications in 3D configurations and are not realistic in the thermal cavitation conditions typical of cryogenic propellants of rocket motors, where travelling bubble cavitation prevails [9]. On the other hand, direct simulation methods are extremely demanding in terms of computational resources and their superior accuracy is eluded in practice by the uncertain knowledge of the initial state of the system, especially the nature, concentration and susceptibility of cavitation nuclei.

This brief overview indicates that the successful choice of a model for simulating cavitation in technical applications must be based on careful consideration of the final objectives and implementation constraints, in order to exploit all opportunities to simplify the formulation of the problem by including only the essential physical phenomena.

1.4 Choice of the cavitation model

The typical requirements of space propulsion applications for the analysis of propellant feed turbo-pumps put especial prize on the suction and dynamic performance of the machine rather than on its resistance to erosion and other long-term effects of cavitation, which are typically a major concern in other applications. Fortunately these properties are essentially related to the large-scale characteristics of the flow field. The local behaviour of the cavities, on the other hand, mainly controls aspects such as erosion, high frequency vibrations and noise, which are less important in liquid propellant rocket engines in view of their limited expected life time.

These considerations indicate the opportunity of choosing an equivalent fluid model where the fine details of the cavity growth and collapse are neglected, and the cavitating flow is visualized in terms of a single fluid, whose properties are obtained by introducing suitable simplifying assumptions.

In cavitating liquids with relatively high vapour pressures (like most cryogenic propellants) the transfer of heat at the interface represents the most important interaction between the two phases because of its influence on the cavity pressure ("thermodynamic effect") and indirectly on the other flow variables. Conversely, in these flows mass and momentum exchanges usually play a comparatively minor role. In view of these considerations, pressure and velocity differences between the two phases can safely be neglected and the choice of an homogeneous flow model explicitly accounting, at least approximately, for thermal cavitation effects seems to be an efficient approach to the simulation of cavitating flows for performance predictions in space propulsion applications [26].

Among the cavitation flow models meeting the above requirements, that one recently proposed by d'Agostino and coworkers [27] deserves special mention. In this model the liquid/vapour mixture behaves isentropically, so that it is possible to use the mixture energy balance in order to evaluate the mass interaction term accounting for evaporation/condensation phenomena between the two phases and ultimately derive a monotonic constitutive relation between the density ρ and the pressure p of the cavitating mixture (i.e. a barotropic thermal state law). In addition, the model naturally accounts for the effects of thermal cavitation by exploiting the specific properties of the thermally-controlled dynamics of cavitating spherical bubbles. Finally, if required, the effects of the active nuclei concentration in the liquid phase can be readily incorporated in the model. Only the essential features of the model are reported below, further details being given in sec. 4.1.2.

In order to keep some degree of generality, the state law provided by the chosen cavitation model is expressed as a generic curve of the form:

$$p = p_{cav}\left(\rho\right) \quad , \quad \rho \in \left[\rho_{min}, \rho_{Lsat}\right] \tag{1}$$

The lower bound of the domain, ρ_{min}, is constrained by some simplifying assumptions on which the cavitation model is based (see sec. 4.1.2) while the upper bound, ρ_{Lsat}, is the liquid saturation density at a given temperature [7] and represents the transition to the pure liquid regime. The physical foundations of the model ensure the strict positivity of ρ and p, which can therefore be interpreted as the "usual" density and pressure of an equivalent fluid in the sense of classical fluid mechanics. Furthermore, the strict positivity of the derivative $\mathrm{d}p/\mathrm{d}\rho$ is also guaranteed, thus satisfying a classical thermodynamic stability requirement [12]

By virtue of the isentropic approximation adopted while deriving the cavitation model, it is possible to interpret the following entity:

$$a := +\sqrt{\frac{\mathrm{d}p}{\mathrm{d}\rho}} \tag{2}$$

as a mixture "sound speed", in analogy with classical fluid mechanics. More in general, (2) is adopted for defining the sound speed of a generic barotropic fluid through the present document.

A common practice is to juxtapose the barotropic state law provided by a homogeneous cavitation model with another one describing the pure liquid [8], in order to obtain a unified model for the working fluid [9], which may or may not cavitate depending on the flow conditions. This approach seems to represent a good compromise between computational cost and accuracy, and it is adopted for the numerical simulation of the industrial test-cases reported in secs. 4 and 6.

Despite their formal simplicity, however, considerable difficulties are still encountered in implementing physically-based unified barotropic models in a workable simulation tool for the prediction of cavitating flows. Indeed, the local presence of two phases dramatically reduces the sound speed of the mixture [10]: both nearly-incompressible zones (pure liquid) and regions where

[7] The liquid is supposed to be at constant temperature; hence, ρ_{Lsat} is properly defined.

[8] Usually, the starting point is a given model for the pure liquid to be coupled with a consistent homogeneous flow cavitation model for the mixture. This perspective has been somehow twisted in the present discussion for ease of presentation.

[9] The chosen cavitation model, in particular, allows a smooth junction (i.e. continuity of p and a, see sec. 4.1.2) to be defined at the transition point ρ_{Lsat}.

[10] For instance, in a water-vapour mixture at 20°C, $a \approx \mathrm{O}\left(10^3\right)$ m/s towards the pure liquid limit, it abruptly decreases to $\mathrm{O}\left(10^{-1} \div 10^0\right)$ m/s in the mixture before rising back to $\mathrm{O}\left(10^2\right)$ m/s towards the pure vapour limit.

the flow may easily become highly supersonic (liquid-vapour mixtures) are present in the flow and need to be solved simultaneously. The sound speed variation at cavitation inception, in particular, is exceedingly abrupt (see sec. 4.1.2); it originates discontinuities within the flow field that, together with the strong shocks occurring in the recondensation at the cavity closure, add to the complexity of the phenomenon. It is therefore evident that specifically designed numerical schemes must be introduced in order to handle this situation.

1.5 Definition of the state law

Let $D_\rho := [\rho_{min}, \rho_{sup})$ denote a density domain. In view of the considerations reported in sec. 1.4, a generic state law is assumed for the working fluid, of the form:

$$p = p(\rho) \quad , \quad \rho \in D_\rho \tag{3}$$

with:

$$\frac{\mathrm{d}p}{\mathrm{d}\rho}(\rho) > 0 \quad , \quad \rho \in D_\rho \tag{4}$$

No loss of generality is introduced by choosing ρ as the independent variable (p would be equivalently acceptable). According to (4) the pressure is allowed to vary within the domain $D_p := [p_{min}, p_{sup})$, with $p_{min} = p(\rho_{min})$ and $p_{sup} = p(\rho_{sup})$. Clearly, for the state law to be meaningful from the viewpoint of classical fluid mechanics, the following conditions must be verified as well:

$$\rho_{min} > 0 \quad , \quad p_{min} > 0 \tag{5}$$

1.6 Statement of the industrial problem

In consideration of the material introduced up to this point, it is possible to state that the present study is aimed at *“developing a numerical method for simulating the flow of a fluid showing the constitutive behaviour defined in sec. 1.5 around 3D geometries of the type of those shown in sec. 1.1”*.

Clearly, it is of primary interest to simulate non-cavitating flows at a first stage, and then to be able to cope with the additional difficulties introduced by cavitation phenomena. In this spirit, the above statement keeps a certain degree of generality: the possibility of simulating a pure liquid or a cavitating mixture (by a homogeneous cavitation model) is completely delegated to the specific state law and, of course, to the actual flow conditions.

2 Mathematical formulation

A natural framework for the mathematical formulation of the industrial problem introduced in sec. 1 is that one of classical fluid mechanics. Within this scope, several sets of governing equations are introduced in sec. 2.2, to be closed by the barotropic state law discussed in sec. 1.5. Each of them is representative of a certain type of approximation and it is exploited at a specific point during the subsequent development of numerical methods. Once recalled some basic issues related to hyperbolic systems and integral solutions in sec. 2.3, some attention is devoted to the Riemann problem in sec. 2.4 because of the key role its solution plays in the set up of modern numerical methods for fluid dynamics. Finally, in sec. 2.5 the ingredients presented in sec. 2.4 are exploited to solve the Riemann problem associated with a generic convex barotropic state law.
General conventions concerning the notation adopted throughout the present document, as well as some relevant definitions, are introduced in sec. 2.1.

2.1 Notation and preliminary definitions

(C1) A bold lowercase symbol like $\mathbf{v}$ denotes a matrix in $\mathbb{R}^{n\times 1}$, hereafter referred to as a vector in $\mathbb{R}^n$. A bold uppercase symbol like $\mathbf{M}$ denotes a matrix in $\mathbb{R}^{n\times n}$. For the present purposes $n \in \{2,3,4\}$. In particular, $\mathbf{0}$ denotes the null vector while $\mathbf{I}$ and $\mathbf{O}$ respectively denote the identity matrix and the null matrix.

(C2) Let a, b, c,... o and p be real numbers. Moreover, let $\mathbf{v}_1 \in \mathbb{R}^3$, $\mathbf{v}_2 \in \mathbb{R}^3$, $\mathbf{M}_1 \in \mathbb{R}^{3\times 3}$, $\mathbf{v}_3 \in \mathbb{R}^4$ and $\mathbf{M}_2 \in \mathbb{R}^{4\times 4}$ admit the following representation:

$$\mathbf{v}_1 = \begin{pmatrix} b \\ c \\ d \end{pmatrix} \quad , \quad \mathbf{v}_2 = \begin{pmatrix} e \\ i \\ m \end{pmatrix} \quad , \quad \mathbf{M}_1 = \begin{pmatrix} f & j & n \\ g & k & o \\ h & l & p \end{pmatrix}$$

$$\mathbf{v}_3 = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \quad , \quad \mathbf{M}_2 = \begin{pmatrix} a & e & i & m \\ b & f & j & n \\ c & g & k & o \\ d & h & l & p \end{pmatrix}$$

Then, the following compact notation is understood:

$$\mathbf{v}_3 = \begin{pmatrix} a \\ \mathbf{v}_1 \end{pmatrix} \quad , \quad \mathbf{M}_2 = \begin{pmatrix} a & \mathbf{v}_2^T \\ \mathbf{v}_1 & \mathbf{M}_1 \end{pmatrix}$$

where $\mathbf{v}_2^T$ denotes the transpose of $\mathbf{v}_2$.

(C3) The "$\cdot$" (centred dot) symbol indicates the common matrix-matrix multiplication.

(C4) Let $\mathbf{M} \in \mathbb{R}^{n \times n}$ be diagonalizable with real eigenvalues λ_h, $h = 1, \ldots, n$. Then:

$$\mathbf{M} = \mathbf{T} \cdot \mathbf{\Lambda} \cdot \mathbf{T}^{-1}$$

where $\mathbf{\Lambda}$ is diagonal:

$$\mathbf{\Lambda} := \mathrm{Diag}\left(\lambda_1, \ldots, \lambda_n\right)$$

and $\mathbf{T}$ is a matrix whose columns are given by the (right) eigenvectors of $\mathbf{M}$. Once defined the application of the usual absolute value $|\cdot|$ to $\mathbf{\Lambda}$ as follows:

$$|\mathbf{\Lambda}| := \mathrm{Diag}\left(|\lambda_1|, \ldots, |\lambda_n|\right)$$

it is possible to introduce the following definitions, extensively used in the sequel:

$$|\mathbf{M}| := \mathbf{T} \cdot |\mathbf{\Lambda}| \cdot \mathbf{T}^{-1} \tag{6}$$

$$\mathbf{M}^{\pm} := \frac{1}{2}\left(\mathbf{M} \pm |\mathbf{M}|\right) \tag{7}$$

(C5) The symbol $\|\mathbf{v}\|$ denotes the L_2 norm of the vector $\mathbf{v}$. Unit vectors (henceforth called versors as well) are marked by a top hat, e.g. $\hat{\mathbf{v}}$.

(C6) Symbol $t \in [0, \infty)$ always denotes time.

(C7) The partial derivative of the scalar f with respect to the scalar v is denoted by $\partial_v f$. The partial derivative of the vector $\mathbf{f} \in \mathbb{R}^n$ with respect to the vector $\mathbf{v} \in \mathbb{R}^n$ is denoted by $\partial_{\mathbf{v}} \mathbf{f}$; it is the usual Jacobian matrix in $\mathbb{R}^{n \times n}$, whose $ij-$th component is given by $\partial_{v_j} f_i$, where f_i and v_j here denote the $i-$th component of $\mathbf{f}$ and the $j-$th component of $\mathbf{v}$, respectively. Consistently, the derivative $\partial_v \mathbf{f}$ is manipulated as a matrix in $\mathbb{R}^{n \times 1}$ (i.e. a vector in $\mathbb{R}^n$) while the derivative $\partial_{\mathbf{v}} f$ is manipulated as a matrix in $\mathbb{R}^{1 \times n}$ (i.e. the transpose of a vector in $\mathbb{R}^n$).

2.2 Governing equations

The Euler equations of classical fluid mechanics, which describe the flow of a compressible and inviscid fluid [88], are chosen as governing equations. The inviscid approximation seems to be justified, at least at a first stage, by the fact that [26]:

- viscous stresses are usually negligible with respect to the huge dynamic actions typical of modern hydraulic turbo-machinery for space propulsion systems;

- in these applications, viscous dissipation plays a minor role in the energy balance, if compared to the contribution due to heat conduction.

More in detail, the Euler equations for a force-free flow are considered, since also the body forces are usually negligible with respect to the dynamic actions under consideration. Furthermore, by virtue of the barotropic state law (3), the energy balance is decoupled from the others (i.e. mass and momentum) [88]; hence, a "reduced" set of equations is considered.

Once introduced the main system of governing equations in both inertial and rotating frames in secs. 2.2.1 and 2.2.2, respectively, several simplified systems are concisely reported in secs. 2.2.3 to 2.2.5. The hierarchical structure of the presented systems is finally discussed in sec. 2.2.6, with the aim of highlighting the degree of approximation of each of them. Neither boundary nor initial conditions are considered at this stage of the discussion.

2.2.1 3D equations

Let $\hat{\mathbf{e}}^{(k)}$, $k = 1, 2, 3$ be the $k-$th versor of a chosen Cartesian orthogonal frame associated with the physical (Euclidean) space. Moreover, let $\mathbf{u} \in \mathbb{R}^3$ indicate the flow velocity, with $k-$th component u_k. Finally, let $\mathcal{V}$ be an arbitrary (regular) space domain having (regular) boundary $\mathcal{S}$ with unit outer normal $\hat{\mathbf{n}}$. Then, the conservation of mass and momentum [88] within $\mathcal{V}$ can be expressed as follows:

$$\partial_t \int_{\mathcal{V}} \mathbf{q} \; \mathrm{dV} + \int_{\mathcal{S}} \left(\sum_{k=1}^{3} \hat{n}_k \, \mathbf{f}^{(k)} \right) \; \mathrm{dS} = \mathbf{0} \tag{8}$$

where $\hat{n}_k$ represents the $k-$th component of $\hat{\mathbf{n}}$, the vectors $\mathbf{q}$ and $\mathbf{f}^{(k)}$ are defined as follows:

$$\mathbf{q} := \begin{pmatrix} \rho \\ \rho\mathbf{u} \end{pmatrix} \tag{9}$$

$$\mathbf{f}^{(k)} := u_k \, \mathbf{q} + p \left(\begin{array}{c} 0 \\ \hat{\mathbf{e}}^{(k)} \end{array} \right) \tag{10}$$

and the pressure p is related to the density ρ by means of the barotropic law (3). Regular solutions of (8) also satisfy its corresponding differential form, namely:

$$\partial_t \, \mathbf{q} + \sum_{k=1}^{3} \partial_{x_k} \mathbf{f}^{(k)} = \mathbf{0} \tag{11}$$

where x_k denotes the $k-$th Cartesian coordinate. The system (11) is referred to as a "system of conservation laws" (see e.g. [34]). The equations in it are said to be written in "conservation" or "divergence" form since they directly descend from the conservation principles (8) by applying the divergence theorem (in consideration of the arbitrariness of $\mathcal{V}$) [88].

The vector $\mathbf{q}$ defined in (9), commonly referred to as the "conservative" state vector, is chosen as the independent state vector [11]. Clearly, u_k can be recast as follows:

$$u_k = \frac{\hat{\mathbf{e}}^{(k)T} \cdot (\rho \mathbf{u})}{\rho} \tag{12}$$

and therefore $\mathbf{f}^{(k)}$ admits the following representation as a function of $\mathbf{q}$:

$$\mathbf{f}^{(k)} (\mathbf{q}) = \frac{\hat{\mathbf{e}}^{(k)T} \cdot (\rho \mathbf{u})}{\rho} \mathbf{q} + p \, (\rho) \left(\begin{array}{c} 0 \\ \hat{\mathbf{e}}^{(k)} \end{array} \right)$$

It is of interest to explicitly compute the following Jacobian:

$$\mathbf{J}^{(k)} := \partial_{\mathbf{q}} \mathbf{f}^{(k)} (\mathbf{q}) \tag{13}$$

in view of the fact that, for smooth solutions, the system (11) is equivalent to the following first-order quasi-linear one [55]:

$$\partial_t \, \mathbf{q} + \sum_{k=1}^{3} \mathbf{J}^{(k)} \cdot \partial_{x_k} \mathbf{q} = \mathbf{0} \tag{14}$$

Then, the following expression is obtained by deriving (10) (by virtue of the relevant definitions introduced in sec. 2.1):

$$\mathbf{J}^{(k)} = \mathbf{q} \cdot \partial_{\mathbf{q}} u_k + u_k \, \partial_{\mathbf{q}} \mathbf{q} + \left(\begin{array}{c} 0 \\ \hat{\mathbf{e}}^{(k)} \end{array} \right) \cdot \partial_{\mathbf{q}} p \tag{15}$$

[11] As pointed out, the "conservation" character of the system (11) is connected with its mathematical structure and it is by no means due to the specific choice of the "conservative" state vector as dependent variable.

By recalling (2)-(4), the derivative $\partial_{\mathbf{q}} p$ is given by:

$$\partial_{\mathbf{q}} p = \begin{pmatrix} a^2 \\ \mathbf{0} \end{pmatrix}^T$$

where a denotes the sound speed and therefore:

$$\begin{pmatrix} 0 \\ \hat{\mathbf{e}}^{(k)} \end{pmatrix} \cdot \partial_{\mathbf{q}} p = \begin{pmatrix} 0 & \mathbf{0}^T \\ a^2 \hat{\mathbf{e}}^{(k)} & \mathbf{O} \end{pmatrix}$$

Moreover, by differentiating (12), the following relation is obtained:

$$\partial_{\mathbf{q}} u_k = \rho^{-1} \begin{pmatrix} -u_k \\ \hat{\mathbf{e}}^{(k)} \end{pmatrix}^T$$

and thus ($\mathbf{q}$ can be easily divided by ρ):

$$\mathbf{q} \cdot \partial_{\mathbf{q}} u_k = \begin{pmatrix} 1 \\ \mathbf{u} \end{pmatrix} \cdot \begin{pmatrix} -u_k \\ \hat{\mathbf{e}}^{(k)} \end{pmatrix}^T = \begin{pmatrix} -u_k & \hat{\mathbf{e}}^{(k)T} \\ -u_k \mathbf{u} & \mathbf{u} \cdot \hat{\mathbf{e}}^{(k)T} \end{pmatrix}$$

Finally, $\partial_{\mathbf{q}} \mathbf{q}$ is trivially equal to the identity matrix. Then, by exploiting the usual compact notation, it is possible to write the following equality:

$$u_k \, \partial_{\mathbf{q}} \mathbf{q} = \begin{pmatrix} u_k & \mathbf{0}^T \\ \mathbf{0} & u_k \mathbf{I} \end{pmatrix}$$

By substituting the relevant entities into (15), the following representation is finally obtained:

$$\mathbf{J}^{(k)} = \begin{pmatrix} 0 & \hat{\mathbf{e}}^{(k)T} \\ \\ a^2 \, \hat{\mathbf{e}}^{(k)} - u_k \mathbf{u} & \mathbf{u} \cdot \hat{\mathbf{e}}^{(k)T} + u_k \, \mathbf{I} \end{pmatrix} \tag{16}$$

Note 3 *It is well-known that the system (8), or its differential counterpart (11), does not explicitly involve any similarity parameter [88]. This means that the considered governing equations can be thought to involve either dimensional or non-dimensional entities (flow, space and time variables). In the latter case, the non-dimensional form is obtained from the dimensional one by a standard technique [88], once introduced the following reference entities:*

- x_{ref}*: a reference length;*

- u_{ref}*: a reference speed;*

- ρ_{ref}*: a reference density;*

- $t_{ref} := x_{ref}\, u_{ref}^{-1}$*: a reference time;*

- $p_{ref} := \rho_{ref}\, u_{ref}^{2}$*: a reference pressure.*

Of course, this observation immediately applies to any subsequent system of equations which is derived from (8) or (11) by means of simplifying assumptions. In particular, it directly applies to the Riemann problem introduced in sec. 2.4 (the similarity character of its solution being preserved by the aforementioned non-dimensionalization procedure). Moreover, the considered observation holds true also when recasting the governing equations in a rotating frame (see sec. 2.2.2) at the only cost of introducing an additional reference entity, namely:

- $\omega_{ref} := x_{ref}^{-1}\, u_{ref}$*: a reference rotational speed.*

2.2.2 3D equations in rotating frames

With respect to a Cartesian frame having the same origin as that one introduced in sec. 2.2.1 and rotating with constant angular velocity $\boldsymbol{\omega}$, the mass and momentum balances (8) read:

$$\partial_t \int_{\mathcal{V}} \mathbf{q}\ \mathrm{dV} + \int_{\mathcal{S}} \left(\sum_{k=1}^{3} \hat{n}_k\, \mathbf{f}^{(k)} \right) \mathrm{dS} = \int_{\mathcal{V}} \mathbf{s}\ \mathrm{dV} \tag{17}$$

with (relevant definitions from sec. 2.2.1 are recalled):

$$\mathbf{s} := - \begin{pmatrix} 0 \\ 2\,\boldsymbol{\omega} \wedge \rho\mathbf{u} + \rho\,\boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \mathbf{x}) \end{pmatrix} \tag{18}$$

where $\mathbf{x}$ denotes the position of the generic fluid particle with respect to the considered rotating frame, $\mathbf{u}$ consistently represents the relative velocity and the symbol $\wedge$ indicates the usual vector product. The vector $\mathbf{s}$ accounts for the non-inertial effects related to the frame rotation; indeed, the terms $2\,\boldsymbol{\omega} \wedge \rho\mathbf{u}$ and $\rho\,\boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \mathbf{x})$ in (18) respectively represent the analogue of the well-known Coriolis and centrifugal forces of classical rational mechanics [68]. Regular solutions of (17) also satisfy its differential counterpart, namely:

$$\partial_t\, \mathbf{q} + \sum_{k=1}^{3} \partial_{x_k} \mathbf{f}^{(k)} = \mathbf{s} \tag{19}$$

2.2.3 Basic-1D equations

Let u be the unique component of the velocity vector in a purely 1D motion along a certain direction, associated with the coordinate x; moreover, let $[\alpha, \beta]$ denote an arbitrary control volume along the x-axis. Once defined the 1D counterparts of (9) and (10) as follows:

$$\mathbf{q}^{(x)} := \begin{pmatrix} \rho \\ \rho u \end{pmatrix} \tag{20}$$

$$\mathbf{f}^{(x)} := \begin{pmatrix} \rho u \\ \rho u^2 + p \end{pmatrix} \tag{21}$$

the 1D balances corresponding to (8) and (11) respectively read:

$$\partial_t \int_\alpha^\beta \mathbf{q}^{(x)} \, \mathrm{d}x + \mathbf{f}^{(x)}|_\beta - \mathbf{f}^{(x)}|_\alpha = \mathbf{0} \tag{22}$$

$$\partial_t \, \mathbf{q}^{(x)} + \partial_x \mathbf{f}^{(x)} = \mathbf{0} \tag{23}$$

2.2.4 Augmented-1D equations

It is possible to extend the balances (22) and (23) so as to also describe the mass conservation of a certain substance merely advected with the flow (commonly referred to as a "passive scalar"). Indeed, once extended the definitions (20) and (21) as follows:

$$\mathbf{q}^{(A)} := \begin{pmatrix} \rho \\ \rho u \\ \rho \xi \end{pmatrix} \tag{24}$$

$$\mathbf{f}^{(A)} := \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho u \xi \end{pmatrix} \tag{25}$$

where ξ denotes the concentration of the passive scalar, the "augmented" versions of the systems (22) and (23) respectively read (see e.g. [98]):

$$\partial_t \int_\alpha^\beta \mathbf{q}^{(A)} \, \mathrm{d}x + \mathbf{f}^{(A)}|_\beta - \mathbf{f}^{(A)}|_\alpha = \mathbf{0} \tag{26}$$

$$\partial_t \, \mathbf{q}^{(A)} + \partial_x \mathbf{f}^{(A)} = \mathbf{0} \tag{27}$$

Note 4 *Clearly, the conservation of the passive scalar is decoupled from the basic 1D system (as, for instance, the energy balance which has been deliberately dropped out at the beginning of sec. 2.2 for the sake of simplicity). Nevertheless, it is added to the basic 1D system in order to prepare the ground for the introduction of the 1D sweeps of the original 3D governing equations (see sec. 2.2.5).*

Note 5 *It is straightforward to extend the "augmentation" procedure described in the present section to the case of* m *passive scalars, with* $m > 1$. *For instance, let* ξ *and* η *be two passive scalars; it is possible to formally keep (26) and (27) as they are, at the only cost of extending (24) and (25) as follows:*

$$\mathbf{q}^{(A)} := \begin{pmatrix} \rho \\ \rho u \\ \rho\xi \\ \rho\eta \end{pmatrix} \tag{28}$$

$$\mathbf{f}^{(A)} := \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho u\xi \\ \rho u\eta \end{pmatrix} \tag{29}$$

It should be noticed that the m *passive scalars, while not affecting the basic 1D flow field (i.e.* ρ *and* u*), neither interact with one another even. Therefore, the structure of the solution is identical for all of them, differences only arising due to the specific initial and boundary conditions. In view of this, it is reasonable to study the case* $m = 1$, *for the sake of simplicity.*

2.2.5 1D sweeps of the 3D equations

The "$k-$th sweep" of the 3D system (11) is obtained by neglecting the summation in it, namely (see e.g. [98]):

$$\partial_t \mathbf{q} + \partial_{x_k}\mathbf{f}^{(k)} = \mathbf{0} \quad , \quad k \in \{1, 2, 3\} \tag{30}$$

By respectively comparing the definitions of $\mathbf{q}$ and $\mathbf{f}^{(k)}$ in (9) and (10) with those of $\mathbf{q}^{(A)}$ and $\mathbf{f}^{(A)}$ in (28) and (29), it is clear that (apart from the order of the components) the k-th sweep (30) is formally equal to the augmented-1D system (27), at the cost of considering:

- the $k-$th coordinate direction (i.e. $\hat{\mathbf{e}}^{(k)}$) as the direction along which a basic-1D flow takes place;

- the velocity components u_h, with $h \in \{1, 2, 3\}$ and $h \neq k$, as advected passive scalars.

This observation is exploited in sec. 5.1.2 in order to discretize the surface integral appearing in the 3D balances (8) and (17).

2.2.6 Hierarchical structure of the presented equations

The numerical discretization of the 3D problems (8) and (17) is discussed in sec. 5. The rotating case, in particular, is treated as a generalization of the non-rotating one and therefore the problem (8) is considered at a preliminary stage.

Currently, a good mathematical understanding of the problems (8) and (11) is largely unavailable [34]. For this reason, the numerical discretization of (8) is based on some numerical techniques which are applied to the 1D sweeps of the original 3D problem. As mentioned in sec. 2.2.5, the 1D sweeps are regarded to as augmented-1D systems and therefore the balances (26) and (27) are considered, in particular, for developing most of the proposed 1D numerical ingredients (see sec. 3).

The basic-1D systems (22) and (23) are exploited in sec. 3.4 to tackle some difficulties which are essentially related to the numerical discretization of the mass and momentum balances appearing in every augmented-1D system.

2.3 Hyperbolicity and integral solutions

In secs. 2.3.1 and 2.3.2, the hyperbolic character of the relevant governing equations introduced in sec. 2.2 is concisely discussed. The concept of integral solution is then presented in sec. 2.3.3 (together with some related issues like the Rankine-Hugoniot condition and the notion of entropy condition), in order to pave the way for discussing shock waves and contact discontinuities in secs. 2.4 and 2.5.

2.3.1 Hyperbolicity of the 3D equations

Let $\mathbf{J}$ represent the following matrix:

$$\mathbf{J} := \sum_{k=1}^{3} \hat{n}_k \, \mathbf{J}^{(k)} \tag{31}$$

where $\hat{n}_k$ denotes the $k-$th component of the normal $\hat{\mathbf{n}}$ appearing in the 3D balances (8) and $\mathbf{J}^{(k)}$ is defined in (13). Once substituted the expression of

$\mathbf{J}^{(k)}$ provided in (16), the following representation is obtained:

$$\mathbf{J} = \begin{pmatrix} 0 & \hat{\mathbf{n}}^T \\ & \\ a^2\,\hat{\mathbf{n}} - \left(\hat{\mathbf{n}}^T\cdot\mathbf{u}\right)\mathbf{u} & \mathbf{u}\cdot\hat{\mathbf{n}}^T + \left(\hat{\mathbf{n}}^T\cdot\mathbf{u}\right)\mathbf{I} \end{pmatrix} \tag{32}$$

The quasi-linear system (14) is said to be hyperbolic (at a certain point of the flow field) if $\mathbf{J}$ has real eigenvalues λ_j and a corresponding set of linearly independent (right) eigenvectors $\mathbf{r}_j$ ($j = 1, \ldots, 4$), for every versor $\hat{\mathbf{n}}$ on the unit sphere $\mathbb{S}^2$. Furthermore, the system is said to be "strictly hyperbolic" if the eigenvalues are all distinct [55].

It is straightforward to verify that the system (14) is actually hyperbolic, with eigenvalues (the eigenvectors are not reported for the sake of conciseness):

$$\lambda_1 = \hat{\mathbf{n}}\cdot\mathbf{u} - a \quad , \quad \lambda_2 = \lambda_3 = \hat{\mathbf{n}}\cdot\mathbf{u} \quad , \quad \lambda_4 = \hat{\mathbf{n}}\cdot\mathbf{u} + a \tag{33}$$

Note 6 *The following equation:*

$$\det\left(\,\mathbf{J} - \lambda\,\mathbf{I}\,\right) = 0 \tag{34}$$

with $\mathbf{J}$ *given by (31), can be regarded to as a partial differential equation where the unknown is a certain function* $z = z\,(x_1, x_2, x_3, t)$ *such that:*

$$\hat{\mathbf{n}} = \frac{\mathbf{z}}{\|\mathbf{z}\|} \quad , \quad \lambda = -\frac{\partial_t z}{\|\mathbf{z}\|}$$

where $\mathbf{z}$ *denotes the (spatial) gradient of* z*. Manifolds* $z = const$*, where* z *is an integral solutions of (34), are "characteristic manifolds" having normal* $\hat{\mathbf{n}}$ *and moving with a normal component of the velocity equal to* λ*:*

$$\lambda = \hat{\mathbf{n}}\cdot\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} \tag{35}$$

where $\mathbf{x}$ *here denotes the position of the generic point on the manifold (see e.g. [19], [20], [55], [56], [65] and [115]). By comparing (33) and (35), it is clear that the manifolds associated with* λ_2 *and* λ_3 *for the system (14) are simply advected with the flow (i.e. they behave like material surfaces) while those associated with* λ_1 *and* λ_4 *isotropically propagate along* $\hat{\mathbf{n}}$ *with the sound speed* a*.*

Characteristic manifolds associated with the speeds (33) can transport discontinuities of the derivatives of the solution of (14) within the flow field (see e.g. [65]). This point, together with the well-known result that discontinuities can arise during the evolution of the solution also by starting from smooth data (see e.g. [34], [55] and [60]), clearly shows that it is not possible, in general, to find smooth solutions of the differential problem (11). Therefore, some way to interpret less regular solutions somehow "solving" (11), or its simplified counterpart (27), must be devised (see sec. 2.3.3).

2.3.2 Hyperbolicity of the augmented-1D equations

The quasi-linear form of the system (27) reads:

$$\partial_t \, \mathbf{q}^{(A)} + \mathbf{J}^{(A)} \cdot \partial_x \mathbf{q}^{(A)} = \mathbf{0} \tag{36}$$

where $\mathbf{J}^{(A)}$ is the Jacobian of the function $\mathbf{f}^{(A)} \left(\mathbf{q}^{(A)} \right)$ defined by (24)-(25):

$$\mathbf{J}^{(A)} := \partial_{\mathbf{q}^{(A)}} \mathbf{f}^{(A)} \tag{37}$$

By analogy with the 3D case discussed in sec. 2.3.1, the hyperbolicity of the system (36) depends on the eigenstructure of $\mathbf{J}^{(A)}$. In particular, it is straightforward to verify that it is strictly hyperbolic, with the following pairs of eigenvalue-eigenvector:

$$\begin{cases} \lambda_1 = u - a & , \quad \mathbf{r}_1 = (1, u - a, \xi)^T \\ \\ \lambda_2 = u & , \quad \mathbf{r}_2 = (0, 0, 1)^T \\ \\ \lambda_3 = u + a & , \quad \mathbf{r}_3 = (1, u + a, \xi)^T \end{cases} \tag{38}$$

2.3.3 Integral solutions

By following [34], a certain field $\mathbf{z} \in L^\infty \left(\mathbb{R} \times (0, \infty) \, ; \mathbb{R}^m \right)$ is said to be an integral solution of the following initial-value problem:

$$\begin{cases} \partial_t \, \mathbf{z} + \partial_x \mathbf{f} & = & \mathbf{0} & \text{in} & \mathbb{R} \times (0, \infty) \\ \mathbf{z} & = & \mathbf{z}^{(0)} & \text{on} & \mathbb{R} \times \{t = 0\} \end{cases} \tag{39}$$

with $\mathbf{f} = \mathbf{f}\left(\mathbf{z}\right)$, once provided the following equality:

$$\int_0^\infty \int_{-\infty}^\infty \left(\mathbf{z} \cdot \partial_t \mathbf{v} + \mathbf{f} \cdot \partial_x \mathbf{v} \right) \mathrm{dx} \, \mathrm{dt} + \int_{-\infty}^\infty \mathbf{z}^{(0)} \cdot \mathbf{v}|_{t=0} \, \mathrm{dx} = 0 \tag{40}$$

holds for all the test functions $\mathbf{v}$ such that $\mathbf{v}$ is smooth and has compact support.

Note 7 *The relation (40) is obtained integrating by parts the dot product between the p.d.e. in (39) and* $\mathbf{v}$*. Even if (40) is obtained by assuming that* $\mathbf{z}$ *is a smooth solution of (39), it makes sense if* $\mathbf{z}$ *is merely bounded. In consideration of the fact that the solution set of (40) contains that one of (39), the integral solutions are also called "weak" or "generalized" solutions (see e.g. [63] and [98]).*

Rankine-Hugoniot condition

Let $\{(x,t) \mid x = s(t)\}$, for some smooth function $s(\cdot) : [0,\infty) \to \mathbb{R}$, represent a curve γ dividing a certain domain within $\mathbb{R} \times (0,\infty)$ into a "left" and a "right" sub-domain. Let $\mathbf{z}$ be an integral solution of (39) which is smooth on either sides of γ, along which $\mathbf{z}$ has simple jump discontinuities. Then, the integral solution must verify the classical Rankine-Hugoniot (hereafter RH as well) condition across γ (see e.g. [34]):

$$[\mathbf{f}] = \sigma \; [\mathbf{z}] \tag{41}$$

where $\sigma := \mathrm{d}s/\mathrm{d}t$ and $[\psi]$ denotes the difference between the "left" and "right" limits of ψ across γ (or vice-versa, consistently on both sides of (41)). By analogy with classical fluid mechanics, the discontinuity along γ is commonly referred to as a "shock wave" or, briefly, "shock".

Entropy conditions

It is well known that integral solutions need not be unique and additional requirements for properly defining generalized solutions of (39) must be introduced [34]. A key issue, in particular, is the definition of "admissible shocks", i.e. discontinuities subjected to (41) which link a certain state $\mathbf{z}_2$ to a given state $\mathbf{z}_1$ in such a way that the evolution from $\mathbf{z}_1$ towards $\mathbf{z}_2$ is acceptable from a certain, say "physical", point of view while the reciprocal path is not. Criteria for selecting admissible shocks are called entropy conditions by analogy with classical gas dynamics, where the admissible shocks (from supersonic to subsonic regimes) are selected by exploiting the second principle of thermodynamics (i.e. the non-decreasing trend of the thermodynamic entropy) [66]. While allowing for the identification of the relevant physical evolution, the entropy conditions generally permit to obtain a unique solution of the mathematical problem.

Classical criteria like the Lax entropy condition [61] or the Liu entropy criterion (see e.g. [34]) provide restrictions on a possible couple of states joined by a shock. However, it is possible to widen the entropy criteria so as to apply to more general integral solutions of the considered conservation laws. In particular, it is possible to define so-called entropy solutions, i.e. solutions obeying certain requirements of the type of the Oleinik condition [72] (see e.g. [34]). The fundamental idea upon which the aforementioned entropy criteria are based is that physically and mathematically correct solutions of the p.d.e. in (39) should arise as the limit of the solutions to the following parabolic system (which admits travelling wave solutions, see e.g. [34]):

$$\partial_t \tilde{\mathbf{z}} + \partial_x \mathbf{f}(\tilde{\mathbf{z}}) = \varepsilon\, \partial_x (\partial_x \tilde{\mathbf{z}}) \quad , \quad \varepsilon > 0 \tag{42}$$

as the "viscosity" term on the right hand side vanishes (i.e. for $\varepsilon \to 0$) [12].

Another approach exploited for selecting relevant integral solutions consists in introducing suitable functions, called entropy functions, for which an additional conservation law holds for smooth solutions that becomes an inequality for discontinuous solutions (see e.g. [34] and [63]).

There is great ongoing interest in studying entropy conditions. Indeed, it is very difficult to assess criteria holding for general conservation laws, and in particular for generic relations $\mathbf{f}(\mathbf{z})$ [13]; a vast number of results is currently available only for simplified systems or for scalar conservation laws [34]. In consideration of this, the classical Lax entropy condition is adopted (see sec. 2.4.2), for the sake of simplicity, in order to determine the solution of the so-called Riemann problem (see sec. 2.4) involving the system (27), closed by a convex barotropic state law (see sec. 2.5.1).

2.4 The Riemann problem

In this section, the following system of conservation laws is considered:

$$\left\{ \begin{array}{rcll} \partial_t \mathbf{z} + \partial_x \mathbf{f} & = & \mathbf{0} & \text{in} \quad \mathbb{R} \times (0, \infty) \\ \\ \mathbf{z} & = & \left\{ \begin{array}{lll} \mathbf{z}_L & if & x < 0 \\ \mathbf{z}_R & if & x > 0 \end{array} \right. & \text{on} \quad \mathbb{R} \times \{t = 0\} \end{array} \right. \tag{43}$$

with $\mathbf{z} \in \mathbb{R}^m$ and $\mathbf{f} = \mathbf{f}(\mathbf{z}) \in \mathbb{R}^m$. This system, characterized by a step-like piece-wise constant initial data, is commonly referred to as the Riemann problem (hereafter RP as well).

The solution to (43) also depends, in general, on the state law closing the relevant p.d.e. through the specific relation $\mathbf{f}(\mathbf{z})$. This solution (when available, since for sophisticated state laws it is very difficult to be obtained) plays an important role in the set up of modern numerical methods for fluid dynamics (see e.g. [63], [98] and [99]) and therefore a vast class of RPs has been studied within this context, even involving complex state laws (see e.g. [13] and [69] amongst many others).

[12] For this reason, an entropy solution is also called a vanishing-viscosity solution.

[13] The specific form of the state law which closes the considered differential problem clearly affects $\mathbf{f}(\mathbf{z})$; in general, it can render it very difficult to define suitable entropy criteria. (see e.g. [3] and the cited references).

Once introduced some relevant definitions and results in sec. 2.4.1, the basic wave solutions of (43) are presented in sec. 2.4.2 and, finally, an important theorem related to the local solution of (43) is mentioned in sec. 2.4.3. All the material concisely presented in secs. 2.4.1 to 2.4.3, essentially taken from [34] and [55], is aimed at preparing the ground for the solution of the specific RP studied in sec 2.5.

2.4.1 Preliminary definitions and results

The system in (43) is supposed to be strictly hyperbolic, with pairs $(\lambda_k, \mathbf{r}_k)$ $(k = 1, \ldots, m)$ of eigenvalue-eigenvector associated with the Jacobian $\partial_{\mathbf{z}}\mathbf{f}$ appearing in its quasi-linear form:

$$\partial_t \, \mathbf{z} + \partial_{\mathbf{z}}\mathbf{f} \cdot \partial_x \mathbf{z} = \mathbf{0}$$

Characteristics

The following differential equations:

$$\frac{\mathrm{d}x}{\mathrm{d}\alpha} = \lambda_k\left(\mathbf{z}\right) \quad , \quad \frac{\mathrm{d}t}{\mathrm{d}\alpha} = 1 \tag{44}$$

where α is an abscissa, define the $k-$th family of characteristic curves (briefly: the $k-$th characteristics) in the $x - t$ plane, associated with the hyperbolic system in (43).

Genuinely non-linear and linearly-degenerate pairs

The pair $(\lambda_k, \mathbf{r}_k)$, with $\lambda_k = \lambda_k\left(\mathbf{z}\right)$ and $\mathbf{r}_k = \mathbf{r}_k\left(\mathbf{z}\right)$ is called genuinely non-linear (briefly: g.n.) provided:

$$\partial_{\mathbf{z}}\lambda_k \cdot \mathbf{r}_k \neq 0 \quad , \quad \forall \mathbf{z} \in \mathbb{R}^m$$

Conversely, it is said to be linearly-degenerate (briefly: l.d.) if:

$$\partial_{\mathbf{z}}\lambda_k \cdot \mathbf{r}_k = 0 \quad , \quad \forall \mathbf{z} \in \mathbb{R}^m$$

Rarefaction curves

Given a fixed state $\mathbf{z}_0 \in \mathbb{R}^m$, the $k-$th rarefaction curve $R_k\left(\mathbf{z}_0\right)$ is defined as the path in $\mathbb{R}^m$ of the solution to the following ordinary differential equation (hereafter o.d.e. as well):

$$\frac{\mathrm{d}\mathbf{v}\left(\xi\right)}{\mathrm{d}\xi} = \mathbf{r}_k\left(\, \mathbf{v}\left(\xi\right)\,\right) \tag{45}$$

which passes through $\mathbf{z}_0$. If $(\lambda_k, \mathbf{r}_k)$ is g.n., then (45) shows that λ_k monotonically increases or decreases along $R_k(\mathbf{z}_0)$ and therefore:

$$R_k(\mathbf{z}_0) = R_k^+(\mathbf{z}_0) \cup \{\mathbf{z}_0\} \cup R_k^-(\mathbf{z}_0)$$

with:

$$\left\{ \begin{array}{lcl} R_k^+(\mathbf{z}_0) & := & \{\mathbf{z} \in R_k(\mathbf{z}_0) \mid \lambda_k(\mathbf{z}_0) < \lambda_k(\mathbf{z})\} \\ \\ R_k^-(\mathbf{z}_0) & := & \{\mathbf{z} \in R_k(\mathbf{z}_0) \mid \lambda_k(\mathbf{z}) < \lambda_k(\mathbf{z}_0)\} \end{array} \right. \tag{46}$$

Simple waves

A simple wave is a solution of the p.d.e. in (43) having the following structure:

$$\mathbf{z}(x,t) = \mathbf{v}(\,\eta(x,t)\,) \quad \text{in} \quad \mathbb{R} \times (0,\infty) \tag{47}$$

It is possible to show that $\mathbf{v}$ in (47) necessarily satisfies (45) for some k. In addition, η in (47) must satisfy the following p.d.e.:

$$\partial_t \eta + \lambda_k(\,\mathbf{v}(\eta)\,)\,\partial_x \eta = 0 \tag{48}$$

The simple wave $\mathbf{z}$ given by (47) is consequently called a $k-$simple wave.

Shock set

Given a fixed state $\mathbf{z}_0 \in \mathbb{R}^m$, the so-called shock set is defined as follows:

$$S(\mathbf{z}_0) := \{\mathbf{z} \in \mathbb{R}^m \mid \mathbf{f}(\mathbf{z}) - \mathbf{f}(\mathbf{z}_0) = \sigma\,(\mathbf{z} - \mathbf{z}_0)\} \tag{49}$$

where σ depends on the states $\mathbf{z}$ and $\mathbf{z}_0$: $\sigma = \sigma(\mathbf{z}, \mathbf{z}_0)$. It is possible to show that, in some neighbourhood of $\mathbf{z}_0$, $S(\mathbf{z}_0)$ consists of the union of m smooth curves $S_k(\mathbf{z}_0)$ $(k = 1, \ldots, m)$ with the following properties:

- $S_k(\mathbf{z}_0)$ passes through $\mathbf{z}_0$ with tangent $\mathbf{r}_k(\mathbf{z}_0)$;
- $\sigma(\mathbf{z} \in S_k(\mathbf{z}_0), \mathbf{z}_0) = \dfrac{\lambda_k(\mathbf{z}) + \lambda_k(\mathbf{z}_0)}{2} + \mathrm{O}\left(|\mathbf{z} - \mathbf{z}_0|^2\right)$ as $\mathbf{z} \to \mathbf{z}_0$.

Furthermore, it is possible to show that, if $(\lambda_k, \mathbf{r}_k)$ is g.n., then (provided $\mathbf{z}$ is close enough to $\mathbf{z}_0$):

$$S_k(\mathbf{z}_0) = S_k^+(\mathbf{z}_0) \cup \{\mathbf{z}_0\} \cup S_k^-(\mathbf{z}_0) \tag{50}$$

with:

$$\left\{ \begin{array}{lcl} S_k^+(\mathbf{z}_0) & := & \{\mathbf{z} \in S_k(\mathbf{z}_0) \mid \lambda_k(\mathbf{z}_0) < \sigma(\mathbf{z}, \mathbf{z}_0) < \lambda_k(\mathbf{z})\} \\ \\ S_k^-(\mathbf{z}_0) & := & \{\mathbf{z} \in S_k(\mathbf{z}_0) \mid \lambda_k(\mathbf{z}) < \sigma(\mathbf{z}, \mathbf{z}_0) < \lambda_k(\mathbf{z}_0)\} \end{array} \right. \tag{51}$$

Linear degeneracy

If $(\lambda_k, \mathbf{r}_k)$ is l.d. for some $k \in \{1, \ldots, m\}$, then it is possible to show that, for each $\mathbf{z}_0 \in \mathbb{R}^m$:

- $S_k(\mathbf{z}_0) = R_k(\mathbf{z}_0)$;
- $\sigma(\mathbf{z}, \mathbf{z}_0) = \lambda_k(\mathbf{z}) = \lambda_k(\mathbf{z}_0)$, $\forall \mathbf{z} \in S_k(\mathbf{z}_0)$.

Note 8 *It should be noticed that the curves $S_k(\mathbf{z}_0)$ and $R_k(\mathbf{z}_0)$, which in general agree at least to the first order at $\mathbf{z}_0$, coincide in the linearly-degenerate case.*

2.4.2 Basic wave solutions

The basic wave solutions of the RP (43) are considered in the sequel.

Rarefactions

It is possible to show that there exists a continuous integral solution $\mathbf{z}$ of the RP (43) which is a $k-$simple wave constant along lines [14] through the origin $(x = 0, t = 0)$, for $\lambda_k(\mathbf{z}_L) \leq x/t \leq \lambda_k(\mathbf{z}_R)$, provided that:

- $(\lambda_k, \mathbf{r}_k)$ is g.n.;
- $\mathbf{z}_R \in R_k^+(\mathbf{z}_L)$.

For this solution, $\eta(x,t)$ in (47) is given by $\eta(x,t) = \bar{\eta}(x/t)$, for a suitable function $\bar{\eta}$. Since η is constant along a $k-$th characteristic by virtue of (48) and (44), it follows that also the $k-$th characteristics, for $\lambda_k(\mathbf{z}_L) \leq x/t \leq \lambda_k(\mathbf{z}_R)$, are lines through the origin and therefore (44), in particular, becomes:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{x}{t} = \lambda_k(\mathbf{z}) \quad , \quad \lambda_k(\mathbf{z}_L) \leq \frac{x}{t} \leq \lambda_k(\mathbf{z}_R) \tag{52}$$

The $k-$th characteristics for the solution under consideration, also referred to as a "$k-$rarefaction (wave)" by analogy with classical gas dynamics [66], are sketched in Fig. 3.

Note 9 *The relation (45), in particular, holds for the solution at hand; once recast in differential form, it states that the solution $\mathbf{z} \in \mathbb{R}^m$, while moving*

[14] In the present document, the word "line" denotes a straight curve while the word "curve" stands for a generic curve.

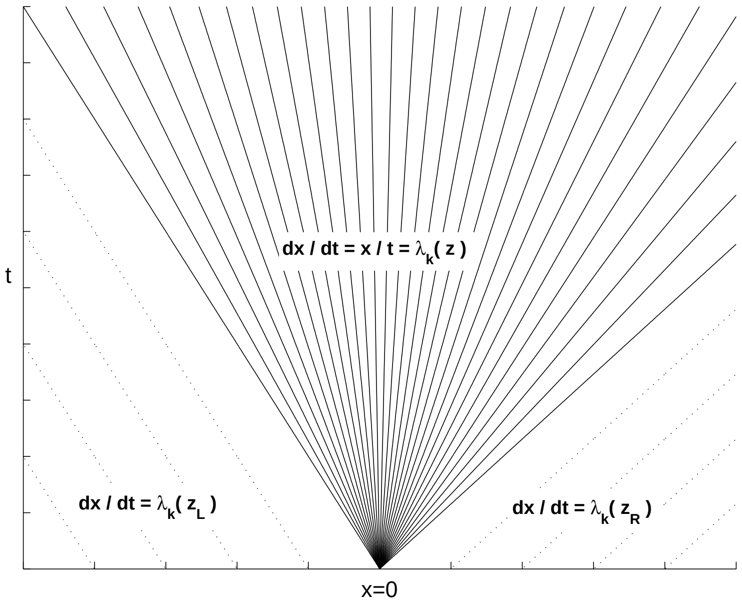

Figure 3: Schematic representation of the $k-$th characteristics for a generic $k-$rarefaction.

between two infinitesimally close lines within the wave region sketched in Fig. 3, must satisfy the following condition:

$$\mathrm{d}\mathbf{z} \propto \mathbf{r}_k\left(\mathbf{z}\right) \tag{53}$$

The above condition provides $m-1$ independent scalar differential relations which can be integrated, thus providing $m-1$ prime integrals across the wave [15]*. The prime integrals under consideration are sometimes called generalized Riemann invariants (see e.g. [55] and [98])* [16]*.*

Shocks - Lax entropy condition

If $\mathbf{z}_R \in S_k\left(\mathbf{z}_L\right)$, the following field:

$$\mathbf{z}\left(x,t\right) = \begin{cases} \mathbf{z}_L & \text{if} \quad x < \sigma t \\ \mathbf{z}_R & \text{if} \quad x > \sigma t \end{cases} \tag{54}$$

[15] Since $\mathbf{r}_k$ is only determined up to an arbitrary multiplicative constant, it might be necessary to choose the multiplicative factor tacitly appearing in (53) so as to be a suitable integrating factor.

[16] Indeed they generalize the classical Riemann invariants, which do not exist, in general, for $m > 2$ (see e.g. [34]).

with $\sigma = \sigma\left(\mathbf{z}_R, \mathbf{z}_L\right)$, is an integral solution of (43). By virtue of (49), the solution at hand satisfies the RH condition (41) and it is consequently referred to as a shock (wave).

If $\left(\lambda_k, \mathbf{r}_k\right)$ is g.n. then $\mathbf{z}_R$ (provided it is close enough to $\mathbf{z}_L$) can belong either to $S_k^+\left(\mathbf{z}_L\right)$ or to $S_k^-\left(\mathbf{z}_L\right)$, due to (50). By adopting the classical Lax entropy condition [61] (hereafter LEC as well), only the latter possibility is considered acceptable; once recalled the definition of S_k^- in (51), it is possible to recast the LEC as follows:

$$\lambda_k\left(\mathbf{z}_R\right) < \sigma\left(\mathbf{z}_R, \mathbf{z}_L\right) < \lambda_k\left(\mathbf{z}_L\right) \tag{55}$$

Then, the shock (54) is accepted as an integral solution of (43) if and only if the pair $\left(\mathbf{z}_R, \mathbf{z}_L\right)$ satisfies (55).

Note 10 *According to the LEC (55), the $k-$th characteristics from left and right (lines on both sides) run into the shock, as sketched in Fig. 4. Since the characteristics act as information carriers (see e.g. [55] and [65]), some information is lost when they reach the shock, thus increasing a suitably defined entropy of the system (see e.g. [34]). This, in turn, is considered as a proper criterion for assessing the physical representativeness of the shock, in analogy with the case of classical gas dynamics in which the LEC paraphrases the non-decreasing character of the thermodynamic entropy, i.e. the second principle of thermodynamics [66]. In view of this consideration, it is clear why the LEC is regarded to as an entropy condition.*

Contact discontinuities

The expression (54) is an integral solution of (43) also when $\left(\lambda_k, \mathbf{r}_k\right)$ is l.d., at the obvious cost of choosing $\sigma = \sigma\left(\mathbf{z}_R, \mathbf{z}_L\right) = \lambda_k\left(\mathbf{z}_L\right) = \lambda_k\left(\mathbf{z}_R\right)$, as imposed by the linear degeneracy (see sec. 2.4.1). The left and right $k-$th characteristics (lines on both sides) are then parallel to the discontinuity, as sketched in Fig. 5. This solution is called a $k-$contact discontinuity.

Note 11 *In consideration of the fact that $S_k\left(\mathbf{z}\right) = R_k\left(\mathbf{z}\right)$ for the contact discontinuity, as imposed by the linear degeneracy (see sec. 2.4.1), the jump $\left[\mathbf{z}\right]$ across the discontinuity can be computed by exploiting either the RH condition (41) or by integrating (53) (as for a rarefaction).*

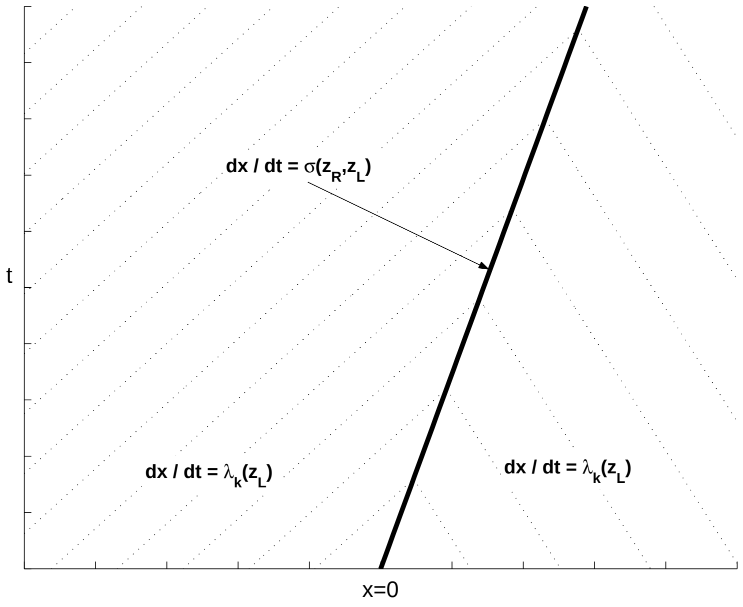
dx / dt = $\sigma(z_R,z_L)$
t
dx / dt = $\lambda_k(z_L)$
dx / dt = $\lambda_k(z_L)$
x=0

Figure 4: Schematic representation of the $k-$th characteristics for a generic shock satisfying the LEC (55).

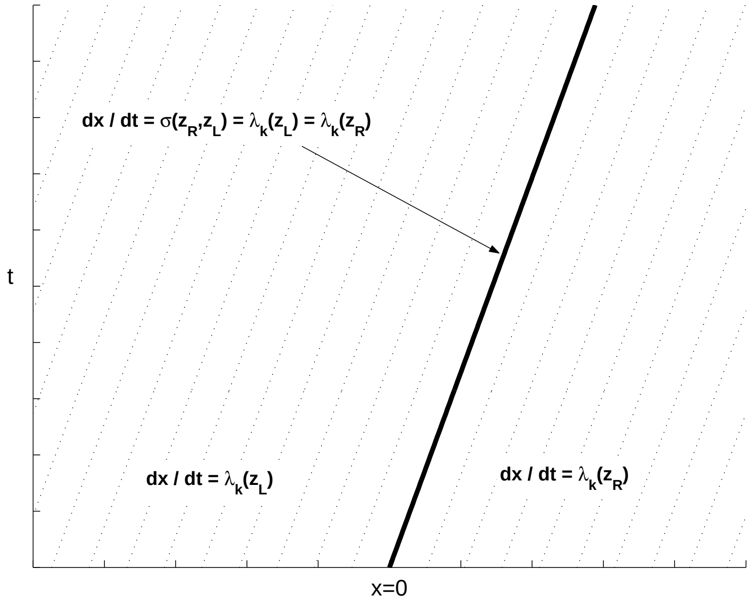
dx / dt = $\sigma(z_R,z_L) = \lambda_k(z_L) = \lambda_k(z_R)$
t
dx / dt = $\lambda_k(z_L)$
dx / dt = $\lambda_k(z_R)$
x=0

Figure 5: Schematic representation of the $k-$th characteristics for a generic $k-$contact discontinuity.

2.4.3 Local solution of the Riemann problem

An important theorem shows that, if $(\lambda_k, \mathbf{r}_k)$ is either g.n. or l.d. for each $k \in \{1, \ldots, m\}$, then there exists an integral solution of the RP (43) which is constant on lines through the origin $(x = 0, t = 0)$, provided the initial states $\mathbf{z}_L$ and $\mathbf{z}_R$ are sufficiently close to each other [34]. While proving the aforementioned statement, it is possible to construct the solution by connecting $m + 1$ states $\mathbf{z}_h$ $(h = 0, \ldots, m)$ by means of m waves of the type of those discussed in sec. 2.4.2 (i.e. rarefactions, shocks and contact discontinuities). More in detail, let $\lambda_k < \lambda_{k+1}$ (which is legitimate, due to the strict hyperbolicity assumed at the beginning of sec. 2.4.1); then, once chosen $\mathbf{z}_0 = \mathbf{z}_L$ and $\mathbf{z}_m = \mathbf{z}_R$, the wave joining $\mathbf{z}_{k-1}$ to $\mathbf{z}_k$ is a $k-$th rarefaction, a shock or a $k-$th contact discontinuity, provided:

$$\mathbf{z}_k \in T_k\left(\mathbf{z}_{k-1}\right)$$

where the curve $T_k(\mathbf{z})$ is defined, in some neighbourhood of $\mathbf{z}$, as follows:

$$T_k(\mathbf{z}) := \begin{cases} R_k^+(\mathbf{z}) \cup \{\mathbf{z}\} \cup S_k^-(\mathbf{z}) & \text{if} \quad (\lambda_k, \mathbf{r}_k) \quad \text{g.n.} \\ \\ R_k(\mathbf{z}) = S_k(\mathbf{z}) & \text{if} \quad (\lambda_k, \mathbf{r}_k) \quad \text{l.d.} \end{cases} \tag{56}$$

This constructive procedure is applied in sec. 2.5.3 in order to solve the RP (43) when associated with a generic but convex barotropic state law.

2.5 The Riemann problem for a convex barotropic state law

In this section, the following RP is considered:

$$\begin{cases} \partial_t \mathbf{q}^{(A)} + \partial_x \mathbf{f}^{(A)} = \mathbf{0} & \text{in} \quad \mathbb{R} \times (0, \infty) \\ \\ \mathbf{q}^{(A)} = \begin{cases} \mathbf{q}_L^{(A)} & if \quad x < 0 \\ \mathbf{q}_R^{(A)} & if \quad x > 0 \end{cases} & \text{on} \quad \mathbb{R} \times \{t = 0\} \end{cases} \tag{57}$$

with $\mathbf{q}^{(A)}$ and $\mathbf{f}^{(A)}$ given by (24) and (25), respectively. The p.d.e. in (57) (i.e. the system (27)) is supposed to be closed by a generic state law like that one defined in sec. 1.5, subjected to an additional constraint that is discussed in sec. 2.5.1. In sec. 2.5.2 basic wave solutions are investigated, which are exploited in sec. 2.5.3 for constructing the solution of the RP (57).

2.5.1 Convexity of the state law

In [69] general constitutive relations involving several thermodynamic entities are investigated and the convexity of a given state law is defined within a quite general context. To the purposes of the present work, it suffices to mention that the generic barotropic state law (3) is said to be convex if:

$$\frac{\mathrm{d}^2 p}{\mathrm{d}v^2} > 0 \tag{58}$$

where $v := \rho^{-1}$ is the so-called specific volume. As a function of ρ, the term on the left-hand side of the condition (58) can be recast as follows:

$$\frac{\mathrm{d}^2 p}{\mathrm{d}v^2} = 2\, a\, \rho^4\, c(\rho) \tag{59}$$

where:

$$c(\rho) := \frac{a}{\rho} + \frac{\mathrm{d}a}{\mathrm{d}\rho} \tag{60}$$

Hence, the condition (58) can be equivalently expressed as follows:

$$c(\rho) > 0 \tag{61}$$

Some considerations can be drawn from the convexity condition (61), namely:

- let $\chi = \chi(\rho)$ be defined as follows:

  $$\chi(\rho) := \rho\; a(\rho) \tag{62}$$

  Clearly $\mathrm{d}\chi/\mathrm{d}\rho = \rho\, c(\rho)$ and therefore for a convex barotropic state law χ is a monotonically increasing function of ρ;

- let $\Theta\,(\rho_0, \rho)$ be defined as follows (a prolongation by continuity is considered):

  $$\Theta\,(\rho_0, \rho) := \begin{cases} \rho_0\, \rho\, \dfrac{p - p_0}{\rho - \rho_0} & \text{if} \quad \rho \neq \rho_0 \\[2ex] \chi_0^2 & \text{if} \quad \rho = \rho_0 \end{cases} \tag{63}$$

  where $\chi_0 := \chi\,(\rho_0)$, with χ defined in (62). In consideration of the stability constraint (4), $\Theta > 0$. Moreover, for a convex barotropic state law it is possible to show that:

  $$\begin{cases} \Theta\,(\rho_0, \rho) > \chi_0^2 & \text{if} \quad \rho > \rho_0 \\[2ex] \Theta\,(\rho_0, \rho) < \chi_0^2 & \text{if} \quad \rho < \rho_0 \end{cases} \tag{64}$$

To the purpose, it suffices to consider the following function:

$$\Gamma(\rho) := (\rho - \rho_0)\left(\Theta - \chi_0^2\right)$$

in which ρ_0 is regarded to as a fixed parameter. Indeed, by virtue of the convexity condition (61), $\mathrm{d}^2\Gamma/\mathrm{d}\rho^2 > 0$ while $\mathrm{d}\Gamma/\mathrm{d}\rho = 0$ for $\rho = \rho_0$. Hence, Γ has one and only one minimum in correspondence of $\rho = \rho_0$; since $\Gamma(\rho = \rho_0) = 0$, it follows that $\Gamma > 0$ for $\rho \neq \rho_0$ and therefore $(\Theta - \chi_0^2)$ has the same sign as $(\rho - \rho_0)$ (thus obtaining (64)).
Once introduced the following definition:

$$\zeta(\rho_0, \rho) := +\,\Theta(\rho_0, \rho)^{\frac{1}{2}} \tag{65}$$

with Θ given by (63), it is possible to recast the inequalities (64) (which only involve positive entities) as follows:

$$\begin{cases} \zeta(\rho_0, \rho) > \chi_0 & \text{if} \quad \rho > \rho_0 \\ \\ \zeta(\rho_0, \rho) < \chi_0 & \text{if} \quad \rho < \rho_0 \end{cases} \tag{66}$$

Then, in consideration of the symmetry $\zeta(\rho_0, \rho) = \zeta(\rho, \rho_0)$ and by inverting the role of ρ_0 and ρ in (66), it follows that for a convex barotropic state law:

$$\chi_0 < \zeta(\rho_0, \rho) = \zeta(\rho, \rho_0) < \chi \quad , \quad \rho_0 < \rho \tag{67}$$

where, of course, $\chi = \chi(\rho)$ according to (62);

- let $\Phi = \Phi(\rho)$ be defined as follows:

$$\Phi(\rho) := \Psi(\rho) + a(\rho) \tag{68}$$

where:

$$\Psi(\rho) := \int_{\rho_0}^{\rho} \frac{a(s)}{s}\,\mathrm{d}s \tag{69}$$

Clearly $\mathrm{d}\Phi/\mathrm{d}\rho = c(\rho)$ and therefore for a convex barotropic state law Φ is a monotonically increasing function of ρ.

Contrarily to the thermodynamic stability constraint (4), the convexity condition (61) is not imposed by physical requirements. Nevertheless, it is attractive from a mathematical point of view since it makes it possible to construct a local solution to the RP (57) by juxtaposing the wave solutions introduced in sec. 2.4.2 (i.e. rarefactions, shocks and contact discontinuities). Indeed, the convexity condition (61) renders all the pairs $(\lambda_k, \mathbf{r}_k)$ $(k \in \{1,2,3\})$ defined in (38) either g.n. or l.d., since:

$$\begin{cases} \partial_{\mathbf{q}^{(A)}}\lambda_1 \cdot \mathbf{r}_1 &= \; -c(\rho) < 0 \quad g.n. \\ \\ \partial_{\mathbf{q}^{(A)}}\lambda_2 \cdot \mathbf{r}_2 &= \; 0 \qquad\qquad\quad l.d. \\ \\ \partial_{\mathbf{q}^{(A)}}\lambda_3 \cdot \mathbf{r}_3 &= \; +c(\rho) > 0 \quad g.n. \end{cases} \tag{70}$$

and thus permits to exploit the results reported in sec. 2.4.3. A generic convex barotropic state law is therefore assumed in secs. 2.5.2 and 2.5.3, in order to construct the solution to the RP (57).

Note 12 *As an example, it is straightforward to verify that the following law:*

$$p = p_{model}(\rho) := \kappa\, \rho^{\varkappa} + \gamma \tag{71}$$

with $\kappa > 0$, $\varkappa > 0$ and γ given (real) constants, is convex. It should be noticed that:

- *for $\gamma = 0$ the classical polytropic gas state law is obtained;*
- *for $\kappa = 2^{-1}$, $\varkappa = 2$ and $\gamma = 0$ the conservation laws (27) become formally identical to the well-known homogeneous shallow water equations (of course, augmented by the advection of the passive scalar), for which there exists a vast literature also investigating RPs (see e.g. [99]);*
- *for $\kappa = \varepsilon\, \rho_0^{-\varkappa}$ and $\gamma = -\varepsilon$, with $\varepsilon > 0$ and $\rho_0 > 0$ given constants, the classical Tait law (which is used for describing isentropic compressible liquids) is formally recovered; the corresponding RP is studied, for instance, in [51].*

Note 13 *The rheological behaviour of a large variety of real-world materials cannot be represented by convex laws and therefore non-convex state laws have been studied as well, for incorporation into classical systems of equations for fluid dynamics (see e.g. [69], [113] and [114]). The solution of the RP associated with a non-convex state law is, in general, more difficult than that one associated with a convex one since it admits, besides the basic waves discussed in sec. 2.4.2, more complex wave solutions (see the aforementioned references).*

2.5.2 Basic wave solutions

As already noticed in sec. 2.3.2, the system (27) is strictly hyperbolic, with pairs of eigenvalue-eigenvector given by (38). In this section, basic k-waves (i.e. wave solutions associated with the pair $(\lambda_k, \mathbf{r}_k)$) are investigated, of the type of those discussed in sec. 2.4.2. In view of the material presented in the aforementioned section, the relation (70) clearly implies that the $2-$waves are necessarily contact discontinuities while the others can be either rarefactions or shocks.

In the rest of this section, the considered waves are supposed to separate a "left" state $\mathbf{q}_l^{(A)}$ and a "right" state $\mathbf{q}_r^{(A)}$. The subscripts l and r are also exploited for concisely representing entities related to the aforementioned states.

$1-$**rarefaction**

Let $q_h^{(A)}$ $(h \in \{1, 2, 3\})$ denote the $h-$th component of $\mathbf{q}^{(A)}$. Once recalled the eigenstructure (38) of the system at hand, the definition of the generalized Riemann invariants (53) leads to the following differential relations:

$$\begin{cases} \mathrm{d}\left(\dfrac{q_2^{(A)}}{q_1^{(A)}}\right) & + & \dfrac{a(q_1^{(A)})}{q_1^{(A)}}\,\mathrm{d}q_1^{(A)} & = & 0 \\[2ex] \dfrac{\mathrm{d}q_3^{(A)}}{q_3^{(A)}} & - & \dfrac{\mathrm{d}q_1^{(A)}}{q_1^{(A)}} & = & 0 \end{cases}$$

which integrate to:

$$\begin{cases} u_r - u_l & = & \Psi_l - \Psi_r \\[2ex] \xi_r - \xi_l & = & 0 \end{cases} \tag{72}$$

where Ψ is given by (69).

Proposition 1 *The rarefaction under consideration is a wave solution of the RP (57) if and only if:*

$$\rho_r < \rho_l \tag{73}$$

Proof *Since the pair $(\lambda_1, \mathbf{r}_1)$ is g.n., it is necessary and sufficient for the wave under consideration to be a solution that (see the relevant paragraph in sec. 2.4.2):*

$$\mathbf{q}_r^{(A)} \in R_1^+\left(\mathbf{q}_l^{(A)}\right)$$

For the present case, the aforementioned condition reads $\lambda_{1l} < \lambda_{1r}$ or, equivalently, $a_r - a_l < u_r - u_l$. Then, by substituting the first relation in (72), the inequality under consideration can be recast as follows:

$$\Phi_r < \Phi_l \tag{74}$$

with Φ defined in (68). By recalling the fact that, for a convex barotropic state law, $\Phi(\rho)$ is a monotonically increasing function, the conditions (73) and (74) are equivalent to each other. This concludes the proof. ■

1−**shock**

By introducing the relevant definitions into the RH condition (41), the following relations are obtained:

$$\begin{cases} \rho_r \bar{u}_r - \rho_l \bar{u}_l &= 0 \\ \rho_r \bar{u}_r^2 - \rho_l \bar{u}_l^2 &= p_l - p_r \\ \xi_r - \xi_l &= 0 \end{cases} \tag{75}$$

where:

$$\bar{u}_j := u_j - \sigma \quad , \quad j \in \{l, r\} \tag{76}$$

and $\sigma = \sigma(u_l, u_r)$ denotes the shock speed [17]. By manipulating the first two equations in (75) the following relations are obtained:

$$\rho_l \bar{u}_l = \rho_r \bar{u}_r = \zeta\left(\rho_l, \rho_r\right) \tag{77}$$

$$u_r - u_l = \frac{\rho_l - \rho_r}{\rho_l\, \rho_r}\, \zeta\left(\rho_l, \rho_r\right) \tag{78}$$

with ζ given by (65). Moreover, by substituting (77) into (76), the following expression is obtained for the shock speed σ:

$$\sigma = u_j - \rho_j^{-1}\, \zeta\left(\rho_l, \rho_r\right) \quad , \quad j \in \{l, r\}$$

Proposition 2 *The shock under consideration is a wave solution of the RP (57) if and only if:*

$$\rho_l < \rho_r \tag{79}$$

Proof *Since the pair $(\lambda_1, \mathbf{r}_1)$ is g.n., the LEC (55) is a necessary and sufficient condition for the shock under consideration to be admissible (see the*

[17] The first relation in (75), directly derived from the first component of (41), is exploited to obtain the representation of the others.

relevant paragraphs in sec. 2.4.2). By exploiting (77), the LEC (55) can be recast as follows:

$$\chi_l < \zeta\left(\rho_l, \rho_r\right) < \chi_r \tag{80}$$

with χ defined in (62). Then, by recalling (67), it is clear that for a convex barotropic state law the condition (80) (i.e. the LEC (55)) and the condition (79) are equivalent to each other. This concludes the proof. ■

2−**contact discontinuity**

Let $q_h^{(A)}$ $(h \in \{1, 2, 3\})$ denote the $h-$th component of $\mathbf{q}^{(A)}$. By applying (53) to the present case (see Note 11 in sec. 2.4.2), the following differential relations are obtained:

$$\left\{\begin{array}{rcl} \mathrm{d}q_1^{(A)} & = & 0 \\ \\ \mathrm{d}q_2^{(A)} & = & 0 \end{array}\right.$$

which trivially integrate to:

$$\left\{\begin{array}{rcl} \rho_r - \rho_l & = & 0 \\ \\ u_r - u_l & = & 0 \end{array}\right. \tag{81}$$

Moreover, the speed σ of the contact discontinuity is straightforwardly given by (see sec. 2.4.2):

$$\sigma = u_l = u_r$$

3−**rarefaction**

Once noticed that $\mathbf{r}_3$ reduces to $\mathbf{r}_1$ by inverting the sign of the sound speed a, (72) directly implies that across the waves under consideration the following relations hold:

$$\left\{\begin{array}{rcl} u_r - u_l & = & \Psi_r - \Psi_l \\ \\ \xi_r - \xi_l & = & 0 \end{array}\right. \tag{82}$$

where Ψ is given by (69).

Proposition 3 *The rarefaction under consideration is a wave solution of the RP (57) if and only if:*

$$\rho_l < \rho_r \tag{83}$$

Proof *Analogous to the that one of Proposition 1 above.* ■

3−shock

By introducing the relevant definitions into the RH condition (41), the following relations are obtained (identical to those in (75)):

$$\left\{ \begin{array}{rcl} \rho_r \bar{u}_r - \rho_l \bar{u}_l & = & 0 \\ \rho_r \bar{u}_r^2 - \rho_l \bar{u}_l^2 & = & p_l - p_r \\ \xi_r - \xi_l & = & 0 \end{array} \right. \tag{84}$$

where $\bar{u}_j$ $(j \in \{l, r\})$ is defined in (76). By manipulating the first two equations in (84) the following relations are obtained:

$$\rho_l \bar{u}_l = \rho_r \bar{u}_r = -\,\zeta\,(\rho_l, \rho_r) \tag{85}$$

$$u_r - u_l = \frac{\rho_r - \rho_l}{\rho_l\, \rho_r}\, \zeta\,(\rho_l, \rho_r) \tag{86}$$

with ζ given by (65) [18]. Moreover, by substituting (85) into (76), the following expression is obtained for the shock speed σ:

$$\sigma = u_j + \rho_j^{-1}\, \zeta\,(\rho_l, \rho_r) \quad , \quad j \in \{l, r\} \tag{87}$$

Proposition 4 *The shock under consideration is a wave solution of the RP (57) if and only if:*

$$\rho_r < \rho_l \tag{88}$$

Proof *Analogous to the that one of Proposition 2 above.* ■

2.5.3 Local solution of the Riemann problem

As remarked in sec. 2.5.1, by adopting a convex barotropic state law it is possible to exploit the constructive procedure outlined in sec. 2.4.3 in order to define a local solution of the RP (57). Moreover, since two generic adjacent states appearing in the solution are only connected to each other by means of a basic wave (i.e. a rarefaction, a shock or a contact discontinuity), it is possible to use the relations obtained in sec. 2.5.2, as explained below. The solution strategy outlined in this section is then exploited in secs. 3 and 4 in order to validate 1D numerical methods.

[18] The difference between (77) and (85) arises from the LEC (55). Indeed, in the former case it must be (in particular) $\lambda_{1l} = u_l - a_l > \sigma$, i.e. $\bar{u}_l > a_l$ while in the latter one it must be (in particular) $\sigma > \lambda_{3r} = u_r + a_r$, i.e. $\bar{u}_r < -\,a_r$. Since the sound speed is always positive, in the former case $\rho_j \bar{u}_j = \pi_1$ while in the latter one $\rho_j \bar{u}_j = -\,\pi_3$, with π_1 and π_3 positive entities. Straightforward computations show that $\pi_1 = \pi_3 = \zeta\,(\rho_l, \rho_r)$.

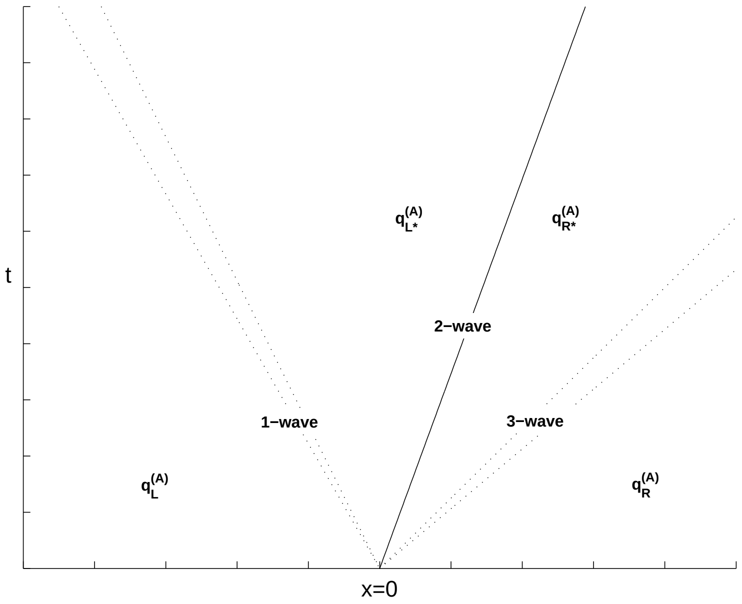

Figure 6: Schematic representation of the solution of the RP (57). The 2−wave is a contact discontinuity while the others can be either a shock or a rarefaction.

Structure of the solution

By recalling the theorem mentioned in sec. 2.4.3, it is clear that the solution of the RP (57) in general consists of three waves of the type of those discussed in sec. 2.5.2. These waves separate four states, $\mathbf{q}_L^{(A)}$, $\mathbf{q}_{L\star}^{(A)}$, $\mathbf{q}_{R\star}^{(A)}$ and $\mathbf{q}_R^{(A)}$, amongst which $\mathbf{q}_L^{(A)}$ and $\mathbf{q}_R^{(A)}$ are given by the initial condition associated with the RP (57) while the others must be determined. The structure of the wave solution is sketched in Fig. 6; in particular, the solid line represents the 2−contact discontinuity while each couple of dotted lines denotes either a shock or a rarefaction.

In consideration of (72), (75), (81), (82) and (84) it is evident that ξ only varies across the contact discontinuity while ρ and u (which are continuous across the contact discontinuity) change, in general, across the other waves. Hence, the state vectors under consideration admit the following representation:

$$\mathbf{q}_L^{(A)} = \begin{pmatrix} \rho_L \\ \rho_L u_L \\ \rho_L \xi_L \end{pmatrix} \quad , \quad \mathbf{q}_{L\star}^{(A)} = \begin{pmatrix} \rho_\star \\ \rho_\star u_\star \\ \rho_\star \xi_L \end{pmatrix}$$

$$\mathbf{q}_{R\star}^{(A)} = \begin{pmatrix} \rho_\star \\ \rho_\star u_\star \\ \rho_\star \xi_R \end{pmatrix} \quad , \quad \mathbf{q}_{R}^{(A)} = \begin{pmatrix} \rho_R \\ \rho_R u_R \\ \rho_R \xi_R \end{pmatrix}$$

where $\rho_\star$ and $u_\star$ need to be defined in order to completely determine the solution.

Determination of the solution

By combining the expressions which give the variation of u across the waves, namely (72), (78), (82) and (86), it is straightforward to express $u_\star$ as a function of $\rho_\star$ as follows:

$$u_\star = u_L - \Omega\left(\rho_\star, \rho_L\right) = u_R + \Omega\left(\rho_\star, \rho_R\right) \tag{89}$$

where (a prolongation by continuity for $\rho = \rho_j$ is considered):

$$\Omega\left(\rho, \rho_j\right) := \begin{cases} \Psi\left(\rho\right) - \Psi\left(\rho_j\right) & \text{if} \quad \rho \leq \rho_j \\ \\ \dfrac{\rho - \rho_j}{\rho\,\rho_j}\,\zeta\left(\rho, \rho_j\right) & \text{if} \quad \rho > \rho_j \end{cases} \quad , \quad j \in \{L, R\}$$

with Ψ and ζ respectively given by (69) and (65). Let Δu be defined as follows:

$$\Delta u := u_R - u_L$$

Then, in view of (89), the identity $\left(u_L - u_\star\right) + \left(u_\star - u_R\right) + \left(u_R - u_L\right) = 0$ can be recast as follows:

$$\Omega\left(\rho_\star, \rho_L\right) + \Omega\left(\rho_\star, \rho_R\right) + \Delta u = 0 \tag{90}$$

Clearly, for the relation (90) to hold, $\rho_\star$ must be a zero of the following function:

$$\Omega_{(L,R)}\left(\rho\right) := \Omega_L\left(\rho\right) + \Omega_R\left(\rho\right) + \Delta u \tag{91}$$

where:

$$\Omega_j\left(\rho\right) := \Omega\left(\rho, \rho_j\right) \quad , \quad j \in \{L, R\}$$

The existence and the uniqueness of such a zero is ensured by the following:

Proposition 5 *Let $D_\rho = [\rho_{min}, \rho_{sup})$ denote the density domain of the considered barotropic state law, as defined in sec. 1.5. There exists a unique solution $\rho_\star \in D_\rho$ of the equation (90) if and only if:*

$$\Delta^{inf} u < \Delta u \leq \Delta^{max} u \tag{92}$$

with:

$$\begin{cases} \Delta^{inf} u & := & -\dfrac{\rho_{sup} - \rho_L}{\rho_{sup}\, \rho_L}\, \zeta\left(\rho_{sup}, \rho_L\right) - \dfrac{\rho_{sup} - \rho_R}{\rho_{sup}\, \rho_R}\, \zeta\left(\rho_{sup}, \rho_R\right) \\ \\ \Delta^{max} u & := & \Psi\left(\rho_L\right) + \Psi\left(\rho_R\right) - 2\, \Psi\left(\rho_{min}\right) \end{cases}$$

Proof *The first derivative of $\Omega_j\left(\rho\right)$ is given by the following (continuous) function:*

$$\frac{\mathrm{d}}{\mathrm{d}\rho}\, \Omega_j\left(\rho\right) = \begin{cases} \dfrac{a}{\rho} & \text{if} \quad \rho \leq \rho_j \\ \\ \dfrac{1}{2}\left(\dfrac{\zeta\left(\rho, \rho_j\right)}{\rho^2} + \dfrac{a^2}{\zeta\left(\rho, \rho_j\right)}\right) & \text{if} \quad \rho > \rho_j \end{cases} \quad , \quad j \in \{L, R\}$$

which is clearly positive. Hence, $\Omega_{(L,R)}\left(\rho\right)$ in (91) is a monotonically increasing function and admits a unique zero, which moves towards lower values of the density as Δu increases. By continuity, there exists a maximum value of Δu, denoted by $\Delta^{max} u$, for which $\rho_\star = \rho_{min}$ as well as an inferior one, denoted by $\Delta^{inf} u$, in correspondence of which $\rho_\star = \rho_{sup}$. Clearly, $\Delta^{max} u$ can be determined by evaluating (90) in correspondence of $\rho = \rho_{min}$. In particular, since $\rho_{min} \leq \rho_j$, $j \in \{L, R\}$, it follows (from the definitions) that:

$$\Omega_L\left(\rho_{min}\right) + \Omega_R\left(\rho_{min}\right) = 2\, \Psi\left(\rho_{min}\right) - \Psi\left(\rho_L\right) - \Psi\left(\rho_R\right)$$

and therefore (90) in the present case reads:

$$2\, \Psi\left(\rho_{min}\right) - \Psi\left(\rho_L\right) - \Psi\left(\rho_R\right) + \Delta^{max} u = 0$$

Similar considerations can be exploited for deriving the expression of $\Delta^{inf} u$. It is evident that (92) represents a necessary and sufficient condition for determining a solution $\rho_\star \in D_\rho$ of the non-linear equation (90). This completes the proof. ■

Note 14 *As Δu transitions between $\Delta^{inf} u$ and $\Delta^{max} u$, the wave structure of the solution of the RP (57) changes. Once introduced the following definitions:*

$$\begin{cases} \Delta^{2s} u & := & -\dfrac{\rho_M - \rho_m}{\rho_M\, \rho_m}\, \zeta\left(\rho_M, \rho_m\right) \\ \\ \Delta^{2r} u & := & \Psi\left(\rho_M\right) - \Psi\left(\rho_m\right) \end{cases}$$

where:

$$\rho_M := \max(\rho_L, \rho_R) \quad , \quad \rho_m := \min(\rho_L, \rho_R)$$

and by recalling the monotonicity of $\Omega_{(L,R)}$*, it is possible to identify the following sequence of wave solutions:*

- *for* $\Delta^{inf}u < \Delta u < \Delta^{2s}u$, $\rho_M < \rho_\star < \rho_{sup}$ *and both the* $1-wave$ *and the* $3-wave$ *are shocks;*
- *for* $\Delta^{2s}u \leq \Delta u < \Delta^{2r}u$, $\rho_m < \rho_\star \leq \rho_M$*; there is a shock between* $\rho_\star$ *and* ρ_m*, and a rarefaction between* $\rho_\star$ *and* ρ_M*;*
- *for* $\Delta^{2r}u \leq \Delta u \leq \Delta^{max}u$, $\rho_{min} \leq \rho_\star \leq \rho_m$ *and both the* $1-wave$ *and the* $3-wave$ *are rarefactions.*

The aforementioned statements can be straightforwardly verified by introducing considerations of the same kind of those reported in the proof of Proposition 5 (above) for determining $\Delta^{max}u$*.*

Note 15 *The solution for* ξ *depends on* $\rho_\star$ *and* $u_\star$ *(since* $u_\star$ *determines the location of the contact discontinuity) 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 (see Note 4 in sec. 2.2.4), permits to straightforwardly extend the structure of the considered solution to the case of an arbitrary number* $m > 1$ *of passive scalars. Indeed, since there is no interaction between them (see Note 5 in sec. 2.2.4), it suffices to make all of them simultaneously jump across the contact discontinuity. Even if for* $m > 1$ *the system (27) ceases to be strictly hyperbolic (the multiplicity of the eigenvalue associated with the contact discontinuity being in general equal to* m*), the fact that the additional waves do not interact with the starting system, neither with one another even, makes the loss of strict hyperbolicity purely formal. In other words, the augmented system with* $m > 1$ *behaves like that one having* $m = 1$ *and it is possible to keep the proposed solution strategy.*

Note 16 *Clearly, the solution of the considered RP is essentially constructed by solving (90). It is therefore evident that it is possible to keep the proposed solution procedure also for formulations adopting* p *instead of* ρ *as the independent variable, at the only cost of straightforward changes in the notation.*

Note 17 *It is worth mentioning that the material presented in secs. 2.5.2 and 2.5.3 generalizes the solution procedure reported in [99] for the RP associated with the homogeneous shallow water equations (see Note 12 in sec. 2.5.1).*

3 1D Numerical method

A 1D numerical method (hereafter also referred to as numerical scheme) for discretizing the augmented-1D equations (27) is developed in the present section. In particular, the integral form of the considered conservation law, namely (26), is considered in order to allow integral solutions (see sec. 2.3.3) to be taken into account. A shock-capturing approach (see e.g. [39], [64], [80] and [98]) is chosen, in order to approximate possibly discontinuous solutions.

Once introduced some basic material in sec. 3.1, a Godunov scheme for (generic) convex barotropic state laws is proposed in sec. 3.2. A Roe scheme is then proposed in sec. 3.3, which can be applied when dealing with generic barotropic state laws. In sec. 3.4 the behaviour of this scheme in the nearly-incompressible limit is investigated and a suitable preconditioning technique for low Mach number flows is introduced. Finally, in sec. 3.5 a linearization of a generic Roe numerical flux function is proposed, only relying on its algebraic properties and therefore applicable to a variety of problems. The proposed linearization is then applied to the barotropic case under consideration, in order to define a linearized implicit time-advancing strategy.

3.1 Generalities on the 1D discretization

In the sequel, some basic concepts and definitions which are related to the numerical discretization of the considered 1D conservation law are introduced, to be exploited within the rest of sec. 3. The concise introduction under consideration does not lay claim to yield a rigorous and complete treatment of the subject; a detailed presentation can be found in a number of textbooks (e.g [39], [64] and [98] amongst many others).

3.1.1 Space discretization

A finite volume approach is adopted for the spatial discretization of the problem (27). The $x-$domain is divided into N_c cells, indexed by $i \in \mathcal{I} := \{1, \dots, N_c\}$. The $i-$th cell spans the interval $C_i := (x_{i-1/2}, x_{i+1/2})$, with $x_{i-1/2} < x_{i+1/2}$, having measure μ_i. On C_i the exact solution $\mathbf{q}^{(A)}(x,t)$ is approximated by a semi-discrete function $\mathbf{q}_i^{(A)}(t)$, which is considered as an approximation of the mean value of $\mathbf{q}^{(A)}(x,t)$ over C_i:

$$\mathbf{q}_i^{(A)}(t) \approx \frac{1}{\mu_i} \int_{C_i} \mathbf{q}^{(A)}(x,t)\,\mathrm{d}x \tag{93}$$

The differential system defining $\mathbf{q}_i^{(A)}$ is obtained by discretizing the integral balance (26) over the control volume C_i. Indeed, by virtue of (93), the time-derivative in (26) is naturally approximated as follows:

$$\frac{\partial}{\partial t}\int_{C_i}\mathbf{q}^{(A)}(x,t)\,\mathrm{d}x \approx \mu_i\,\frac{\mathrm{d}}{\mathrm{d}t}\,\mathbf{q}_i^{(A)} \tag{94}$$

while the inter-cell flux $\mathbf{f}^{(A)}$ defined in (25) is approximated by introducing a suitable numerical flux function (hereafter numerical flux, as well) $\boldsymbol{\phi}^{(A)}$, depending on the semi-discrete solution.

Numerical flux

Let π_i denote the set of indexes identifying the cells in the neighbourhood of C_i, namely:

$$\pi_i := \{i-1, i+1\} \tag{95}$$

A certain degree of locality is usually assumed for $\boldsymbol{\phi}^{(A)}$ and the flux crossing the boundary between C_i and C_j towards C_j is commonly approximated by means of the following expression:

$$\boldsymbol{\phi}^{(A)}\left(\mathbf{q}_i^{(A)},\mathbf{q}_j^{(A)},\hat{\boldsymbol{\nu}}_{ij}\right) \quad , \quad j\in\pi_i \tag{96}$$

where $\hat{\boldsymbol{\nu}}_{ij}$ is the versor associated with a generic vector mapping an arbitrary point of C_i to an arbitrary point of C_j. If $\hat{\boldsymbol{e}}$ denotes the versor associated with the $x-$axis, then clearly:

$$\hat{\boldsymbol{\nu}}_{ij} = (j-i)\,\hat{\boldsymbol{e}} \quad , \quad j\in\pi_i \tag{97}$$

The following definition is consequently introduced, to be exploited in the sequel:

$$s_{ij} := \hat{\boldsymbol{\nu}}_{ij}\cdot\hat{\boldsymbol{e}} = \mathrm{sign}(j-i) \quad , \quad j\in\pi_i \tag{98}$$

Note 18 *The explicit dependence of $\boldsymbol{\phi}^{(A)}$ on $\hat{\boldsymbol{\nu}}_{ij}$ in (96) may seem somewhat redundant in the present context. Indeed, a 1D case is intrinsically structured: any internal cell C_i has two and only two neighbours, C_{i-1} and C_{i+1}, and $\hat{\boldsymbol{\nu}}_{ij} = \pm\hat{\boldsymbol{e}}$ according to (97). However, the proposed formulation (96), allows for an extension to 2D and 3D -possibly unstructured- spatial discretizations to be obtained (see e.g. sec. 5.1.2), since it does not a priori incorporate any structure.*

In general, the numerical flux must satisfy the following basic requirements:

$$\boldsymbol{\phi}^{(A)}\left(\mathbf{q}_j^{(A)},\mathbf{q}_i^{(A)},\hat{\boldsymbol{\nu}}_{ji}=-\,\hat{\boldsymbol{\nu}}_{ij}\right) \;=\; -\,\boldsymbol{\phi}^{(A)}\left(\mathbf{q}_i^{(A)},\mathbf{q}_j^{(A)},\hat{\boldsymbol{\nu}}_{ij}\right) \tag{99}$$

$$\boldsymbol{\phi}^{(A)}\left(\mathbf{q}_i^{(A)},\mathbf{q}_j^{(A)}=\mathbf{q}_i^{(A)},\hat{\boldsymbol{\nu}}_{ij}\right) \;=\; s_{ij}\;\mathbf{f}^{(A)}\left(\mathbf{q}_i^{(A)}\right) \tag{100}$$

The property (99) is directly inherited from the continuous flux; it permits to associate the numerical flux with the inter-cell boundary in a well-defined way, thus allowing for the definition of "conservative" numerical schemes (see Note 19 in sec. 3.1.2). The property (100), instead, enforces a natural consistency requirement.

Godunov approach

By following the well-known approach originally proposed by Godunov [40], the piece-wise constant approximant $\mathbf{q}_i^{(A)}$ $(i \in \mathcal{I})$ can be considered as defining a local RP like (57) at each interface $x_{(i+j)/2}$ $(j \in \pi_i)$. The numerical flux between C_i and C_j can therefore be constructed by properly exploiting the solution either of the RP of interest (see sec. 3.2.1) or of a suitable approximation of it (see sec. 3.3.1).

Semi-discrete formulation and boundary conditions

By exploiting (94) and (96) it is straightforward to obtain the following class of semi-discrete approximations of (26), depending on the specific choice of the numerical flux:

$$\mu_i \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{q}_i^{(A)} + \sum_{j \in \pi_i} \boldsymbol{\phi}^{(A)} \left(\mathbf{q}_i^{(A)}, \mathbf{q}_j^{(A)}, \hat{\boldsymbol{\nu}}_{ij} \right) = \mathbf{0} \quad , \quad i \in \mathcal{I} \tag{101}$$

At the present stage of the discussion, the scheme (101) is not properly defined for $i = 1$ and $i = N_c$: suitable boundary conditions (hereafter BCs as well) must be introduced in order to completely define the semi-discrete formulation. To the purpose, two fictitious state vectors, $\mathbf{q}_0^{(A)}$ and $\mathbf{q}_{N_c+1}^{(A)}$, are introduced. Once these vectors have been given a value (modelling the chosen BCs), it is possible to directly apply (101) to every cell C_i $(i \in \mathcal{I})$.

3.1.2 Time discretization: basic discrete schemes

A fully-discrete (hereafter discrete) numerical scheme approximating (26) is defined by considering (101) as an ordinary differential equation [19]. The discrete solution at time-level $n+1$ (corresponding to $t = t^{n+1}$), denoted by $\mathbf{q}_i^{(A)n+1}$ within cell C_i, can therefore be obtained from that one at time-level n by exploiting a variety of integration techniques. Basic discrete schemes are presented below.

[19] This approach, keeping space and time discretizations separate, is sometimes referred to as a "method of lines" (see e.g. [39]).

Explicit time-advancing

An explicit discrete scheme can be obtained from (101) by considering, for instance, the classical "forward Euler" integration technique (see e.g. [79]):

$$\mathbf{q}_i^{(A)n+1} = \mathbf{q}_i^{(A)n} - \frac{\delta^n t}{\mu_i} \sum_{j \in \pi_i} \boldsymbol{\phi}^{(A)} \left(\mathbf{q}_i^{(A)n}, \mathbf{q}_j^{(A)n}, \hat{\boldsymbol{\nu}}_{ij} \right) \quad , \quad i \in \mathcal{I} \tag{102}$$

where:

$$\delta^n(\cdot) := (\cdot)^{n+1} - (\cdot)^n \tag{103}$$

Note 19 *Due to its specific form, the scheme (102) is said to be "conservative" (see e.g. [39] and [98]). It is very important to exploit conservative schemes in order to compute possibly discontinuous integral solutions; indeed, non-conservative formulations do not converge* [20] *to the correct solution of the problem if it involves shock waves [48]. A classical result, on the other hand, states that conservative numerical methods, if convergent, do converge to an integral solution of the considered conservation law [62]. Hence, it is practically compulsory to exploit conservative schemes when adopting a shock-capturing numerical approach* [21]*.*

Implicit time-advancing

An implicit discrete scheme can be obtained from (101) approximating the time derivative by means of a backward finite difference as follows:

$$\frac{\mu_i}{\delta^n t}\, \delta^n \mathbf{q}_i^{(A)} + \sum_{j \in \pi_i} \boldsymbol{\phi}^{(A)} \left(\mathbf{q}_i^{(A)n+1}, \mathbf{q}_j^{(A)n+1}, \hat{\boldsymbol{\nu}}_{ij} \right) = \mathbf{0} \quad , \quad i \in \mathcal{I} \tag{104}$$

It is well known that implicit schemes like (104) permit a more efficient time-advancing than explicit ones, because they do not suffer from time-step limitations caused by CFL-like stability constraints (see e.g. [44]). However, the scheme (104) can be demanding from a computational point of view, since it requires the solution of a non-linear system at each time-level. Indeed, the flux function is, in general, non-linear and the specific form of the considered state law can add to the complexity of the algorithm. As a matter of fact, 3D numerical schemes based on the extension of (104) are exceedingly

[20] The notion of convergence, even if not formally introduced, is assumed to be understood at this point of the discussion.

[21] Different choices can be considered when adopting other techniques (e.g. shock-fitting or adaptive primitive-conservative numerical methods) [98].

intensive from a computational point of view [22], especially if they are applied to industrial problems involving very complex geometries.

Linearized implicit time-advancing

In view of the above considerations, a reasonable compromise between a purely explicit and a purely implicit scheme seems to be provided by a linearized implicit time-advancing strategy. This technique is based on the following approximate linearization of the numerical flux (which, in general, is not differentiable), assuming it exists:

$$\delta^n \boldsymbol{\phi}_{ij}^{(A)} \approx \mathbf{A}_{ij}^{(A)n} \cdot \delta^n \mathbf{q}_i^{(A)} + \mathbf{B}_{ij}^{(A)n} \cdot \delta^n \mathbf{q}_j^{(A)} \tag{105}$$

where:

$$\boldsymbol{\phi}_{ij}^{(A)n} := \boldsymbol{\phi}^{(A)} \left(\mathbf{q}_i^{(A)n}, \mathbf{q}_j^{(A)n}, \hat{\boldsymbol{\nu}}_{ij} \right)$$

and:

$$\left\{ \begin{array}{lcl} \mathbf{A}_{ij}^{(A)n} & := & \mathbf{A}^{(A)} \left(\mathbf{q}_i^{(A)n}, \mathbf{q}_j^{(A)n}, \hat{\boldsymbol{\nu}}_{ij} \right) \\ & & \\ \mathbf{B}_{ij}^{(A)n} & := & \mathbf{B}^{(A)} \left(\mathbf{q}_i^{(A)n}, \mathbf{q}_j^{(A)n}, \hat{\boldsymbol{\nu}}_{ij} \right) \end{array} \right. \tag{106}$$

with $\mathbf{A}^{(A)}$ and $\mathbf{B}^{(A)}$ suitably defined matrices. By substituting (105) into (104), the following scheme is obtained:

$$\left(\frac{\mu_i}{\delta^n t} \mathbf{I} + \sum_{j \in \pi_i} \mathbf{A}_{ij}^{(A)n} \right) \cdot \delta^n \mathbf{q}_i^{(A)} + \sum_{j \in \pi_i} \mathbf{B}_{ij}^{(A)n} \cdot \delta^n \mathbf{q}_j^{(A)} = - \sum_{j \in \pi_i} \boldsymbol{\phi}_{ij}^{(A)n} \quad , \quad i \in \mathcal{I} \tag{107}$$

The scheme (107) represents a linear system (in particular, a block tridiagonal system in the considered 1D case) for the unknowns $\delta^n \mathbf{q}_i^{(A)}$; once it has been solved, the unknowns at time-level $n+1$ are trivially given by $\mathbf{q}_i^{(A)n+1} = \mathbf{q}_i^{(A)n} + \delta^n \mathbf{q}_i^{(A)}$. Clearly, the linearized scheme (107) involves an additional degree of approximation with respect to the implicit scheme (104) (due to the approximate linearization [23]) but it is less demanding from a computational point of view. For this reason, a linearized implicit time-advancing is proposed in sec. 3.5.

[22] Unless exploiting specific supercomputing resources which are usually not available for common research or even industrial projects.

[23] Of course, the effects the approximate linearization has on the numerical solution may be relatively less important for simulations marching towards a steady-state.

The scheme (107) can be regarded to as a particular instance of a more general linearized implicit formulation which is described below. The semi-discrete formulation (101) can be rewritten, in a more general way, as follows:

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathbf{z}_h + \boldsymbol{\psi}_h^{(p)}\left(\mathbf{z}_h\right) = \mathbf{0} \tag{108}$$

where $\mathbf{z}_h$ denotes a suitable state vector representing the semi-discrete solution and $\boldsymbol{\psi}_h^{(p)}(\cdot)$ denotes a vector operator whose components are spatial difference operators (the superscript p is discussed below). It is possible to approximate the time derivative as follows:

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathbf{z}_h\left(t^{n+1}\right) \approx \frac{\alpha_k\,\mathbf{z}_h^{n+1} - \mathbf{z}_h^{(n,k)}}{\delta^n t} \tag{109}$$

where $\mathbf{z}_h^{(n,k)}$ denotes a linear combination of $\mathbf{z}_h^n, \ldots, \mathbf{z}_h^{n+1-k}$ and α_k is a scalar. Clearly, by suitably defining α_k and $\mathbf{z}_h^{(n,k)}$ it is possible to obtain a certain order of accuracy (in the sense of the finite difference schemes) for the approximation of the time derivative. Once the approximation (109) has been substituted into the semi-discrete scheme (108), its discrete counterpart reads:

$$\boldsymbol{\mu}_h^{(p,k)}\left(\mathbf{z}_h^{n+1}\right) = \mathbf{0} \tag{110}$$

where the operator $\boldsymbol{\mu}_h^{(p,k)}(\cdot)$ is defined as follows:

$$\boldsymbol{\mu}_h^{(p,k)}\left(\bar{\mathbf{z}}_h\right) := \frac{\alpha_k}{\delta^n t}\,\bar{\mathbf{z}}_h + \boldsymbol{\psi}_h^{(p)}\left(\bar{\mathbf{z}}_h\right) - \frac{1}{\delta^n t}\,\mathbf{z}_h^{(n,k)} \tag{111}$$

The non-linear discrete problem (110)-(111) can then be solved by means of a variety of techniques. In particular, it is possible to:

(i) introduce an approximate linearization of the operator $\boldsymbol{\mu}_h^{(p,k)}$, as follows:

$$\boldsymbol{\mu}_h^{(p,k)}\left(\mathbf{z}_h^{n+1}\right) - \boldsymbol{\mu}_h^{(p,k)}\left(\mathbf{z}_h^{n}\right) \approx \mathbf{J}_h^{(p,k)}\left(\mathbf{z}_h^{n}\right)\cdot\delta^n\mathbf{z}_h \tag{112}$$

where:

$$\mathbf{J}_h^{(p,k)}(\cdot) := \frac{\alpha_k}{\delta^n t}\mathbf{I} + \delta\boldsymbol{\psi}_h^{(p)}(\cdot) \tag{113}$$

and $\delta\boldsymbol{\psi}_h^{(p)}(\cdot)$ denotes an approximation of the Jacobian of $\boldsymbol{\psi}_h^{(p)}(\cdot)$ or, more in general, a term rendering the approximation (112) acceptable. Once (112) has been substituted into (110), it is straightforward to solve the resulting linear problem with respect to $\delta^n\mathbf{z}_h$ (the discrete scheme (107), for instance, is obtained by following this approach). Of course, the computational cost of the considered strategy is mainly determined by the inversion of the linear operator $\mathbf{J}_h^{(p,k)}$, which can be still demanding for complex 3D industrial problems. When dealing with structured grids, for instance, it is possible to contain the computational cost by applying an approximate factorization technique to $\mathbf{J}_h^{(p,k)}$ [11]; indeed, the introduction of an additional discretization error due to the factorization may be acceptable, in consideration of the fact that simpler linear systems must be solved.

(ii) iteratively solve (110) by determining a fixed-point of the following relation, which implicitly maps $\mathbf{z}_h^{\lambda}$ to $\mathbf{z}_h^{\lambda+1}$ (with $\mathbf{z}_h^{\lambda=0} = \mathbf{z}_h^n$ as starting point) [24]:

$$\boldsymbol{\mu}_h^{(q,k)} \left(\mathbf{z}_h^{\lambda+1}\right) = \boldsymbol{\mu}_h^{(q,k)} \left(\mathbf{z}_h^{\lambda}\right) - \boldsymbol{\mu}_h^{(p,k)} \left(\mathbf{z}_h^{\lambda}\right) \tag{114}$$

where $\boldsymbol{\mu}_h^{(q,k)}$ is formally defined by (111), with q in place of p. Once introduced the following linearization (in the spirit of (112)):

$$\boldsymbol{\mu}_h^{(q,k)} \left(\mathbf{z}_h^{\lambda+1}\right) - \boldsymbol{\mu}_h^{(q,k)} \left(\mathbf{z}_h^{\lambda}\right) \approx \mathbf{J}_h^{(q,k)} \left(\mathbf{z}_h^{\lambda}\right) \cdot \delta^{\lambda}\mathbf{z}_h$$

the map (114) can be explicitly approximated as follows:

$$\mathbf{z}_h^{\lambda+1} \approx \mathbf{z}_h^{\lambda} - \left(\mathbf{J}_h^{(q,k)} \left(\mathbf{z}_h^{\lambda}\right)\right)^{-1} \cdot \left(\boldsymbol{\mu}_h^{(p,k)} \left(\mathbf{z}_h^{\lambda}\right)\right) \tag{115}$$

Obviously, for the considered strategy to be computationally attractive with respect to that one described in the point (i) above, the inversion of $\mathbf{J}_h^{(q,k)}$ must be cheaper than that one of $\mathbf{J}_h^{(p,k)}$ [25]. In particular, as for the point (i) above, it is possible to contain the considered computational cost when dealing with structured grids by applying an approximate factorization technique to $\mathbf{J}_h^{(q,k)}$ [11]. For practical purposes, the solution is advanced only for a limited number, say λ_{max}^n, of iterations and the considered scheme globally reads:

$$\left\{\begin{array}{lcl} \mathbf{z}_h^{\lambda=0} & = & \mathbf{z}_h^n \\ \mathbf{z}_h^{\lambda+1} & = & \mathbf{z}_h^{\lambda} - \left(\mathbf{J}_h^{(q,k)} \left(\mathbf{z}_h^{\lambda}\right)\right)^{-1} \cdot \left(\boldsymbol{\mu}_h^{(p,k)} \left(\mathbf{z}_h^{\lambda}\right)\right) \quad \lambda = 0, \ldots, (\lambda_{max}^n - 1) \\ \mathbf{z}_h^{n+1} & = & \mathbf{z}_h^{\lambda=\lambda_{max}^n} \end{array}\right. \tag{116}$$

[24] The considered iterations advance the solution with respect to the counter λ and the fixed-point $\mathbf{z}_h^{\lambda=\bar{\lambda}}$ represents the discrete solution at time-level $n+1$ (i.e. $\mathbf{z}_h^{n+1}$) independently of the specific value of $\bar{\lambda}$. In order to emphasize this concept, the iterations at hand are sometimes referred to as "internal iterations" or "pseudo-iterations".

[25] Typically, $\mathbf{J}_h^{(q,k)}$ is sparser than $\mathbf{J}_h^{(p,k)}$.

If the superscripts p and q denote a formal order of accuracy of the corresponding spatial operators, it suffices, in general, to choose $q < p$ for making the inversion of $\mathbf{J}_h^{(q,k)}$ cheaper than that one of $\mathbf{J}_h^{(p,k)}$. Let k denote the formal order of accuracy of the discretization of the time derivative in (109); in particular, let $k = p$ for the sake of simplicity. Under these assumptions, the solution of the discrete scheme (110) as well as the fixed-point solution of (114) are formally of order p. On the other hand, the solution obtained by taking a single step of (116) is of order (q, p) (i.e. of order $q < p$ for the space discretization and of order p for the time discretization). Nevertheless, it is possible to recover a $p-$order accuracy within a certain number of time-steps, without fully converging to the fixed-point solution. This consideration, which is at the basis of the "Defect Correction" methods (see e.g. [67]), renders the iterative scheme (116) appealing from a computational point of view. The definition of a suitable DeC scheme is mentioned in sec. 3.5.3;

(iii) adopt a dual time-stepping approach [52], according to which the solution $\mathbf{z}_h^{n+1}$ in (110) is obtained by advancing the following problem:

$$\left\{ \begin{array}{rcl} \dfrac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{y}_h + \boldsymbol{\mu}_h^{(p,k)}\left(\mathbf{y}_h\right) & = & \mathbf{0} \\ \\ \mathbf{y}_h\left(\tau = 0\right) & = & \mathbf{z}_h^n \end{array} \right. \tag{117}$$

with respect to the pseudo-time τ, up to a steady-state. Of course, the numerical scheme specifically adopted for discretizing the pseudo-time derivative in (117) characterizes the considered "artificial" evolution between time-level n and time-level $n + 1$.

3.2 A Godunov scheme for convex barotropic state laws

In sec. 3.2.1 a Godunov numerical flux function applicable to generic convex barotropic state laws is defined. In sec. 3.2.2, the scheme (102) exploiting the considered numerical flux is validated against an exact solution.

3.2.1 Godunov numerical flux

Let

$$\mathbf{q}_{RP}^{(A)}\left(\mathbf{q}_L^{(A)},\mathbf{q}_R^{(A)},\zeta\right)$$

denote the solution of a RP having "left" and "right" initial states $\mathbf{q}_L^{(A)}$ and $\mathbf{q}_R^{(A)}$, respectively, in correspondence of $x/t=\zeta$. The Godunov numerical flux at the interface $x_{(i+j)/2}$, with $j\in\pi_i$ and π_i given by (95), is constructed by evaluating the analytical flux $\mathbf{f}^{(A)}$ in correspondence of the following state vector:

$$\mathbf{q}_{RP}^{(A)}\left(\mathbf{q}_{L_{ij}}^{(A)},\mathbf{q}_{R_{ij}}^{(A)},0\right)$$

where:

$$L_{ij}:=\min(i,j)\quad,\quad R_{ij}:=\max(i,j)$$

and $\zeta=0$ (i.e. $x=0$) is chosen for correctly picking out the considered interface with respect to the local $x-$coordinate system to which the initial states are referred. The orientation defined by $\hat{\boldsymbol{\nu}}_{ij}$ is straightforwardly taken into account by defining the Godunov numerical flux as follows:

$$\boldsymbol{\phi}^{(A)GOD}\left(\mathbf{q}_i^{(A)},\mathbf{q}_j^{(A)},\hat{\boldsymbol{\nu}}_{ij}\right):=s_{ij}\,\mathbf{f}^{(A)}\left(\,\mathbf{q}_{RP}^{(A)}\left(\mathbf{q}_{L_{ij}}^{(A)},\mathbf{q}_{R_{ij}}^{(A)},0\right)\,\right)\quad,\quad j\in\pi_i \tag{118}$$

with s_{ij} given by (98).

Note 20 *It should be noticed that the numerical flux (118) satisfies the conservation property (99); indeed, $L_{ji}=L_{ij}$ and $R_{ji}=R_{ij}$, while $s_{ji}=-s_{ij}$. The consistency property (100) is satisfied as well, by virtue of the consistency of $\mathbf{q}_{RP}^{(A)}(\cdot,\cdot,\cdot)$ which does not perturb a uniform (trivial) initial condition.*

For the case of a generic convex barotropic state law (see sec. 2.5.1), it is possible to define the numerical flux (118) by exploiting the solution $\mathbf{q}_{RP}^{(A)}(\cdot,\cdot,\cdot)$ proposed in sec. 2.5.

Benchmark	κ	$\varkappa$	γ	ρ_L	u_L	ξ_L	ρ_R	u_R	ξ_R	t_{eval}
B1	10^6	1	0	1.02	10	2	1	20	4	1
B2	10^6	2	0	1.02	10	2	1	20	4	1

Table 1: Considered benchmarks.

3.2.2 Numerical results

The solution of a chosen RP is considered as a quantitative benchmark for validating the discrete scheme (102), based on the proposed numerical flux (118).

Benchmarks

The considered benchmarks are summarized in Tab. 1. In this table, κ, $\varkappa$ and γ refer to the chosen convex state law (71), ρ_L, u_L, ξ_L, ρ_R, u_R and ξ_R characterize the initial condition (hereafter IC as well) associated with the RP and t_{eval} denotes the time at which the considered solution is picked. The instances of the convex state law (71) incorporated in Tab. 1 are simple power laws which permit to control the characteristic sound speed and the wave structure (hence, the flow compressibility) of the solution to the considered RP, by tuning the IC. In particular:

- for the state law considered in the benchmark B1, the sound speed is constant: $a = \sqrt{\kappa} = 10^3$ and therefore $\tilde{a} = 10^3$ represents the characteristic sound speed of the flow. Once chosen $\rho_R = 1$, $u_L = \tilde{M}\,\tilde{a}$ with $\tilde{M} = 10^{-2}$ and $u_R = 2\,u_L$, the value of ρ_L is tuned so as to obtain a left rarefaction and a right shock, with $u_\star/\tilde{a} = \mathrm{O}\left(\tilde{M}\right)$ (in particular $\rho_L/\rho_R = 1 + \mathrm{O}\left(\tilde{M}\right)$). Hence, $\tilde{M} = 10^{-2}$ represents a characteristic Mach number for the whole flow field under consideration;

- for the state law considered in the benchmark B2, the sound speed varies with the density as follows: $a = \sqrt{2\kappa\rho}$. Once chosen $\rho_R = 1$, $u_L = \zeta\, a_R/\sqrt{2}$ with $\zeta = 10^{-2}$ and $u_R = 2\,u_L$, the value of ρ_L is tuned so as to obtain a left rarefaction and a right shock, with $u_\star/a_R = \mathrm{O}\left(\zeta\right)$ (in particular $\rho_L/\rho_R = 1 + \mathrm{O}\left(\zeta\right)$). In the present case, the characteristic sound speed and the characteristic Mach number of the flow are $\tilde{a} = a_R$ and $\tilde{M} = \zeta$ (in particular, $\tilde{M} = 10^{-2}$ as for the benchmark B1).

Note 21 *The structure of the solution to the considered RPs (i.e. rarefaction, contact discontinuity and shock wave) is that one of the classical "Sod test-case" [95]. Usually (see e.g. [98] and [99]), the data are chosen for the Sod test-case so as to get a "sonic rarefaction" (i.e. a rarefaction for which* $\|\mathbf{u}\| = a$ *along a certain characteristic line, see sec. 2.4.2), since this wave is a representative benchmark for evaluating the entropic behaviour of a considered numerical scheme (see e.g. Note 32 in sec. 3.3.1). Nevertheless, no sonic rarefactions are present in the solution of the considered benchmarks, since a first target of the present work is the simulation of non-cavitating, nearly-incompressible, liquid flows (see sec. 1.6) in which sonic conditions can not take place. Conversely, it is of interest here to investigate the behaviour of numerical schemes dealing with low Mach number flows (e.g.* $\tilde{M} = \mathrm{O}\left(10^{-3}\right) \div \mathrm{O}\left(10^{-2}\right)$*), like those considered in the aforementioned benchmarks. However, it must be remarked that the application of the proposed numerical techniques is by no means restricted to low Mach number flows.*

Initial and boundary conditions

The initial discontinuity of the considered RP is located at $x = 0$. Moreover, the space discretization is built in such a way that the right boundary of the cell $C_{\bar{s}}$ ($\bar{s} \in \mathcal{I}$, $\bar{s} < N_c$) is systematically located at $x = 0$. Hence the following IC is directly derived from that one of the considered RP:

$$\mathbf{q}_i^{(A)0} := \begin{cases} \mathbf{q}_L^{(A)} & i = 1, \ldots, \bar{s} \\ \mathbf{q}_R^{(A)} & i = \bar{s}+1, \ldots, N_c \end{cases} \tag{119}$$

As far as the BCs are concerned, transmissive conditions are chosen (see e.g. [98]), obtained by defining the fictitious state vectors $\mathbf{q}_0^{(A)n}$ and $\mathbf{q}_{N_c+1}^{(A)n}$ (introduced in the relevant paragraph of sec. 3.1.1) as follows:

$$\mathbf{q}_0^{(A)n} = \mathbf{q}_1^{(A)n} \quad , \quad \mathbf{q}_{N_c+1}^{(A)n} = \mathbf{q}_{N_c}^{(A)n} \quad , \quad n = 0, 1, 2, \ldots \tag{120}$$

Test-cases

For the sake of simplicity, a uniform space discretization as well as a constant time-step is adopted in (102), namely:

$$\begin{array}{rcl} \mu_i & = & \mu \quad , \quad i \in \mathcal{I} \\ \delta^n t & = & \tau \quad , \quad n = 0, 1, 2, \ldots \end{array}$$

Then, the following CFL-like stability constraint should be enforced when adopting the basic explicit scheme under consideration:

$$\tau \leq c^{(CFL)} \frac{\mu}{s^n_{max}} \quad , \quad n = 0, 1, 2, \ldots \tag{121}$$

where:

- s^n_{max} represents the largest wave speed present throughout the computational domain at time-level n;
- $c^{(CFL)}$ denotes a suitable safety coefficient. A possible choice, originally proposed by Godunov, is $c^{(CFL)} = 0.5$, which prevents any wave interaction from taking place within the generic cell C_i. This choice seems to be a little bit strict and a coefficient $0 < c^{(CFL)} \leq 1.0$ is commonly adopted, by assuming that no wave acceleration occurs as a consequence of wave interaction (see e.g. [98]).

For the sake of simplicity, both μ and τ are chosen at the beginning of the simulation and the CFL condition (121) is only checked during the simulation, at each time-level (in particular, s^n_{max} is exactly computed by exploiting the relevant relations introduced in sec. 2.5.2). The considered test-cases are summarized in Tab. 2, where n_L and n_R respectively represent the number of cells introduced within the "left" and "right" sub-domains (i.e $n_L = \bar{s}$ and $n_R = N_c - \bar{s}$, with $\bar{s}$ appearing in (119)); the corresponding numerical solutions are shown in Figs. 7-14.

Test-case	Benchmark	μ	(n_L, n_R)	τ
EG1-1	B1	100	$(2,2)\cdot 10^1$	$5\cdot 10^{-2}$
EG1-2	B1	10	$(2,2)\cdot 10^2$	$5\cdot 10^{-3}$
EG1-3	B1	1	$(2,2)\cdot 10^3$	$5\cdot 10^{-4}$
EG1-4	B1	0.1	$(2,2)\cdot 10^4$	$5\cdot 10^{-5}$
EG2-1	B2	100	$(2,2)\cdot 10^1$	$5\cdot 10^{-2}$
EG2-2	B2	10	$(2,2)\cdot 10^2$	$5\cdot 10^{-3}$
EG2-3	B2	1	$(2,2)\cdot 10^3$	$5\cdot 10^{-4}$
EG2-4	B2	0.1	$(2,2)\cdot 10^4$	$5\cdot 10^{-5}$

Table 2: Considered test-cases for the discrete scheme (102), based on the numerical flux (118).

It should be noticed that the left rarefaction appearing in the aforementioned figures is very steep. Such a behaviour is typical of low Mach number flows (see e.g. [49]) and can be justified as follows. For the considered flows, the head of the left rarefaction [26] travels with a speed $u_L - a_L \approx u_L - \tilde{a} \approx -\tilde{a}$ where $\tilde{a}$ is the characteristic sound speed ($\tilde{a} \gg |u|$ since $\tilde{M} \ll 1$). Moreover, according to (87) the speed of the right shock is $\sigma = u_R + \rho_R^{-1}\zeta\left(\rho_\star, \rho_R\right)$ and, in consideration of (67), $\sigma > u_R + a_R$. However, since $\rho_\star$ turns out to be close to ρ_R: $\rho_\star/\rho_R = 1 + \mathrm{O}\left(\tilde{M}\right)$, it follows that $\sigma \approx u_R + a_R\left(1 + \mathrm{O}\left(\tilde{M}\right)\right) \approx \tilde{a}$ and therefore $\tilde{a}$ can be considered as an acceptable estimate for the shock speed as well [27]. It is therefore clear that, during a unit time interval ($t_{eval} = 1$), the flow perturbation extends over an interval having width w_{domain} of the order of $\tilde{a}$:

$$\frac{w_{domain}}{\tilde{a}} = \mathrm{O}\left(1\right) \tag{122}$$

As far as the left rarefaction fan is concerned, it is delimited by the following characteristics (compare with (52)):

$$\frac{x}{t} = u_L - a_L \quad , \quad \frac{x}{t} = u_\star - a_\star$$

[26] The head of the wave is the extreme of the wave region which is in contact with the unperturbed state while the tail is the extreme of the wave region adjacent to the star region [98].

[27] These considerations incidentally show that, at low Mach numbers, the rarefactions and the shocks approximately travel at the same speed (in absolute value), as confirmed e.g. by Figs. 7-9 and Figs. 11-13 (in which the shock and the rarefaction are roughly located at the same distance from the position of the initial discontinuity, i.e. $x = 0$).

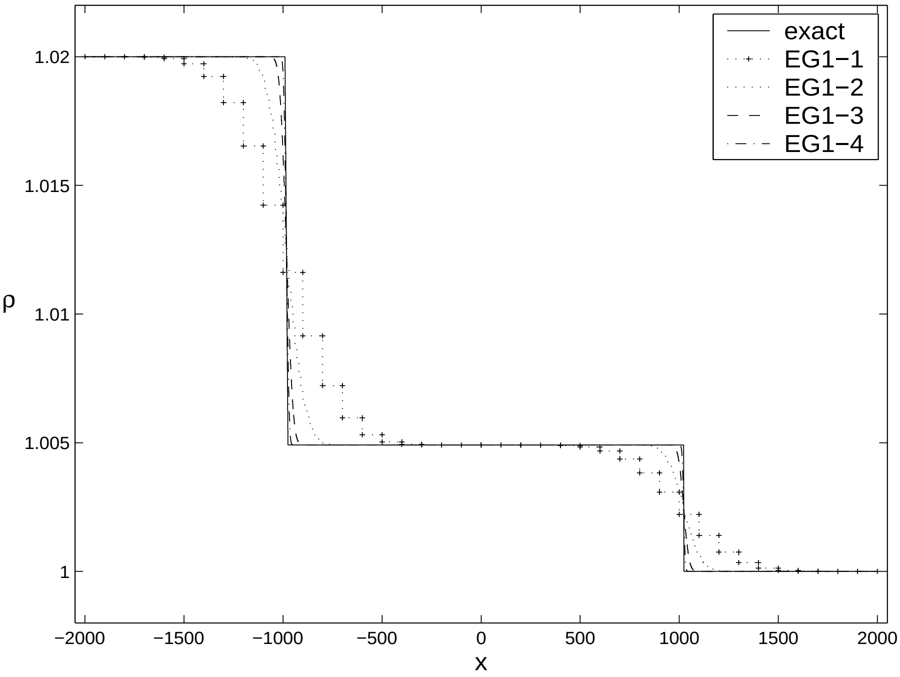

Figure 7: Approximation of ρ for the test-cases EG1-1 to EG1-4.

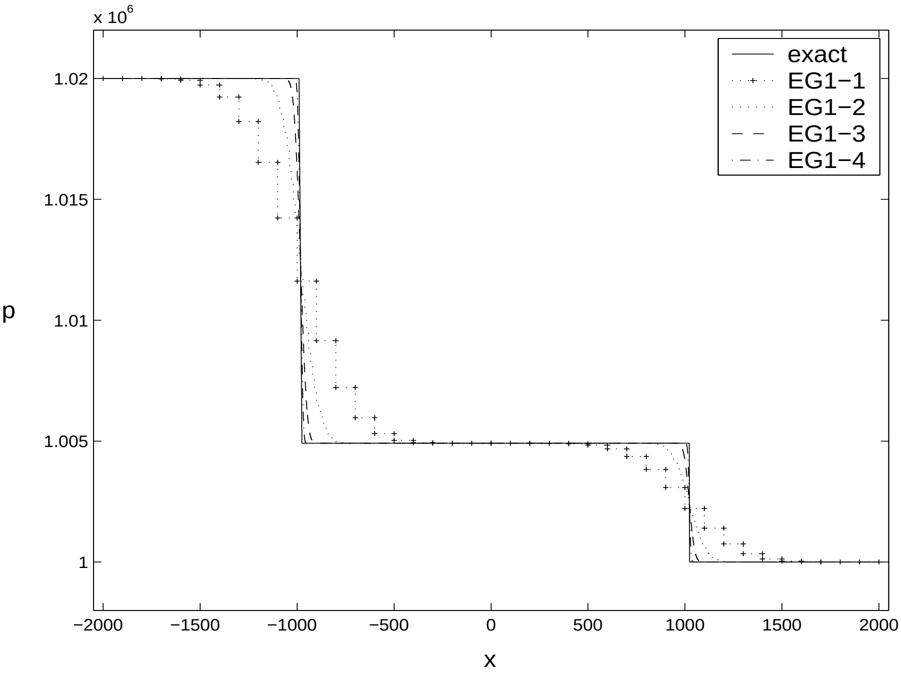

Figure 8: Approximation of p for the test-cases EG1-1 to EG1-4.

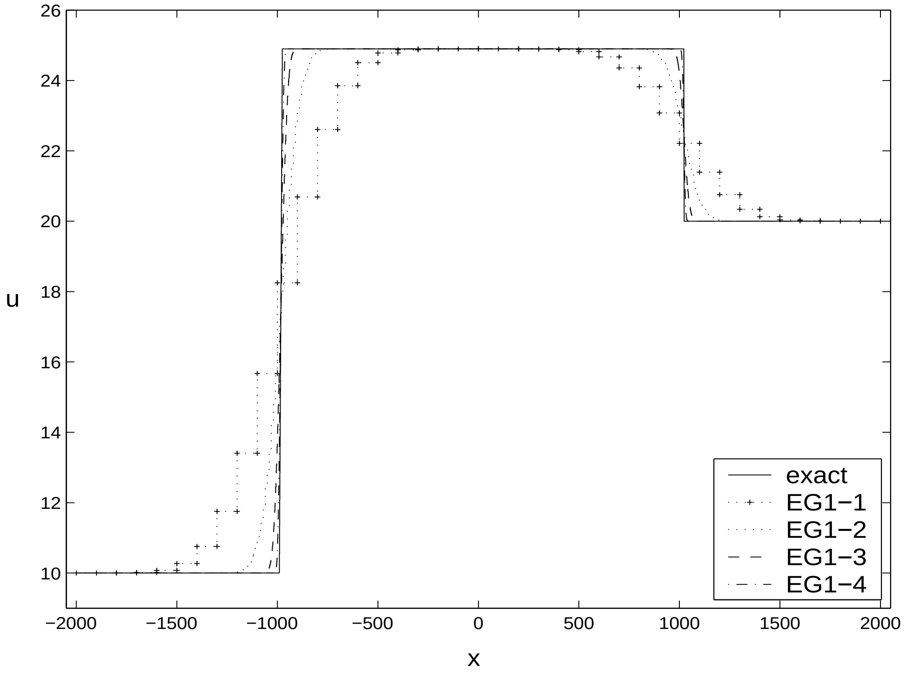

Figure 9: Approximation of u for the test-cases EG1-1 to EG1-4.

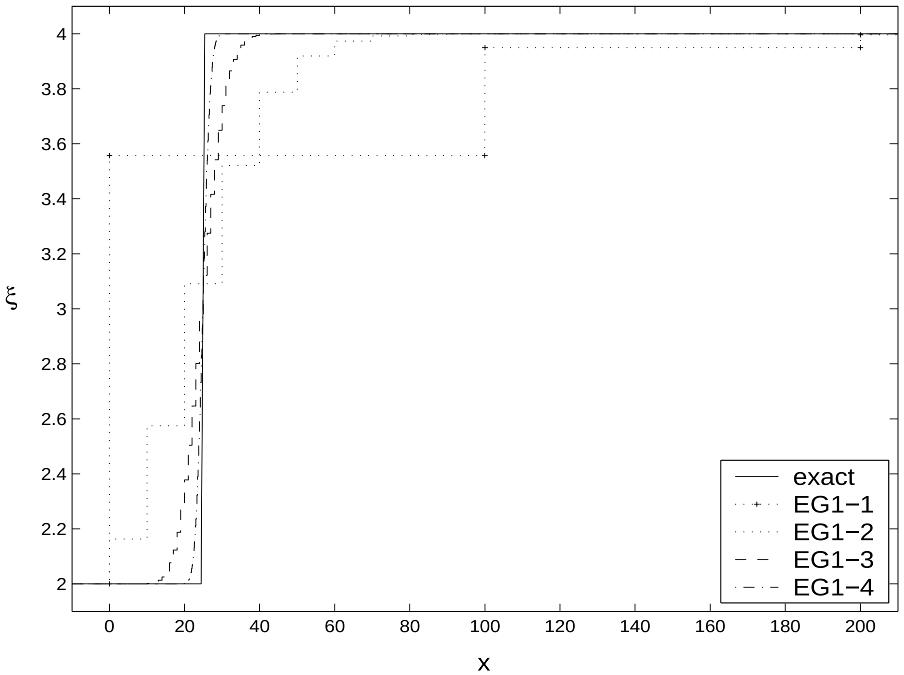

Figure 10: Approximation of ξ for the test-cases EG1-1 to EG1-4. The $x-$range is cut for ease of readability.

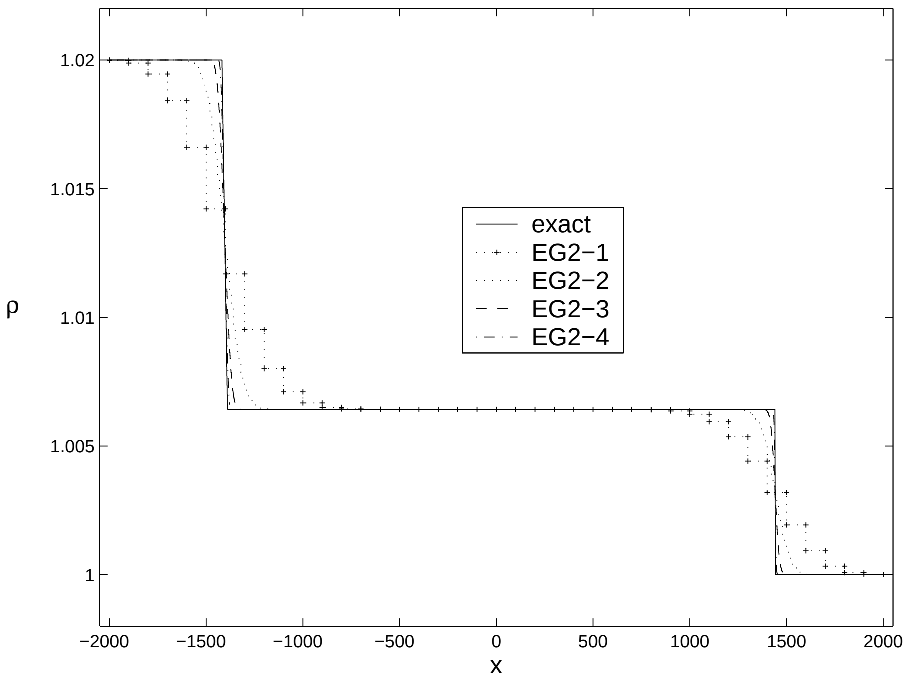

Figure 11: Approximation of ρ for the test-cases EG2-1 to EG2-4.

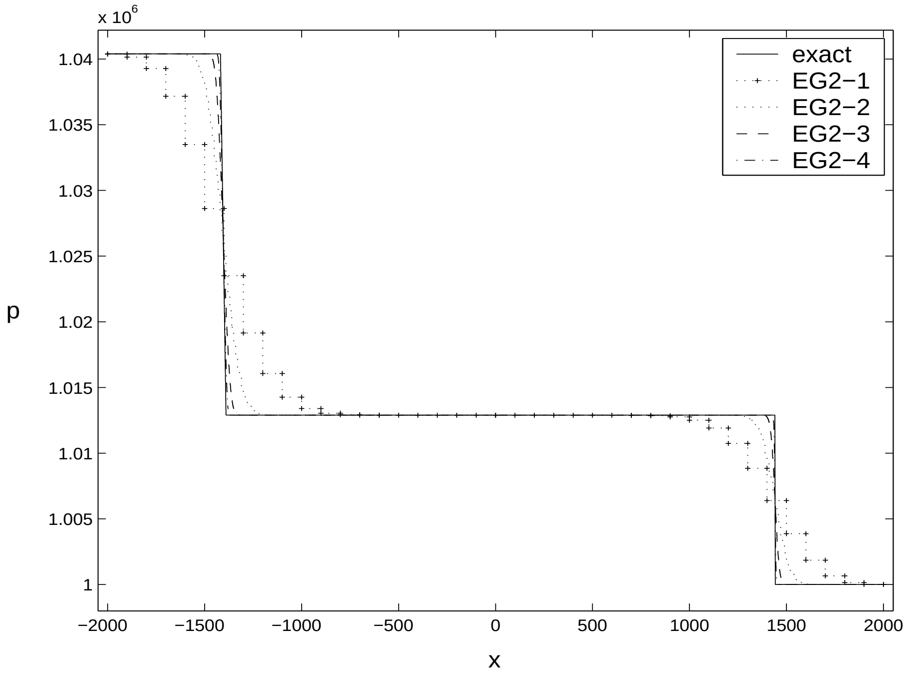

Figure 12: Approximation of p for the test-cases EG2-1 to EG2-4.

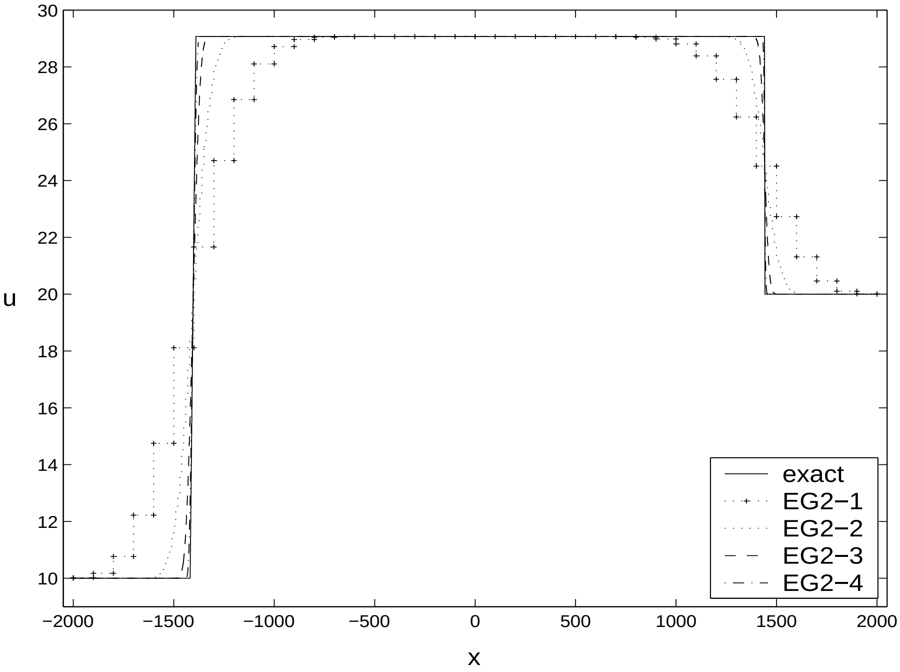

Figure 13: Approximation of u for the test-cases EG2-1 to EG2-4.

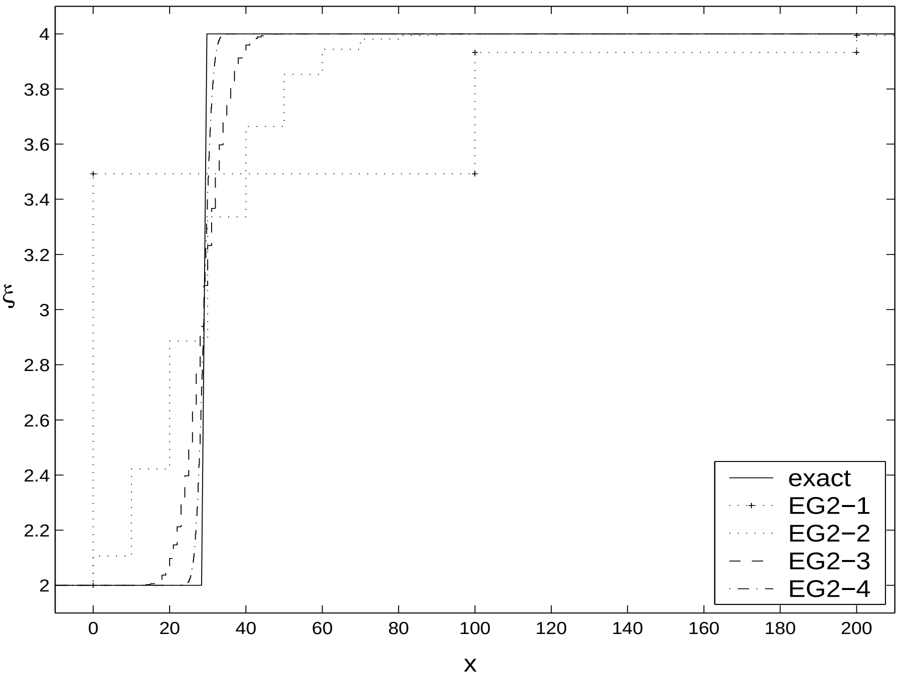

Figure 14: Approximation of ξ for the test-cases EG2-1 to EG2-4. The $x-$range is cut for ease of readability.

and thus, for $t_{eval} = 1$, its width w_{fan} can be expressed as follows:

$$w_{fan} = (u_\star - u_L) + (a_L - a_\star)$$

In particular:

- for the benchmark B1, $a_L = a_\star = \tilde{a}$, $u_L = \alpha_L \tilde{M}\,\tilde{a}$ and $u_\star = \alpha_\star \tilde{M}\,\tilde{a}$, with α_L and $\alpha_\star$ of the order of the unity and such that $(\alpha_\star - \alpha_L) > 0$. Hence, in the present case w_{fan} reads:

$$w_{fan} = (\alpha_\star - \alpha_L)\ \tilde{M}\,\tilde{a} \tag{123}$$

- for the benchmark B2, $u_L = \gamma_L \tilde{M}\,\tilde{a}$ and $u_\star = \gamma_\star \tilde{M}\,\tilde{a}$, with γ_L and $\gamma_\star$ of the order of the unity and such that $(\gamma_\star - \gamma_L) > 0$. Moreover, $\rho_L = \rho_R \left(1 + \beta_L \tilde{M}\right)$ and $\rho_\star = \rho_R \left(1 + \beta_\star \tilde{M}\right)$, with β_L and $\beta_\star$ of the order of the unity and such that $(\beta_L - \beta_\star) > 0$. Then, since $a = \sqrt{2\kappa\rho}$ and $\tilde{a} = a_R = \sqrt{2\kappa}$, it follows that $a_L \approx \tilde{a}\left(1 + \beta_L \tilde{M}/2\right)$ and $a_\star \approx \tilde{a}\left(1 + \beta_\star \tilde{M}/2\right)$. Hence, w_{fan} for the present case reads:

$$w_{fan} \approx \left[(\gamma_\star - \gamma_L) + \frac{1}{2}\,(\beta_L - \beta_\star)\right] \tilde{M}\,\tilde{a} \tag{124}$$

In light of (123) and (124),

$$\frac{w_{fan}}{\tilde{a}} = \mathrm{O}\left(\tilde{M}\right)$$

for both the considered benchmarks and, by recalling (122), it is clear that:

$$\frac{w_{fan}}{w_{domain}} = \mathrm{O}\left(\tilde{M}\right)$$

thus motivating the aforementioned observation.

Some entities which can be exploited in order to evaluate the accuracy as well as the computational cost of each simulation are finally reported in Tab. 3, namely:

- an estimate $\tilde{c}^{(CFL)}$ of the CFL coefficient, defined as follows (compare with (121)):

$$\tilde{c}^{(CFL)} := \frac{\tau\,\tilde{s}_{max}}{\mu} \tag{125}$$

  where $\tilde{s}_{max}$ denotes the largest wave speed of the RP associated with the relevant benchmark;

Test-case	$\tilde{c}^{(CFL)}$	t_{CPU}	$e(\rho)$	$e(p)$	$e(u)$	$e(\xi)$
EG1-1	0.51	$\approx$ 0.1 sec.	0.1792	0.1792	8.5733	4.3568
EG1-2	0.51	$\approx$ 1 sec.	0.0967	0.0967	4.6240	2.0348
EG1-3	0.51	$\approx$ 35 sec.	0.0492	0.0492	2.3530	1.0299
EG1-4	0.51	$\approx$ 35 min.	0.0211	0.0211	1.0110	0.4740
EG2-1	0.72	$\approx$ 0.1 sec.	0.1587	0.3185	8.6763	4.5884
EG2-2	0.72	$\approx$ 1 sec.	0.0837	0.1679	4.5757	2.0057
EG2-3	0.72	$\approx$ 35 sec.	0.0392	0.0786	2.1438	1.0515
EG2-4	0.72	$\approx$ 35 min.	0.0141	0.0282	0.7703	0.4493

Table 3: CFL estimate, CPU time and error estimates for the test-cases reported in Tab. 2.

- the CPU time t_{CPU}, as required on a laptop having the following characteristics: Intel P4 CPU 2.66GHz, 512kB L2 cache, 512MB RAM;

- some error estimates concerning the numerical solution for ρ, p, u and ξ, whose definition is discussed below by considering the generic entity ψ. Let ψ_k^{bench} denote a sequence obtained by sampling the exact solution ψ of the relevant RP in correspondence of the sequence x_k^{bench}, with $k = 1, 2, \ldots, N_{bench}$, which is fine enough to reproduce the variation of the considered solution almost exactly. Moreover, let ψ_j^{num} denote a discrete approximation of ψ which is constant within each interval $\left(x_j^{num}, x_{j+1}^{num}\right)$, with $j = 1, 2, \ldots, (N_{num} - 1)$. Finally, let D_x represent the following interval:

  $$D_x := (-n_L\,\mu, n_R\,\mu)$$

  Clearly, it is always possible to define the aforementioned sequences over D_x by adjusting the parameters controlling the space discretization in such a way that:

  $$\left(x_1^{bench}, x_{N_{bench}}^{bench}\right) = \left(x_1^{num}, x_{N_{num}}^{num}\right) = D_x$$

  Furthermore, it is possible to merge x_k^{bench} and x_j^{num} into a new sequence of abscissae, say x_h^{merge} (with $h = 1, 2, \ldots, N_{merge}$) and to define a linear interpolation of ψ_k^{bench} and ψ_j^{num}, respectively denoted by $\hat{\psi}_h^{bench}$ and

$\hat{\psi}_h^{num}$, over the new sequence [28]. Then, the following definition can be introduced in order to estimate the error $e\left(\psi\right)$ connected with the numerical approximation of ψ:

$$e\left(\psi\right) := \left(\frac{\displaystyle\int_{D_x} \left(\hat{\psi}_h^{num} - \hat{\psi}_h^{bench}\right)^2 \mathrm{d}x}{\displaystyle\int_{D_x} \left(\hat{\psi}_h^{bench}\right)^2 \mathrm{d}x} \right)^{\frac{1}{2}} \cdot 10^2 \tag{126}$$

where the integration can be carried out by means of the trapezoidal rule [79], consistently with the chosen linear interpolation.
According to Tab. 3, the discrete solution correctly converges towards the exact one [29]. Moreover, the convergence is sub-linear and roughly exhibits the same trend for all the considered entities, as shown in Figs. 15 and 16.

[28] The linear interpolation is consistent with the piece-wise constant numerical discretization. Moreover, as far as the exact solution is concerned, it should introduce a negligible error in view of the adopted fine sampling.

[29] For the test-cases EG1-1 to EG1-4, $e\left(\rho\right) = e\left(p\right)$ by virtue of the direct proportionality between ρ and p which is introduced by the state law associated with the benchmark B1.

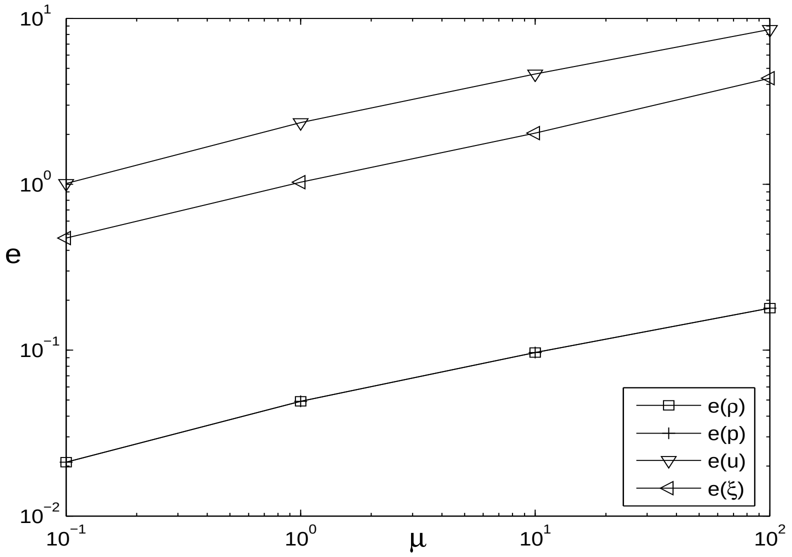

Figure 15: Plot of the error estimates for the test-cases EG1-1 to EG1-4 reported in Tab. 3.

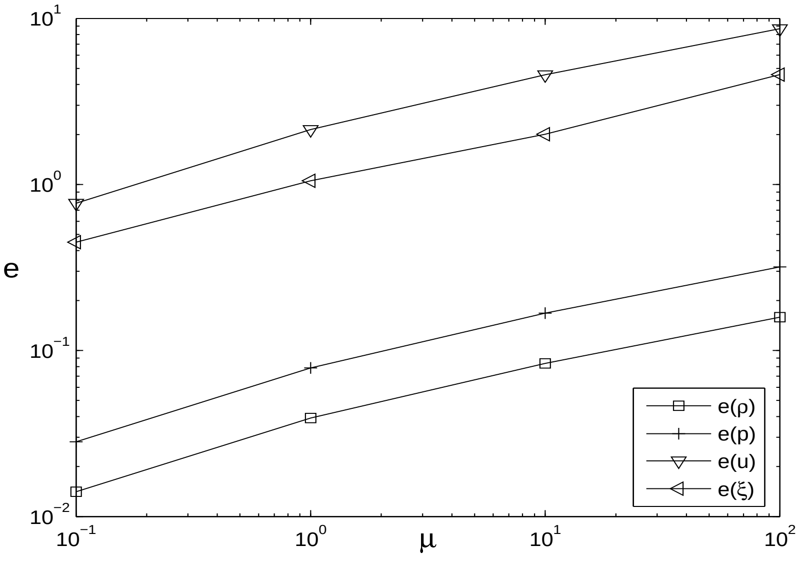

Figure 16: Plot of the error estimates for the test-cases EG2-1 to EG2-4 reported in Tab. 3.

3.3 A Roe scheme for generic barotropic state laws

In sec. 3.3.1 a Roe numerical flux function applicable to generic barotropic state laws is proposed. In sec. 3.3.2 the scheme (102) exploiting the proposed numerical flux is validated against an exact solution.

3.3.1 Roe numerical flux

A common numerical flux suitable for incorporation into the Godunov approach (see sec. 3.1.1) is that one originally proposed by Roe [84]. According to this method, an approximate RP is suitably introduced at each cell interface and the numerical flux is defined by considering the flux -as obtained by exactly solving the approximate RP- which crosses the interface. In the Roe method, the approximation of the inter-cell flux is obtained "directly" (contrarily to other methods, generally referred to as "approximate-state Riemann solvers", which define the numerical flux by evaluating the analytical one $\mathbf{f}^{(A)}$ in correspondence of a suitably defined state vector); further details can be found in a number of textbooks, e.g. [39] and [98] amongst many others.

Definition of the numerical flux $\boldsymbol{\phi}_{LR}^{(A)ROE}$

The non-linear p.d.e. in (57) is locally replaced with the following linear one:

$$\partial_t \, \mathbf{q}^{(A)} + \tilde{\mathbf{J}}_{LR}^{(A)} \cdot \partial_x \, \mathbf{q}^{(A)} = \mathbf{0} \tag{127}$$

where $\tilde{\mathbf{J}}_{LR}^{(A)}$ represents a suitable matrix, called "Roe matrix", depending on the "left" and "right" states $\mathbf{q}_L^{(A)}$ and $\mathbf{q}_R^{(A)}$:

$$\tilde{\mathbf{J}}_{LR}^{(A)} := \tilde{\mathbf{J}}^{(A)} \left(\mathbf{q}_L^{(A)}, \mathbf{q}_R^{(A)} \right) \tag{128}$$

The Roe matrix must verify the following conditions for any couple $\left(\mathbf{q}_L^{(A)}, \mathbf{q}_R^{(A)} \right)$:

(RM1) $\tilde{\mathbf{J}}_{LR}^{(A)}$ is diagonalizable with real eigenvalues;

(RM2) $\tilde{\mathbf{J}}^{(A)} \left(\mathbf{q}_L^{(A)} \to \mathbf{q}^{(A)\star}, \mathbf{q}_R^{(A)} \to \mathbf{q}^{(A)\star} \right) \to \mathbf{J}^{(A)} \left(\mathbf{q}^{(A)\star} \right)$

where $\mathbf{J}^{(A)}$ denotes the Jacobian of the original (non linear) flux $\mathbf{f}^{(A)}$ defined in (37);

(RM3) let $\Delta^{LR} \, \mathbf{z}$ denote the variation of the generic vector $\mathbf{z}$ between a "left" state $\mathbf{z}_L$ and a "right" state $\mathbf{z}_R$:

$$\Delta^{LR} \, \mathbf{z} := \mathbf{z}_R - \mathbf{z}_L \tag{129}$$

Then:

$$\Delta^{LR}\,\mathbf{f}^{(A)} = \tilde{\mathbf{J}}^{(A)}_{LR} \cdot \Delta^{LR}\,\mathbf{q}^{(A)} \tag{130}$$

where, of course, $\mathbf{f}^{(A)}_s$ is understood as $\mathbf{f}^{(A)}\left(\mathbf{q}^{(A)}_s\right)$, $s \in \{L, R\}$.

Note 22 *The condition (RM1) ensures that the hyperbolicity of the 1D problem at hand (see sec. 2.3.2) is preserved when replacing the original RP with the approximate, linearized, one. The condition (RM2) enforces a natural consistency requirement. The condition (RM3), instead, is imposed by the fact that a unique value for the flux crossing the interface must be obtained by only considering either the left or the right portion of the solution to the considered linearized RP (of course, due to conservation) [98].*

By recalling the solution of the considered linearized RP (see e.g. [98]), it is possible to express the numerical flux from $\mathbf{q}^{(A)}_L$ to $\mathbf{q}^{(A)}_R$ (along the direction associated with the versor $\hat{\boldsymbol{e}}$ defined in sec. 3.1.1) as follows:

$$\boldsymbol{\phi}^{(A)ROE}_{LR} := \boldsymbol{\phi}^{(A)ROE}_{c,LR} + \boldsymbol{\phi}^{(A)ROE}_{u,LR} \tag{131}$$

where:

$$\boldsymbol{\phi}^{(A)ROE}_{c,LR} := \frac{1}{2}\left(\mathbf{f}^{(A)}_L + \mathbf{f}^{(A)}_R\right) \tag{132}$$

$$\boldsymbol{\phi}^{(A)ROE}_{u,LR} := \mathbf{D}^{(A)}_{LR} \cdot \Delta^{LR}\mathbf{q}^{(A)} \tag{133}$$

$$\mathbf{D}^{(A)}_{LR} := -\frac{1}{2}\left|\tilde{\mathbf{J}}^{(A)}_{LR}\right| \tag{134}$$

As far as the definition (134) is concerned, the operator $|\cdot|$, defined in (6), can be rightfully applied in consideration of the condition (RM1) above.

Note 23 *The expression (133) takes into account the wave structure of the linearized RP (in particular, the sudden variation of the solution across the waves propagating along the characteristics [98]) and it is consequently referred to as the "upwind" component of the numerical flux function. The expression (132), instead, is often referred to as the "centred" component of the numerical flux function, due to its symmetrical form.*

By exploiting the definition (7), together with the property (RM3) above, it is possible to recast the numerical flux $\boldsymbol{\phi}^{(A)ROE}_{LR}$ as follows:

$$\begin{cases} \boldsymbol{\phi}^{(A)ROE}_{LR} = \mathbf{f}^{(A)}_L + \left(\tilde{\mathbf{J}}^{(A)}_{LR}\right)^{-} \cdot \Delta^{LR}\mathbf{q}^{(A)} \\ \boldsymbol{\phi}^{(A)ROE}_{LR} = \mathbf{f}^{(A)}_R - \left(\tilde{\mathbf{J}}^{(A)}_{LR}\right)^{+} \cdot \Delta^{LR}\mathbf{q}^{(A)} \end{cases} \tag{135}$$

A Roe matrix for generic barotropic state laws

Clearly, the Roe matrix depends in general on the specific problem under consideration and, in particular, on the specific state law. In the original paper [84], for instance, a Roe matrix is defined for the Euler equations associated with a perfect gas state law [30]. A crucial constraint on the definition of the Roe matrix is given by the condition (RM3) above. In [84], the fulfilment of this condition is obtained by identifying a suitable vector, called "parameter vector", such that both the state vector and the (analytical) flux are homogeneous quadratic functions of it. In the present (barotropic) case, the approach under discussion would encourage to seek a certain vector $\mathbf{z}$ such that $\mathbf{q}^{(A)}$ and $\mathbf{f}^{(A)}$ are homogeneous quadratic functions of $\mathbf{z}$. Clearly, it is not possible to define such a vector due to the assumed generality of the barotropic curve $p = p(\rho)$. Nevertheless, it is possible to take advantage of the basic idea underlying the considered approach, as described below.

Once the flux $\mathbf{f}^{(A)}$ has been split as follows:

$$\mathbf{f}^{(A)} = \mathbf{f}_H^{(A)} + \mathbf{f}_{NH}^{(A)} \tag{136}$$

with:

$$\mathbf{f}_H^{(A)} := \begin{pmatrix} \rho u \\ \rho u^2 \\ \rho u \xi \end{pmatrix} \quad , \quad \mathbf{f}_{NH}^{(A)} := \begin{pmatrix} 0 \\ p \\ 0 \end{pmatrix} \tag{137}$$

the variation $\Delta^{LR}\mathbf{f}^{(A)}$ clearly reads:

$$\Delta^{LR}\mathbf{f}^{(A)} = \Delta^{LR}\,\mathbf{f}_H^{(A)} + \Delta^{LR}\,\mathbf{f}_{NH}^{(A)} \tag{138}$$

Moreover, once introduced the following vector:

$$\mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} := \begin{pmatrix} \sqrt{\rho} \\ \sqrt{\rho}\,u \\ \sqrt{\rho}\,\xi \end{pmatrix}$$

it is clear that both $\mathbf{q}^{(A)}$ and $\mathbf{f}_H^{(A)}$ are homogeneous quadratic function of $\mathbf{z}$, since:

$$\mathbf{q}^{(A)}(\mathbf{z}) = \begin{pmatrix} z_1^2 \\ z_1 z_2 \\ z_1 z_3 \end{pmatrix} \quad , \quad \mathbf{f}_H^{(A)}(\mathbf{z}) = \begin{pmatrix} z_1 z_2 \\ z_2^2 \\ z_2 z_3 \end{pmatrix}$$

[30] Different extensions to more complex cases have been proposed in the literature (see e.g. [37], [111], [1] and [41]) amongst many others.

and therefore the following relations hold (by a well-known property of the homogeneous quadratic functions):

$$\Delta^{LR}\mathbf{q}^{(A)} = \mathbf{Q}_{LR} \cdot \Delta^{LR}\mathbf{z} \quad , \quad \Delta^{LR}\mathbf{f}_H^{(A)} = \mathbf{F}_{H,LR} \cdot \Delta^{LR}\mathbf{z} \tag{139}$$

where:

$$\mathbf{Q}_{LR} := \partial_{\mathbf{z}}\mathbf{q}^{(A)} \left(\mathbf{z} = \frac{\mathbf{z}_L + \mathbf{z}_R}{2}\right) \quad , \quad \mathbf{F}_{H,LR} := \partial_{\mathbf{z}}\mathbf{f}_H^{(A)} \left(\mathbf{z} = \frac{\mathbf{z}_L + \mathbf{z}_R}{2}\right)$$

Then, by combining the equations in (139), the following relation is immediately obtained:

$$\Delta^{LR}\mathbf{f}_H^{(A)} = \hat{\mathbf{J}}_{LR}^{(A)} \cdot \Delta^{LR}\mathbf{q}^{(A)}$$

with:

$$\hat{\mathbf{J}}_{LR}^{(A)} := \mathbf{F}_{H,LR} \cdot \mathbf{Q}_{LR}^{-1}$$

and (138) can be finally recast as follows:

$$\Delta^{LR}\mathbf{f}^{(A)} = \hat{\mathbf{J}}_{LR}^{(A)} \cdot \Delta^{LR}\mathbf{q}^{(A)} + \Delta^{LR}\,\mathbf{f}_{NH}^{(A)} \tag{140}$$

Straightforward computations lead to the following representation, in particular, for $\hat{\mathbf{J}}_{LR}^{(A)}$:

$$\hat{\mathbf{J}}_{LR}^{(A)} = \begin{pmatrix} 0 & 1 & 0 \\ -u_{LR}^2 & 2\,u_{LR} & 0 \\ -u_{LR}\,\xi_{LR} & \xi_{LR} & u_{LR} \end{pmatrix} \tag{141}$$

where (subscripts “L” and “R”, as applied to vector components, are understood in the sequel):

$$\begin{cases} u_{LR} := \dfrac{\sqrt{\rho_L}\,u_L + \sqrt{\rho_R}\,u_R}{\sqrt{\rho_L} + \sqrt{\rho_R}} \\[2ex] \xi_{LR} := \dfrac{\sqrt{\rho_L}\,\xi_L + \sqrt{\rho_R}\,\xi_R}{\sqrt{\rho_L} + \sqrt{\rho_R}} \end{cases} \tag{142}$$

By starting from (140), it is possible to to match the condition (130) and, consequently, to define a Roe matrix for the barotropic case under consideration, as shown in the following:

Proposition 6 *A Roe matrix* $\tilde{\mathbf{J}}^{(A)}_{LR}$ *applicable when considering a generic barotropic state law reads:*

$$\tilde{\mathbf{J}}^{(A)}_{LR} = \begin{pmatrix} 0 & 1 & 0 \\ \\ a^2_{LR} - u^2_{LR} & 2\,u_{LR} & 0 \\ \\ -u_{LR}\,\xi_{LR} & \xi_{LR} & u_{LR} \end{pmatrix} \tag{143}$$

where:

$$a_{LR} := \begin{cases} \left(\dfrac{\Delta^{LR}\,p}{\Delta^{LR}\,\rho}\right)^{\frac{1}{2}} & \text{if} \quad \rho_R \neq \rho_L \\ \\ a(\rho_\star) & \text{if} \quad \rho_R = \rho_L = \rho_\star \end{cases} \tag{144}$$

Proof *At a first step, a matrix* $\check{\mathbf{J}}^{(A)}_{LR}$ *is sought such that:*

$$\begin{pmatrix} 0 \\ \Delta^{LR}\,p \\ 0 \end{pmatrix} = \Delta^{LR}\mathbf{f}^{(A)}_{NH} = \check{\mathbf{J}}^{(A)}_{LR} \cdot \Delta^{LR}\mathbf{q}^{(A)} \tag{145}$$

Let α_{mn} $(m, n \in \{1, 2, 3\})$ *denote the* $mn-th$ *component of* $\check{\mathbf{J}}^{(A)}_{LR}$. *Then,* $\alpha_{1n} = \alpha_{3n} = 0$ *due to the mutual independence of the state vector components (i.e.* ρ, ρu *and* $\rho\xi$*), while* $\alpha_{22} = \alpha_{23} = 0$ *by virtue of the barotropic state law (3). Hence, (145) reduces to the following scalar equation:*

$$\Delta^{LR}\,p = \alpha_{21}\,\Delta^{LR}\,\rho \tag{146}$$

When $\rho_R = \rho_L$ *the above equation is trivially verified regardless of the specific value of* α_{21} *while, for* $\rho_R \neq \rho_L$ *it necessarily follows that:*

$$\alpha_{21} = \frac{\Delta^{LR}\,p}{\Delta^{LR}\,\rho} \quad , \quad \rho_R \neq \rho_L$$

where the divided difference is positive, due to the strict monotonicity of $p\,(\rho)$ *assumed in (4). Hence, by choosing* $\alpha_{21} = a^2_{LR}$ *with* a_{LR} *defined in (144), a continuous (i.e. prolongated by continuity) solution is obtained. As a result, the expression of* $\check{\mathbf{J}}^{(A)}_{LR}$ *reads:*

$$\check{\mathbf{J}}^{(A)}_{LR} = \begin{pmatrix} 0 & 0 & 0 \\ a^2_{LR} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{147}$$

By substituting (145) into (140) it is evident that the following matrix:

$$\tilde{\mathbf{J}}_{LR}^{(A)} = \hat{\mathbf{J}}_{LR}^{(A)} + \check{\mathbf{J}}_{LR}^{(A)} \tag{148}$$

satisfies the condition (130) (i.e. the condition (RM3) above). Furthermore, it is straightforward to verify that the considered matrix $\tilde{\mathbf{J}}_{LR}^{(A)}$ *also satisfies the aforementioned conditions (RM1) and (RM2) and, therefore, it is a suitable Roe matrix for the generic barotropic case at hand. As far as its representation is concerned, by substituting (141) and (147) into (148), the expression (143) is immediately obtained. This completes the proof.* ■

Note 24 *While* u_{LR} *and* ξ_{LR} *in (142) are well-known "Roe averages" [84],* a_{LR} *in (144) represents an average value (hereafter referred to as Roe average as well) which is specific to the present (generic) barotropic case. For instance, it can be also exploited when considering the well-known homogeneous shallow water equations, since they can be derived from the considered conservation laws (see Note 12 in sec. 2.5.1). Indeed, the expression (144) generalizes the relevant one defined in [99] for the shallow water case.*

Note 25 *It may be worth mentioning that, as far as its numerical implementation is concerned,* a_{LR} *should be defined as follows:*

$$a_{LR} := \begin{cases} \left(\dfrac{\Delta^{LR} p}{\Delta^{LR} \rho} \right)^{1/2} & \text{if} \quad \mid \rho_R - \rho_L \mid > \epsilon_\rho \\ a\left(\rho = \varrho\left(\rho_L, \rho_R\right) \right) & \text{if} \quad \mid \rho_R - \rho_L \mid < \epsilon_\rho \end{cases}$$

where ϵ_ρ *is a suitable numerical threshold and* $\varrho\left(\rho_L, \rho_R\right)$ *is an average value (e.g. a geometrical mean) such that* $\varrho\left(\rho_L \to \rho^\star, \rho_R \to \rho^\star\right) \to \rho^\star$.

Note 26 *The expression of the Jacobian* $\mathbf{J}^{(A)}$ *defined in (37) reads:*

$$\mathbf{J}^{(A)} = \begin{pmatrix} 0 & 1 & 0 \\ a^2 - u^2 & 2\,u & 0 \\ -u\,\xi & \xi & u \end{pmatrix} \tag{149}$$

Once noticed the formal similarity between (149) and (143) and by interpreting $\mathbf{J}^{(A)}$ *in (149) as a function of* a*,* u *and* ξ*, the following relation clearly holds* [31]*:*

$$\tilde{\mathbf{J}}_{LR}^{(A)} = \mathbf{J}^{(A)}\left(a = a_{LR}, u = u_{LR}, \xi = \xi_{LR} \right) \tag{150}$$

[31] An analogous relation holds for the Euler equations associated with a perfect gas state law [84].

Note 27 *The Roe matrix (143) keeps the same representation even when the "left" and "right" states are interchanged, due to the "symmetrical" definition of its components, namely (142) and (144).*

Note 28 *By neglecting the third row and the third column of (143), a Roe matrix for the basic-1D equations (see sec. 2.2.3) is obtained, which has been previously introduced in [38].*

Note 29 *In [91], a Roe matrix for the basic-1D equations (see sec. 2.2.3) completed with the energy balance is defined. The averages appearing in [91] can be derived from those obtained in [37] for the case of a generic state law of the form $p = p(\rho, e_i)$, e_i denoting the internal energy per unit mass.*

Note 30 *It is straightforward to extend the Roe matrix (143) to the case of $m > 1$ passive scalars. When adopting, for instance, the definition (28) for the state vector $\mathbf{q}^{(A)}$ ($m = 2$), the considered Roe matrix reads:*

$$\tilde{\mathbf{J}}_{LR}^{(A)} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ a_{LR}^2 - u_{LR}^2 & 2\,u_{LR} & 0 & 0 \\ -u_{LR}\,\xi_{LR} & \xi_{LR} & u_{LR} & 0 \\ -u_{LR}\,\eta_{LR} & \eta_{LR} & 0 & u_{LR} \end{pmatrix} \tag{151}$$

with η_{LR} defined analogously to ξ_{LR} in (142).

Definition of the numerical flux $\boldsymbol{\phi}_{ij}^{(A)ROE}$

Let $\mathbf{q}_i^{(A)}$, $\mathbf{q}_j^{(A)}$ and $\hat{\boldsymbol{\nu}}_{ij}$, with $j \in \pi_i$, be defined as in sec. 3.1.1. In particular, $\mathbf{q}_i^{(A)}$ and $\mathbf{q}_j^{(A)}$ can respectively represent either a couple of "left" and "right" states or vice-versa. It is possible to exploit the numerical flux (131)-(134) in order to define a Roe numerical flux from $\mathbf{q}_i^{(A)}$ to $\mathbf{q}_j^{(A)}$, as described below.

As far as the Roe matrix is concerned, in view of the "symmetry" already pointed out in Note 27 above, it is straightforward to generalize (143) as follows:

$$\tilde{\mathbf{J}}_{ij}^{(A)} := \begin{pmatrix} 0 & 1 & 0 \\ \\ a_{ij}^2 - u_{ij}^2 & 2\,u_{ij} & 0 \\ \\ -u_{ij}\;\xi_{ij} & \xi_{ij} & u_{ij} \end{pmatrix} \tag{152}$$

with:

$$\begin{cases} u_{ij} := \dfrac{\sqrt{\rho_i}\,u_i + \sqrt{\rho_j}\,u_j}{\sqrt{\rho_i} + \sqrt{\rho_j}} \\ \\ \xi_{ij} := \dfrac{\sqrt{\rho_i}\,\xi_i + \sqrt{\rho_j}\,\xi_j}{\sqrt{\rho_i} + \sqrt{\rho_j}} \end{cases} \tag{153}$$

$$a_{ij} := \begin{cases} \left(\dfrac{\Delta^{ij}\,p}{\Delta^{ij}\,\rho}\right)^{\frac{1}{2}} & \text{if} \quad \rho_j \neq \rho_i \\ \\ a(\rho_\star) & \text{if} \quad \rho_j = \rho_i = \rho_\star \end{cases} \tag{154}$$

and, of course:

$$\Delta^{ij}\,(\cdot) := (\cdot)_j - (\cdot)_i \tag{155}$$

Then, once recalled the definition of s_{ij} given in (98), it is possible to define the Roe numerical under consideration as follows:

$$\boldsymbol{\phi}^{(A)ROE}\left(\mathbf{q}_i^{(A)}, \mathbf{q}_j^{(A)}, \hat{\boldsymbol{\nu}}_{ij}\right) := \boldsymbol{\phi}_{ij}^{(A)ROE} \quad , \quad j \in \pi_i \tag{156}$$

with:

$$\boldsymbol{\phi}_{ij}^{(A)ROE} \;:=\; \boldsymbol{\phi}_{c,ij}^{(A)ROE} + \boldsymbol{\phi}_{u,ij}^{(A)ROE} \tag{157}$$

$$\boldsymbol{\phi}_{c,ij}^{(A)ROE} \;:=\; \frac{1}{2}\,s_{ij}\left(\mathbf{f}_i^{(A)} + \mathbf{f}_j^{(A)}\right) \tag{158}$$

$$\boldsymbol{\phi}_{u,ij}^{(A)ROE} \;:=\; \mathbf{D}_{ij}^{(A)} \cdot \Delta^{ij}\mathbf{q}^{(A)} \tag{159}$$

$$\mathbf{D}_{ij}^{(A)} \;:=\; -\frac{1}{2}\left|\,s_{ij}\,\tilde{\mathbf{J}}_{ij}^{(A)}\right| \tag{160}$$

where, of course, $\mathbf{f}_i^{(A)}$ is understood as $\mathbf{f}^{(A)}\left(\mathbf{q}_i^{(A)}\right)$.

The expressions (158)-(160) are defined by considering the following RP which generalizes (57):

$$\left\{\begin{array}{rclcl} \partial_t\,\mathbf{q}^{(A)}+\partial_x\,\left(s_{ij}\,\mathbf{f}^{(A)}\right) & = & \mathbf{0} & \text{in} & \mathbb{R}\times(0,\infty) \\ \mathbf{q}^{(A)} & = & \left\{\begin{array}{lll}\mathbf{q}_i^{(A)} & if & x<0 \\ \mathbf{q}_j^{(A)} & if & x>0\end{array}\right. & \text{on} & \mathbb{R}\times\{t=0\} \end{array}\right. \tag{161}$$

with $i\in\mathcal{I}$ and $j\in\pi_i$. Indeed:

- for $j=i+1$ (161) reduces to (57), with $\mathbf{q}_L^{(A)}=\mathbf{q}_i^{(A)}$ and $\mathbf{q}_R^{(A)}=\mathbf{q}_j^{(A)}$;
- also for $j=i-1$ (161) reduces to (57), at the only cost of reversing the orientation of the $x-$axis (accordingly to $\hat{\boldsymbol{\nu}}_{ij}=-\,\hat{\boldsymbol{e}}$); in this case, $\mathbf{q}_L^{(A)}=\mathbf{q}_j^{(A)}$ and $\mathbf{q}_R^{(A)}=\mathbf{q}_i^{(A)}$.

In both cases, the generalized analytical flux $s_{ij}\,\mathbf{f}^{(A)}$ must be considered when defining the Roe linearization and the corresponding Roe matrix is clearly given by $s_{ij}\,\tilde{\mathbf{J}}_{ij}^{(A)}$, with $\tilde{\mathbf{J}}_{ij}^{(A)}$ defined in (152).

Note 31 *It should be noticed that the Roe numerical flux (156) trivially satisfies the consistency property (100). The conservation property (99) is satisfied as well. Indeed, while $s_{ji}=-s_{ij}$ and $\tilde{\mathbf{J}}_{ji}^{(A)}=\tilde{\mathbf{J}}_{ij}^{(A)}$, the following relation clearly holds, due to the definition of the operator $|\cdot|$ introduced in (6):*

$$\left|\,s_{ji}\,\tilde{\mathbf{J}}_{ji}^{(A)}\,\right| = \left|\,s_{ij}\,\tilde{\mathbf{J}}_{ij}^{(A)}\,\right| \tag{162}$$

In consideration of the equality (162), the introduction of s_{ij} into (160) may appear not necessary. Nevertheless, for the sake of consistency with (134), the operator $|\cdot|$ should be applied to the Roe matrix of the considered, namely $s_{ij}\,\tilde{\mathbf{J}}_{ij}^{(A)}$. Moreover, it is compulsory to work with the proper Roe matrix $s_{ij}\,\tilde{\mathbf{J}}_{ij}^{(A)}$ if alternative formulations of the type of (135) are sought (as, for instance, in sec. 3.5.1), since:

$$\left(s_{ij}\,\tilde{\mathbf{J}}_{ij}^{(A)}\right)^{\pm}\neq\left(\tilde{\mathbf{J}}_{ij}^{(A)}\right)^{\pm}$$

In particular, the following relations hold:

$$\left(s_{ji}\,\tilde{\mathbf{J}}_{ji}^{(A)}\right)^{\pm}=-\,\left(s_{ij}\,\tilde{\mathbf{J}}_{ij}^{(A)}\right)^{\mp} \tag{163}$$

where the definition of the operators $(\cdot)^{+}$ and $(\cdot)^{-}$ is introduced in (7).

Note 32 *It is known from the literature (see e.g. [39], [84] and [98]) that non-physical results may arise when exploiting the Roe scheme (156)-(160), due to the fact that the solution to the linearized problem, always consisting of discontinuities (see e.g. [98]), does not provide a correct approximation of continuous waves, like rarefactions. In a practical computational set up, however, problems generally arise when dealing with sonic rarefactions (i.e. rarefactions for which $\|\mathbf{u}\| = a$ along a certain characteristic line, see sec. 2.4.2): these show up in the form of discontinuities violating the RH condition (41). Besides the classical correction technique introduced in [43], various "entropy fixes" have been proposed in the literature to counteract this problem (see [39], [64] and [98] for a comprehensive list of references).*

It must be stated in advance that, despite the importance of the issue under consideration, no entropy fixes are considered in the present document, to be applied to the proposed Roe numerical flux (156)-(160). Indeed, the time-schedule of the industrial project this work was based on, imposed to directly concentrate on the simulation of non-cavitating flows, as a compulsory intermediate step towards the simulation of cavitation, as mentioned in sec. 1.6. Then, in consideration of the fact that pure liquid flows are nearly-incompressible (and therefore far from allowing for sonic conditions to take place), the investigation of the entropic behaviour of the considered numerical schemes was initially postponed and the numerical method developed for non-cavitating flows has been exploited for cavitating simulations as well (see sec. 6). On the other hand, when cavitation occurs a transonic regime is systematically encountered (see sec. 1.4) and the entropic behaviour of the considered Roe scheme must be assessed; according to the author, a further study should be devoted to this issue.

3.3.2 Numerical results

Benchmarks

The benchmarks introduced and discussed in sec. 3.2.2 (see Tab. 1) are considered here for validating the discrete scheme (102), based on the proposed numerical flux (156)-(160).

Initial and boundary conditions

The IC and the BCs introduced in sec. 3.2.2, namely (119) and (120), are adopted here.

Test-case	Benchmark	μ	(n_L, n_R)	τ
ER1-1	B1	100	$(2,2)\cdot 10^1$	$5\cdot 10^{-2}$
ER1-2	B1	10	$(2,2)\cdot 10^2$	$5\cdot 10^{-3}$
ER1-3	B1	1	$(2,2)\cdot 10^3$	$5\cdot 10^{-4}$
ER1-4	B1	0.1	$(2,2)\cdot 10^4$	$5\cdot 10^{-5}$
ER2-1	B2	100	$(2,2)\cdot 10^1$	$5\cdot 10^{-2}$
ER2-2	B2	10	$(2,2)\cdot 10^2$	$5\cdot 10^{-3}$
ER2-3	B2	1	$(2,2)\cdot 10^3$	$5\cdot 10^{-4}$
ER2-4	B2	0.1	$(2,2)\cdot 10^4$	$5\cdot 10^{-5}$

Table 4: Considered test-cases for the discrete scheme (102), based on the numerical flux (156)-(160).

Test-cases

In order to directly compare the proposed Roe numerical flux with the Godunov one proposed in sec. 3.2.1, the test-cases introduced in sec. 3.2.2 (see Tab. 2) are considered here. They are reported in Tab. 4, where the labels in the first column (different from those in Tab. 2) remind that the Roe numerical flux is exploited here. The corresponding numerical solutions are shown in Figs. 17-24.

As for the case of the Godunov flux, some entities which can be exploited for evaluating the accuracy and the computational cost of the considered simulations are reported in Tab. 5 (the definition of the relevant entities is reported in sec. 3.2.2, in correspondence of the introduction of Tab. 3). It should be noticed that:

- the column reporting the estimate $\tilde{c}^{(CFL)}$ of the CFL coefficient is clearly identical to the corresponding one in Tab. 3 since all the parameters involved in the definition of $\tilde{c}^{(CFL)}$, namely τ, μ and $\tilde{s}_{max}$ in (125), have the same value for corresponding test-cases;
- the CPU time (on a laptop with Intel P4 CPU 2.66GHz, 512kB L2 cache, 512MB RAM) is similar to that one required when adopting the Godunov flux. According to the author, the discrepancy between the test-cases ER1-4 and EG1-4 (or ER2-4 and EG2-4) may be due to some differences in the implementation of the considered schemes [32];

[32] The implementation of the Roe scheme, in particular, has been developed by repeat-

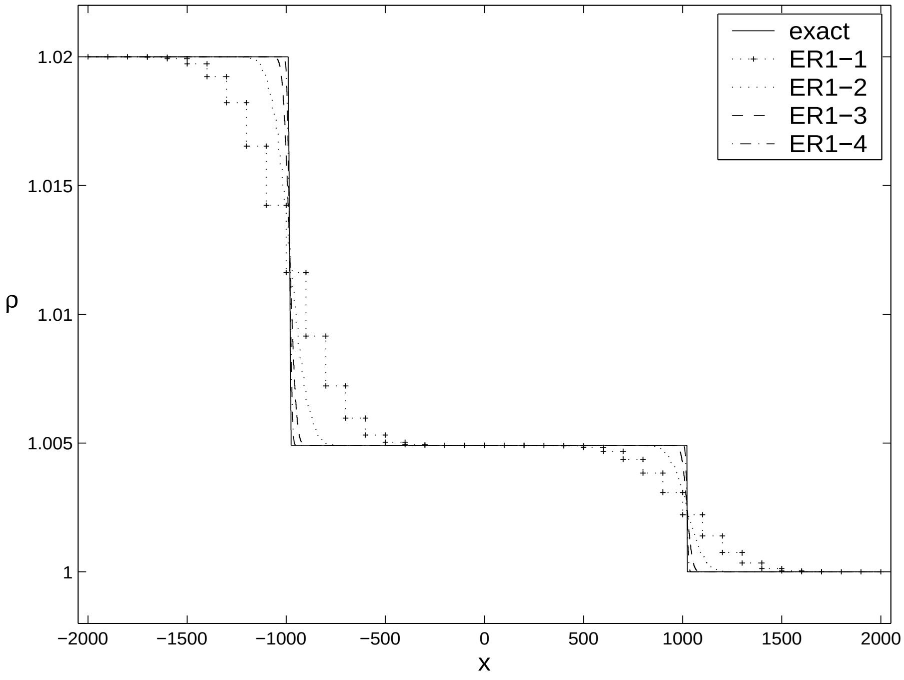

Figure 17: Approximation of ρ for the test-cases ER1-1 to ER1-4.

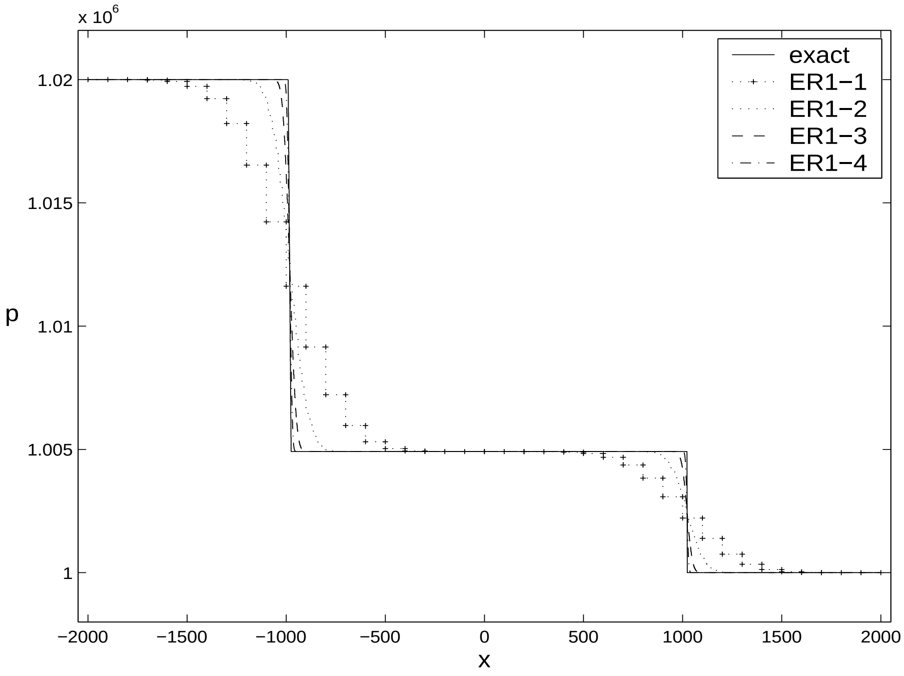

Figure 18: Approximation of p for the test-cases ER1-1 to ER1-4.

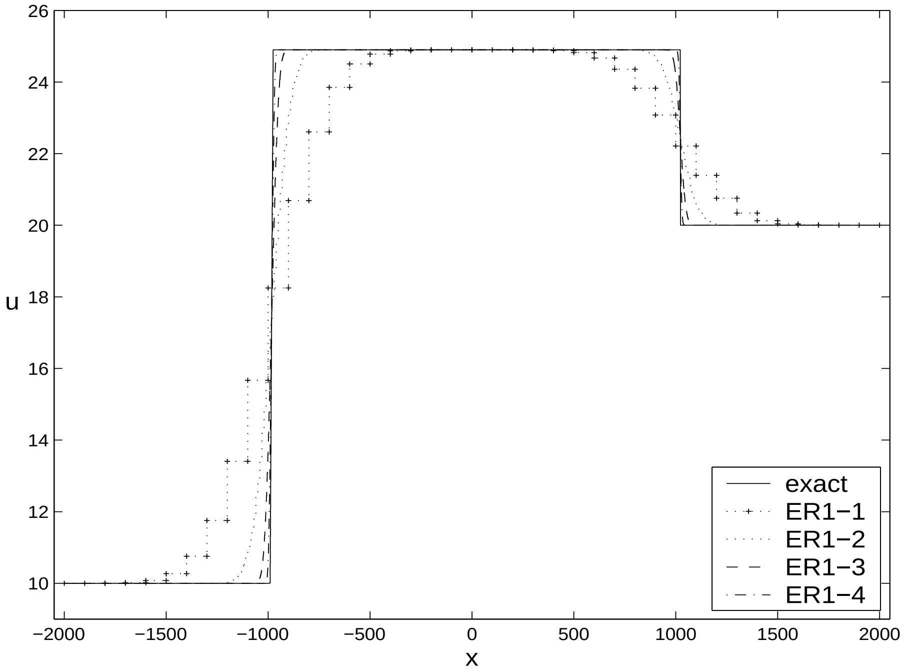

Figure 19: Approximation of u for the test-cases ER1-1 to ER1-4.

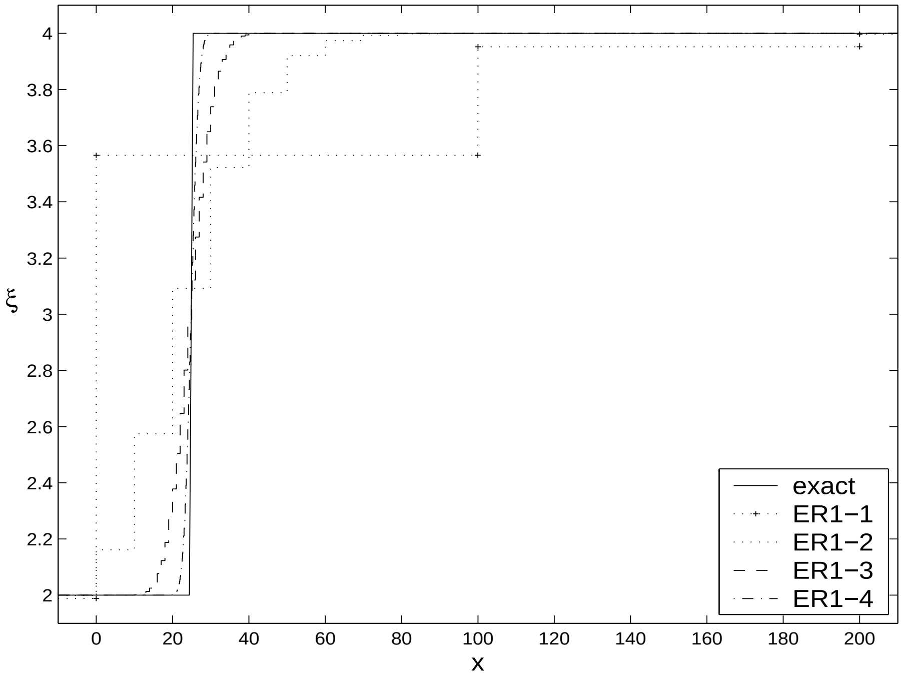

Figure 20: Approximation of ξ for the test-cases ER1-1 to ER1-4. The $x-$range is cut for ease of readability.

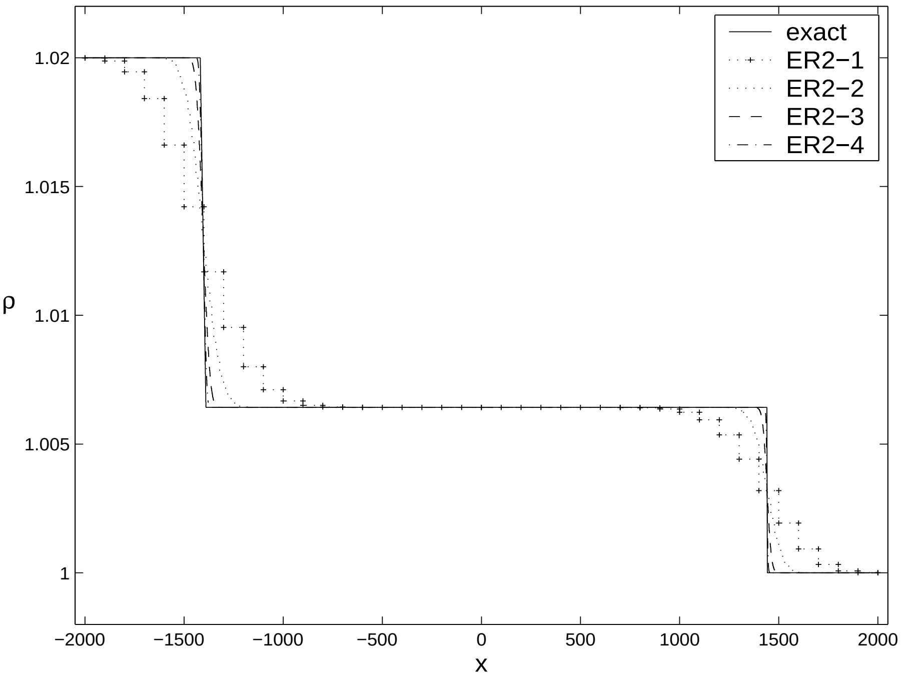

Figure 21: Approximation of ρ for the test-cases ER2-1 to ER2-4.

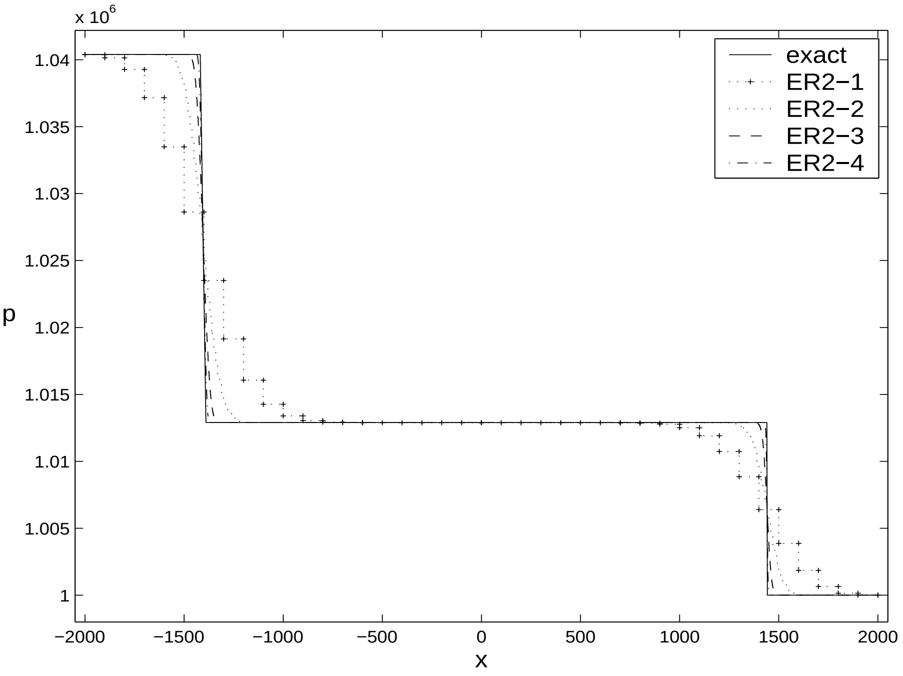

Figure 22: Approximation of p for the test-cases ER2-1 to ER2-4.

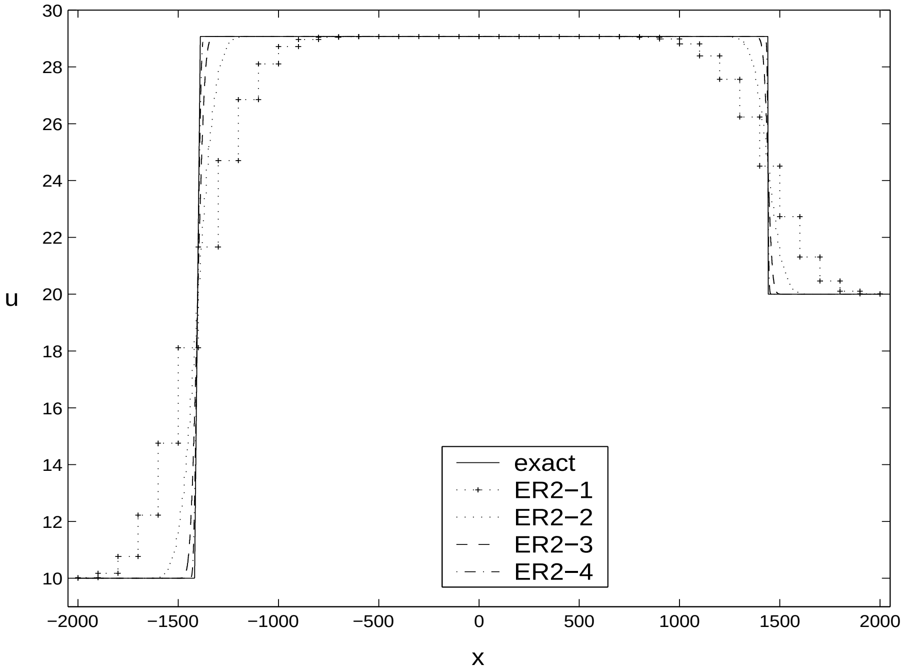

Figure 23: Approximation of u for the test-cases ER2-1 to ER2-4.

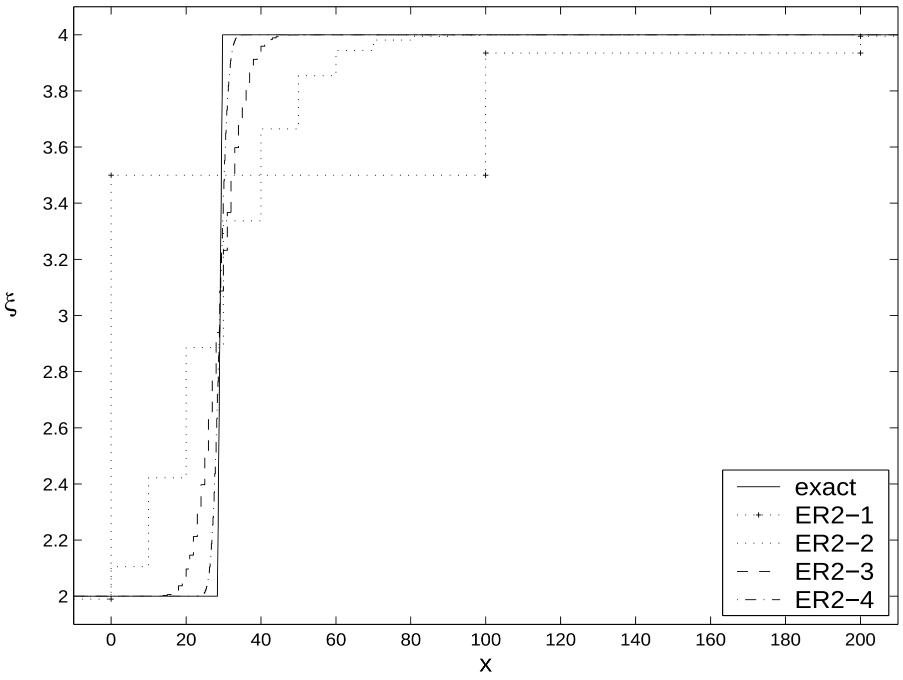

Figure 24: Approximation of ξ for the test-cases ER2-1 to ER2-4. The $x-$range is cut for ease of readability.

Test-case	$\tilde{c}^{(CFL)}$	t_{CPU}	$e(\rho)$	$e(p)$	$e(u)$	$e(\xi)$
ER1-1	0.51	$\approx$ 0.1 sec.	0.1792	0.1792	8.5737	4.3596
ER1-2	0.51	$\approx$ 1 sec.	0.0967	0.0967	4.6240	2.0329
ER1-3	0.51	$\approx$ 35 sec.	0.0492	0.0492	2.3530	1.0297
ER1-4	0.51	$\approx$ 40 min.	0.0211	0.0211	1.0110	0.4740
ER2-1	0.72	$\approx$ 0.1 sec.	0.1587	0.3185	8.6766	4.5910
ER2-2	0.72	$\approx$ 1 sec.	0.0837	0.1679	4.5757	2.0039
ER2-3	0.72	$\approx$ 35 sec.	0.0392	0.0786	2.1438	1.0514
ER2-4	0.72	$\approx$ 40 min.	0.0141	0.0282	0.7703	0.4492

Table 5: CFL estimate, CPU time and error estimates for the test-cases reported in Tab. 4.

- by comparing the error estimates in Tabs. 3 and 5, the discrete scheme based on the Roe flux turns out to behave similarly to that one exploiting the Godunov flux. This result could be partly related to the fact that, for low Mach numbers, the shocks and the rarefactions appearing in the solution of the original (non-linear) RP tend to be close to the corresponding discontinuities in the Roe-linearized RP. This observation can be applied locally, at the generic cell interface between the state vectors $\mathbf{q}_i^{(A)}$ and $\mathbf{q}_j^{(A)}$. For the sake of illustration, let $\tilde{a}$ and $\tilde{M} \ll 1$ respectively denote a characteristic sound speed and a characteristic Mach number of the considered flow field. Then, the aforementioned discontinuities in the Roe-linearized RP travel with a speed $u_{ij} \pm a_{ij}$. Since for low Mach numbers $a_{ij} \approx \tilde{a}$ and $u_{ij} \leq \max(u_i, u_j) \approx \pm\tilde{a}\tilde{M}$, it follows that the speed under consideration can be approximated by $\pm\,\tilde{a}\left(1 + \mathrm{O}\left(\tilde{M}\right)\right)$, as for the shocks and the rarefactions -which originate, in practice, discontinuities like the shocks- of the non-linear problem (see the relevant paragraph in sec. 3.2.2). As far as the contact discontinuity is concerned, its speed is given by u_{ij} for the Roe-linearized problem while for the non-linear one is given by $u_\star$. Let the distance of

edly calling some BLAS (Basic Linear Algebra Subprograms, see www.netlib.org/blas) routines, even when dealing with very small (i.e. 2-4 components) arrays. This point, together with the fact that no *ad-hoc* tuning has been performed for the aforementioned external library (in consideration of the underlying computing platform), may have introduced a certain amount of computational overhead, which becomes more evident for the longest simulations.

both ρ_i and ρ_j from a certain reference value, say $\tilde{\rho}$, be of the order of $\tilde{\rho}\tilde{M}$ (consistently with the fact that small density variations take place in nearly-incompressible flows); then, from the definitions (153) and (89) it follows that for low Mach numbers:

$$u_{ij} \approx \frac{u_i + u_j}{2} \left(1 + \mathrm{O}\left(\tilde{M}\right)\right)$$

$$u_\star \approx \frac{u_i + u_j}{2} \left(1 + \mathrm{O}\left(1\right)\right)$$

and therefore the asymptotic behaviour of the considered speeds is different. Furthermore, while u_{ij} is always contained between $\min\left(u_i, u_j\right)$ and $\max\left(u_i, u_j\right)$ since the relations in (153) are convex combinations, $u_\star$ does not necessarily belong to the aforementioned interval, as shown e.g. in Figs. 19 and 23. Nevertheless, the difference under consideration may not be accurately perceived by the error estimate $e(\xi)$ defined in (126), since the numerator of (126) for the present case is small with respect to the denominator (indeed the variation of ξ across the contact discontinuity is abrupt) for both cases.

According to Tab. 5, the discrete solution correctly converges towards the exact one [33]. Moreover, the convergence is sub-linear and roughly exhibits the same trend for all the considered entities, as shown in Figs. 25 and 26.

[33] For the test-cases ER1-1 to ER1-4, $e\left(\rho\right) = e\left(p\right)$ by virtue of the direct proportionality between ρ and p which is introduced by the state law associated with the benchmark B1.

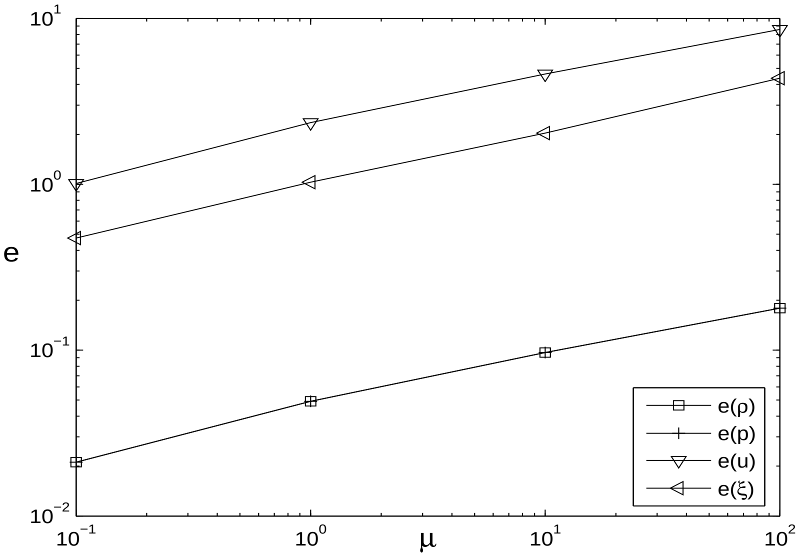

Figure 25: Plot of the error estimates for the test-cases ER1-1 to ER1-4 reported in Tab. 5.

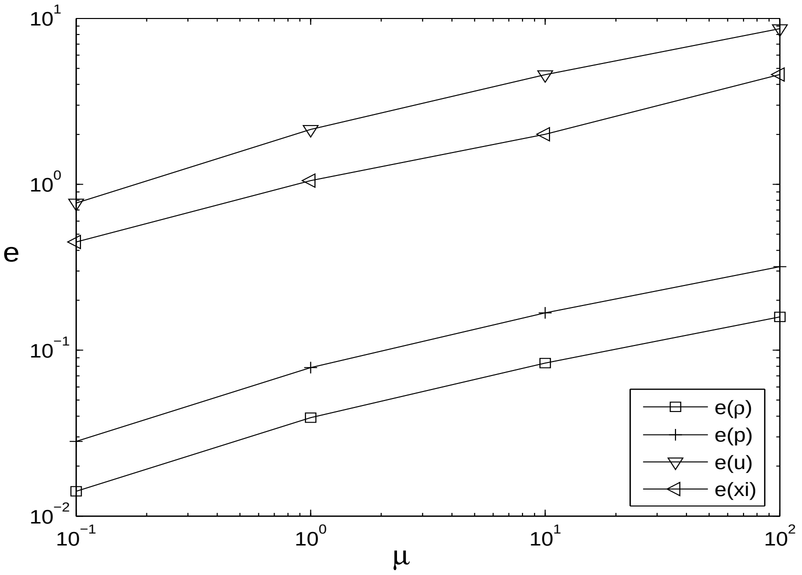

Figure 26: Plot of the error estimates for the test-cases ER2-1 to ER2-4 reported in Tab. 5.

3.4 Preconditioning of the Roe scheme for low Mach number flows

It is known from the literature that classical numerical schemes conceived for compressible flows (in particular, a variety of finite volume methods) exhibit accuracy problems when dealing with nearly-incompressible ones. In [42], for instance, the compressible Euler equations coupled with a perfect gas state law are considered. More in detail, the low Mach number asymptotic solution is investigated, as obtained by starting from the continuous formulation as well as from a semi-discrete one, of the type of (101) and based on a Roe flux function. It is shown, in particular, that the semi-discrete solution can exhibit pressure variations in space higher than those associated with the analytical one. Furthermore, a suitably modified numerical flux function is consequently introduced in order to counteract this discrepancy.

An investigation of the same type as that one in [42] is performed in the present section, by considering the basic-1D partial differential system (23) for the sake of simplicity [34]. More in detail, consistently with the fact that nearly-incompressible flow regions generally do not exhibit discontinuities, smooth (i.e. differentiable enough) solutions to (23) are considered [35].

The semi-discrete formulation, based on the Roe numerical flux proposed in sec. 3.3.1, which corresponds to (23) reads (compare with (101)):

$$\mu_i \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{q}_i^{(x)} + \sum_{j \in \pi_i} \boldsymbol{\phi}^{(x)ROE} \left(\mathbf{q}_i^{(x)}, \mathbf{q}_j^{(x)}, \hat{\boldsymbol{\nu}}_{ij} \right) = \mathbf{0} \quad , \quad i \in \mathcal{I} \tag{164}$$

where $\mathbf{q}^{(x)}$ is given by (20) and $\boldsymbol{\phi}^{(x)ROE}(\cdot, \cdot, \cdot)$ is straightforwardly derived from (156)-(160) as follows:

$$\boldsymbol{\phi}^{(x)ROE} \left(\mathbf{q}_i^{(x)}, \mathbf{q}_j^{(x)}, \hat{\boldsymbol{\nu}}_{ij} \right) := \boldsymbol{\phi}_{ij}^{(x)ROE} \quad , \quad j \in \pi_i \tag{165}$$

[34] Indeed, as remarked in Note 4 (sec. 2.2.4), the presence of the passive scalar does not affect the underlying 1D flow field. Hence, in principle, it suffices to consider the mass and momentum balances in order to highlight the problem under consideration.

[35] By virtue of this position, the solution to a RP associated with (23) is not of interest in the present case, even when it involves two rarefactions. Indeed, it is not difficult to see that both ρ -and therefore p- and u are strictly monotonic functions of x -for a fixed time- within the rarefaction fans (their derivative being proportional to the convexity marker $c(\rho)$ defined in (60)); consequently, the considered solution is continuous (since there is no contact discontinuity in the basic-1D case) but it is not differentiable.

with:

$$\begin{aligned}
\boldsymbol{\phi}_{ij}^{(x)ROE} &:= \boldsymbol{\phi}_{c,ij}^{(x)ROE} + \boldsymbol{\phi}_{u,ij}^{(x)ROE} \\
\boldsymbol{\phi}_{c,ij}^{(x)ROE} &:= \frac{1}{2}\, s_{ij} \left(\mathbf{f}_i^{(x)} + \mathbf{f}_j^{(x)} \right) \qquad (166) \\
\boldsymbol{\phi}_{u,ij}^{(x)ROE} &:= \mathbf{D}_{ij}^{(x)} \cdot \Delta^{ij} \mathbf{q}^{(x)} \\
\mathbf{D}_{ij}^{(x)} &:= -\frac{1}{2} \left| s_{ij}\, \tilde{\mathbf{J}}_{ij}^{(x)} \right| \qquad (167)
\end{aligned}$$

where s_{ij} is defined in (98) and, of course, $\mathbf{f}_i^{(x)}$ is understood as $\mathbf{f}^{(x)}\left(\mathbf{q}_i^{(x)}\right)$, with $\mathbf{f}^{(x)}(\cdot)$ defined in (21). Moreover, the Roe matrix $\tilde{\mathbf{J}}_{ij}^{(x)}$ in (167) is defined as follows (see Note 28 in sec. 3.3.1):

$$\tilde{\mathbf{J}}_{ij}^{(x)} := \begin{pmatrix} 0 & 1 \\ a_{ij}^2 - u_{ij}^2 & 2\, u_{ij} \end{pmatrix} \qquad (168)$$

where u_{ij} and a_{ij} are respectively given by (153) and (154).

By following [42], an asymptotic study is performed in sec. 3.4.1; it is shown, in particular, that also in the present (barotropic) case, for nearly-incompressible flows, there is a discrepancy between the behaviour of the solution of the continuous problem (23) and that one of the semi-discrete one (164). In sec. 3.4.2 a concise introduction to preconditioning techniques for low Mach flows is reported. Then, in sec. 3.4.3 the preconditioning technique originally proposed in [42] is applied to the proposed numerical flux (156)-(160) (in particular to its purely 1D counterpart (165)-(167)), with the aim of eliminating the discrepancy under consideration. Finally, in sec. 3.4.4 a discrete scheme based on the proposed preconditioned numerical flux is validated against a nearly-exact solution.

3.4.1 Low Mach number asymptotic study

Non-dimensionalization

In order to determine the behaviour of low Mach number asymptotic solutions, both the continuous problem (23) and the semi-discrete one (164) are non-dimensionalized by means of the following reference entities:

$$\begin{cases} x_{ref} & \\ u_{ref} & := \max\limits_{x \in D_x} u(x, t = 0) \\ \rho_{ref} & := \max\limits_{x \in D_x} \rho(x, t = 0) \\ a_{ref} & := \max\limits_{x \in D_x} a(x, t = 0) \\ t_{ref} & := x_{ref} \, u_{ref}^{-1} \\ p_{ref} & := \rho_{ref} \, a_{ref}^2 \end{cases} \tag{169}$$

where x_{ref} denotes a suitable reference length and D_x represents the $x-$domain (the remaining entities being understood). More in detail, each non-dimensional entity (namely the flow variables, the Roe averages, any relevant function like the state law, etc...) is defined dividing its dimensional counterpart by the proper reference value.
It must be noticed that a reference sound speed a_{ref} is explicitly introduced in (169) in order to directly take into account the compressibility effects [36].

Note 33 *No specific symbols are introduced in order to distinguish between the non-dimensional entities and their dimensional counterparts, for the sake of simplicity.*

The non-dimensional form of the continuous system (23), which is introduced in sec. A.1 for ease of presentation, reads:

$$\begin{cases} \partial_t (\rho) & = & & \Psi_c^{(0)} \\ \partial_t (\rho u) & = & M_\star^{-2} \, \Theta_c^{(-2)} & + \quad \Theta_c^{(0)} \end{cases} \tag{170}$$

[36] This position is not in contrast with that one mentioned in Note 3 (sec. 2.2).

where the relevant coefficients are defined in the aforementioned section [37]. The non-dimensional parameter $M_\star$ appearing in (170) is defined as follows:

$$M_\star := \frac{u_{ref}}{a_{ref}} \tag{171}$$

and plays a key role in the asymptotic study under consideration. Indeed, the nearly-incompressible limit of the considered equations is obtained for $M_\star \to 0$ (and therefore $M_\star$ is considered as a characteristic Mach number of the flow field).

For $M_\star \to 0$ the non-dimensional form of the semi-discrete system (164), which is derived in sec. A.2 for ease of presentation, reads (as usual, $i \in \mathcal{I}$):

$$\begin{cases} 2\mu_i \dfrac{\mathrm{d}}{\mathrm{d}t}(\rho_i) & = & & & M_\star^{-1}\,\Psi_{sd}^{(-1)} & + & \hat{\Psi}_{sd}^{(0)} \\ \\ 2\mu_i \dfrac{\mathrm{d}}{\mathrm{d}t}(\rho_i u_i) & = & M_\star^{-2}\,\Theta_{sd}^{(-2)} & + & M_\star^{-1}\,\Theta_{sd}^{(-1)} & + & \hat{\Theta}_{sd}^{(0)} \end{cases} \tag{172}$$

where the relevant coefficients are defined in the aforementioned section [38].

Asymptotic analysis

Following [42], solutions are sought to the continuous problem (170) and the semi-discrete one (172) in the nearly-incompressible limit (i.e. for $M_\star \to 0$), in the form of asymptotic expansions in power of $M_\star$, namely:

$$\begin{cases} \rho(x,t) & = & \rho_0(x,t) & + & M_\star\,\rho_1(x,t) & + & M_\star^2\,\rho_2(x,t) & + & \cdots \\ \\ u(x,t) & = & u_0(x,t) & + & M_\star\,u_1(x,t) & + & M_\star^2\,u_2(x,t) & + & \cdots \\ \\ p(x,t) & = & p_0(x,t) & + & M_\star\,p_1(x,t) & + & M_\star^2\,p_2(x,t) & + & \cdots \end{cases} \tag{173}$$

for the continuous system and:

$$\begin{cases} \rho_i(t) & = & \rho_{0i}(t) & + & M_\star\,\rho_{1i}(t) & + & M_\star^2\,\rho_{2i}(t) & + & \cdots \\ \\ u_i(t) & = & u_{0i}(t) & + & M_\star\,u_{1i}(t) & + & M_\star^2\,u_{2i}(t) & + & \cdots \\ \\ p_i(t) & = & p_{0i}(t) & + & M_\star\,p_{1i}(t) & + & M_\star^2\,p_{2i}(t) & + & \cdots \end{cases} \tag{174}$$

[37] In particular by (336), once understood the same symbol for corresponding dimensional and non-dimensional entities, as declared in Note 33 above.

[38] In particular by (346) and (349), once understood the same symbol for corresponding dimensional and non-dimensional entities, as declared in Note 33 above.

for the semi-discrete one. All the entities appearing in the expansions above are supposed to be regular enough (for any further manipulations).

Note 34 *By virtue of the considered barotropic state law (3), it is possible to derive some relations between the coefficients $\rho_k(\cdot,\cdot)$ and $p_h(\cdot,\cdot)$ appearing in (173). For instance,*

$$\begin{aligned} p = p(\rho) &= p\left(\rho_0 + M_\star \rho_1 + \cdots\right) \\ &= p\left(\rho_0\right) + M_\star\, a^2\left(\rho_0\right)\, \rho_1 + \cdots \end{aligned}$$

and therefore:

$$\begin{cases} p_0 &= p(\rho_0) \\ p_1 &= a^2(\rho_0)\,\rho_1 \end{cases}$$

Similar considerations can be introduced for the coefficients $\rho_{ki}(\cdot)$ and $p_{hi}(\cdot)$ in (174); thus, for instance, the following relations hold as well:

$$\begin{cases} p_{0i} &= p(\rho_{0i}) \\ p_{1i} &= a^2(\rho_{0i})\,\rho_{1i} \end{cases} \tag{175}$$

Note 35 *It is possible to exploit the equalities in (174) in order to also expand the Roe averages. For instance, the coefficient a_{0ij} appearing in the following expansion of a_{ij}:*

$$a_{ij}(t) = a_{0ij}(t) + M_\star\, a_{1ij}(t) + M_\star^2\, a_{2ij}(t) + \cdots$$

can be obtained as follows:

- *if $\Delta^{ij}\rho = 0$ then $\rho_i = \rho_j = \bar{\rho}$ and, according to (154) and (174), the following relation holds:*

  $$\begin{aligned} a_{ij} &= a\left(\bar{\rho}\right) \\ &= a\left(\bar{\rho}_0 + M_\star\,\bar{\rho}_1 + \cdots\right) \\ &= a\left(\bar{\rho}_0\right) + M_\star \frac{\mathrm{d}a}{\mathrm{d}\rho}\left(\bar{\rho}_0\right)\,\bar{\rho}_1 + \cdots \end{aligned}$$

  Then, clearly:

  $$a_{0ij} = a\left(\bar{\rho}_0\right)$$

- *if* $\Delta^{ij}\rho \neq 0$ *then, by only considering the zero-order terms in the expansion of the following equality (which is directly obtained from (154)):* $\Delta^{ij}p = a_{ij}^2\, \Delta^{ij}\rho$*, the subsequent relation is obtained:* $\Delta^{ij}p_0 = a_{0ij}^2\, \Delta^{ij}\rho_0$*. The coefficient* a_{0ij} *is positive since* a_{ij} *is positive (by definition) and* $a_{ij}\,(M_\star \to 0) \to a_{0ij}$*. Hence,* $\Delta^{ij}p_0 = 0 \Leftrightarrow \Delta^{ij}\rho_0 = 0$*. As a consequence,*
  - *if* $\Delta^{ij}\rho_0 \neq 0$*, then:*

$$\begin{aligned} a_{ij} &= \left(\frac{\Delta^{ij}p}{\Delta^{ij}\rho}\right)^{\frac{1}{2}} \\ &= \left(\frac{\Delta^{ij}p_0}{\Delta^{ij}\rho_0}\right)^{\frac{1}{2}} \left(1 + M_\star \frac{\Delta^{ij}p_1}{\Delta^{ij}p_0} + \cdots\right)^{\frac{1}{2}} \left(1 + M_\star \frac{\Delta^{ij}\rho_1}{\Delta^{ij}\rho_0} + \cdots\right)^{-\frac{1}{2}} \\ &= \left(\frac{\Delta^{ij}p_0}{\Delta^{ij}\rho_0}\right)^{\frac{1}{2}} + M_\star\, \frac{1}{2} \left(\frac{\Delta^{ij}p_0}{\Delta^{ij}\rho_0}\right)^{\frac{1}{2}} \left(\frac{\Delta^{ij}p_1}{\Delta^{ij}p_0} - \frac{\Delta^{ij}\rho_1}{\Delta^{ij}\rho_0}\right) + \cdots \end{aligned}$$

    In this case:

$$a_{0ij} = \left(\frac{\Delta^{ij}p_0}{\Delta^{ij}\rho_0}\right)^{\frac{1}{2}}$$

  - *if* $\Delta^{ij}\rho_0 = 0$*, then by exploiting the same kind of linearization as above, the following expression is obtained:*

$$a_{0ij} = \left(\frac{\Delta^{ij}p_k}{\Delta^{ij}\rho_k}\right)^{\frac{1}{2}}$$

    where k *denotes the first integer such that* $\Delta^{ij}\rho_k \neq 0$*.*

Of course, once defined the relevant coefficients appearing in the expansion of the Roe averages, it is possible to expand all the derived entities as well. For instance, once noticed that (as reminded above, $a_{0ij} > 0$*):*

$$a_{ij}^{-1} = a_{0ij}^{-1} \left(1 + M_\star \frac{a_{1ij}}{a_{0ij}} + \cdots\right)^{-1} = a_{0ij}^{-1} \left(1 - M_\star \frac{a_{1ij}}{a_{0ij}} + \cdots\right)$$

the parameter M_{ij} *introduced in (347) (sec. A.2) and reported below for the sake of clarity:*

$$M_{ij} := \frac{u_{ij}}{a_{ij}}$$

admits the following asymptotic expression:

$$M_{ij} = \frac{u_{0ij}}{a_{0ij}} + M_\star \left(\frac{u_{1ij}}{a_{0ij}} - \frac{u_{0ij}\, a_{1ij}}{a_{0ij}^2} \right) + \cdots$$

Hence, in particular, $M_{0ij} = \dfrac{u_{0ij}}{a_{0ij}}$*.*

By exploiting the aforementioned expansions, it is possible to state the following:

Proposition 7 *For $M_\star \to 0$, the pressure associated with the solution of the continuous problem (170) is of the form:*

$$p(x,t) = \bar{p}_0(t) + M_\star\, \bar{p}_1(t) + M_\star^2\, p_2(x,t) + \cdots \tag{176}$$

while that one relative to the semi-discrete problem (172) admits the following representation:

$$p_i(t) = \tilde{p}_0(t) + M_\star\, p_{1i}(t) + \cdots \tag{177}$$

Proof *The proof is reported in sec. A.3, for ease of presentation.* ■

By comparing the expansions (176) and (177) it is clear that, in the nearly-incompressible limit, the semi-discrete solution admits pressure variations in space higher than those associated with the continuous one. As a consequence, for $M_\star \to 0$ the discrete schemes based on the proposed Roe numerical flux (165)-(167) (i.e. (156)-(160)) may provide a numerical solution remarkably different from the continuous one. In other words, the accuracy of the considered compressible solvers can be dramatically reduced when the flow tends to become (even locally) nearly-incompressible.

Note 36 *An asymptotic behaviour of the same kind of that one described by the expansions (176) and (177) is obtained in [42], when considering the compressible Euler equations coupled with a perfect gas state law.*

3.4.2 A brief introduction to preconditioning techniques for the low speed Euler and Navier-Stokes equations

The considered preconditioning techniques originate from the "artificial compressibility" method proposed by Chorin [18] for determining a steady-state solution to the incompressible Navier-Stokes equations. When considering the two-dimensional incompressible equations written in terms of the so-called "primitive" variables p, u_1 and u_2 (where, of course, u_1 and u_2 denote the components of the velocity vector), the continuity equation reads:

$$\partial_{x_1} u_1 + \partial_{x_2} u_2 = 0$$

Then, by following the artificial compressibility method a fictitious pressure time-derivative is added to the above equation, as follows:

$$\kappa^{-1}\, \partial_t p + \partial_{x_1} u_1 + \partial_{x_2} u_2 = 0$$

where κ is a constant. Once completed the set of the governing equations by also considering the proper momentum balance, the original and the modified system only differ from each other as for the time-derivative term, which in the former case reads:

$$\partial_t \begin{pmatrix} 0 \\ u_1 \\ u_2 \end{pmatrix}$$

while in the latter one can be expressed as follows:

$$\mathbf{P}_{Chorin}^{-1} \cdot \partial_t \begin{pmatrix} p \\ u_1 \\ u_2 \end{pmatrix}$$

with:

$$\mathbf{P}_{Chorin}^{-1} := \begin{pmatrix} \kappa^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{178}$$

By introducing the aforementioned pressure time-derivative term, the decoupling between the pressure and the velocity field, which represents an important issue for the numerical discretization of the incompressible equations, is avoided and the hyperbolicity of the governing system is restored. However, the modified system is not consistent in time and therefore the considered formulation can only be exploited in order to march towards a steady-state solution, hopefully by guaranteeing a stable and efficient convergence through the definition of the parameter κ in $\mathbf{P}_{Chorin}^{-1}$.

Note 37 *In consideration of the fact that the artificial compressibility formulation results in a hyperbolic system of equations, its discretization is carried out in [5] by exploiting classical techniques conceived for compressible flows (namely a finite volume method based on upwind schemes and Riemann solvers). In particular, a linearized implicit time-advancing strategy is defined in which the parameter κ, originally associated with the fictitious pressure time-derivative, appears in such a way that the consistency in time is preserved. Hence, the considered scheme can be exploited for unsteady simulations as well.*

The basic idea of pre-multiplying the time-derivative term by a suitable matrix gave birth to a class of numerical methods designed for improving the convergence of the compressible Euler and Navier-Stokes equations to a steady-state. In particular, the time-derivatives are modulated (again, at the cost of loosing the consistency in time) in order to achieve a stable and efficient time-marching; in this spirit, the matrix under consideration is referred to as a preconditioner. The resulting schemes are generally referred to as "pseudo-unsteady" methods (see e.g. [75]) and the considered preconditioning technique is often indicated as "iterative preconditioning" (see e.g. [112]).

The numerical technique under consideration has been also exploited for the numerical simulation of flows in which there is a significant discrepancy between the convective and the acoustic speeds (i.e. time-scales). In particular, it has been used for reducing the numerical stiffness of compressible solvers dealing with flows at low Mach numbers. In this context, a well-known preconditioner has been put forward by Turkel [100]. Once introduced a state vector $\mathbf{z}$ such that $\mathrm{d}\mathbf{z} = \left((\rho\, a)^{-1}\, \mathrm{d}p, \mathrm{d}u_1, \mathrm{d}u_2, a\, c_P^{-1}\, \mathrm{d}s\right)^T$ where s indicates the entropy per unit mass of the fluid and c_P its specific heat at constant pressure (a denoting the sound speed), the expression of the considered preconditioner which appears in the compressible equations written in terms of $\mathbf{z}$ reads [101]:

$$\mathbf{P}_{Turkel}^{-1} := \begin{pmatrix} \dfrac{1}{\beta^2} & 0 & 0 & \delta \\ \dfrac{\alpha u_1}{a\,\beta^2} & 1 & 0 & 0 \\ \dfrac{\alpha u_2}{a\,\beta^2} & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{179}$$

where α, β and δ indicate suitable non-dimensional parameters. In particular β is chosen of the order of $\tilde{M}$, where $\tilde{M}$ denotes the characteristic Mach number of the flow field, for the considered preconditioning to be effective. For $\alpha = 0$ and $\delta = 1$ the matrix (179) reduces to another classical preconditioner introduced by Choi and Merkle [17] while for $\delta = 0$ it clearly generalizes that matrix of Chorin (178), by also altering the time-derivative appearing in the momentum balance.

The iterative preconditioning has been also introduced when considering upwind schemes like, for instance, the Roe scheme. The upwinding strategy, in particular, can be applied to the preconditioned formulation (see e.g. [109]). The resulting scheme, in general, may be not consistent in time. However, by confining the effect of the preconditioner within a portion of the numerical scheme which does not affect its consistency in time (as, for instance, that one associated with the upwind component of the Roe numerical flux when dealing with the corresponding scheme), it is possible to exploit the resulting scheme for unsteady simulations as well. A time-consistent preconditioning strategy is defined, in particular, in [42] and [112], which is recalled in the following sec. 3.4.3.

It may be worth mentioning that, as an alternative to the aforementioned approach, a dual time-step strategy is usually adopted in order to overcome the time-consistency problem (see e.g. [23], [58], [73] and [102]). More in detail, by starting from the following system:

$$\partial_t \mathbf{z} + \mathbf{r}\left(\mathbf{z}\right) = \mathbf{0}$$

in which $\mathbf{r}$ denotes the steady-state residual as a function of the chosen state vector $\mathbf{z}$, an additional term is added, as follows:

$$\mathbf{P}^{-1} \cdot \partial_\tau \mathbf{z} + \partial_t \mathbf{z} + \mathbf{r}\left(\mathbf{z}\right) = \mathbf{0}$$

where τ denotes a fictitious time and $\mathbf{P}^{-1}$ is a suitable matrix. More in detail, by advancing the solution of the latter system with respect to τ up to a steady-state [39], the solution of the former one is recovered; the consistency in time is clearly preserved. The matrix $\mathbf{P}^{-1}$ is designed for optimizing the aforementioned convergence and therefore represents a preconditioner (in the sense of the iterative preconditioning).

[39] Of course, in a practical set-up only a limited number of time-advancing steps are performed.

3.4.3 Preconditioning of the Roe numerical flux

As pointed out in Note 36 above, the asymptotic behaviours obtained for the perfect gas state law and for a generic barotropic state law are similar to each other. Hence, the preconditioning technique proposed in [42] is also considered for the barotropic case. Basically, it consists in replacing the Roe flux (165)-(167) with the following expression:

$$\boldsymbol{\phi}^{(x)ROE,p}\left(\mathbf{q}_i^{(x)},\mathbf{q}_j^{(x)},\hat{\boldsymbol{\nu}}_{ij}\right) := \boldsymbol{\phi}_{ij}^{(x)ROE,p} \quad , \quad j \in \pi_i \tag{180}$$

where:

$$\boldsymbol{\phi}_{ij}^{(x)ROE,p} := \boldsymbol{\phi}_{c,ij}^{(x)ROE} + \boldsymbol{\phi}_{u,ij}^{(x)ROE,p} \tag{181}$$

$$\boldsymbol{\phi}_{u,ij}^{(x)ROE,p} := \mathbf{D}_{ij}^{(x),p} \cdot \Delta^{ij}\mathbf{q}^{(x)}$$

$$\mathbf{D}_{ij}^{(x),p} := -\frac{1}{2}\left(\mathbf{P}_{ij}^{(x)}\right)^{-1} \cdot \left|\mathbf{P}_{ij}^{(x)} \cdot \left(s_{ij}\,\tilde{\mathbf{J}}_{ij}^{(x)}\right)\right| \tag{182}$$

and $\boldsymbol{\phi}_{c,ij}^{(x)ROE}$ in (181) is defined by (166). It should be noticed that only the upwind component of the flux function is modified, by means of the preconditioning matrix $\mathbf{P}_{ij}^{(x)}$ defined below. Let $\mathbf{w}_p^{(x)}$ denote the following basic-1D primitive state vector:

$$\mathbf{w}_p^{(x)} := \begin{pmatrix} p \\ u \end{pmatrix}$$

Then, the following matrix:

$$\mathbf{P}_{\mathbf{q}}^{(x)} := \partial_{\mathbf{w}_p^{(x)}}\mathbf{q}^{(x)} \cdot \mathbf{P}_{\mathbf{w}_p}^{(x)} \cdot \partial_{\mathbf{q}^{(x)}}\mathbf{w}_p^{(x)} \tag{183}$$

where $\mathbf{q}^{(x)}$ denotes the basic-1D conservative state vector (20) and $\mathbf{P}_{\mathbf{w}_p}^{(x)}$ is defined as follows:

$$\mathbf{P}_{\mathbf{w}_p}^{(x)} := \begin{pmatrix} \beta^2 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad \beta = const \tag{184}$$

defines a function of u, namely:

$$\mathbf{P}_{\mathbf{q}}^{(x)}(u) = \mathbf{I} + \left(\beta^2 - 1\right)\begin{pmatrix} 1 & 0 \\ u & 0 \end{pmatrix} \tag{185}$$

The matrix $\mathbf{P}_{ij}^{(x)}$ is finally defined by evaluating (185) in correspondence of the proper Roe average, as follows:

$$\mathbf{P}_{ij}^{(x)} := \mathbf{P}_{\mathbf{q}}^{(x)}(u = u_{ij}) \tag{186}$$

It is worth remarking that the matrix $\mathbf{P}_{ij}^{(x)} \cdot \left(s_{ij}\, \tilde{\mathbf{J}}_{ij}^{(x)} \right)$ appearing in the preconditioned upwind component (182) is diagonalizable with real eigenvalues (see sec. A.4 for details) and therefore the operator $|\cdot|$, defined by (6), can be rightfully applied.

Note 38 *The numerical flux (180)-(182) is generally referred to as the Roe-Turkel numerical flux (see e.g. [42]). Indeed, the matrix (184) can be derived from the 1D counterpart of the preconditioner of Turkel (179) for $\alpha = \delta = 0$, by a change of variables* [40]*.*

Note 39 *The preconditioner is usually introduced by considering a quasi-linear formulation, consistently with the fact that regular solutions (e.g. without shocks) are investigated in the nearly-incompressible limit. Several sets of independent variables can be chosen in order to derive the preconditioner (see e.g. [102] and [110])* [41]*. For the present case, the preconditioner is introduced in terms of the primitive variables through the matrix (184) and then it is converted to the conservative variables by means of the expression (183). The specific form of the adopted state law comes into play at this point of the derivation; the expression (185), in particular, is valid for a (generic) barotropic state law.*

In order to assess the effects the considered preconditioning technique produces on the asymptotic semi-discrete solution, the following semi-discrete system is considered:

$$\mu_i \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{q}_i^{(x)} + \sum_{j \in \pi_i} \boldsymbol{\phi}^{(x)ROE,p} \left(\mathbf{q}_i^{(x)}, \mathbf{q}_j^{(x)}, \hat{\boldsymbol{\nu}}_{ij} \right) = \mathbf{0} \quad , \quad i \in \mathcal{I} \tag{187}$$

where the numerical flux $\boldsymbol{\phi}^{(x)ROE,p}(\cdot,\cdot,\cdot)$ is given by (180)-(182). For $M_\star \to 0$ the non-dimensional form of the semi-discrete system (187), which is derived in sec. A.4 for ease of presentation, reads (as usual, $i \in \mathcal{I}$):

$$\left\{ \begin{array}{lcrcrcr} 2\mu_i \dfrac{\mathrm{d}}{\mathrm{d}t}(\rho_i) & = & & & M_\star^{-1}\, \Psi_{sd,p}^{(-1)} & + & \hat{\Psi}_{sd,p}^{(0)} \\ & & & & & & \\ 2\mu_i \dfrac{\mathrm{d}}{\mathrm{d}t}(\rho_i u_i) & = & M_\star^{-2}\, \Theta_{sd,p}^{(-2)} & + & M_\star^{-1}\, \Theta_{sd,p}^{(-1)} & + & \hat{\Theta}_{sd,p}^{(0)} \end{array} \right. \tag{188}$$

[40] For $\alpha = \delta = 0$, the preconditioner of Turkel reduces to that one of Chorin, i.e. (178).

[41] The specific choice affects the convergence to a steady-state and the accuracy of the numerical solutions for low Mach number steady and unsteady flows [102].

where the relevant coefficients are defined in the aforementioned section [42].

The expansion (188) is obtained by assuming that the parameter β in (184) is formally of the order of the unity. However, by following [42], the parameter β is hereafter assumed of the order of the characteristic Mach number $M_\star$, namely:

$$\beta = \beta_{ref}\, M_\star \tag{189}$$

where β_{ref} is a given constant of the order of the unity. The position (189) renders the considered preconditioning technique effective, as shown in the sequel. First of all, the fact that now β explicitly introduces the factor $M_\star$ leads to a non-dimensional system different from (188). In particular, it is possible to show that for $M_\star \to 0$ the non-dimensional form of (187) now reads:

$$\left\{\begin{array}{lclclcl} 2\mu_i \dfrac{\mathrm{d}}{\mathrm{d}t}\left(\rho_i\right) & = & M_\star^{-2}\,\check{\Psi}^{(-2)}_{sd,p} & + & M_\star^{-1}\,\check{\Psi}^{(-1)}_{sd,p} & + & \ddot{\Psi}^{(0)}_{sd,p} \\ 2\mu_i \dfrac{\mathrm{d}}{\mathrm{d}t}\left(\rho_i u_i\right) & = & M_\star^{-2}\,\check{\Theta}^{(-2)}_{sd,p} & + & M_\star^{-1}\,\check{\Theta}^{(-1)}_{sd,p} & + & \ddot{\Theta}^{(0)}_{sd,p} \end{array}\right. \tag{190}$$

The definition of the coefficients appearing in (190) is not reported here because inessential to the present purposes. However, it should be noticed that the coefficient $\check{\Psi}^{(-2)}_{sd,p}$ in (190) has no counterpart in (188). Indeed, it derives from the position (189), by a mechanism of the type of that one mentioned in sec. A.5. The behaviour of the the system (190) in the nearly-incompressible limit is described by the following:

Proposition 8 *For $M_\star \to 0$, the pressure associated with the solution of the semi-discrete problem (190) admits the following representation:*

$$p_i(t) = \hat{p}_0(t) + M_\star\,\hat{p}_1(t) + M_\star^2\, p_{2i}(t) + \cdots \tag{191}$$

Proof *The proof is reported in sec. A.6, for ease of presentation.* ■

By comparing the expansions (191) and (176) it is clear that, in the nearly-incompressible limit, the solution associated with the preconditioned semi-discrete formulation exhibits a behaviour which is qualitatively similar to that of the continuous one. This should result, in principle, in a more accurate discrete solution for $M_\star \to 0$, as confirmed by the numerical results reported in the following sec. 3.4.4.

[42] In particular by (365) and (368), once understood the same symbol for corresponding dimensional and non-dimensional entities, as declared in Note 33 (sec. 3.4.1).

Note 40 *It is straightforward to extend the considered preconditioning technique to the augmented-1D systems defined in sec. 2.2.4. To the purpose, the following numerical flux function is introduced:*

$$\boldsymbol{\phi}^{(A)ROE,p}\left(\mathbf{q}_i^{(A)}, \mathbf{q}_j^{(A)}, \hat{\boldsymbol{\nu}}_{ij}\right) := \boldsymbol{\phi}_{ij}^{(A)ROE,p} \quad , \quad j \in \pi_i \tag{192}$$

with:

$$\begin{aligned}
\boldsymbol{\phi}_{ij}^{(A)ROE,p} &:= \boldsymbol{\phi}_{c,ij}^{(A)ROE} + \boldsymbol{\phi}_{u,ij}^{(A)ROE,p} \\
\boldsymbol{\phi}_{u,ij}^{(A)ROE,p} &:= \mathbf{D}_{ij}^{(A),p} \cdot \Delta^{ij}\mathbf{q}^{(A)} \\
\mathbf{D}_{ij}^{(A),p} &:= -\frac{1}{2}\left(\mathbf{P}_{ij}^{(A)}\right)^{-1} \cdot \left|\mathbf{P}_{ij}^{(A)} \cdot \left(s_{ij}\,\tilde{\mathbf{J}}_{ij}^{(A)}\right)\right|
\end{aligned} \tag{193}$$

where $\boldsymbol{\phi}_{c,ij}^{(A)ROE}$ *is formally given by (158) and:*

- *if* $\mathbf{q}^{(A)}$ *is defined by (24) then* $\mathbf{f}^{(A)}$ *is given by (25),* $\tilde{\mathbf{J}}_{ij}^{(A)}$ *is given by (152) and the preconditioner reads:*

$$\mathbf{P}_{ij}^{(A)} := \mathbf{I} + \left(\beta^2 - 1\right)\begin{pmatrix} 1 & 0 & 0 \\ u_{ij} & 0 & 0 \\ \xi_{ij} & 0 & 0 \end{pmatrix}$$

- *if* $\mathbf{q}^{(A)}$ *is given by (28) then* $\mathbf{f}^{(A)}$ *is given by (29),* $\tilde{\mathbf{J}}_{ij}^{(A)}$ *is defined analogously to (151) and the preconditioner reads:*

$$\mathbf{P}_{ij}^{(A)} := \mathbf{I} + \left(\beta^2 - 1\right)\begin{pmatrix} 1 & 0 & 0 & 0 \\ u_{ij} & 0 & 0 & 0 \\ \xi_{ij} & 0 & 0 & 0 \\ \eta_{ij} & 0 & 0 & 0 \end{pmatrix} \tag{194}$$

3.4.4 Numerical results

Benchmark

A quasi-1D, inviscid, barotropic flow within a duct having variable cross-sectional area $A = A(x)$ (e.g. a convergent-divergent nozzle) is considered. Let $\mathbf{q}^{(x)}$ and $\mathbf{f}^{(x)}$ be defined by (20) and (21), respectively. The relevant mass and momentum balances read (compare with (23)):

$$\partial_t \, \mathbf{q}^{(x)} + \partial_x \, \mathbf{f}^{(x)} = \mathbf{s}^{(x)} \left(\mathbf{q}^{(x)} \right) \tag{195}$$

with:

$$\mathbf{s}^{(x)} \left(\mathbf{q}^{(x)} \right) := - \frac{\mathrm{d}}{\mathrm{d}x} \ln \left(\, A(x) \, \right) \left(\begin{array}{c} \rho \, u \\ \rho \, u^2 \end{array} \right)$$

As declared at the beginning of sec. 3.4, a smooth solution to (195) is sought in the present context. In particular, the steady, nearly-incompressible solution to (195) described below is considered in order to define a benchmark for the proposed preconditioning strategy.

Let $M_\star \ll 1$ denote the characteristic Mach number of a flow field exhibiting a roughly constant density:

$$\rho \approx \rho_\infty \tag{196}$$

where the subscript ∞ hereafter refers to the inlet conditions, associated with the section located at $x = x_{min} \in [x_{min}, x_{max}]$. In consideration of (196), the conservation of the mass approximately reduces to the following relation:

$$u \, A \approx u_\infty \, A_\infty \tag{197}$$

Moreover, by invoking the well-known Bernoulli theorem (for incompressible, non-dissipative steady flows) [88], the momentum balance can be approximated as follows:

$$p + \frac{1}{2} \, \rho_\infty \, u^2 \approx p_\infty + \frac{1}{2} \, \rho_\infty \, u_\infty^2 \tag{198}$$

Then, once defined:

$$\alpha(x) := \frac{A(x)}{A_\infty} \tag{199}$$

it is possible to respectively recast (197) and (198) as follows:

$$\frac{u}{u_\infty} \approx \alpha^{-1} \tag{200}$$

$$\frac{p}{p_\infty} \approx 1 + \frac{1}{2} \, \frac{\rho_\infty \, u_\infty^2}{p_\infty} \left(1 - \alpha^{-2} \right) \tag{201}$$

Benchmark	κ	$\varkappa$	γ	ρ_∞	u_∞	x_{min}	x_1	x_2	x_{max}	σ
BN	10^6	1	0	1	1	-2000	-1000	1000	2000	$2.5 \cdot 10^{-2}$

Table 6: Considered benchmark.

The relations (200) and (201) provide a nearly-exact, steady solution to (195) (they tend to be exact for $M_\star \to 0$) which can be exploited for validating the discrete scheme (205), based on the preconditioned numerical flux (180)-(182).

As far as the variation of the cross-sectional area is concerned, the following definition is adopted, in particular, for $\alpha(x)$:

$$\alpha(x) := \begin{cases} 1 & \text{if} \quad x_{min} \leq x \leq x_1 \\ 1 - \sigma \left(1 - \cos\left(2\pi \dfrac{x - x_1}{x_2 - x_1}\right)\right) & \text{if} \quad x_1 < x < x_2 \\ 1 & \text{if} \quad x_2 \leq x \leq x_{max} \end{cases} \tag{202}$$

where x_1, x_2 and $0 < \sigma < 1/2$ are adjustable parameters. The function (202) represents a sinusoidal reduction of the cross-sectional area between x_1 and x_2. In particular, the minimum value of α is given by:

$$\alpha_{min} := 1 - 2\,\sigma \tag{203}$$

and it is obtained for $x = (x_1 + x_2)/2$.

The considered benchmark is summarized in Tab. 6. In this table, κ, $\varkappa$ and γ characterize an instance of the convex barotropic state law (71) which introduces, in particular, a constant sound speed $\tilde{a} = \sqrt{\kappa} = 10^3$ for the flow. The variation of the cross-sectional area $A/A_\infty = \alpha$ is shown in Fig. 27. In consideration of (200), the maximum value of u is given by $u \approx \alpha_{min}^{-1} \approx 1.05$ and therefore $M_\star = 10^{-3}$ can be regarded to as a characteristic Mach number of the considered flow.

Discrete scheme, initial and boundary conditions

By starting from the system (164), the following semi-discrete formulation is introduced, based on the proposed preconditioned numerical flux (180)-(182):

$$\mu_i \frac{\mathrm{d}}{\mathrm{d}t}\, \mathbf{q}_i^{(x)} + \sum_{j \in \pi_i} \boldsymbol{\phi}^{(x)ROE,p}\left(\mathbf{q}_i^{(x)}, \mathbf{q}_j^{(x)}, \hat{\boldsymbol{\nu}}_{ij}\right) = \mathbf{s}_i^{(x)} \quad , \quad i \in \mathcal{I} \tag{204}$$

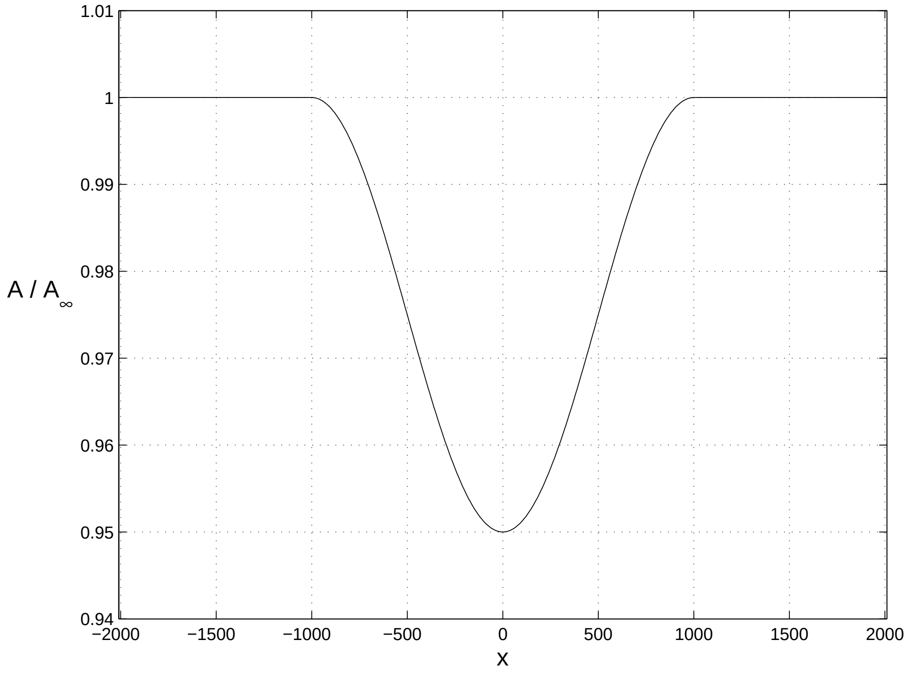

Figure 27: Variation of the cross-sectional area for the test-case under consideration.

where:

$$\mathbf{s}_i^{(x)} \approx \int_{C_i} \mathbf{s}^{(x)}\, \mathrm{d}x$$

In particular, the following definition is adopted:

$$\mathbf{s}_i^{(x)} := \gamma_i \begin{pmatrix} \rho_i\, u_i \\ \rho_i\, u_i^2 \end{pmatrix}$$

with:

$$\gamma_i := \ln\left(\frac{A(x_{i-1/2})}{A(x_{i+1/2})}\right) = \ln\left(\frac{\alpha(x_{i-1/2})}{\alpha(x_{i+1/2})}\right)$$

Then, in the spirit of the basic explicit scheme (102), the following discretization is considered:

$$\mathbf{q}_i^{(x)n+1} = \mathbf{q}_i^{(x)n} + \frac{\delta^n t}{\mu_i} \left(- \sum_{j\in\pi_i} \boldsymbol{\phi}^{(x)ROE,p} \left(\mathbf{q}_i^{(x)n}, \mathbf{q}_j^{(x)n}, \hat{\boldsymbol{\nu}}_{ij} \right) + \mathbf{s}_i^{(x)n} \right) \quad , \quad i \in \mathcal{I} \tag{205}$$

The following uniform field is chosen as IC:

$$\mathbf{q}_i^{(x)0} = \mathbf{q}_\infty^{(x)} \quad , \quad i \in \mathcal{I} \tag{206}$$

where:

$$\mathbf{q}_{\infty}^{(x)} := \begin{pmatrix} \rho_{\infty} \\ \rho_{\infty}\, u_{\infty} \end{pmatrix} \tag{207}$$

and both ρ_{∞} and u_{∞} are introduced in the previous paragraph.

As far as the BCs are concerned, a Dirichlet-like inlet BC is enforced by defining the fictitious state vector $\mathbf{q}_0^{(x)n}$ as follows:

$$\mathbf{q}_0^{(x)n} = \mathbf{q}_{\infty}^{(x)} \quad , \quad n = 0, 1, 2, \ldots \tag{208}$$

while the following transmissive BC is adopted at the outlet:

$$\mathbf{q}_{N_c+1}^{(x)n} = \mathbf{q}_{N_c}^{(x)n} \quad , \quad n = 0, 1, 2, \ldots \tag{209}$$

A remark on linear stability

The explicit discrete scheme (205) is subjected to a CFL-like constraint of the type of (121), as the basic explicit scheme (102) from which it is derived. Independently of $\mathbf{s}_i^{(x)}$, which accounts for the specific geometry of the duct, it makes sense to investigate the effect the preconditioning strategy has on the wave structure of the linearized problem and in particular on the corresponding maximum wave speed s_{max}, since it directly affects the CFL-like constraint. To the purpose, some considerations are reported below; a rather informal presentation is adopted for the sake of simplicity.

An estimate of s_{max} can be obtained by considering the spectral radius of the matrix $\mathbf{D}_{ij}^{(x),p}$ defined in (182). In particular, it is possible to linearize the flow field in the neighbourhood of a certain point $\tilde{\mathbf{q}}^{(x)}$ and, by virtue of the property (RM2) reported in sec. 3.3.1, to evaluate the spectral radius of $\mathbf{D}_{ij}^{(x),p}$ in correspondence of $\tilde{\mathbf{q}}^{(x)}$ (see e.g. [8]). Straightforward algebraic manipulations (not reported here for the sake of conciseness) show that:

$$M_{\star} \ll 1 \Rightarrow \mathbf{D}_{ij}^{(x),p} \approx \begin{pmatrix} \mathrm{O}\left(M_{\star}^{-1}\, \tilde{a}\right) & \mathrm{O}\left(M_{\star}^{-2}\right) \\ \\ \mathrm{O}\left(\tilde{a}^2\right) & \mathrm{O}\left(M_{\star}^{-1}\, \tilde{a}\right) \end{pmatrix}$$

where $\tilde{a}$ denote the characteristic sound speed associated with $\tilde{\mathbf{q}}^{(x)}$. Clearly, $s_{max} = \mathrm{O}\left(M_{\star}^{-1}\, \tilde{a}\right)$ for the present, preconditioned case. On the other hand, the non-preconditioned case can be analysed exactly in the same manner (in particular, by simply choosing $\beta^2 = 1$ where appropriate), thus obtaining:

$$M_{\star} \ll 1 \Rightarrow \mathbf{D}_{ij}^{(x)} \approx \begin{pmatrix} \mathrm{O}\left(\tilde{a}\right) & \mathrm{O}\left(M_{\star}\right) \\ \\ \mathrm{O}\left(M_{\star}\, \tilde{a}^2\right) & \mathrm{O}\left(\tilde{a}\right) \end{pmatrix}$$

Test-case	Benchmark	μ	β^2	τ
ER-NOPREC	BN	10	1	$5 \cdot 10^{-3}$
ER-PREC	BN	10	10^{-6}	$5 \cdot 10^{-6}$

Table 7: Considered test-cases for the discrete scheme (205), based on the preconditioned numerical flux (180)-(182).

In this case, $s_{max} = \mathrm{O}\left(\tilde{a}\right)$ (as already noticed, for instance, in sec. 3.3.2). On the basis of the aforementioned analysis, it is clear that the largest wave speed increases of $\mathrm{O}\left(M_\star^{-1}\right) \gg 1$ when switching the preconditioning technique on. In other words, in consideration of the CFL-like constraint (121), it should be necessary to reduce the time-step of $\mathrm{O}\left(M_\star\right) \ll 1$ in order to keep the explicit time-advancing stable:

$$\tau_{prec} = \mathrm{O}\left(M_\star\right) \cdot \tau_{noprec} \tag{210}$$

As a result, the considered (explicit preconditioned) discrete scheme (205) should exhibit very severe efficiency limitations. This point is confirmed by the numerical results reported in the following paragraph.

Test cases

The considered test-cases are summarized in Tab. 7. In particular:

- the $x-$domain $[x_{min}, x_{max}]$ is uniformly discretized (with cells having size μ) for both cases;
- for the test-case ER-NOPREC the proposed preconditioning strategy is not activated (indeed for $\beta^2 = 1$ the preconditioner (185) reduces to the identity matrix and therefore it does not modify the numerical flux function). Conversely, the test-case ER-PREC exploits the preconditioning strategy at hand, by choosing β_{ref} in (189) exactly equal to 1;
- the time-step τ for the test-case ER-NOPREC is chosen equal to that one associated with the test-case ER1-2 in Tab. 4 (sec. 3.3.2), since both the considered (non-preconditioned) test-cases are based on the same state law as well as the same space discretization. The value which is chosen for the test-case ER-PREC, instead, represents the maximum

time-step which can be adopted, as a matter of fact, in order to obtain a stable time-advancing (up to the benchmark steady-state).

The behaviour of the corresponding numerical solutions is shown in Figs. 28 and 29. It should be noticed that:

- the approximation of u does not suffer from the accuracy problems related to the low Mach number flow (its main driver being the area variation, according to (197)), while the approximation of p exhibits the problems highlighted in Proposition 7 (sec. 3.4.1). The proposed preconditioning strategy, however, seems to effectively counteract these problems, as shown in Fig. 29;

- the time-steps reported in Tab. 7 clearly satisfies the relation (210), thus supporting the considerations regarding the stability of the considered scheme which are introduced in the previous paragraph. The extremely small time-step required by the preconditioned scheme clearly indicates that an extension of the considered explicit time-advancing strategy to more complex test-cases (e.g. 3D industrial geometries) could be hardly affordable, due to the high computational cost of the simulation [43]. This result is not restricted to the very simple time-advancing strategy considered in (205); indeed, a similar time-step reduction can be observed, for instance, when adopting a classical $4-$th-order Runge-Kutta scheme [91].

[43] For instance, the test-case ER-PREC approximately requires 12 hours (CPU time on a laptop with Intel P4 CPU 2.66GHz, 512kB L2 cache and 512MB RAM) for reaching the steady-state.

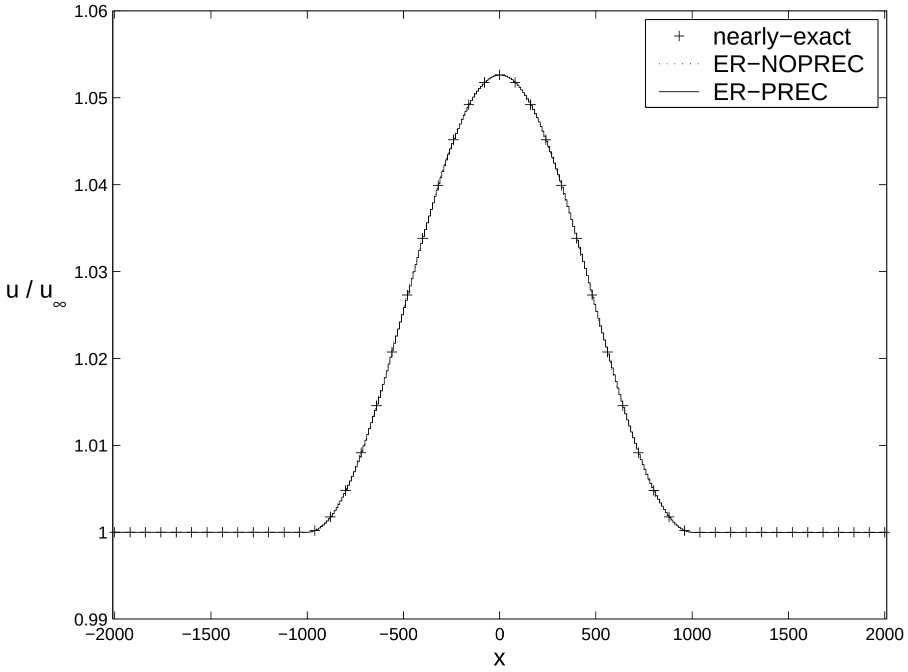

Figure 28: Approximation of u for the test-cases in Tab. 7.

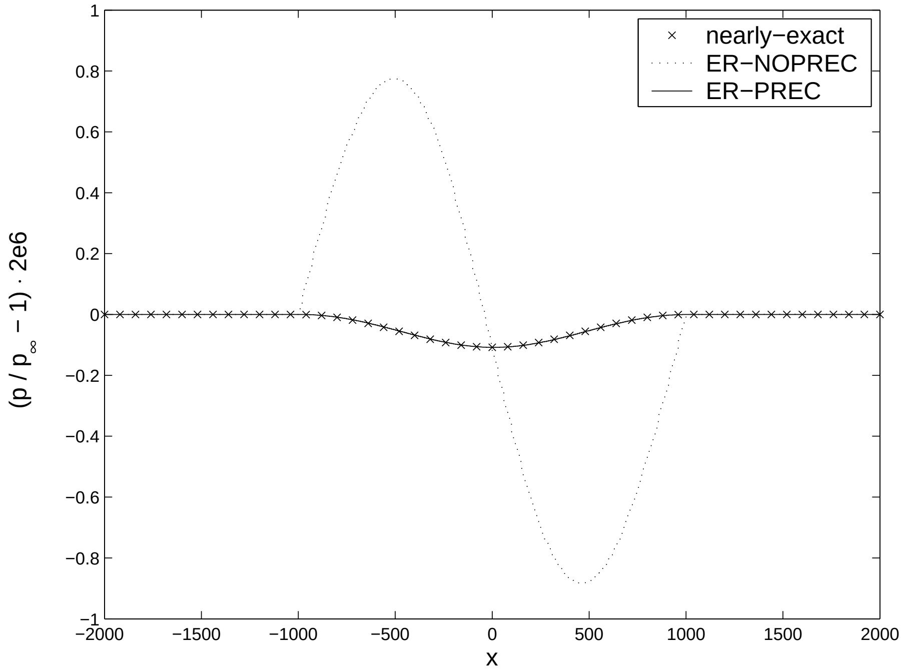

Figure 29: Approximation of p for the test-cases in Tab. 7. The pressure variation on the $y-$axis is scaled for ease of readability.

3.5 Linearized implicit time-advancing

An approximate linearization of the type of (105) is proposed in sec. 3.5.1, which can be applied to generic Roe numerical flux functions and in particular to (156)-(160). Then, in sec. 3.5.2 the proposed linearization is generalized so as to be applied to preconditioned Roe numerical flux functions of the type of (192)-(193). Furthermore, a second-order accurate scheme is briefly introduced in sec. 3.5.3, defined through a "Defect Correction" technique based on the proposed linearization. Finally, in secs. 3.5.4 and 3.5.5, the discrete solution obtained by exploiting the proposed linearization strategy is respectively validated against a nearly-exact and an exact benchmark.

3.5.1 A linearization of a generic Roe numerical flux function

Linearization of a Roe numerical flux $\boldsymbol{\phi}_{LR}^{(g)ROE}$

Let $\boldsymbol{\phi}_{LR}^{(g)ROE}$ denote a Roe numerical flux of the type of (131)-(134), associated with a generic hyperbolic problem hereafter reminded by superscript (g) (where appropriate).

Note 41 *No specific assumptions on the considered state law are introduced at this stage of the discussion. In particular, the application of the linearization strategy proposed below is not restricted to problems associated with a (generic) barotropic state law.*

In order to derive an approximate linearization of $\boldsymbol{\phi}_{LR}^{(g)ROE}$ of the type of (105), two matrices $\mathbf{A}_{LR}^{(g)n}$ and $\mathbf{B}_{LR}^{n}$ are sought, such that the following relation holds:

$$\delta^n \boldsymbol{\phi}_{LR}^{(g)ROE} \approx \mathbf{A}_{LR}^{(g)n} \cdot \delta^n \mathbf{q}_L^{(g)} + \mathbf{B}_{LR}^{(g)n} \cdot \delta^n \mathbf{q}_R^{(g)} \tag{211}$$

where the increment $\delta^n(\cdot)$ is defined in (103). Clearly, the Roe numerical flux is not differentiable and therefore the linearization (211) cannot be obtained by a first-order Taylor expansion. Moreover, if the analytical flux $\mathbf{f}^{(g)}$ is a first-order homogeneous function [44], the following relation is satisfied (by definition):

$$\mathbf{f}^{(g)} = \mathbf{J}^{(g)} \cdot \mathbf{q}^{(g)} \tag{212}$$

where, of course:

$$\mathbf{J}^{(g)} := \partial_{\mathbf{q}^{(g)}} \mathbf{f}^{(g)} \tag{213}$$

and it is possible to recast $\boldsymbol{\phi}_{LR}^{(g)}$ as follows:

$$\boldsymbol{\phi}_{LR}^{(g)} = \mathbf{F} \cdot \mathbf{q}_L^{(g)} + \mathbf{G} \cdot \mathbf{q}_R^{(g)} \tag{214}$$

[44] As, for instance, for the Euler equations associated with a perfect gas state law.

where $\mathbf{F} = \mathbf{F}\left(\mathbf{q}_L^{(g)}, \mathbf{q}_R^{(g)}\right)$ and $\mathbf{G} = \mathbf{G}\left(\mathbf{q}_L^{(g)}, \mathbf{q}_R^{(g)}\right)$ are suitably defined matrices (see below). Then, by assuming $\mathbf{F}$ and $\mathbf{G}$ to be weakly dependent on their arguments, it is possible to choose $\mathbf{A}_{LR}^{(g)n}$ and $\mathbf{B}_{LR}^{(g)n}$ in (211) as follows:

$$\mathbf{A}_{LR}^{(g)n} = \mathbf{F}\left(\mathbf{q}_L^{(g)n}, \mathbf{q}_R^{(g)n}\right) \quad , \quad \mathbf{B}_{LR}^{(g)n} = \mathbf{G}\left(\mathbf{q}_L^{(g)n}, \mathbf{q}_R^{(g)n}\right)$$

This is a rather classical approach for obtaining an approximate linearization of the type of (211) (see e.g. [36]). However, as pointed out in [91], there is no uniqueness as far as the choice of $\mathbf{F}$ and $\mathbf{G}$ is concerned. Indeed, by substituting (212) into the equalities which are obtained by formally replacing the superscript (A) with (g) in (135), it is straightforward to identify the following choices for $\mathbf{F}$ and $\mathbf{G}$:

$$\left\{ \begin{array}{lllll} \mathbf{F} & = & \mathbf{F}^{(1)}\left(\mathbf{q}_L^{(g)}, \mathbf{q}_R^{(g)}\right) & := & \mathbf{J}_L^{(g)} - \left(\tilde{\mathbf{J}}_{LR}^{(g)}\right)^{-} \\ & & & & \\ \mathbf{G} & = & \mathbf{G}^{(1)}\left(\mathbf{q}_L^{(g)}, \mathbf{q}_R^{(g)}\right) & := & \left(\tilde{\mathbf{J}}_{LR}^{(g)}\right)^{-} \end{array} \right. \tag{215}$$

$$\left\{ \begin{array}{lllll} \mathbf{F} & = & \mathbf{F}^{(2)}\left(\mathbf{q}_L^{(g)}, \mathbf{q}_R^{(g)}\right) & := & \left(\tilde{\mathbf{J}}_{LR}^{(g)}\right)^{+} \\ & & & & \\ \mathbf{G} & = & \mathbf{G}^{(2)}\left(\mathbf{q}_L^{(g)}, \mathbf{q}_R^{(g)}\right) & := & \mathbf{J}_R^{(g)} - \left(\tilde{\mathbf{J}}_{LR}^{(g)}\right)^{+} \end{array} \right. \tag{216}$$

where $\mathbf{J}_s^{(g)}$ ($s \in \{L, R\}$) is naturally understood as $\mathbf{J}^{(g)}\left(\mathbf{q}_s^{(g)}\right)$ and, of course, $\tilde{\mathbf{J}}_{LR}^{(g)}$ denotes the relevant Roe matrix. Then, by exploiting (215) and (216), the following class of approximate linearizations can be introduced [91]:

$$(\mathbf{F}, \mathbf{G}) = \gamma \left(\mathbf{F}^{(1)}, \mathbf{G}^{(1)}\right) + (1 - \gamma)\left(\mathbf{F}^{(2)}, \mathbf{G}^{(2)}\right) \tag{217}$$

where γ is a free parameter.

It is possible to propose a linearization of the type of (211) even when the first-order homogeneity condition (212) is not verified (as, for instance, for the case of the state vectors introduced in sec. 2.2, associated with a generic barotropic state law), by virtue of the following:

Proposition 9 *The Roe numerical flux function $\boldsymbol{\phi}_{LR}^{(g)ROE}$ satisfies the following relation:*

$$\delta^n \boldsymbol{\phi}_{LR}^{(g)ROE} = \left(\tilde{\mathbf{J}}_{LR}^{(g)n}\right)^+ \cdot \delta^n \mathbf{q}_L^{(g)} + \left(\tilde{\mathbf{J}}_{LR}^{(g)n}\right)^- \cdot \delta^n \mathbf{q}_R^{(g)} + \frac{1}{2}\, \mathbf{r}_{LR}^{(g),n,n+1} \tag{218}$$

where $\mathbf{r}_{LR}^{(g),n,n+1}$ is defined as follows:

$$\begin{aligned}
\mathbf{r}_{LR}^{(g),n,n+1} := \quad & \left(\Delta_L \left(\tilde{\mathbf{J}}_{LR}^{(g)}\right)^+ + \Delta_R \left(\tilde{\mathbf{J}}_{LR}^{(g)}\right)^+\right) \cdot \delta^n \mathbf{q}_L^{(g)} && + \\
& \left(\Delta_L \left(\tilde{\mathbf{J}}_{LR}^{(g)}\right)^- + \Delta_R \left(\tilde{\mathbf{J}}_{LR}^{(g)}\right)^-\right) \cdot \delta^n \mathbf{q}_R^{(g)} && + \\
& \left(\Delta_R \left(\tilde{\mathbf{J}}_{LR}^{(g)}\right)^- - \Delta_L \left(\tilde{\mathbf{J}}_{LR}^{(g)}\right)^+\right) \cdot \Delta^{LR} \mathbf{q}^{(g)n} && + \\
& \left(\bar{\Delta}_L \left(\tilde{\mathbf{J}}_{LR}^{(g)}\right)^+ - \bar{\Delta}_R \left(\tilde{\mathbf{J}}_{LR}^{(g)}\right)^-\right) \cdot \Delta^{LR} \mathbf{q}^{(g)n+1} &&
\end{aligned} \tag{219}$$

with $\Delta^{LR}(\cdot)$ introduced in (129) and:

$$\begin{cases}
\Delta_L \left(\tilde{\mathbf{J}}_{LR}^{(g)}\right)^\pm := \tilde{\mathbf{J}}^\pm\left(\mathbf{q}_L^{(g)n+1}, \mathbf{q}_R^{(g)n}\right) - \tilde{\mathbf{J}}^\pm\left(\mathbf{q}_L^{(g)n}, \mathbf{q}_R^{(g)n}\right) \\
\Delta_R \left(\tilde{\mathbf{J}}_{LR}^{(g)}\right)^\pm := \tilde{\mathbf{J}}^\pm\left(\mathbf{q}_L^{(g)n}, \mathbf{q}_R^{(g)n+1}\right) - \tilde{\mathbf{J}}^\pm\left(\mathbf{q}_L^{(g)n}, \mathbf{q}_R^{(g)n}\right) \\
\bar{\Delta}_L \left(\tilde{\mathbf{J}}_{LR}^{(g)}\right)^\pm := \tilde{\mathbf{J}}^\pm\left(\mathbf{q}_L^{(g)n}, \mathbf{q}_R^{(g)n+1}\right) - \tilde{\mathbf{J}}^\pm\left(\mathbf{q}_L^{(g)n+1}, \mathbf{q}_R^{(g)n+1}\right) \\
\bar{\Delta}_R \left(\tilde{\mathbf{J}}_{LR}^{(g)}\right)^\pm := \tilde{\mathbf{J}}^\pm\left(\mathbf{q}_L^{(g)n+1}, \mathbf{q}_R^{(g)n}\right) - \tilde{\mathbf{J}}^\pm\left(\mathbf{q}_L^{(g)n+1}, \mathbf{q}_R^{(g)n+1}\right)
\end{cases} \tag{220}$$

where, finally, $\tilde{\mathbf{J}}^\pm\left(\mathbf{q}_L^{(g)}, \mathbf{q}_R^{(g)}\right)$ is obtained by applying the operators defined in (7) to the considered Roe matrix $\tilde{\mathbf{J}}_{LR}^{(g)} = \tilde{\mathbf{J}}\left(\mathbf{q}_L^{(g)}, \mathbf{q}_R^{(g)}\right)$.
Proof *The proof is reported in sec. A.7, for ease of presentation.* ■

In view of the aforementioned proposition, it is possible to state that, if for all $\left(\mathbf{q}_L^{(g)n}, \mathbf{q}_R^{(g)n}, \mathbf{q}_L^{(g)n+1}, \mathbf{q}_R^{(g)n+1}\right)$ in a same neighbourhood:

$$\left\|\mathbf{r}_{LR}^{(g),n,n+1}\right\| \ll \left\|\left(\tilde{\mathbf{J}}_{LR}^{(g)n}\right)^+ \cdot \delta^n \mathbf{q}_L^{(g)} + \left(\tilde{\mathbf{J}}_{LR}^{(g)n}\right)^- \cdot \delta^n \mathbf{q}_R^{(g)}\right\| \tag{221}$$

then, the following approximation can be considered:

$$\delta^n \boldsymbol{\phi}_{LR}^{(g)ROE} \approx \left(\tilde{\mathbf{J}}_{LR}^{(g)n}\right)^+ \cdot \delta^n \mathbf{q}_L^{(g)} + \left(\tilde{\mathbf{J}}_{LR}^{(g)n}\right)^- \cdot \delta^n \mathbf{q}_R^{(g)} \tag{222}$$

It is worth emphasizing that, since the relation (218) is obtained by only exploiting the algebraic properties of the Roe numerical flux function (see sec. A.7), the linearization (222) is independent of the specific state law. Hence, it can be applied to a variety of problems.

Note 42 *Let μ and τ denote the characteristic sizes of the space and time discretizations, respectively. If a certain degree of regularity is assumed for the discrete solution, then $\delta^n \mathbf{q}_s^{(g)}$, $s \in \{L, R\}$, is of the order of τ while $\Delta^{LR}\mathbf{q}^{(g)n}$ is of the order of μ. Furthermore, if the matrices $\left(\tilde{\mathbf{J}}_{LR}^{(g)}\right)^{\pm}$ are functions regular enough (e.g. Lipschitzian) with respect to their arguments, then the entities in (220) are of the order of τ. Hence, the condition (221) is verified since $\left\|\mathbf{r}_{LR}^{(g),n,n+1}\right\| = \mathrm{O}\left(\tau^2, \tau\mu\right)$ while the right-hand side term of (221) is of the order of τ. In this spirit, the proposed linearization (222) is thought to introduce an error which is formally of the order of $\mathrm{O}\left(\tau^2, \tau\mu\right)$.*

Linearization of a Roe numerical flux $\boldsymbol{\phi}_{ij}^{(A)ROE}$

As far as the linearization of the augmented-1D system of interest is concerned, it is possible to directly exploit the approximate linearization (222) at the only cost of formally replacing (g) with (A). Then, by recalling the considerations introduced when deriving $\boldsymbol{\phi}_{ij}^{(A)ROE}$ from $\boldsymbol{\phi}_{LR}^{(A)ROE}$ in sec. 3.3.1, it is straightforward to generalize the proposed linearization so as to take the orientation of $\hat{\boldsymbol{\nu}}_{ij}$ into account. To the purpose, it suffices to choose the right-hand terms of (106) as follows:

$$\left\{\begin{array}{lcl} \mathbf{A}^{(A)}\left(\mathbf{q}_i^{(A)n}, \mathbf{q}_j^{(A)n}, \hat{\boldsymbol{\nu}}_{ij}\right) & = & \left(s_{ij}\, \tilde{\mathbf{J}}_{ij}^{(A)n}\right)^+ \\ \\ \mathbf{B}^{(A)}\left(\mathbf{q}_i^{(A)n}, \mathbf{q}_j^{(A)n}, \hat{\boldsymbol{\nu}}_{ij}\right) & = & \left(s_{ij}\, \tilde{\mathbf{J}}_{ij}^{(A)n}\right)^- \end{array}\right. \tag{223}$$

In consideration of (223), the linear system (107) which is associated with the proposed linearized implicit scheme reads:

$$\mathbf{M}_{(-1)}^{i,n} \cdot \delta^n \mathbf{q}_{i-1}^{(A)} + \mathbf{M}_{(0)}^{i,n} \cdot \delta^n \mathbf{q}_i^{(A)} + \mathbf{M}_{(+1)}^{i,n} \cdot \delta^n \mathbf{q}_{i+1}^{(A)} = \mathbf{m}^{i,n} \quad , \quad i \in \mathcal{I} \tag{224}$$

where:

$$\left\{\begin{array}{lll}
\mathbf{M}^{i,n}_{(-1)} & := & \left(s_{i(i-1)}\,\tilde{\mathbf{J}}^{(A)n}_{i(i-1)}\right)^{-} \\
\mathbf{M}^{i,n}_{(0)} & := & \dfrac{\mu_i}{\delta^n t}\,\mathbf{I} \\
 & + & \left(s_{i(i-1)}\,\tilde{\mathbf{J}}^{(A)n}_{i(i-1)}\right)^{+} \\
 & + & \left(s_{i(i+1)}\,\tilde{\mathbf{J}}^{(A)n}_{i(i+1)}\right)^{+} \\
\mathbf{M}^{i,n}_{(+1)} & := & \left(s_{i(i+1)}\,\tilde{\mathbf{J}}^{(A)n}_{i(i+1)}\right)^{-} \\
\mathbf{m}^{i,n} & := & \boldsymbol{\phi}^{(A)ROE\,n}_{(i-1)i} - \boldsymbol{\phi}^{(A)ROE\,n}_{i(i+1)}
\end{array}\right. \tag{225}$$

Moreover, by recalling the definition of s_{ij} given in (98) as well as the relation (163), it is possible to simplify the representation of the coefficients on the right-hand side of (225) as follows:

$$\left\{\begin{array}{lll}
\mathbf{M}^{i,n}_{(-1)} & = & -\left(\tilde{\mathbf{J}}^{(A)n}_{(i-1)i}\right)^{+} \\
\mathbf{M}^{i,n}_{(0)} & = & \dfrac{\mu_i}{\delta^n t}\,\mathbf{I} \\
 & - & \left(\tilde{\mathbf{J}}^{(A)n}_{(i-1)i}\right)^{-} \\
 & + & \left(\tilde{\mathbf{J}}^{(A)n}_{i(i+1)}\right)^{+} \\
\mathbf{M}^{i,n}_{(+1)} & = & \left(\tilde{\mathbf{J}}^{(A)n}_{i(i+1)}\right)^{-} \\
\mathbf{m}^{i,n} & := & \boldsymbol{\phi}^{(A)ROE\,n}_{(i-1)i} - \boldsymbol{\phi}^{(A)ROE\,n}_{i(i+1)}
\end{array}\right. \tag{226}$$

3.5.2 Incorporation of the preconditioning strategy

Let $\boldsymbol{\phi}_{ij}^{(g)ROE,p}$ denote a generic Roe numerical flux function, formally obtained by replacing (A) with (g) in the definition (192)-(193) [45]. As for the non-preconditioned case, by exploiting the definition (7) together with the property (RM3) introduced in sec. 3.3.1, it is possible to recast $\boldsymbol{\phi}_{ij}^{(g)ROE,p}$ as follows (the superscript (g) is correctly introduced for the preconditioner as well):

$$\left\{ \begin{array}{lcl} \boldsymbol{\phi}_{ij}^{(g)ROE,p} & = & s_{ij}\,\mathbf{f}_i^{(g)} + \left(\mathbf{P}_{ij}^{(g)}\right)^{-1} \cdot \left(\mathbf{P}_{ij}^{(g)} \cdot \left(s_{ij}\,\tilde{\mathbf{J}}_{ij}^{(g)}\right)\right)^{-} \cdot \Delta^{ij}\mathbf{q}^{(g)} \\ \\ \boldsymbol{\phi}_{ij}^{(g)ROE,p} & = & s_{ij}\,\mathbf{f}_j^{(g)} - \left(\mathbf{P}_{ij}^{(g)}\right)^{-1} \cdot \left(\mathbf{P}_{ij}^{(g)} \cdot \left(s_{ij}\,\tilde{\mathbf{J}}_{ij}^{(g)}\right)\right)^{+} \cdot \Delta^{ij}\mathbf{q}^{(g)} \end{array} \right. \tag{227}$$

There is a close formal analogy between the relation (227) and the relation (383) introduced in sec. A.7 for proving the Proposition 9. In view of this point, it is possible to extend the relevant passages reported in the aforementioned section to the considered preconditioned numerical flux, thus obtaining the following relation:

$$\left\{ \begin{array}{lcl} \mathbf{A}^{(g)}\left(\mathbf{q}_i^{(g)n}, \mathbf{q}_j^{(g)n}, \hat{\boldsymbol{\nu}}_{ij}\right) & = & \left(\mathbf{P}_{ij}^{(g)n}\right)^{-1} \cdot \left(\mathbf{P}_{ij}^{(g)n} \cdot \left(s_{ij}\,\tilde{\mathbf{J}}_{ij}^{(g)n}\right)\right)^{+} \\ \\ \mathbf{B}^{(g)}\left(\mathbf{q}_i^{(g)n}, \mathbf{q}_j^{(g)n}, \hat{\boldsymbol{\nu}}_{ij}\right) & = & \left(\mathbf{P}_{ij}^{(g)n}\right)^{-1} \cdot \left(\mathbf{P}_{ij}^{(g)n} \cdot \left(s_{ij}\,\tilde{\mathbf{J}}_{ij}^{(g)n}\right)\right)^{-} \end{array} \right. \tag{228}$$

As for the non-preconditioned case, the proposed linearization (228) is only based on the algebraic properties of the Roe numerical flux function and therefore it can be applied to a variety of problems.

The formulation corresponding to the augmented-1D problem considered in the present document is straightforwardly obtained from (228) by a trivial change of notation (superscript (A) in place of (g)) and it is reported below for the sake of completeness:

$$\left\{ \begin{array}{lcl} \mathbf{A}^{(A)}\left(\mathbf{q}_i^{(A)n}, \mathbf{q}_j^{(A)n}, \hat{\boldsymbol{\nu}}_{ij}\right) & = & \left(\mathbf{P}_{ij}^{(A)n}\right)^{-1} \cdot \left(\mathbf{P}_{ij}^{(A)n} \cdot \left(s_{ij}\,\tilde{\mathbf{J}}_{ij}^{(A)n}\right)\right)^{+} \\ \\ \mathbf{B}^{(A)}\left(\mathbf{q}_i^{(A)n}, \mathbf{q}_j^{(A)n}, \hat{\boldsymbol{\nu}}_{ij}\right) & = & \left(\mathbf{P}_{ij}^{(A)n}\right)^{-1} \cdot \left(\mathbf{P}_{ij}^{(A)n} \cdot \left(s_{ij}\,\tilde{\mathbf{J}}_{ij}^{(A)n}\right)\right)^{-} \end{array} \right. \tag{229}$$

In consideration of the relation (229) (which clearly generalizes (223)), the linear system (107) associated with the proposed preconditioned linearized

[45] A generic state law is assumed at this stage of the discussion. Hence, for instance, besides the barotropic case specifically treated in the present document it is possible to consider the preconditioned numerical flux discussed in [42].

implicit scheme reads:

$$\mathbf{L}^{i,n}_{(-1)} \cdot \delta^n \mathbf{q}^{(A)}_{i-1} + \mathbf{L}^{i,n}_{(0)} \cdot \delta^n \mathbf{q}^{(A)}_{i} + \mathbf{L}^{i,n}_{(+1)} \cdot \delta^n \mathbf{q}^{(A)}_{i+1} = \mathbf{l}^{i,n} \quad , \quad i \in \mathcal{I} \tag{230}$$

where the relevant coefficients are straightforward generalizations of those reported in (226), namely:

$$\begin{cases} \mathbf{L}^{i,n}_{(-1)} & := & - & \left(\mathbf{P}^{(A)n}_{(i-1)i}\right)^{-1} \cdot \left(\mathbf{P}^{(A)n}_{(i-1)i} \cdot \tilde{\mathbf{J}}^{(A)n}_{(i-1)i}\right)^{+} \\ \mathbf{L}^{i,n}_{(0)} & := & & \dfrac{\mu_i}{\delta^n t}\mathbf{I} \\ & & - & \left(\mathbf{P}^{(A)n}_{(i-1)i}\right)^{-1} \cdot \left(\mathbf{P}^{(A)n}_{(i-1)i} \cdot \tilde{\mathbf{J}}^{(A)n}_{(i-1)i}\right)^{-} \\ & & + & \left(\mathbf{P}^{(A)n}_{i(i+1)}\right)^{-1} \cdot \left(\mathbf{P}^{(A)n}_{i(i+1)} \cdot \tilde{\mathbf{J}}^{(A)n}_{i(i+1)}\right)^{+} \\ \mathbf{L}^{i,n}_{(+1)} & := & & \left(\mathbf{P}^{(A)n}_{i(i+1)}\right)^{-1} \cdot \left(\mathbf{P}^{(A)n}_{i(i+1)} \cdot \tilde{\mathbf{J}}^{(A)n}_{i(i+1)}\right)^{-} \\ \mathbf{l}^{i,n} & := & & \boldsymbol{\phi}^{(A)ROE,p\ n}_{(i-1)i} - \boldsymbol{\phi}^{(A)ROE,p\ n}_{i(i+1)} \end{cases}$$

3.5.3 A second-order defect-correction scheme

A generalization of the considered linearized implicit scheme, i.e. (107) coupled with (223) or (229), is concisely introduced in the present section, based on the "Defect Correction" technique [67] (hereafter DeC) mentioned in the relevant paragraph of sec. 3.1.2.

The discrete scheme at hand can be regarded to as an instance of the iterative scheme (116). More in detail (of course, the discrete solution is here $\mathbf{z}_h = \mathbf{q}^{(A)}_h$):

- a single iteration is considered: $\lambda^n_{max} = 1$;
- the time discretization is obtained from the expression (109) by considering:
$$k = 1 \quad , \quad \alpha_1 = 1 \quad , \quad \mathbf{z}^{(n,k)}_h = \mathbf{q}^{(A)n}_h$$
This approximation is first-order accurate and the corresponding truncation error is formally $\mathrm{O}(\tau)$, where τ represents the characteristic size of the time discretization;

- the spatial component $\boldsymbol{\psi}_h^{(p)}$ of the non-linear operator $\boldsymbol{\mu}_h^{(p,k)}$ defined in (111) is based on the Roe numerical flux. This leads to a first-order accurate discretization (see e.g. [39]): $p = 1$. The corresponding error is formally $\mathrm{O}(\mu)$, μ representing the characteristic size of the space discretization;

- the linear operator $\mathbf{J}_h^{(q,k)}$ is defined by choosing $q = 1$. The term $\delta\boldsymbol{\psi}_h^{(q)}$ appearing in the relevant definition (113) is constructed, in particular, by exploiting the approximate linearization (222) which, as pointed out in Note 42 (sec. 3.5.1), formally introduces a discretization error $\mathrm{O}(\tau^2, \tau\mu)$.

The resulting scheme clearly introduces a discretization error $\mathrm{O}(\mu, \tau)$ and therefore it is only first-order accurate.

It is possible to increase the accuracy of the aforementioned scheme up to the second order by adopting a DeC strategy, as briefly outlined in sec. 3.1.2. To the purpose:

- a second-order backward finite difference approximation is derived from (109) by means of the following settings:

$$k = 2 \quad , \quad \alpha_2 = \frac{1 + 2\theta}{1 + \theta} \quad , \quad \mathbf{z}_h^{(n,k)} = (1 + \theta)\; \mathbf{q}_h^{(A)n} - \frac{\theta^2}{1 + \theta}\, \mathbf{q}_h^{(A)n-1}$$

  where:

$$\theta := \frac{\delta^n t}{\delta^{n-1} t}$$

  The truncation error associated with the approximation at hand is formally $\mathrm{O}(\tau^2)$;

- a second-order spatial discretization $\boldsymbol{\psi}_h^{(p)}$, with $p = 2$, is introduced by performing a MUSCL reconstruction [106] [107] [108] before evaluating the Roe numerical flux. According to this strategy, the Roe numerical flux between the cells C_i and C_j (towards C_j) is computed as follows:

$$\boldsymbol{\phi}^{(A)ROE}\left(\mathbf{q}_{[i]j}^{(A)}, \mathbf{q}_{i[j]}^{(A)}, \hat{\boldsymbol{\nu}}_{ij}\right)$$

  where $\boldsymbol{\phi}^{(A)ROE}$ represents the usual Roe flux function [46] while $\mathbf{q}_{[i]j}^{(A)}$ and $\mathbf{q}_{i[j]}^{(A)}$ denote suitably extrapolated values at the interface between C_i and C_j, respectively on the side of C_i and C_j. The considered extrapolation

[46] Any additional superscript, like that one denoting the preconditioning technique discussed in the previous sections, is here dropped, for the sake of simplicity.

is constrained (by exploiting the starting, piece-wise constant, discrete solution and therefore in a non-linear fashion) so as to avoid spurious oscillations (see e.g. [39], [64], [98] and many references cited therein);

- in the spirit of the DeC approach [67], a value $q < p$ is chosen for containing the computational cost associated with the inversion of the linear operator $\mathbf{J}_h^{(q,k)}$ (see sec. 3.1.2). Hence, in particular, the value $q = 1$ is adopted, as for the case discussed in the corresponding point of the previous list. In other words, the term $\delta\boldsymbol{\psi}_h^{(q)}$ is constructed by applying the proposed approximate linearization (222) to the starting, piece-wise constant, numerical solution. Of course, the corresponding discretization error is still $\mathrm{O}\left(\tau^2, \tau\mu\right)$.

In view of the aforementioned points, a single iteration (i.e. $\lambda_{max}^n = 1$) of the scheme (116) yields a discretization error $\mathrm{O}\left(\mu^2, \tau^2, \tau\mu\right) = \mathrm{O}\left(\epsilon^2\right)$, with $\epsilon := \max\left(\mu, \tau\right)$ and therefore a second-order accurate solution is obtained. Moreover, on the basis of some preliminary carried out numerical experiments [90], a sensible improvement in the solution behaviour is observed by slightly increasing λ_{max}^n, e.g. by performing 2 or 3 iterations. Hence, the DeC seems to be a promising strategy for defining high-order, efficient schemes based on the proposed linearization (222); further investigations on this subject is definitely recommended.

3.5.4 Numerical results for smooth flows

Benchmarks

The benchmark already introduced in sec. 3.4.4 (namely the quasi-1D flow within a duct having variable cross-sectional area) and summarized, in particular, in Tab. 6 is considered here, with the aim of directly comparing the proposed linearized implicit time-advancing strategy with the explicit one given by (205). The relevant definitions/considerations are tacitly recalled from the aforementioned section.

Discrete scheme, initial and boundary conditions

A linearized implicit discrete scheme can be derived from (204) by following the procedure which permits to obtain (107) from (101), at the only cost of extending the linearization to the term $\mathbf{s}^{(x)}$ as follows:

$$\mathbf{s}_i^{(x)n+1} \approx \mathbf{s}_i^{(x)n} + \mathbf{S}_i^{(x)n} \cdot \delta^n \mathbf{q}_i^{(x)} \tag{231}$$

where:

$$\mathbf{S}_i^{(x)n} := \partial_{\mathbf{q}^{(x)}} \partial \mathbf{s}^{(x)} \left(\mathbf{q}_i^{(x)n} \right) = \begin{pmatrix} 0 & 1 \\ \\ -\left(u_i^n\right)^2 & 2\, u_i^n \end{pmatrix}$$

In particular, it suffices to respectively incorporate $\mathbf{s}_i^{(x)n}$ and $\mathbf{S}_i^{(x)n}$ into the right-hand side term and the diagonal coefficient of a linear system of the type of (230), namely:

$$\hat{\mathbf{L}}_{(-1)}^{i,n} \cdot \delta^n \mathbf{q}_{i-1}^{(x)} + \hat{\mathbf{L}}_{(0)}^{i,n} \cdot \delta^n \mathbf{q}_{i}^{(x)} + \hat{\mathbf{L}}_{(+1)}^{i,n} \cdot \delta^n \mathbf{q}_{i+1}^{(x)} = \hat{\mathbf{l}}^{i,n} \quad , \quad i \in \mathcal{I} \tag{232}$$

with:

$$\left\{ \begin{array}{lcl} \hat{\mathbf{L}}_{(-1)}^{i,n} & := & - \quad \left(\mathbf{P}_{(i-1)i}^{(x)n} \right)^{-1} \cdot \left(\mathbf{P}_{(i-1)i}^{(x)n} \cdot \tilde{\mathbf{J}}_{(i-1)i}^{(x)n} \right)^{+} \\ \hat{\mathbf{L}}_{(0)}^{i,n} & := & \quad \dfrac{\mu_i}{\delta^n t} \mathbf{I} \\ & & - \quad \mathbf{S}_i^{(x)n} \\ & & - \quad \left(\mathbf{P}_{(i-1)i}^{(x)n} \right)^{-1} \cdot \left(\mathbf{P}_{(i-1)i}^{(x)n} \cdot \tilde{\mathbf{J}}_{(i-1)i}^{(x)n} \right)^{-} \\ & & + \quad \left(\mathbf{P}_{i(i+1)}^{(x)n} \right)^{-1} \cdot \left(\mathbf{P}_{i(i+1)}^{(x)n} \cdot \tilde{\mathbf{J}}_{i(i+1)}^{(x)n} \right)^{+} \\ \hat{\mathbf{L}}_{(+1)}^{i,n} & := & \quad \left(\mathbf{P}_{i(i+1)}^{(x)n} \right)^{-1} \cdot \left(\mathbf{P}_{i(i+1)}^{(x)n} \cdot \tilde{\mathbf{J}}_{i(i+1)}^{(x)n} \right)^{-} \\ \hat{\mathbf{l}}^{i,n} & := & \quad \mathbf{s}_i^{(x)n} + \boldsymbol{\phi}_{(i-1)i}^{(x)ROE,p\;n} - \boldsymbol{\phi}_{i(i+1)}^{(x)ROE,p\;n} \end{array} \right.$$

The uniform flow field given by (206) is chosen as IC while, as far as the BCs are concerned:

- the Dirichlet-like BC (208) clearly implies that:

$$\delta^n \mathbf{q}_0^{(x)} = \mathbf{0}$$

  and therefore it is naturally implemented as follows:

$$\hat{\mathbf{L}}_{(0)}^{1,n} \cdot \delta^n \mathbf{q}_1^{(x)} + \hat{\mathbf{L}}_{(+1)}^{1,n} \cdot \delta^n \mathbf{q}_2^{(x)} = \hat{\mathbf{l}}^{1,n}$$

- the transmissive BC (209) clearly implies that:

$$\delta^n \mathbf{q}_{N_c+1}^{(x)} = \delta^n \mathbf{q}_{N_c}^{(x)}$$

  and therefore it is naturally implemented as follows:

$$\hat{\mathbf{L}}_{(-1)}^{N_c,n} \cdot \delta^n \mathbf{q}_{N_c-1}^{(x)} + \left(\hat{\mathbf{L}}_{(0)}^{N_c,n} + \hat{\mathbf{L}}_{(+1)}^{N_c,n} \right) \cdot \delta^n \mathbf{q}_{N_c}^{(x)} = \hat{\mathbf{l}}^{N_c,n} \tag{233}$$

Test-case	Benchmark	μ	β^2	τ
IR-NOPREC	BN	10	1	$\approx\infty$
IR-PREC	BN	10	10^{-6}	$\approx\infty$

Table 8: Considered test-cases for the discrete scheme (232), based on the preconditioned numerical flux (180)-(182).

Test-cases

The considered test-cases are summarized in Tab. 8, in which BN denotes the considered benchmark (described in Tab. 6, sec. 3.4.4). In particular:

- the $x-$domain $[x_{min}, x_{max}]$ (with x_{min} and x_{max} defined in the aforementioned Tab. 6) is uniformly discretized (with cells having size μ) for both cases;
- for the test-case IR-NOPREC the proposed preconditioning strategy is not activated (indeed for $\beta^2 = 1$ the preconditioner (185) reduces to the identity matrix and therefore it does not modify the numerical flux function). Conversely, the test-case IR-PREC exploits the preconditioning strategy at hand, by choosing β_{ref} in (189) exactly equal to 1;
- it turns out that, for both the considered test-cases, a practically "unbounded" time-step can be adopted for advancing the numerical solution by means of the proposed linearized implicit strategy. In the carried out numerical experiments τ has been increased up to 10^5, thus reaching the steady-state solution in a very few (namely 2 to 5) iterations; the required CPU time is practically negligible.

The corresponding numerical solutions, shown in Figs. 30 and 31, are indistinguishable from their counterparts obtained by the explicit time-advancing (reported in Figs. 28 and 29, sec. 3.4.4).

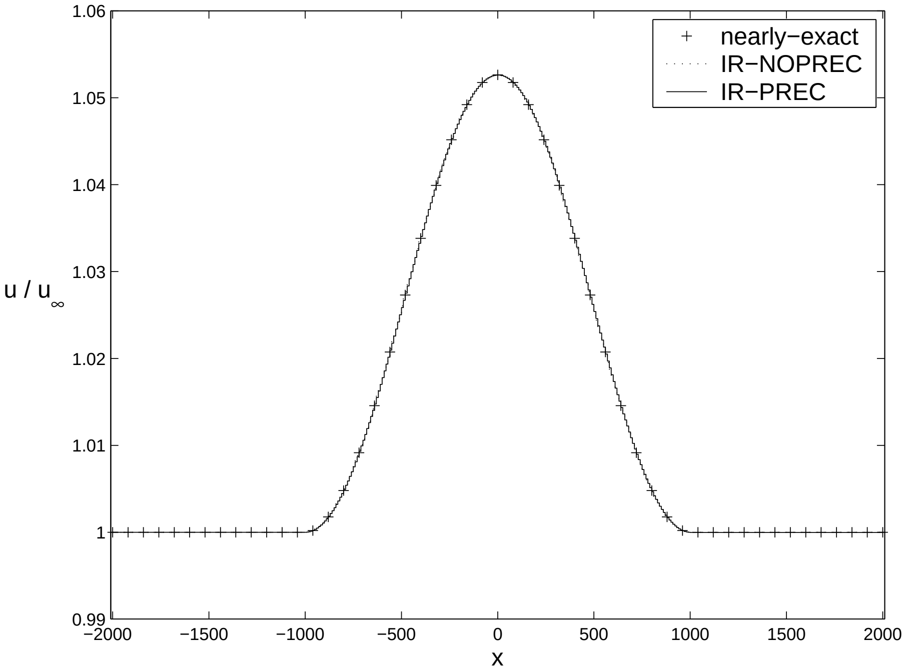

Figure 30: Approximation of u for the test-cases in Tab. 8.

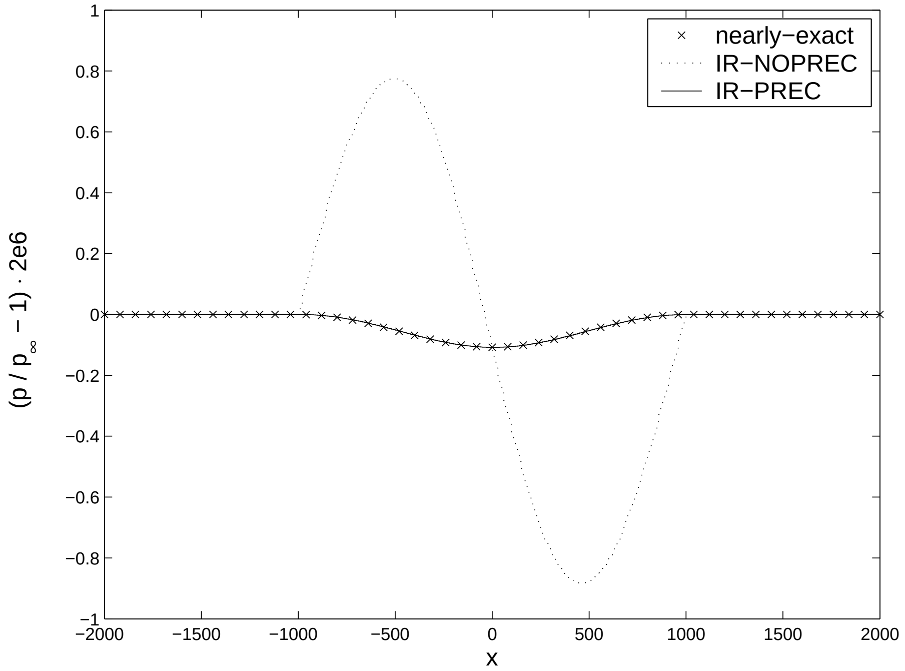

Figure 31: Approximation of p for the test-cases in Tab. 8. The pressure variation on the $y-$axis is scaled for ease of readability.

A local preconditioning strategy

It may be worth investigating the behaviour of the preconditioned numerical flux as the local Mach number of the flow field undergoes non-negligible variations. Indeed, under this circumstance, it may be difficult to identify a unique Mach number $M_\star$ which is representative of the entire flow field and, consequently, the definition of the preconditioning parameter β^2 in (189) may be not straightforward.

To the purpose, it is possible to consider the nozzle flow introduced in sec. 3.4.4. More precisely, by varying the parameter σ (which determines the variation of cross-sectional area through (202)), it is possible to control the velocity u in the duct as well as the corresponding (local) Mach number, since the sound speed is constant for the considered benchmark ($a = 10^3$). Thus, for instance, by choosing $\sigma = 4.5 \cdot 10^{-1}$ (while keeping the other settings in Tab. 6, sec. 3.4.4), the minimum value of α provided by (203) is $\alpha_{min} = 0.1$. The maximum value of u can be obtained from (200), since the flow turns out to be nearly-incompressible for the present case as well. In particular, the maximum value of u (taken in correspondence of the minimum cross-sectional area of the duct, hereafter referred to as throat as well) is roughly $10 \cdot u_\infty$. Hence, if M_∞ denotes the inlet Mach number ($M_\infty = 10^{-3}$ for the present case), the Mach number at the throat is $M_{throat} = 10 \cdot M_\infty$. Then, by choosing M_∞ as representative of the whole flow field: $M_\star = M_\infty$ and by choosing $\beta^2 = M_\star^2 \, (= 10^{-6})$ as preconditioning parameter, it follows that $\beta^2 = M_{throat}^3 \neq M_{throat}^2$. As a consequence, it is reasonable to expect a less accurate numerical solution near the throat. This is confirmed by the curves labelled with “GLOBAL-PREC” in Figs. 32 and 33, which are computed using the settings of the test-case IR-PREC reported in Tab. 8 above. It may be worth noticing that the discrepancy between the numerical and the nearly-exact solution in the aforementioned figures does not appreciably propagate towards the inlet section, since the adopted Dirichlet-like BC does not allow for a substantial variation of the state vector to take place.

In view of the aforementioned considerations, it makes sense to investigate the effects which are produced on the numerical solution by replacing the original preconditioning parameter β^2 with a new one, say β_{ij}^2, taking into account the local Mach number. More in detail, since the preconditioner acts (as the numerical flux, of course) at the cell interface, it seems reasonable to relate β_{ij}^2 to a certain Mach number $\bar{M}_{ij}$ which can be considered representative of both the adjacent state vectors $\mathbf{q}_i^{(x)}$ and $\mathbf{q}_j^{(x)}$. In consideration of the fact that the Roe flux between $\mathbf{q}_i^{(x)}$ and $\mathbf{q}_j^{(x)}$ is essentially based on the averaging defined, as the name suggests, by the Roe averages (see Note 24 in sec. 3.3.1), the following choice seems to be quite natural:

$$\bar{M}_{ij} := \frac{|u_{ij}|}{a_{ij}}$$

Then, a definition for β_{ij}^2 may be the following:

$$\beta_{ij}^2 := 1 - \exp\left(-\kappa_{ij} \cdot \left(\bar{M}_{ij}\right)^2\right) \tag{234}$$

where $\kappa_{ij} = \mathrm{O}(1)$ is a free parameter. Indeed, according to the definition above:

- for $\bar{M}_{ij} \to 0$, $\beta_{ij}^2 \to \kappa_{ij} \cdot \left(\bar{M}_{ij}\right)^2$, somehow (locally) recovering the original relation (189);
- as $\bar{M}_{ij}$ increases, $\beta_{ij}^2 \to 1$ and the effects of the preconditioning strategy correctly disappear. The parameter κ_{ij}, in particular, can be modelled in order to control the transition under consideration.

The numerical solution which is obtained by adopting (234) with $\kappa_{ij} = 1$ (while keeping the remaining settings of the aforementioned test-case IR-PREC) is shown by the curves labelled with "LOCAL-PREC" in Figs. 32 and 33. The considered solution turns out to be more accurate than that one obtained by the global preconditioning technique, even if there are still discrepancies with respect to the nearly-exact benchmark. In view of this result, it seems reasonable to further investigate (even heuristic) generalizations of the considered preconditioning technique, like (234), in order to accurately simulate flow fields characterized by non-negligible variations of the local Mach number. Such a study is postponed to a subsequent research activity.

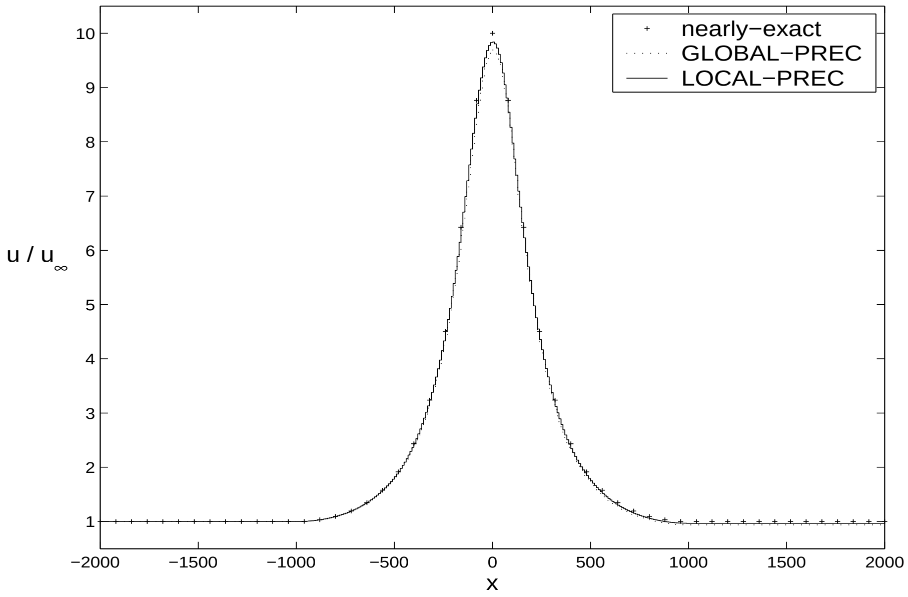

Figure 32: Comparison between the original ("global") preconditioning strategy and the modified ("local") one: effects on u.

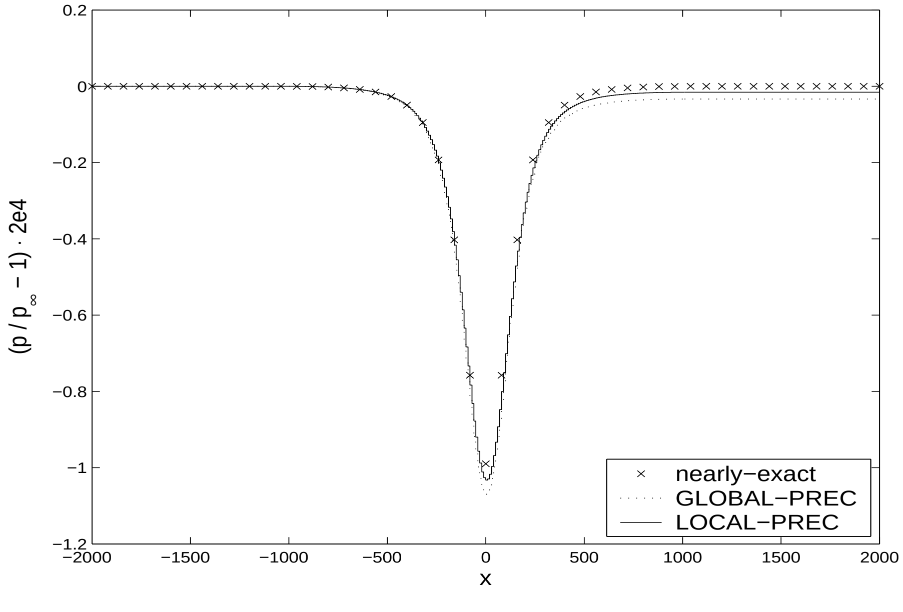

Figure 33: Comparison between the original ("global") preconditioning strategy and the modified ("local") one: effects on p. The pressure variation on the $y-$axis is scaled for ease of readability.

Test-case	Benchmark	μ	(n_L, n_R)	τ
IR2-3-1	B2	1	$(2,2)\cdot 10^3$	$5\cdot 10^{-4}$
IR2-3-2	B2	1	$(2,2)\cdot 10^3$	$5\cdot 10^{-3}$
IR2-3-3	B2	1	$(2,2)\cdot 10^3$	$5\cdot 10^{-2}$

Table 9: Considered test-cases for the discrete scheme (224), based on the numerical flux (156)-(160).

3.5.5 Numerical results for non-smooth flows at low Mach numbers

In order to directly compare the proposed linearized implicit time-advancing strategy with an explicit one, the implicit counterpart of the test-case ER-2-3 described in Tab. 4 (sec. 3.3.2) is considered.

The benchmark description as well as any relevant definition is tacitly recalled from the aforementioned section. The discrete scheme (224) is considered, associated with a constant time-step τ. The implementation of the assumed BCs (120) reads (compare with (233)):

$$\begin{cases} \left(\mathbf{M}^{1,n}_{(-1)} + \mathbf{M}^{1,n}_{(0)}\right)\cdot \delta^n \mathbf{q}^{(A)}_1 + \mathbf{M}^{1,n}_{(+1)}\cdot \delta^n \mathbf{q}^{(A)}_2 & = \quad \mathbf{m}^{1,n} \\ \mathbf{M}^{N_c,n}_{(-1)}\cdot \delta^n \mathbf{q}^{(A)}_{N_c-1} + \left(\mathbf{M}^{N_c,n}_{(0)} + \mathbf{M}^{N_c,n}_{(+1)}\right)\cdot \delta^n \mathbf{q}^{(A)}_{N_c} & = \quad \mathbf{m}^{N_c,n} \end{cases} \tag{235}$$

The considered test-cases are reported in Tab. 9; the behaviour of the corresponding numerical solutions is shown in Figs. 34-37.

Some entities which can be exploited for evaluating the accuracy as well as the computational cost of the considered simulations are reported in Tab. 10 (to be compared with the relevant row of Tab. 5 in sec. 3.3.2). It should be noticed that:

- the estimate $\tilde{c}^{(CFL)}$ is directly proportional to τ since its definition (125) is based on the largest wave speed of the benchmark RP (which, of course, is not affected by the numerical set-up) for all the considered test-cases. In particular, $\tilde{c}^{(CFL)}$ assumes the same value for the test-cases ER2-3 and IR2-3-1, for which the same time-step is adopted. The linearized implicit scheme does not suffer from the stability restriction encountered in the explicit case (coefficients $\tilde{c}^{(CFL)} > 1$ can be adopted) and it is therefore more efficient than the explicit one. However, as

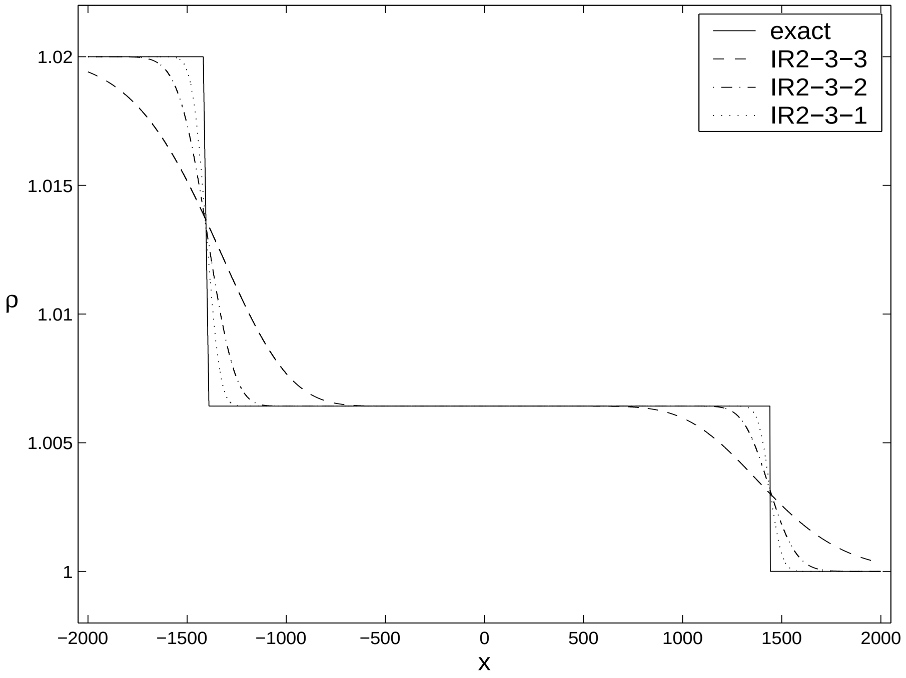

Figure 34: Approximation of ρ for the test-cases reported in Tab. 9.

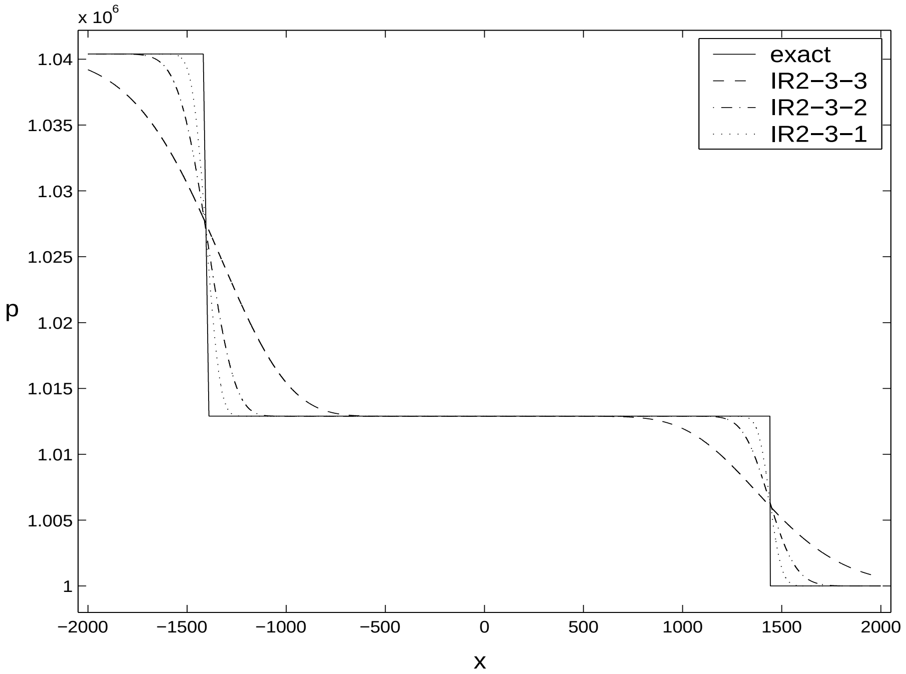

Figure 35: Approximation of p for the test-cases reported in Tab. 9.

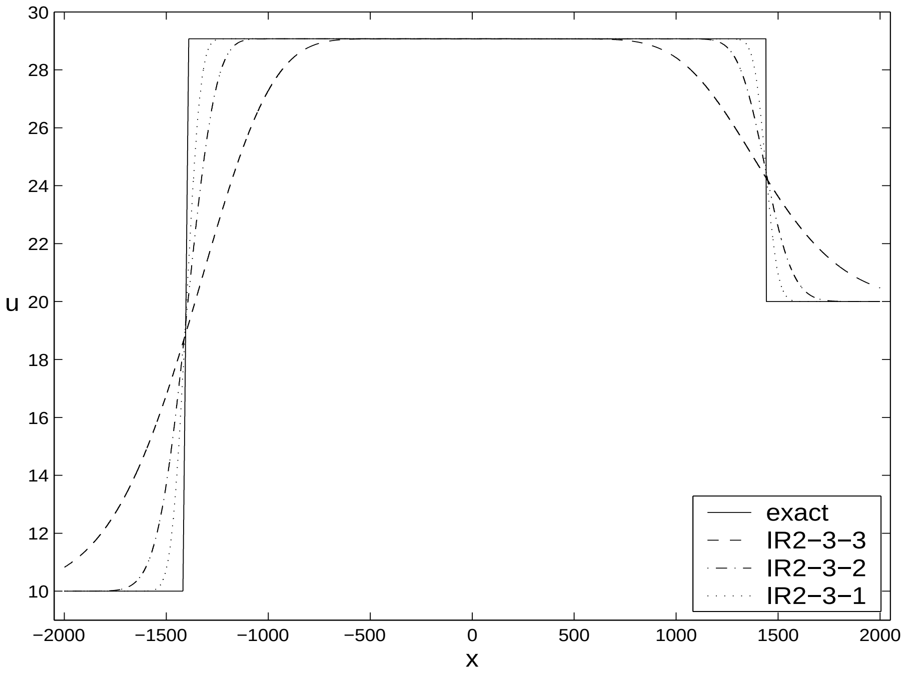

Figure 36: Approximation of u for the test-cases reported in Tab. 9.

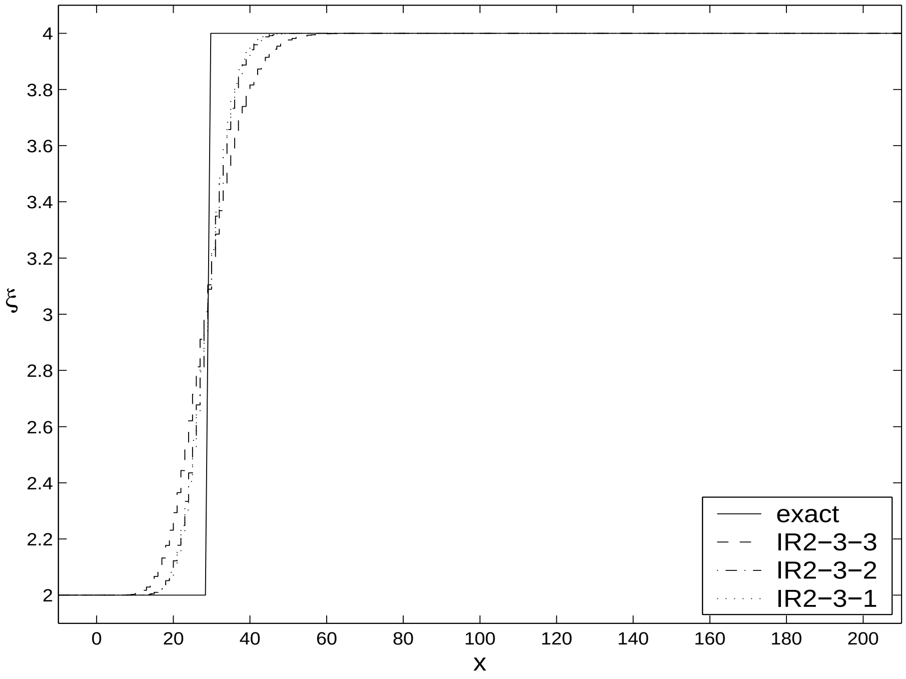

Figure 37: Approximation of ξ for the test-cases reported in Tab. 9. The $x-$range is cut for ease of readability.

Test-case	$\tilde{c}^{(CFL)}$	t_{CPU}	$e(\rho)$	$e(p)$	$e(u)$	$e(\xi)$
IR2-3-1	7.2e-1	≈ 5 min.	0.0707	0.1419	3.8683	1.0599
IR2-3-2	7.2	≈ 30 sec.	0.1111	0.2231	6.0785	1.0956
IR2-3-3	7.2e1	≈ 4 sec.	0.1992	0.3997	10.8928	1.3401

Table 10: CFL estimate, CPU time and error estimates for the test-cases reported in Tab. 9.

discussed in more detail in sec. 3.5.6, the proposed linearized implicit scheme is not unconditionally stable and the largest time-step which can be adopted seems to be somehow related to the magnitude of the discontinuity introduced by the IC of the underlying benchmark RP;

- the CPU time (on a laptop with Intel P4 CPU 2.66GHz, 512kB L2 cache, 512MB RAM) turns out to be inversely proportional to τ. The considered implementation of the implicit time-advancing turns out to be slower than the explicit one of about one order of magnitude [47];

- despite small differences (which may be partly addressed to the implementation), the numerical solution obtained by the considered implicit scheme turns out to be as accurate as that one obtained by the considered explicit scheme (compare the test-case IR2-3-1 in Tab. 10 with the test-case ER2-3 in Tab. 5, sec. 3.3.2). Moreover, the accuracy of the numerical solution provided by the linearized implicit scheme rapidly degrades as the time-step is increased (for a chosen space discretization), as shown in Fig. 38 [48]. Hence, the largest time-step which can be adopted when using the proposed linearized implicit scheme for unsteady simulations could be bounded by chosen accuracy requirements even before reaching the aforementioned stability limit.

[47] This result does not seem to be closely related to the specific implementation of the implicit solver. Indeed, comparable CPU times have been obtained by considering two different solvers for the linear system of interest (namely a library routine for banded matrices and the block version of the Thomas algorithm [79] for tri-diagonal systems).

[48] It should be noticed that, with the only exception of $e(\xi)$ (whose measure is, in general, more susceptible to errors due to the specific shape of the relevant curve), the remaining curves exhibit a similar trend.

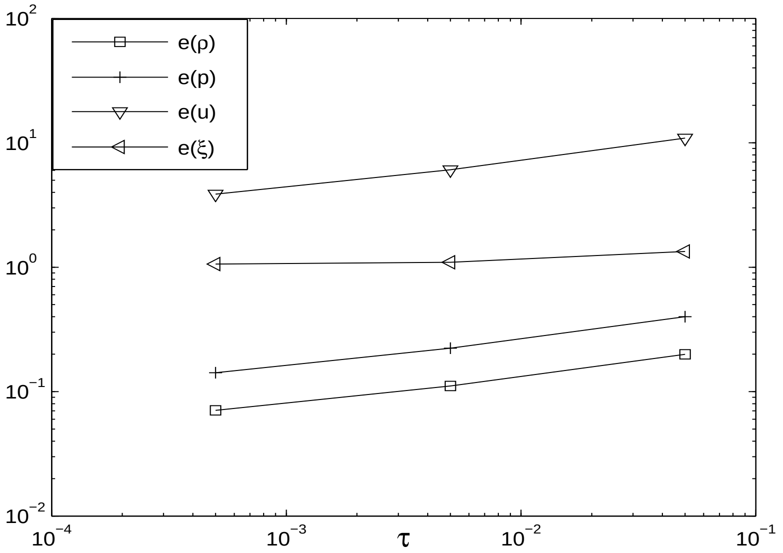

Figure 38: Plot of the error estimates for the test-cases reported in Tab. 10.

Benchmark	κ	$\varkappa$	γ	ρ_L	u_L	ξ_L	ρ_R	u_R	ξ_R	t_{eval}
B3	1	1	0	1	0.9	2	1	-0.9	4	1

Table 11: Considered benchmark.

3.5.6 Numerical results for non-smooth flows at generic Mach numbers

Up to the present section, all the considered numerical experiments have focused attention on low Mach number flows, because of widely discussed reasons. However, since the proposed linearized implicit time-advancing can be applied to generic Roe numerical flux functions and therefore to a variety of problems, it is of interest to also investigate the behaviour of the discrete solution when the characteristic Mach number of the flow (if any) is not close to zero.

To the purpose, the benchmark summarized in Tab. 11 is considered. In this table, κ, $\varkappa$ and γ refer to the chosen convex state law (71), ρ_L, u_L, ξ_L, ρ_R, u_R and ξ_R characterize the initial condition (IC) of a Riemann problem (RP) and t_{eval} denotes the time at which the considered solution is picked. The adopted state law introduces a constant sound speed $a = \sqrt{\kappa} = 1$ and

Test-case	Benchmark	μ	(n_L, n_R)	τ
IR-M09-1	B3	10^{-2}	$(2,2)\cdot 10^2$	$5\cdot 10^{-3}$
IR-M09-2	B3	10^{-2}	$(2,2)\cdot 10^2$	$1\cdot 10^{-2}$
IR-M09-3	B3	10^{-2}	$(2,2)\cdot 10^2$	$1\cdot 10^{-1}$

Table 12: Considered test-cases.

therefore the unperturbed "left" and "right" regions of the relevant RP are characterized by a local Mach number equal to 0.9. On the other hand, due to the symmetry of the chosen IC, $u_\star = 0$ (easily derived by averaging the expressions in (89) for $\rho_L = \rho_R$) and consequently the local Mach number undergoes a noticeable variation through the domain [49]. In this sense, the present benchmark introduces discontinuities which are stronger than those considered in previous numerical experiments. Besides the (stationary, since $u_\star = 0$) contact discontinuity, two symmetrical shock waves appear as part of the solution, travelling with speed $\tilde{s} \approx 0.65$.

The discrete scheme (224) is considered, associated with a uniform space discretization having measure $\mu = 10^{-2}$ and a constant time-step τ. It may be worth noticing that the adopted space discretization is only apparently finer than that one considered in previous numerical experiments, e.g. those reported in Tab. 9. Indeed, a relevant parameter is $\mu/\tilde{\mu}$, where $\tilde{\mu}$ denotes a characteristic length scale of the problem. For the RP at hand, $\tilde{\mu} \approx \tilde{s}\, t_{eval} = \mathrm{O}\,(1)$ and $\mu/\tilde{\mu} = \mathrm{O}\,(10^{-2})$ while for the test-cases reported in Tab. 9 (for which the largest wave speed is of the order of 10^3) $\mu/\tilde{\mu} = \mathrm{O}\,(10^{-3})$. In a similar manner, a relevant parameter for the time discretization is τ/t_{eval}. The boundary conditions (235) are adopted for closing the problem.

The considered test-cases are reported in Tab. 12. The numerical approximation of ρ (i.e. p, since $p = \rho$ according to the considered state law) and u are shown in Figs. 39 and 40, respectively. An example of the numerical approximation of ξ is reported in Fig. 41.
It is worth remarking that:

- by increasing the time-step (for a fixed space discretization) the numerical scheme becomes unstable. More in detail:
    - for the test-cases IR-M09-1 the coefficient $\tilde{c}^{(CFL)}$ defined in (125) is approximately equal to 0.33 and an explicit time-advancing run-

[49] Sonic conditions are deliberately avoided, see Note 32 in sec. 3.3.1.

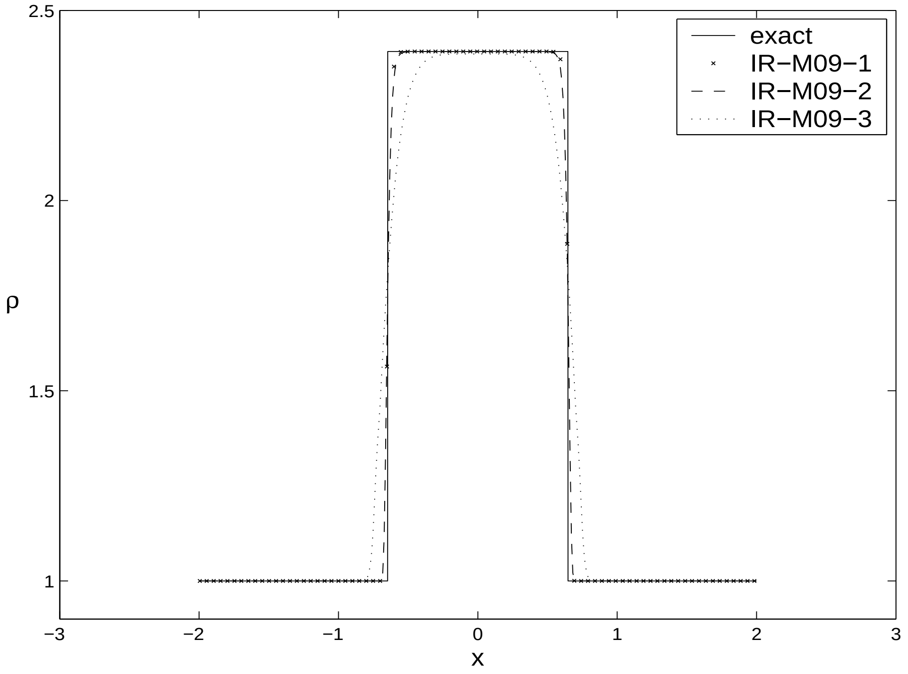

Figure 39: Approximation of ρ (i.e. p) for the test-cases reported in Tab. 12.

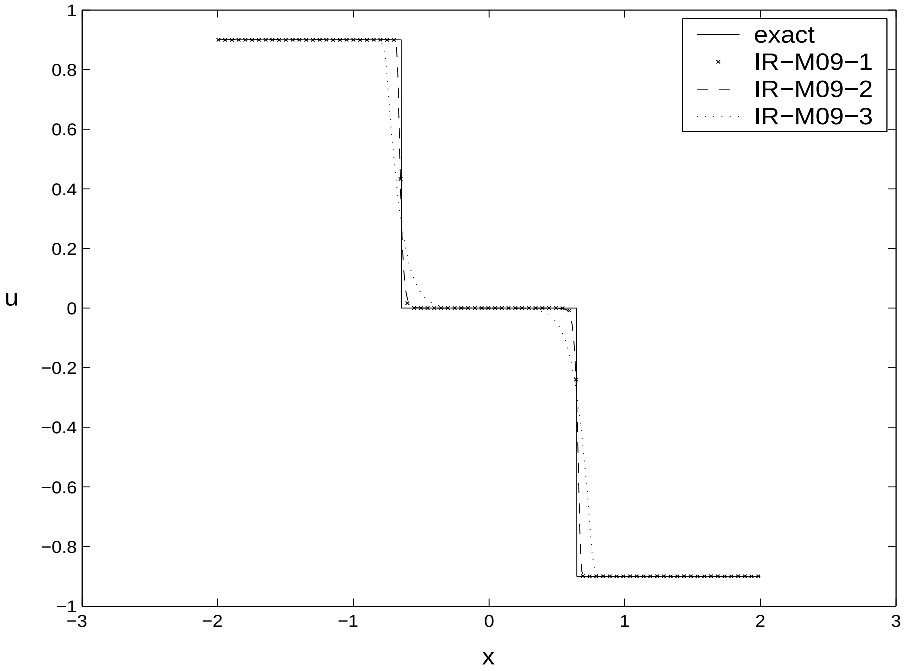

Figure 40: Approximation of u for the test-cases reported in Tab. 12.

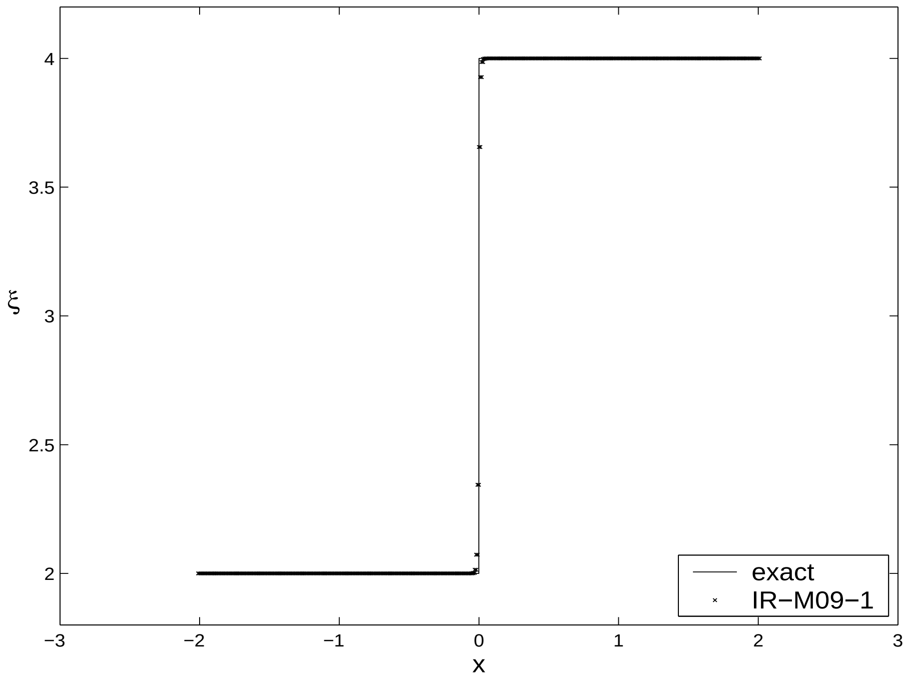

Figure 41: Approximation of ξ for the test-case IR-M09-1 reported in Tab. 12.

ning with the same $\tilde{c}^{(CFL)}$ turns out to be stable as well;

- for the test-cases IR-M09-2 and IR-M09-3, $\tilde{c}^{(CFL)}$ is approximately equal to 0.65 and 6.5, respectively. By further increasing the time-step ($1.25{\cdot}10^{-1} < \tau < 2.00{\cdot}10^{-1} \Rightarrow \tilde{c}^{(CFL)} \approx 10$), a blow-up occurs after a few iterations: the numerical solution becomes unstable since the beginning of the simulation, in correspondence of the discontinuity associated with the benchmark RP.

This observation suggests the existence of a stability limit for the linearized implicit time-advancing. By comparing the maximum $\tilde{c}^{(CFL)}$ which can be adopted in the present case with e.g that one associated with the test-case IR2-3-3 in Tab. 10, it is possible to put forward the hypothesis that the stability limit under consideration somehow becomes more severe as the magnitude of some relevant discontinuities associated with the underlying RP increases [50]. For instance, the variation of the density for the benchmark B3 (associated with the test-cases in Tab. 12) is O (1) (see Fig. 39) while that one for the benchmark B2

[50] The jump of the passive scalar ξ may play a minor role, since it does not directly affect the considered Roe linearization.

(associated with the test-cases in Tab. 10) is $\mathrm{O}\left(10^{-2}\right)$; an even more considerable difference is observed when considering the variation of the Mach number.
The aforementioned hypothesis may be supported by the fact that for smooth solutions this stability problem does not appear (see e.g. the nozzle flow discussed in sec. 3.5.4). On the other hand, some numerical experiments involving stronger initial discontinuities (not reported here for brevity) exhibit an even narrow stability margin. After all, the presence of a discontinuity within the flow field makes it more difficult to apply the proposed linearization strategy (222); in particular, it is likely to violate the condition (221) in the neighbourhood of the discontinuity. In summary, as soon as considerable discontinuities appear within the flow field, a time-step reduction could be required in order to keep the proposed linearized implicit time-advancing algorithm stable [51], thus reducing the efficiency of the corresponding numerical scheme. However, this stability problem does not appear to be related to the specific linearization which is proposed in the present work; indeed, it affects other linearizations as well, as discussed in sec. 3.5.7;

- as for the test-cases presented in sec. 3.5.5, the accuracy of the considered numerical solutions rapidly degrades as the time-step is increased and the largest time-step which can be adopted for unsteady simulations could be bounded by chosen accuracy requirements even before reaching the aforementioned stability limit;

- the considered linearized implicit scheme is able to approximate the contact discontinuity with a reasonably good accuracy (i.e. a few cells), as shown in Fig. 41 for the test-case IR-M09-1. This result is not obvious, since in general it is not a trivial task to approximate slowly moving (in particular, stationary) contact discontinuities [98].

[51] In a practical computational set-up the time-step can be modulated, possibly by an adaptive strategy, so as to mitigate the stability problem under consideration.

3.5.7 A remark on the linearization technique

It seems valuable to address the issue of whether the stability constraint highlighted in sec. 3.5.6 is specific to the proposed linearization technique or not. To the purpose, two linearizations of the Roe flux function are recalled, of the type of (105). As for the proposed linearization (222), the aforementioned ones can be applied to a generic Roe numerical flux function $\boldsymbol{\phi}_{LR}^{(g)ROE}$.

(L1) By adopting the following approximations:

$$\delta^n \mathbf{f}^{(g)} \approx \mathbf{J}^{(g)n} \cdot \delta^n \mathbf{q}^{(g)} \tag{236}$$

$$\delta^n \left| \tilde{\mathbf{J}}_{LR}^{(g)} \right| \approx \mathbf{0} \tag{237}$$

where $\mathbf{J}^{(g)}$ denotes the Jacobian defined by (213) and $\tilde{\mathbf{J}}_{LR}^{(g)}$ represents the relevant Roe matrix, the variation of the centred and upwind components of $\boldsymbol{\phi}_{LR}^{(g)ROE}$ can be expressed as follows:

$$\delta^n \boldsymbol{\phi}_{c,LR}^{(g)ROE} \approx \frac{1}{2} \left(\mathbf{J}_L^{(g)n} \cdot \delta^n \mathbf{q}_L^{(g)} + \mathbf{J}_R^{(g)n} \cdot \delta^n \mathbf{q}_R^{(g)} \right) \tag{238}$$

$$\delta^n \boldsymbol{\phi}_{u,LR}^{(g)ROE} \approx -\frac{1}{2} \left| \tilde{\mathbf{J}}_{LR}^{(g)n} \right| \cdot \left(\delta^n \mathbf{q}_R^{(g)} - \delta^n \mathbf{q}_L^{(g)} \right) \tag{239}$$

Consequently, a linearization of the type of (211) can be straightforwardly introduced, involving the following coefficients:

$$\left\{ \begin{array}{lcl} \mathbf{A}_{LR}^n & = & \frac{1}{2} \left(\mathbf{J}_L^{(g)n} + \left| \tilde{\mathbf{J}}_{LR}^{(g)n} \right| \right) \\ \\ \mathbf{B}_{LR}^n & = & \frac{1}{2} \left(\mathbf{J}_R^{(g)n} - \left| \tilde{\mathbf{J}}_{LR}^{(g)n} \right| \right) \end{array} \right. \tag{240}$$

The linearization (240) is exploited in [31] for defining a linearized implicit time-advancing technique.

(L2) By defining a matrix $\mathbf{J}^{(g)\star} = \mathbf{J}^{(g)\star} \left(\mathbf{q}^{(g)} \right)$ which mimics the first-order homogeneity property (212) as follows:

$$\mathbf{f}^{(g)} = \mathbf{J}^{(g)\star} \cdot \mathbf{q}^{(g)} \tag{241}$$

it is possible to introduce the following linearization (formally similar to (236)):

$$\delta^n \mathbf{f}^{(g)} \approx \mathbf{J}^{(g)\star n} \cdot \delta^n \mathbf{q}^{(g)}$$

Then, it is possible to replace (238) with the following expression:

$$\delta^n \boldsymbol{\phi}^{(g)ROE}_{c,LR} \approx \frac{1}{2}\left(\mathbf{J}^{(g)\star n}_{L} \cdot \delta^n \mathbf{q}^{(g)}_{L} + \mathbf{J}^{(g)\star n}_{R} \cdot \delta^n \mathbf{q}^{(g)}_{R}\right)$$

Finally, by keeping the approximation (237) (and, consequently, (239)) the following additional linearization can be introduced:

$$\left\{\begin{array}{lcl} \mathbf{A}^n_{LR} & = & \frac{1}{2}\left(\mathbf{J}^{(g)\star n}_{L} + \left|\tilde{\mathbf{J}}^{(g)n}_{LR}\right|\right) \\ \\ \mathbf{B}^n_{LR} & = & \frac{1}{2}\left(\mathbf{J}^{(g)\star n}_{R} - \left|\tilde{\mathbf{J}}^{(g)n}_{LR}\right|\right) \end{array}\right. \tag{242}$$

which, of course, is similar to the previous one (240). The definition of the matrix $\mathbf{J}^{(g)\star}$) and, consequently, the linearization (242) are introduced in [4] in order to define a linearized implicit time-advancing technique.

Note 43 *In general, the definition of $\mathbf{J}^{(g)\star}$ is not unique, as shown by the following example based on the basic-1D state vector $\mathbf{q}^{(x)}$ (introduced in sec. 2.2.3). Let α_{mn}, with $m, n \in \{1, 2\}$, denote the $mn-th$ component of $\mathbf{J}^{(x)\star}$. Then, the following relation must be verified, by the definition (241):*

$$\left(\begin{array}{c} \rho\, u \\ \rho\, u^2 + p \end{array}\right) = \left(\begin{array}{cc} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{array}\right) \cdot \left(\begin{array}{c} \rho \\ \rho\, u \end{array}\right)$$

By virtue of the fact that ρ and u are independent of each other, it necessarily follows that $\alpha_{11} = 0$ and $\alpha_{12} = 1$, while the remaining relation leads to the following equation:

$$\alpha_{21}\,\rho + \alpha_{22}\,\rho\, u = \rho\, u^2 + p \tag{243}$$

which admits an infinite number of solutions. By assuming, for instance, $\alpha_{22} = 2\,u$, the matrix $\mathbf{J}^{(x)\star}$ can be expressed as follows:

$$\mathbf{J}^{(x)\star} = \left(\begin{array}{cc} 0 & 1 \\ \\ a^2 - u^2 + \left(\frac{p}{\rho} - a^2\right) & 2\,u \end{array}\right) \tag{244}$$

while, by assuming $\alpha_{21} = a^2 - u^2$, the following representation is obtained:

$$\mathbf{J}^{(x)\star} = \left(\begin{array}{cc} 0 & 1 \\ \\ a^2 - u^2 & 2\,u + u^{-1}\left(\frac{p}{\rho} - a^2\right) \end{array}\right) \tag{245}$$

The matrix (244), in particular, is exploited in [4].

It is worth noticing that the aforementioned linearization techniques (L1) and (L2) may coincide with each other. For instance, if the adopted barotropic state law is first-order homogeneous:

$$p = \frac{\mathrm{d}p}{\mathrm{d}\rho}\,\rho = a^2\,\rho \tag{246}$$

then the expressions (244) and (245) become equal to each other and $\mathbf{J}^{(x)\star}$ reduces to the relevant Jacobian $\mathbf{J}^{(x)}$. As a result, the aforementioned linearization techniques (L1) and (L2) coincide with each other.

A few simulations have been carried out, also involving the aforementioned linearizations (L1) and (L2). In particular, the test-case IR-M09-3 reported in Tab. 12 above has been considered. The state law associated with the corresponding benchmark (namely B3, defined in Tab. 11) verifies the condition (246) and therefore the linearizations (L1) and (L2) coincide for the case at hand. Some relevant behaviours are shown in Figs. 42 and 43, in which the label $L1/L2$ concisely refers to both the aforementioned linearizations while L_{orig} refers to the proposed one (222). It is worth noticing that:

- for the considered test-case the considered discrete solutions turn out to be very similar to each other;
- the stability of the linearized implicit time-advancing based on (L1) i.e (L2) turns out to be comparable with that one based on the proposed linearization. Indeed, as far as the test-cases reported in Tab. 12 are concerned (hence, in particular for IR-M09-3), a blow-up occurs when advancing the simulations with a coefficient $\tilde{c}^{(CFL)} = \mathrm{O}\,(10)$, namely for $1.5 \cdot 10^{-1} < \tau < 5.0 \cdot 10^{-1}$ (compare with the relevant point in sec. 3.5.6).

Similar results have been obtained by exploiting a slightly different state law for which (L1) does not coincide with (L2). On the basis of the carried-out numerical experiments, the proposed linearization technique (222) seems to behave similarly to the aforementioned ones, as far as the accuracy and the efficiency are concerned. In particular, the stability restrictions imposed on the time-advancing by the presence of discontinuities within the flow field seem to affect all the considered linearized algorithms in a similar manner. However, only a preliminary investigation has been performed in this regard and further study is definitely recommended.

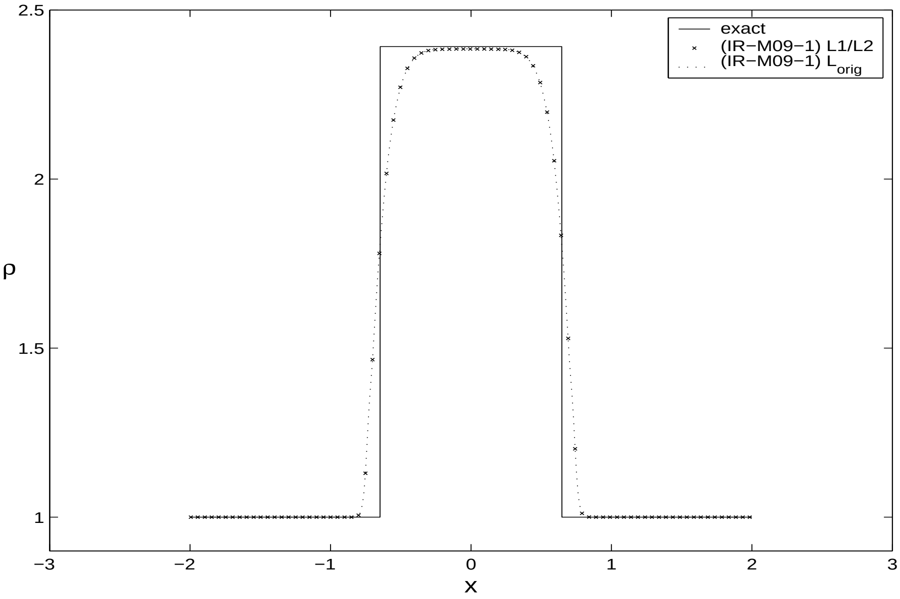

Figure 42: Comparison between the considered approximate linearizations: approximation of ρ (i.e. p) for the test-case IR-M09-3 in Tab. 12.

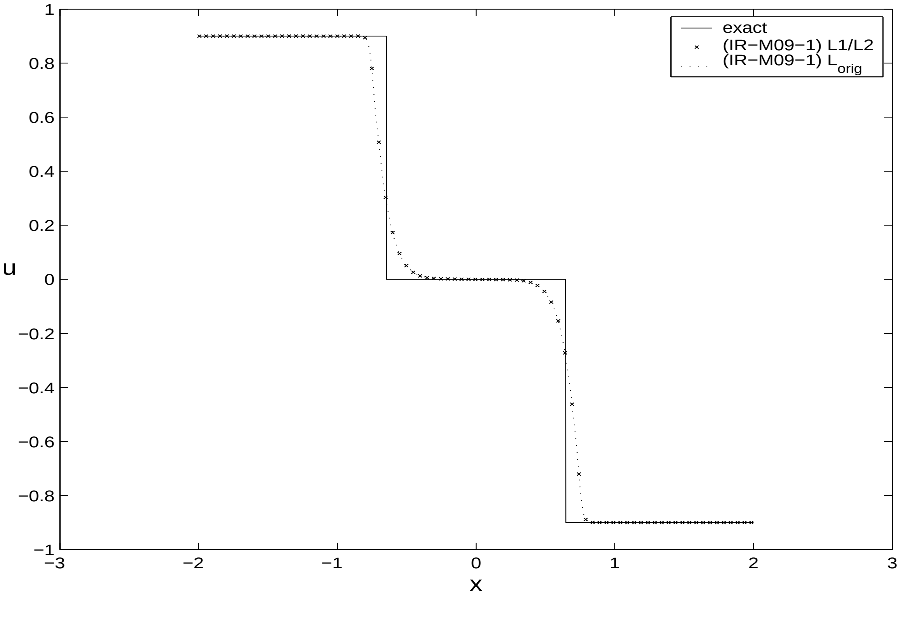

Figure 43: Comparison between the considered approximate linearizations: approximation of u for the test-case IR-M09-3 in Tab. 12.

4 1D Applications to cavitating flows

In sec. 4.1 the considered state law is introduced, together with some details concerning its numerical implementation. In sec. 4.2 some numerical results obtained in [91] are recalled and an illustrative numerical experiment, based on a RP whose IC leads to a cavitating flow, is presented.

4.1 State law of the working fluid

A state law of the type of that one defined in 1.5 is assumed for the working fluid. The density domain, in particular, is split into two adjacent subdomains: an upper one, where the working fluid behaves as a pure liquid, and a lower one, where cavitation phenomena are taken into account by a homogeneous flow cavitation model (see secs. 1.3 and 1.4).

The mathematical definition of both models, whose physical assumptions and implications are widely discussed in [26] and [27], is respectively given in secs. 4.1.1 and 4.1.2. Some issues regarding the numerical implementation of the cavitating mixture state law are then discussed in sec. 4.1.3. Finally, in sec. 4.1.4, the convexity of the chosen barotropic state law is discussed.

4.1.1 Pure liquid model

The working liquid is supposed to be at a constant temperature T_L. Let p_{sat} and ρ_{Lsat} be the saturation pressure and the liquid saturation density, respectively, at temperature T_L. Furthermore, let $\beta_L > 0$ denote the coefficient of isentropic compressibility of the liquid at temperature T_L [12]. The non-dimensional form of the considered liquid barotropic state law reads:

$$\bar{p} = \bar{p}_{liq}(\bar{\rho}) := 1 + \vartheta \ln(\bar{\rho}) \quad , \quad \bar{\rho} \in [1, \infty) \tag{247}$$

where:

$$\bar{\rho} := \frac{\rho}{\rho_{Lsat}} \quad , \quad \bar{p} := \frac{p}{p_{sat}} \quad , \quad \vartheta := (\beta_L p_{sat})^{-1} \tag{248}$$

Note 44 *Common liquids are nearly-incompressible: their non-dimensional compressibility coefficient ϑ, as computed by adopting physically-based values, is very high (e.g.* $\mathrm{O}(10^6)$ *for water at* 20°C*). Consequently, the density is practically constant (*$\bar{\rho} \approx 1$*) for preventing unphysically high pressure values to be produced. In consideration of this point, in many computations involving real fluids under ordinary conditions, the logarithmic state law (247) is replaced with its linearization (see e.g. [96] and [78]), namely:*

$$\bar{p} - 1 \approx \vartheta(\bar{\rho} - 1) \quad , \quad \bar{\rho} \in [1, \infty) \tag{249}$$

The linearized state law may be preferable to the original one in view of its simplicity. It is worth noticing that the expression (249) represents a specific instance of the convex barotropic state law (71), obtained in particular for $\kappa = \vartheta$, $\varkappa = 1$ *and* $\gamma = 1 - \vartheta$.

4.1.2 Cavitation model

The chosen cavitation model provides the following differential relation between the non-dimensional density and pressure -introduced in (248)- within the mixture region:

$$\frac{\bar{p}}{\bar{\rho}}\frac{\mathrm{d}\bar{\rho}}{\mathrm{d}\bar{p}} = \bar{\rho}\left\{ (1-\varepsilon)\,\vartheta^{-1}\,\bar{p} + \sigma_1\,\varepsilon\,\bar{p}^{\sigma_2} \right\} + (1-\bar{\rho})\,\{\sigma_3\} \quad , \quad \bar{\rho} < 1 \tag{250}$$

where:

- the non-dimensional parameters σ_1, σ_2 and σ_3 are defined as follows (p_{sat} being introduced in the previous section):

  $$\sigma_1 := g^\star \left(\frac{p_c}{p_{sat}}\right)^\eta \quad , \quad \sigma_2 := -\eta \quad , \quad \sigma_3 := \frac{1}{\gamma_V}$$

  in which p_c denotes the saturation pressure of the fluid at hand, γ_V represents the specific heat ratio (i.e. specific heat at constant pressure over specific heat at constant volume) of the relevant vapour and $g^\star$ and η are constants depending on the fluid under consideration [10]. It should be noticed, in particular, that σ_2 and σ_3 only depend on the chosen liquid while σ_1 is also affected by the liquid temperature T_L;

- the symbol ε denotes the following non-dimensional function of $\bar{\rho}$:

  $$\varepsilon = \varepsilon_\zeta(\bar{\rho}) := \left\{ \left\{ \left((1+\zeta)^3 - 1\right)\,\frac{1-\bar{\rho}}{\bar{\rho}} \right\}^{-3} + 1 \right\}^{-1/3} \quad , \quad \bar{\rho} < 1 \tag{251}$$

  which describes the liquid volume fraction ($0 \le \varepsilon \le 1$) which is in thermal equilibrium with the cavities. The symbol $\zeta > 0$ in (251) denotes a free model parameter accounting for thermal cavitation effects and, possibly, for the concentration of the active cavitation nuclei [26] [27]. The function $\varepsilon_\zeta(\bar{\rho})$ is monotonically decreasing and, in particular, admits the following asymptotic behaviour:

  $$\varepsilon_\zeta(\bar{\rho} \to 1) \to 0 \tag{252}$$

  which correctly models the fact that a negligible fraction of the liquid participates to the heat exchange at the interface with a vanishing cavity [26].

Note 45 *It should be noticed that, once chosen the working liquid, the pure liquid state law (247) only depends on the chosen temperature T_L while the mixture model (250) also depends on the free parameter ζ.*

The physical foundations of the considered cavitation model ensure, in particular, that the monotonicity requirement (4) is satisfied. Hence, once given T_L and ζ a value, the o.d.e. (250) can be numerically integrated with the following physically based initial condition:

$$\bar{p}\left(\bar{\rho}=1\right)=1 \tag{253}$$

Moreover, due to some approximations that are introduced when deriving the cavitation model [26], the integration can only be extended down to a certain threshold $\bar{\rho}_{min}$ such that:

$$\bar{\rho}_{min} \gg \frac{\rho_{Vsat}}{\rho_{Lsat}} \tag{254}$$

where ρ_{Vsat} is the vapour saturation density at temperature T_L. The condition (254) clearly prevents the model to be applied for describing liquid-vapour mixtures towards the pure vapour limit (hence, the chosen cavitation model is not suitable, as it is, for juxtaposition with a barotropic state law describing the pure vapour).

For consistency with the expression (1), the integral curve defined by (250) and (253) is formally denoted as follows:

$$\bar{p}=\bar{p}_{cav}\left(\bar{\rho}\right) \quad , \quad \bar{\rho}\in\left[\bar{\rho}_{min},1\right) \tag{255}$$

where the half-open density domain must be juxtaposed with that one of the pure liquid. The considered mixture state law (255), in particular, smoothly joins the liquid one (247) at the saturation point $(\bar{\rho}=1,\bar{p}=1)$, up to the first derivative. Indeed, the continuity of p is trivially enforced by the initial condition (253). Moreover, by substituting (253) and (252) into (250), it follows that:

$$\frac{\mathrm{d}\bar{\rho}}{\mathrm{d}\bar{p}}\left(\bar{\rho}\to 1\right)\to\vartheta^{-1}$$

in agreement with the fact that, according to (247), $\mathrm{d}\bar{p}/\mathrm{d}\bar{\rho}\left(\bar{\rho}=1\right)=\vartheta$. In other words, both p and a are continuous across the saturation point.

As an example, water at $T_L = 293.16$ K is considered, leading to the following values for the parameters in (250): $\vartheta \approx 8.55\cdot 10^5$ (see [26] and [83]), $\sigma_1 \approx 1.33\cdot 10^3$, $\sigma_2 \approx -0.73$ and $\sigma_3 \approx 0.78$ (see [10], [26] and [83]). Since $\rho_{Vsat}/\rho_{Lsat} = \mathrm{O}\left(10^{-5}\right)$ for the case under consideration, a threshold $\bar{\rho}_{min} = \mathrm{O}\left(10^{-4}\div 10^{-3}\right)$ is chosen in consideration of (254). Two barotropic curves obtained by choosing different values of the free parameter ζ are shown in Fig. 44. The corresponding (dimensional) sound speed curves are reported in Fig. 45. In this figure, the scale of the y-axis is deliberately cut for ease of readability. Indeed, both limits $a\left(\bar{\rho} \to 1\right) \to \left(\vartheta\, p_{sat}/\rho_{Lsat}\right)^{1/2} \approx 1.41 \cdot 10^3$ m/s and $a\left(\bar{\rho} \to \bar{\rho}_{min}\right) \to \mathrm{O}\left(10^2\right)$ m/s would in practice squash almost all the curve on the x-axis.

The very sharp, step-like, transition of the sound speed occurring near the saturation point in Fig. 45 is typical of the cavitation inception at low temperatures T_L ("cold cavitation"), of the type of the considered one. As already mentioned in sec. 1.4, this abrupt transition is essentially related to the considered physical phenomenon, as modelled by a homogeneous cavitation model, and not to the specific model here adopted. In particular, it is also present when considering the well-known barotropic cavitation model originally proposed by Delannoy (see e.g. [22], [29] and [30]). Clearly, it is very challenging to incorporate state laws like those shown in Fig. 44 (coupled with a suitable liquid model, e.g. (247) or (249)), into state-of-the-art numerical schemes.

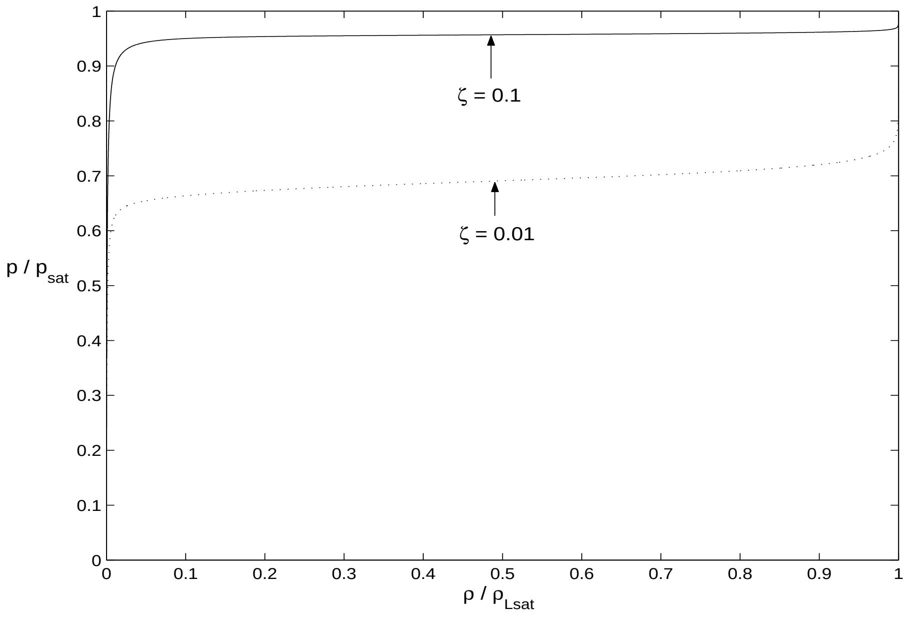

Figure 44: Typical trends of the considered mixture barotropic state law for water at $T_L = 293.16$ K.

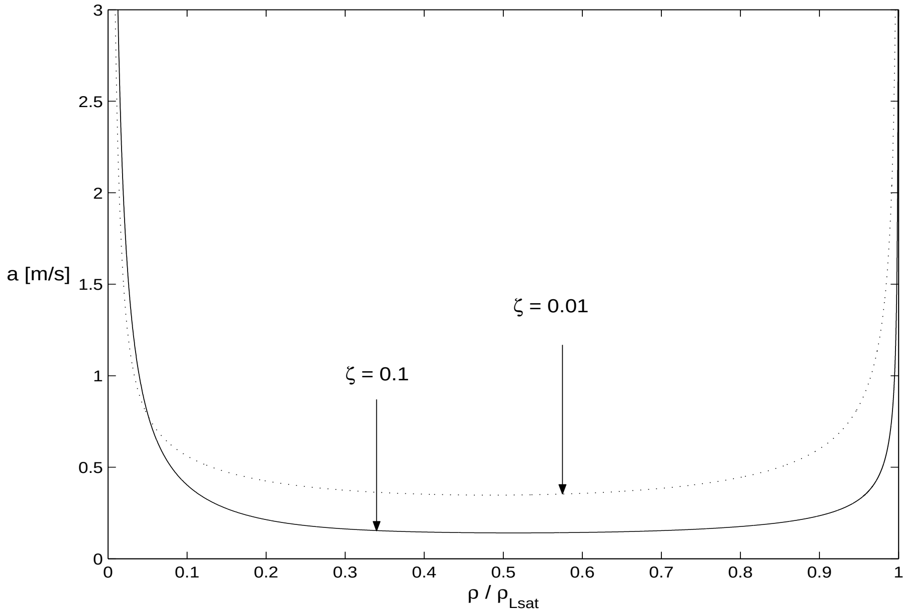

Figure 45: Trend of the mixture sound speed corresponding to the state laws shown in Fig. 44. The y-axis is cut for ease of readability.

4.1.3 Numerical implementation of the mixture state law

The cavitation model (250) does not explicitly provide the output of interest (e.g. p) in correspondence of the chosen input (e.g. ρ): to this purpose, an o.d.e. must be solved in advance. Obviously, when performing a simulation, it is not convenient from a computational standpoint to solve such an o.d.e. at each time-step and at each point of the computational grid in order to obtain the desired output. Thus, it seems convenient to numerically integrate the mixture cavitation model at the beginning of each simulation and then to store a table of the form:

$$(\rho_i\,,\, p_i\,,\, a_i) \quad , \quad i \in \{0, \ldots, n-1\} \tag{256}$$

with, say, $\rho_0 = \rho_{Lsat}$ and $\rho_{n-1} = \rho_{min} := \bar{\rho}_{min} \cdot \rho_{Lsat}$, to be accessed as required by the simulation algorithm. It is therefore necessary to define a fast table look-up strategy in order to efficiently incorporate the cavitating branch (255) of the barotropic model into a suitable numerical solver. To the purpose, it is possible to take advantage of the typical "S-like" shape of the mixture state law (see Fig. 44), as explained below.

A density-based algorithm is assumed for the remainder of the present section. A typical access to the table (256) within such an algorithm is aimed at finding the pressure p and the sound speed a corresponding to a certain input value of the independent variable $\rho < \rho_{Lsat}$. It is possible to define a fast look-up strategy by firstly noticing that the distribution along the x-axis in Fig. 44 of the density "nodes" ρ_i, as provided by an ordinary adaptive integration algorithm (e.g. a classical fourth-order Runge-Kutta scheme [79] with adaptive step-size control), typically exhibits clusters near the extremes $i = 0$ and $i = n-1$ due to the high value of the mixture sound speed, respectively in correspondence of $\rho = \rho_{Lsat}$ and $\rho = \rho_{min}$. It is therefore possible to approximate the original density sequence ρ_i by a new one, say ρ'_j, obtained by juxtaposing two geometric sequences, $\rho_j^{(right)}$ and $\rho_j^{(left)}$, respectively starting from ρ_0 and ρ_{n-1} and joining each other at a certain node $\rho_{i_\star}$ such that $\rho_{i_\star} \approx 0.5 \cdot \rho_{Lsat}$. Let $\gamma_r > 1$ and $\gamma_l > 1$ denote the ratios of $\rho_j^{(right)}$ and $\rho_j^{(left)}$, respectively. Once defined the number of points in each sequence, say n_r and n_l respectively, the following representations are easily obtained:

$$\rho_j^{(right)} := \rho_0 - \frac{\gamma_r^j - 1}{\gamma_r - 1}\,\delta_r \quad , \quad j \in \{0, \ldots, (n_r-1)\} \tag{257}$$

$$\rho_j^{(left)} := \rho_{n-1} + \frac{\gamma_l^{(n_r+n_l-2)-j} - 1}{\gamma_l - 1}\,\delta_l \quad , \quad j \in \{(n_r-1), \ldots, (n_r+n_l-2)\} \tag{258}$$

where:

$$\delta_r := (\rho_0 - \rho_{i_\star})\ \frac{\gamma_r - 1}{\gamma_r^{(n_r-1)} - 1}$$

$$\delta_l := (\rho_{i_\star} - \rho_{n-1})\ \frac{\gamma_l - 1}{\gamma_l^{(n_l-1)} - 1}$$

and the new density sequence finally reads:

$$\rho'_j := \begin{cases} \rho_0 & , \ j = 0 \\ \rho_j^{(right)} & , \ j \in \{1, \ldots, (n_r - 2)\} \\ \rho_{i_\star} & , \ j = (n_r - 1) \\ \rho_j^{(left)} & , \ j \in \{n_r, \ldots, (n_r + n_l - 3)\} \\ \rho_{n-1} & , \ j = (n_r + n_l - 2) \end{cases} \tag{259}$$

The new density sequence has a noticeable advantage over the old one: it permits to analytically identify the nodal span to which a given value of the density ρ belongs by inverting (257) and (258) as follows (the cases $\rho = \rho_0$, $\rho = \rho_{i_\star}$ and $\rho = \rho_{n-1}$ are neglected because trivial):

$$\rho \in \begin{cases} \left(\rho'_{\sigma(\rho)+1}, \rho'_{\sigma(\rho)}\right] & , \ \rho_{i_\star} < \rho < \rho_0 \\ \left[\rho'_{\tau(\rho)}, \rho'_{\tau(\rho)-1}\right) & , \ \rho_{n-1} < \rho < \rho_{i_\star} \end{cases} \tag{260}$$

with:

$$\sigma(\rho) := \left\lfloor \frac{1}{\ln(\gamma_r)} \ln\left\{1 + (\rho_0 - \rho)\ \frac{\gamma_r - 1}{\delta_r}\right\}\right\rfloor \tag{261}$$

$$\tau(\rho) := (n_r + n_l - 2) - \left\lfloor \frac{1}{\ln(\gamma_l)} \ln\left\{1 + (\rho - \rho_{n-1})\ \frac{\gamma_l - 1}{\delta_l}\right\}\right\rfloor \tag{262}$$

where, of course, the symbol $\lfloor\cdot\rfloor$ denotes the floor function.

Once defined the new density sequence ρ'_j, a new table can be built either by solving the o.d.e. (250) once more, now in correspondence of the sequence ρ'_j, or by interpolating the original table. The latter strategy is considered here and the following new table, in particular, is built:

$$\left(\rho'_j\,,\, p'_j\,,\, a'_j\right) \quad , \quad j \in \{0, \ldots, (n_r + n_l - 2)\} \tag{263}$$

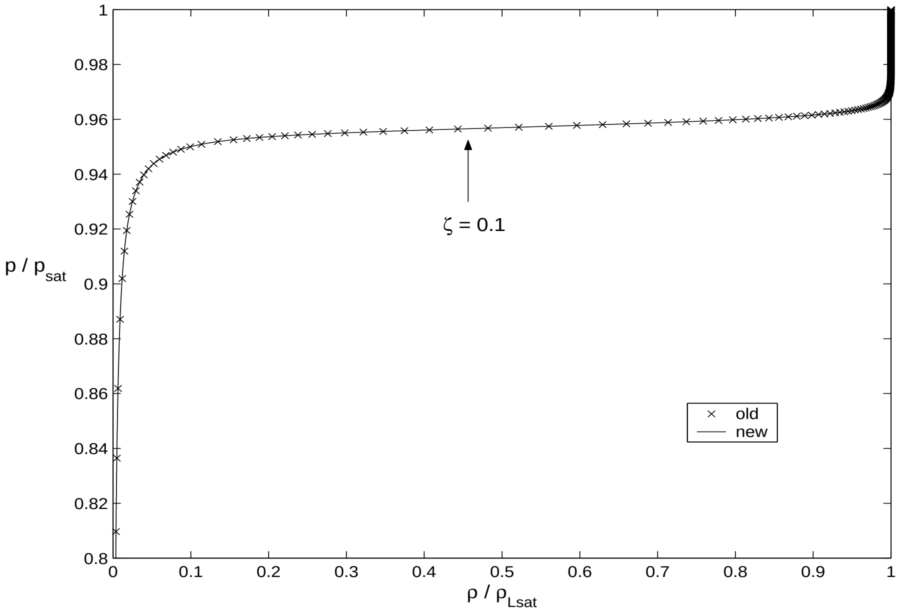

Figure 46: Comparison between the barotropic curves extracted from tables (256) ("old") and (263) ("new") for water at $T_L = 20°\text{C}$. Data: $n = 8127$, $i_\star = 6586$, $\gamma_r = \gamma_l = 1.004$, $n_r = 6587$ and $n_l = 1541$. The y-axis is cut for ease of readability.

by linearly interpolating the original one (256) in correspondence of the new density sequence (259). Clearly, the original table can be discarded at this point, since it is never accessed by the considered algorithm. It may be worth noticing that, besides being attractive for its simplicity, a linear interpolation preserves the strict monotonicity of the p-ρ curve.

For suitable values of the relevant parameters, the new table very well approximates the original one, as shown for instance in Fig. 46. It is therefore natural to define the following two-step access strategy based on the new table (263):

- given an input density ρ (the cases $\rho = \rho_0$, $\rho = \rho_{i_\star}$ and $\rho = \rho_{n-1}$ are not considered here because trivial), the corresponding span within the new table (263) is identified, by means of (260)-(262);
- the values of p and a corresponding to ρ are then defined by linear interpolation within the identified span. Of course, this procedure can be extended to an arbitrary number of dependent variables (e.g. the function Ψ, defined in (69), to be used for solving RPs associated with convex state laws, see sec. 2.5.1).

Evidently, the aforementioned access strategy is more efficient than a crude look-up within the original table (256). A similar technique can be defined for pressure-based algorithms, as outlined in sec. B.

4.1.4 Convexity of the chosen state law

In consideration of the equality (59), the original convexity condition (58) is introduced as follows:

$$2\,a\;c(\rho) > 0$$

where $c(\rho)$ is defined by (60). Then, by substituting the expression of $c(\rho)$, the condition above is recast as follows, for later convenience:

$$2\,\frac{\varphi}{\rho} + \frac{\mathrm{d}\varphi}{\mathrm{d}\rho} > 0 \tag{264}$$

where:

$$\varphi := a^2(\rho)$$

The condition (264) is exploited below for assessing the convexity of the chosen barotropic state law.

For the pure liquid model (247)-(248) the following equality holds:

$$2\,\frac{\varphi}{\rho} + \frac{\mathrm{d}\varphi}{\mathrm{d}\rho} = \frac{1}{\beta_L\,\rho^2}$$

and therefore the convexity condition (264) is clearly satisfied ($\beta_L > 0$ by definition). This holds true also for the linearized liquid model (249), as already noticed in Note 44 (sec. 4.1.1).

In view of the fact that $\varphi = a^2 = \mathrm{d}p/\mathrm{d}\rho$, the cavitation model (250) can be formally written in a more general way as follows:

$$\varphi = \frac{p}{\rho\;\omega\left(\rho, p(\rho)\right)} \tag{265}$$

where the function ω represents the right-hand side of (250). Then, by differentiating (265), the following equality is obtained:

$$2\,\frac{\varphi}{\rho} + \frac{\mathrm{d}\varphi}{\mathrm{d}\rho} = \frac{\varphi}{\rho} + \frac{\varphi^2}{p}\left\{1 - \rho\left(\frac{\partial\omega}{\partial\rho} + \varphi\,\frac{\partial\omega}{\partial p}\right)\right\} \tag{266}$$

The right-hand side of (266) can be exactly computed during the numerical integration of the o.d.e. (250) (of course, the partial derivatives of ω are known functions) and therefore it is possible to assess the convexity of the cavitating branch of specific state law as well.

Both the barotropic laws shown in Fig. 44, for instance, turn out to be convex. Hence, in spite of the fact that the convexity marker $c(\rho)$ defined in (60) exhibits a jump at the junction between the liquid and the cavitating branch (which is due to the discontinuity of $\mathrm{d}a/\mathrm{d}\rho$ across the saturation point (ρ_{Lsat}, p_{sat})), the corresponding unified barotropic curve (i.e. (247) coupled with (255)) can be classified as convex on the whole [69]. However, the aforementioned discontinuity of $\mathrm{d}a/\mathrm{d}\rho$ across the saturation point is not a "pathology" affecting the chosen cavitation model. On the contrary, it reflects, within the limits of the homogeneous flow modelling, the characteristic behaviour of the state law at phase transition. Indeed, in general, "phase transitions in the fluid are a principal cause of non-convexity, since the sound speed in a mixed phase region is smaller than in the pure phase" [69].

It may be worth remarking that the approximations introduced when deriving homogeneous flow cavitation models may affect the convexity of the resulting state laws. Indeed, even small differences between two given homogeneous flow models can lead to substantially different wave solutions of the same system of governing equations. For instance, a non-convex barotropic state law is considered in [103], which is qualitatively similar to those shown in Fig. 44. This law, which is smooth within the mixture region and which allows for smooth junctions with a pure liquid and a pure vapour barotropic models to be defined, is exploited in [103] to solve a RP by following [113]. Besides the classical rarefaction and shock waves presented in sec. 2.4.2, so called "composite" waves appear as part of the solution, which are defined by juxtaposing up to three classical waves in an alternate fashion (i.e. shock-rarefaction, rarefaction-shock, shock-rarefaction-shock and rarefaction-shock-rarefaction [52]). According to the author, the aforementioned sensitivity, besides highlighting the key role that modelling plays in this context, can encourage to also consider cavitation models which expressly take into account additional physical effects, e.g. non-homogeneous models (see sec. 1.3) or models directly incorporating thermodynamic effects related to phase transition. This opinion seems to be somehow supported by the fact that difficulties arise in applying common entropy conditions (see sec. 2.3.3) for selecting numerical solutions to classical p.d.e.s coupled with the state laws provided by classical homogeneous flow models (see e.g. [3] and [69]). Alternative approaches (e.g. the entropy-satisfying procedure based on the mixture thermodynamics which is proposed in [3]) should be carefully considered.

[52] No contact discontinuities are involved in the solution of the system at hand [103].

4.2 Numerical results

The Roe flux function, the preconditioning strategy and the linearized implicit time-advancing respectively presented in secs. 3.3, 3.4 and 3.5 have been originally introduced in [91]. A qualitative appraisal of the considered numerical ingredients is reported in the aforementioned document, based on the quasi-1D water flow within a convergent-divergent nozzle. In particular, the state law reported in Fig. 44 which is associated with $\zeta = 0.1$ is considered in order to numerically simulate both non-cavitating and cavitating flows. The obtained results, simply recalled here for conciseness, show that:

- the semi-discrete scheme based on the proposed Roe flux function exhibits accuracy problems at the low Mach numbers typical of liquid flows. The considered preconditioning strategy effectively overcomes this problem (in particular, a local preconditioning strategy of the type of that one mentioned in sec. 6.1.6 turns out to be effective also when cavitation occurs);

- the preconditioning technique restricts the stability of the considered explicit time-advancing algorithm (a $4-$th order Runge-Kutta scheme). The proposed linearized implicit strategy counteracts this problem: it permits to efficiently advance in time the non-cavitating simulations. However, when cavitation takes place, a noticeable time-step restriction must be accepted; in particular, the allowable time-step turns out to be of the order of that one required by the explicit non-preconditioned scheme.

The relevant numerical experiments reported in secs. 3.3, 3.4 and 3.5 are in agreement with the aforementioned results. The time-step reduction which must be introduced when considering cavitating flows, in particular, is due to the occurrence of noticeable discontinuities -especially as far as the Mach number and the density are concerned- which are associated with the inception of cavitation (see secs. 1.4 and 4.1.2). However, as discussed in sec. 3.5.7, this problem does not seem to be specifically introduced by the proposed linearization technique (222). Furthermore, it has been also observed by performing a rather extensive number of numerical simulations [7], based on the proposed linearization (222) and involving a different homogeneous flow cavitation model (namely the instance of the well-known barotropic cavitation model of Delannoy which is reported in [22]).

Clearly, once introduced a convex instance of the unified barotropic curve (247)-(255), it is possible to recall the material introduced in sec. 2.5.3 in order to exactly solve 1D Riemann problems (RPs) based on the considered

Benchmark	Liquid	T_L	ζ	ρ_L	u_L	ρ_R	u_R	t_{eval}
B4	water	293.16	0.1	998	-0.1	998	0.1	1

Table 13: Considered benchmark.

state law. These, in turn, provide exact benchmarks for validating 1D numerical methods dealing with cavitating flows and permit, in particular, to accurately investigate the behaviour of the considered numerical schemes at cavitation inception (thus addressing most of the difficulties related to the phase transition, as described by a homogeneous flow model). A systematic study of this type is postponed to a subsequent research stage; nevertheless, an illustrative test-case is considered in the sequel, showing some features that characterize the numerical discretization of the phase transition, as described by the unified barotropic model (247)-(255).

Benchmark

The considered benchmark is defined in Tab. 13. The mixture branch of the chosen state law is one of the two curves reported in Fig. 44. The relevant non-dimensional dependent parameters for the expressions (247) and (250) are: $\vartheta \approx 8.55 \cdot 10^5$ (see [26] and [83]), $\sigma_1 \approx 1.33 \cdot 10^3$, $\sigma_2 \approx -0.73$ and $\sigma_3 \approx 0.78$ (see [10], [26] and [83]). At the chosen temperature T_L, the liquid saturation density is $\rho_{Lsat} = 997.95$ and therefore the IC in Tab. 13 defines two liquid states [53] (passive scalars are neglected for the sake of simplicity). Moreover, the speeds u_L and u_R are chosen so as to obtain two rarefactions (symmetrical with respect to the original discontinuity $x = 0$) which lead to a cavitating star region characterized by $\rho_\star \approx 960.47 < \rho_{Lsat}$ and $u_\star = 0$ (by symmetry). The sound speed in the liquid is $a_L = a_R \approx 1415.63$ while in the cavitating region it falls down to approximately $a_{cav} \approx 0.37$; the resulting flow is entirely subsonic [54]. In consideration of the aforementioned variation of the sound speed, it is to be expected that the star region is hardly observable as part of the solution.

[53] The SI units are tacitly understood, see Note 3 in sec. 2.2.

[54] Sonic conditions are deliberately avoided, see Note 32 in sec. 3.3.1.

Test-case	Benchmark	μ	(n_L, n_R)	τ
LdA1	B4	1	$(2,2)\cdot 10^3$	10^{-4}
LdA2	B4	1	$(2,2)\cdot 10^3$	10^{-3}
LdA3	B4	1	$(2,2)\cdot 10^3$	10^{-2}

Table 14: Considered test-cases.

Discretization

The discrete scheme (224) is considered (more precisely, its basic-1D counterpart not involving the passive scalar ξ), associated with a uniform space discretization having measure $\mu = 1$ and a constant time-step τ. Transmissive BCs of the type of (120) are adopted, leading to equations similar to (235). The considered test-cases are reported in Tab. 14. The numerical approximation of ρ, p and u is shown in Figs. 47-51. It should be noticed that:

- the density undergoes a spike-like variation close to the cavitating region. It is practically impossible to distinguish the head as well as the tail of the density waves (see Note 26 in sec. 3.2.2) in Fig. 47, because the density variation close to rarefaction head is squashed by the considerable variation occurring towards the cavitating region. Furthermore, the width of the star region is not resolved by the adopted space discretization.
  When examining in Fig. 48 a narrower sub-domain around $x = 0$ it is evident that, as expected, the accuracy of the numerical solution improves when adopting smaller time-steps. However, it is extremely difficult to accurately describe the cavity, which only occupies a very small region close to the minimum of the "exact" curve shown in Fig. 48 (see the following point). Indeed, the characteristic size of the cavity is $\mathrm{O}\left(a_{cav} \cdot t_{eval}\right) = \mathrm{O}\left(10^{-1}\right)$, clearly finer than the adopted space discretization [55].

- the pressure exhibits a remarkably different trend with respect to the density, as shown in Fig. 49. Indeed:

[55] A uniform space discretization is adopted for consistency with the other 1D numerical experiments reported in the present document. A finer discretization is not considered in order not to introduce a computational overhead within the liquid region (which represents the vast majority of the computational domain).

  - the head of the rarefactions is clearly visible; that one of the left rarefaction, for example, is marked by P1 in the considered figure;
  - most of the pressure variation takes place, in practice, near the head of the rarefactions. For instance, as far as the left rarefaction is concerned, the pressure abruptly reaches the saturation value p_{sat}, marked by P2 in Fig. 49, as well as the “right corner” of the relevant pressure curve in Fig. 44 (very close to the right margin of the figure), marked by P3 in Fig. 49. The transition between P2 and P3 is aligned with that one between P1 and P2 (i.e. no abrupt changes occur when entering the mixture region). Indeed, the rightmost portion of the relevant cavitating curve in Fig. 44 is practically vertical near the saturation point (due to the very weak compressibility of the liquid) and therefore it behaves like a prolongation of the adopted liquid model;
  - the cavity, whose width is $\mathrm{O}\left(10^{-1}\right)$ (see above), is indicated in Fig. 49 by P4. When moving from P3 to P4, the pressure weakly decreases (the variation is not resolved in the figure) and therefore this arc is not part of the star region of the considered RP (where the solution is constant, see sec. 2.5.3). Indeed, most of this arc corresponds to the practically horizontal portion of the relevant pressure curve in Fig. 44 and the pressure decrease occurring near the left extreme of the aforementioned curve originates the variation shown in Fig. 50;

- as shown in Fig. 51, the approximation of u is reasonably good, even near the cavity

In consideration of the previous points, it is clear that an accurate description of the ratrefaction’s tail and, more in general, of the cavity is only possible at the cost of a very fine space discretization. Of course, several numerical investigations of the type of that one reported above can be performed by exploiting the chosen barotropic state law (the discretization of the sound speed, for instance, is considered in Fig. 52). However, as stated above, such an investigation is postponed to a subsequent research stage.

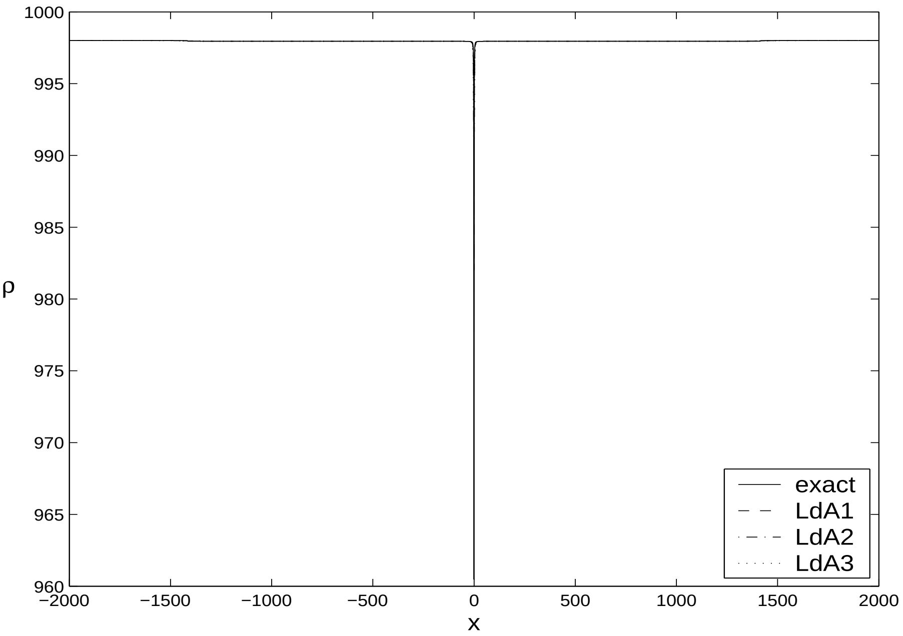

Figure 47: Approximation of ρ for the test-cases reported in Tab. 14.

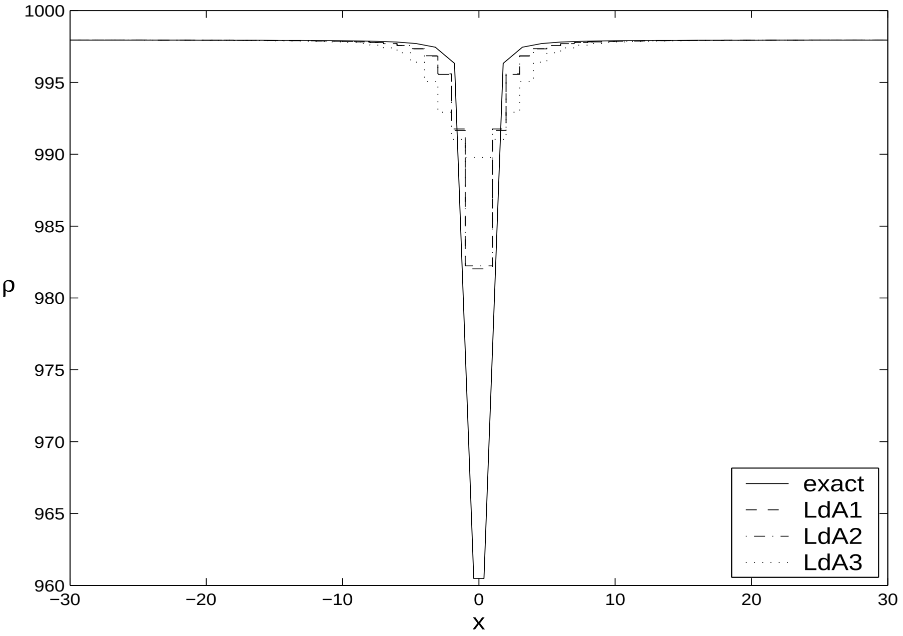

Figure 48: Detail of Fig. 47.

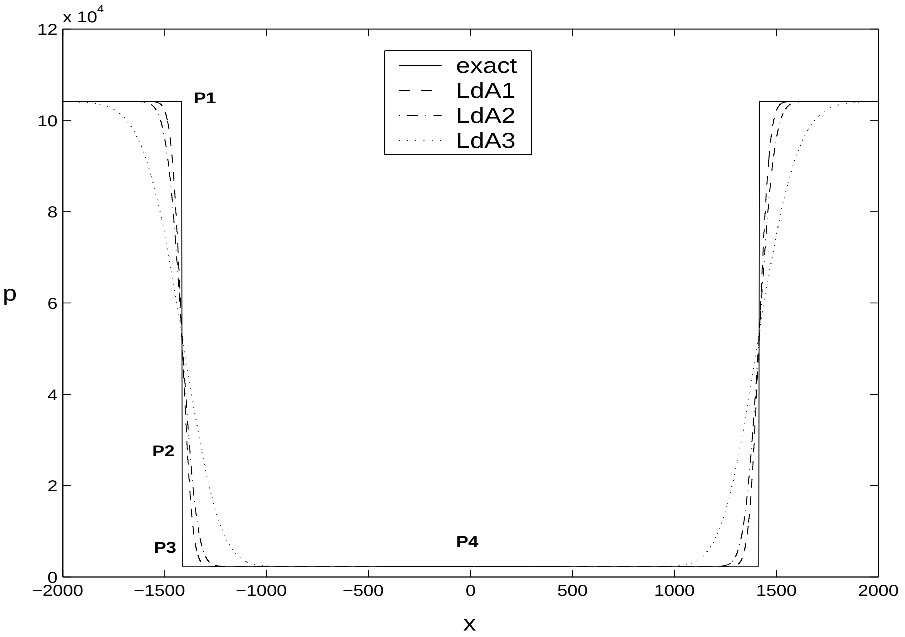

Figure 49: Approximation of p for the test-cases reported in Tab. 14. The labels P1-P4 are added for ease of discussion.

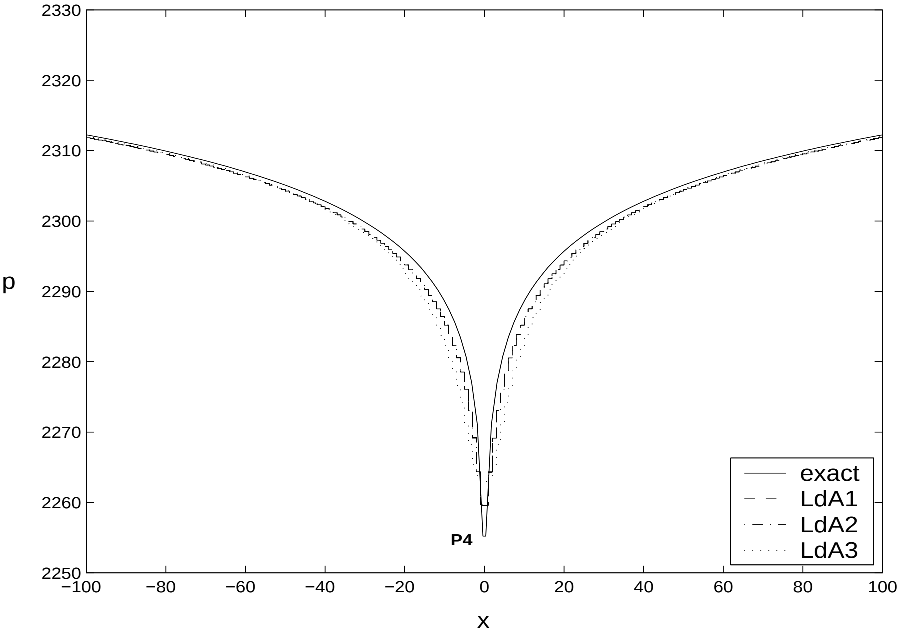

Figure 50: Detail of Fig. 49.

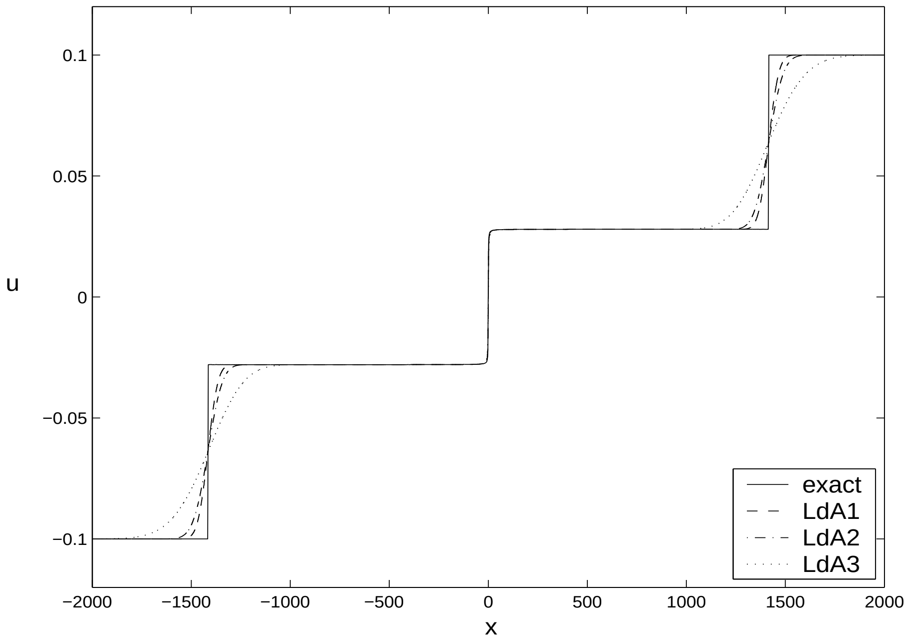

Figure 51: Approximation of u for the test-cases reported in Tab. 14.

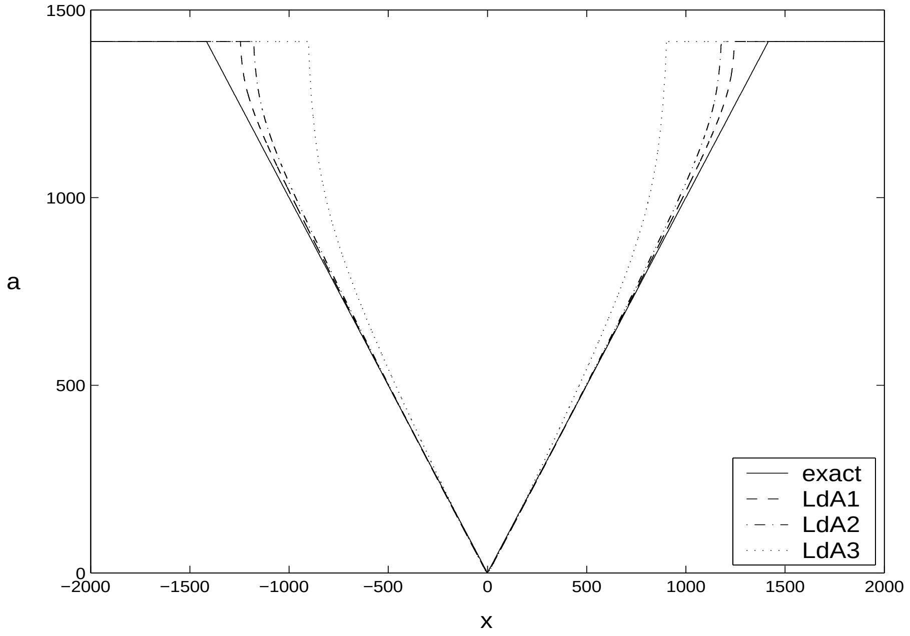

Figure 52: Approximation of a for the test-cases reported in Tab. 14.

5 3D Numerical method

In the present section, a linearized implicit discrete scheme is proposed for solving the 3D governing equations introduced in secs. 2.2.1 and 2.2.2, based on some numerical ingredients introduced in sec. 3. By adopting the architecture of the numerical frame mentioned in the introduction to the present document (namely the AERO code), the space and time discretizations are kept separate from each other.

As far as the space discretization is concerned, some basic issues regarding the considered unstructured grids are recalled in sec. 5.1.1. Then, a generalization of the Roe numerical flux proposed in sec. 3.3 is discussed in sec. 5.1.2. Moreover, in sec. 5.1.3 the preconditioning technique introduced in sec. 3.4 is incorporated into the considered 3D Roe numerical flux. Finally, once specified the discretization of the convective fluxes, the relevant semi-discrete formulation is introduced in sec. 5.1.4 and extended to rotating frames in sec. 5.1.5.

As far as the time discretization is concerned, in sec. 5.2.1 the linearization of the Roe numerical flux function proposed in sec. 3.5 is generalized to the present 3D context. Furthermore, in sec. 5.2.2 a linearized implicit discrete scheme is defined, based on the relevant material introduced in the preceding sections (numerical simulations exploiting this scheme are reported in sec. 6).

5.1 Space discretization

In this section, the main issues regarding the adopted finite volume space discretization are discussed.

5.1.1 Finite volume approximation

The considered space discretization is based on the finite volume approach introduced in sec. 3.1.1; the definition of the finite volume cells for the 3D case at hand is described below. At a preliminary stage, the considered 3D (bounded) computational domain $\mathcal{D} \in \mathbb{R}^3$ is approximated by means of a polyhedral domain $\mathcal{D}^{pol}$ which, in turn, is divided into N_t tetrahedra having vertices $\mathbf{P}_i$, with $i \in \mathcal{I} := \{1, \ldots, N_c\}$. Let T_h, with $h \in \mathcal{H} := \{1, \ldots, N_t\}$, denote the $h-$th tetrahedron; the following relations are (by construction) satisfied:

$$T_{h_1} \cap T_{h_2 \neq h_1} = \{0\} \quad , \quad \mathcal{D}^{pol} = \bigcup_{h \in \mathcal{H}} T_h$$

The $i-$th finite volume cell C_i, associated with $\mathbf{P}_i$, is given by:

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

where:

- $t(i) \subset \mathcal{H}$ is the set of indexes marking those tetrahedra which share $\mathbf{P}_i$ as a vertex;
- $C_r^{(h)}$ represents the subset of T_h which is defined by further dividing T_h into 24 sub-tetrahedra by means of its median planes [56] and subsequently considering those 6 sub-tetrahedra which share $\mathbf{P}_r$ as a vertex.

Clearly, there is a finite volume cell for each vertex [57]. Moreover, the resulting finite volume discretization clearly verifies the following relations:

$$C_{i_1} \cap C_{i_2 \neq i_1} = \{0\} \quad , \quad \mathcal{D}^{pol} = \bigcup_{i \in \mathcal{I}} C_i$$

and it is sometimes referred to as a “dual mesh” (see e.g. [39]), by virtue of the specific procedure which is adopted in order to build the cells by starting from the tetrahedra.

An example of the construction of the finite volume cells is shown in Fig. 53, for the 2D counterpart of the aforementioned 3D case. In this figure the tetrahedra are replaced with triangles in the $x_1 - x_2$ plane (whose vertices and edges are respectively marked by circles and dashed lines) and the median planes reduce to the ordinary medians (marked by dotted lines). Each triangle is then divided into 6 sub-triangles by the medians and 2 of them are associated with each vertex. The boundary of the cell C_i associated with $\mathbf{P}_i$ is identified by a solid line and the portion of this boundary representing, in particular, the interface between C_i and C_h is highlighted by a thicker line.

Let μ_i represent the measure of C_i. On C_i the exact solution $\mathbf{q}(\mathbf{x}, t)$, where $\mathbf{x} \in \mathbb{R}^3$ denotes the position vector and $\mathbf{q}$ represents the conservative state vector defined in (9), is approximated by a semi-discrete function $\mathbf{q}_i(t)$ which is considered as an approximation of the mean value of $\mathbf{q}(\mathbf{x}, t)$ over C_i (in analogy with (93)):

$$\mathbf{q}_i(t) \approx \frac{1}{\mu_i} \int_{C_i} \mathbf{q}(\mathbf{x}, t)\, \mathrm{dV} \tag{267}$$

[56] Each median plane is associated with an edge. The median plane relative to a certain edge $\tilde{e}$ contains $\tilde{e}$ as well as the middle point of the (unique) edge $\bar{e}$ which is not directly connected to $\tilde{e}$.

[57] The considered finite volume discretization can be regarded to as a “cell vertex” one [39], even if it is not necessary -to the purposes of the present study- to associate the quantities defined on C_i with a specific point belonging to C_i (in particular with $\mathbf{P}_i$).

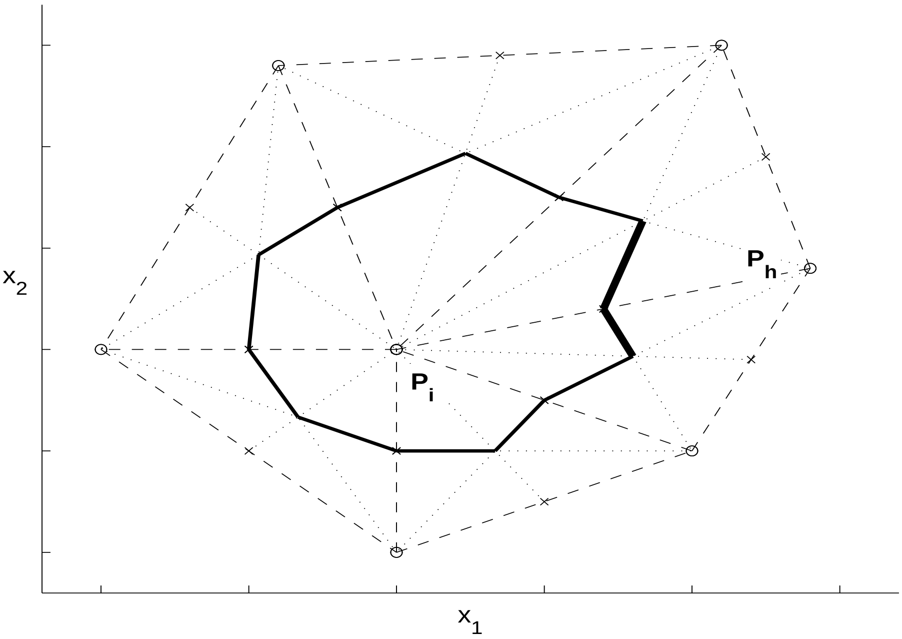

Figure 53: Example of the construction of a 2D finite volume cell by a dual mesh approach based on the medians.

The differential system defining $\mathbf{q}_i$ is obtained by discretizing the integral balance (8) over the control volume C_i. To the purpose, by virtue of (267), the time-derivative in (8) is naturally approximated as follows:

$$\partial_t \int_{C_i} \mathbf{q}(\mathbf{x}, t)\, \mathrm{dV} \approx \mu_i \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{q}_i \tag{268}$$

while the term involving the flux is discretized as described in sec. 5.1.4.

5.1.2 A 3D Roe numerical flux for generic barotropic state laws

In the spirit of (96), let:

$$\boldsymbol{\phi}\left(\mathbf{q}_i, \mathbf{q}_j, \hat{\boldsymbol{\nu}}_{ij}\right) \tag{269}$$

denote a 3D numerical flux from $\mathbf{q}_i$ to $\mathbf{q}_j$, along the direction $\hat{\boldsymbol{\nu}}_{ij}$. An instance of the aforementioned flux function is defined in the present section, based on the proposed augmented-1D Roe numerical flux (156)-(160).

Frame change and rotational invariance

Let $\mathbf{R} \in \mathbb{R}^{3\times 3}$ denote an orthogonal matrix ($\mathbf{R}^{-1} = \mathbf{R}^T$) associated with a rotation of the chosen Cartesian frame (see sec. 2.2.1). By introducing a matrix $\bar{\mathbf{R}} \in \mathbb{R}^{4\times 4}$ defined as follows:

$$\bar{\mathbf{R}} := \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{R} \end{pmatrix}$$

it is straightforward to compactly apply the aforementioned rotation to the considered state vector $\mathbf{q} \in \mathbb{R}^4$ (whose first component is a scalar, obviously invariant with respect to a frame change) as follows:

$$\mathbf{q} \longrightarrow \mathbf{q}' := \bar{\mathbf{R}} \cdot \mathbf{q}$$

In order to correctly discretize the considered balance (8), which is properly formulated as a tensorial relation, the flux function (269) must satisfy the following property (rotational invariance, see e.g. [39] and [98]):

$$\boldsymbol{\phi}\left(\mathbf{q}_i, \mathbf{q}_j, \hat{\boldsymbol{\nu}}_{ij}\right) = \bar{\mathbf{R}}^{-1} \cdot \boldsymbol{\phi}\left(\mathbf{q}'_i, \mathbf{q}'_j, \hat{\boldsymbol{\nu}}'_{ij}\right) \tag{270}$$

where, of course, $\hat{\boldsymbol{\nu}}'_{ij}$ corresponds in the rotated frame to $\hat{\boldsymbol{\nu}}_{ij}$:

$$\hat{\boldsymbol{\nu}}'_{ij} := \mathbf{R} \cdot \hat{\boldsymbol{\nu}}_{ij}$$

Sweep approximation

A frame rotation $\mathbf{R}$ is considered; without any loss of generality, the rotated direction $\hat{\boldsymbol{\nu}}'_{ij}$ is supposed to coincide with the versor $\hat{\mathbf{e}}'^{(k)}$ associated with the $k-$th direction x'_k ($k \in \{1, 2, 3\}$) of the rotated frame:

$$\hat{\boldsymbol{\nu}}'_{ij} = \hat{\mathbf{e}}'^{(k)} \tag{271}$$

Moreover, a basic-1D flow is assumed to take place along x'_k (this assumption plays a fundamental role in the subsequent derivation). Thus, by recalling the $k-$th sweep of the relevant 3D governing equations (see sec. 2.2.5) written in the rotated frame, it is possible to define an instance of the flux function $\boldsymbol{\phi}\left(\mathbf{q}'_i, \mathbf{q}'_j, \hat{\boldsymbol{\nu}}'_{ij}\right)$ appearing in (270). More precisely, due to the formal identity between the augmented-1D equations and the 1D sweeps of the 3D equations (see sec. 2.2.5), the considered instance of $\boldsymbol{\phi}\left(\mathbf{q}'_i, \mathbf{q}'_j, \hat{\boldsymbol{\nu}}'_{ij}\right)$ can be defined by introducing a Roe numerical flux $\boldsymbol{\phi}'^{ROE}_{ij}$ which generalizes the proposed one (156)-(160), as described below.

The centred component of $\boldsymbol{\phi}'^{ROE}_{ij}$ is firstly considered. Let $\mathbf{f}^{(\hat{\boldsymbol{\nu}}'_{ij})}\left(\mathbf{q}'\right)$ denote the augmented-1D analytical flux along x'_k. In consideration of (271), it is straightforward to derive from (10) the following representation:

$$\mathbf{f}^{(\hat{\boldsymbol{\nu}}'_{ij})}\left(\mathbf{q}'\right) = \left(\mathbf{u}'^T \cdot \hat{\boldsymbol{\nu}}'_{ij}\right) \mathbf{q}' + p \begin{pmatrix} 0 \\ \hat{\boldsymbol{\nu}}'_{ij} \end{pmatrix} \tag{272}$$

where, of course:

$$\mathbf{u}' := \mathbf{R} \cdot \mathbf{u}$$

The sought centred component, which generalizes the augmented-1D one (158), can then be defined as follows:

$$\boldsymbol{\phi}'^{ROE}_{c,ij} := \frac{1}{2} \left(\mathbf{f}^{(\hat{\boldsymbol{\nu}}'_{ij})} \left(\mathbf{q}'_i \right) + \mathbf{f}^{(\hat{\boldsymbol{\nu}}'_{ij})} \left(\mathbf{q}'_j \right) \right)$$

The upwind component of $\boldsymbol{\phi}'^{ROE}_{ij}$ is considered in the sequel. According to the sweep approximation, the velocity components associated with the versors $\hat{\mathbf{e}}'^{(h)}$, $h \neq k$, of the rotated frame are treated as passive scalars. Consequently, the Roe averages to be introduced in a Roe matrix for the $k-$th sweep under consideration are a_{ij}, defined in (154), and $\mathbf{u}'_{ij}$, with $\mathbf{u}'_{ij}$ defined as follows:

$$\mathbf{u}'_{ij} := \mathbf{R} \cdot \mathbf{u}_{ij} \tag{273}$$

where:

$$\mathbf{u}_{ij} := \frac{\sqrt{\rho_i}\, \mathbf{u}_i + \sqrt{\rho_j}\, \mathbf{u}_j}{\sqrt{\rho_i} + \sqrt{\rho_j}} \tag{274}$$

The above definition, in particular, extends the Roe averages (153) to the present context. Let $\tilde{\mathbf{J}}'_{ij}$ denote the sought Roe matrix, which clearly generalizes the matrix $s_{ij}\, \tilde{\mathbf{J}}^{(A)}_{ij}$ appearing in (160). By a straightforward extension of (150), it is possible to define $\tilde{\mathbf{J}}'_{ij}$ by evaluating the Jacobian associated with the direction x'_k -which, in turn, can be derived from (16)- in correspondence of the aforementioned Roe averages, namely:

$$\tilde{\mathbf{J}}'_{ij} := \begin{pmatrix} 0 & \hat{\boldsymbol{\nu}}'^T_{ij} \\ \\ a^2_{ij}\, \hat{\boldsymbol{\nu}}'_{ij} - \sigma_{ij}\, \mathbf{u}'_{ij} & \mathbf{u}'_{ij} \cdot \hat{\boldsymbol{\nu}}'^T_{ij} + \sigma_{ij}\, \mathbf{I} \end{pmatrix} \tag{275}$$

where:

$$\sigma_{ij} := \mathbf{u}'^T_{ij} \cdot \hat{\boldsymbol{\nu}}'_{ij} = \mathbf{u}^T_{ij} \cdot \hat{\boldsymbol{\nu}}_{ij} \tag{276}$$

Then, once recalled the definition of Δ^{ij} given in (155), the sought upwind component generalizing (159)-(160) can be defined as follows:

$$\boldsymbol{\phi}'^{ROE}_{u,ij} \ := \ \mathbf{D}'_{ij} \cdot \Delta^{ij} \mathbf{q}' \tag{277}$$

$$\mathbf{D}'_{ij} \ := \ -\frac{1}{2} \left| \tilde{\mathbf{J}}'_{ij} \right|$$

and the resulting Roe flux function:

$$\boldsymbol{\phi}'^{ROE}_{ij} := \boldsymbol{\phi}'^{ROE}_{c,ij} + \boldsymbol{\phi}'^{ROE}_{u,ij} \tag{278}$$

can be considered as an instance of the flux function $\boldsymbol{\phi}\left(\mathbf{q}_i', \mathbf{q}_j', \hat{\boldsymbol{\nu}}_{ij}'\right)$, namely:

$$\boldsymbol{\phi}\left(\mathbf{q}_i', \mathbf{q}_j', \hat{\boldsymbol{\nu}}_{ij}'\right) = \boldsymbol{\phi}_{ij}'^{ROE} \tag{279}$$

3D Roe numerical flux

In view of (279), the representation of the considered instance of $\boldsymbol{\phi}\left(\mathbf{q}_i, \mathbf{q}_j, \hat{\boldsymbol{\nu}}_{ij}\right)$ can be derived from the definitions introduced in the previous paragraph, by a trivial change of notation. Nevertheless, such a representation is reported below for ease of presentation.

Let the considered 3D Roe numerical flux function $\boldsymbol{\phi}_{ij}^{ROE}$ be defined as follows:

$$\begin{aligned}
\boldsymbol{\phi}_{ij}^{ROE} &:= \boldsymbol{\phi}_{c,ij}^{ROE} + \boldsymbol{\phi}_{u,ij}^{ROE} && (280)\\
\boldsymbol{\phi}_{c,ij}^{ROE} &:= \frac{1}{2}\left(\mathbf{f}^{(\hat{\boldsymbol{\nu}}_{ij})}\left(\mathbf{q}_i\right) + \mathbf{f}^{(\hat{\boldsymbol{\nu}}_{ij})}\left(\mathbf{q}_j\right)\right) && (281)\\
\boldsymbol{\phi}_{u,ij}^{ROE} &:= \mathbf{D}_{ij} \cdot \Delta^{ij}\mathbf{q} \\
\mathbf{D}_{ij} &:= -\frac{1}{2}\left|\tilde{\mathbf{J}}_{ij}\right| && (282)
\end{aligned}$$

where the function $\mathbf{f}^{(\hat{\boldsymbol{\nu}}_{ij})}\left(\mathbf{q}\right)$ in (281) is straightforwardly derived from (272) as follows:

$$\mathbf{f}^{(\hat{\boldsymbol{\nu}}_{ij})}\left(\mathbf{q}\right) = \left(\mathbf{u}^T \cdot \hat{\boldsymbol{\nu}}_{ij}\right)\mathbf{q} + p\begin{pmatrix} 0 \\ \hat{\boldsymbol{\nu}}_{ij}\end{pmatrix} \tag{283}$$

and the Roe matrix $\tilde{\mathbf{J}}_{ij}$ in (282) is trivially derived from (275) as follows:

$$\tilde{\mathbf{J}}_{ij} := \begin{pmatrix} 0 & \hat{\boldsymbol{\nu}}_{ij}^T \\ a_{ij}^2\,\hat{\boldsymbol{\nu}}_{ij} - \sigma_{ij}\,\mathbf{u}_{ij} & \mathbf{u}_{ij}\cdot\hat{\boldsymbol{\nu}}_{ij}^T + \sigma_{ij}\,\mathbf{I}\end{pmatrix} \tag{284}$$

with σ_{ij} introduced in (276). From (279) it follows that the considered instance of the 3D flux function (269) reads:

$$\boldsymbol{\phi}\left(\mathbf{q}_i, \mathbf{q}_j, \hat{\boldsymbol{\nu}}_{ij}\right) = \boldsymbol{\phi}_{ij}^{ROE} \tag{285}$$

with $\boldsymbol{\phi}_{ij}^{ROE}$ defined in (280)-(284).

Note 46 *The numerical flux (285) evidently satisfies the following relations:*

$$\begin{aligned}
\boldsymbol{\phi}\left(\mathbf{q}_j, \mathbf{q}_i, \hat{\boldsymbol{\nu}}_{ji} = -\hat{\boldsymbol{\nu}}_{ij}\right) &= -\boldsymbol{\phi}\left(\mathbf{q}_i, \mathbf{q}_j, \hat{\boldsymbol{\nu}}_{ij}\right) && (286)\\
\boldsymbol{\phi}\left(\mathbf{q}_i, \mathbf{q}_j = \mathbf{q}_i, \hat{\boldsymbol{\nu}}_{ij}\right) &= \mathbf{f}^{(\hat{\boldsymbol{\nu}}_{ij})}\left(\mathbf{q}_i\right) && (287)
\end{aligned}$$

with $\mathbf{f}^{(\hat{\boldsymbol{\nu}}_{ij})}(\mathbf{q})$ *given by (283). The relation (286) clearly extends the conservation property (99) while the relation (287) provides a generalization of the consistency property (100).*

Moreover, the numerical flux (285) also satisfies the rotational invariance condition (270). Clearly, in order to verify the previous assertion it suffices to show that the following relation:

$$\boldsymbol{\phi}'^{ROE}_{ij} = \bar{\mathbf{R}} \cdot \boldsymbol{\phi}^{ROE}_{ij} \tag{288}$$

holds true. To the purpose, the centred components are firstly considered. Once noticed that the right-hand side of the expression (272) can be recast as follows (of course, $\mathbf{u}'^T \cdot \hat{\boldsymbol{\nu}}'_{ij} = \mathbf{u}^T \cdot \hat{\boldsymbol{\nu}}_{ij}$*):*

$$\left(\mathbf{u}^T \cdot \hat{\boldsymbol{\nu}}_{ij}\right) \bar{\mathbf{R}} \cdot \mathbf{q} + p\,\bar{\mathbf{R}} \cdot \begin{pmatrix} 0 \\ \hat{\boldsymbol{\nu}}_{ij} \end{pmatrix} = \bar{\mathbf{R}} \cdot \mathbf{f}^{(\hat{\boldsymbol{\nu}}_{ij})}(\mathbf{q})$$

with $\mathbf{f}^{(\hat{\boldsymbol{\nu}}_{ij})}(\mathbf{q})$ *given by (283), it is evident that:*

$$\boldsymbol{\phi}'^{ROE}_{c,ij} = \bar{\mathbf{R}} \cdot \boldsymbol{\phi}^{ROE}_{c,ij} \tag{289}$$

As far as the upwind components are concerned, the following relation (straightforwardly derived from the relevant definitions) can be introduced:

$$\tilde{\mathbf{J}}'_{ij} = \bar{\mathbf{R}} \cdot \tilde{\mathbf{J}}_{ij} \cdot \bar{\mathbf{R}}^{-1} \tag{290}$$

In consideration of the fact that the eigenvalues are invariant with respect to a frame change and by recalling the definition of the operator $|\cdot|$ *given in (6), it follows from (290) that:*

$$|\tilde{\mathbf{J}}'_{ij}| = \bar{\mathbf{R}} \cdot |\tilde{\mathbf{J}}_{ij}| \cdot \bar{\mathbf{R}}^{-1} \quad \Rightarrow \quad \mathbf{D}'_{ij} = \bar{\mathbf{R}} \cdot \mathbf{D}_{ij} \cdot \bar{\mathbf{R}}^{-1}$$

Hence, the right-hand side of (277) can be recast as follows (of course, $\Delta^{ij}\mathbf{q}' = \bar{\mathbf{R}} \cdot \Delta^{ij}\mathbf{q}$*):*

$$\bar{\mathbf{R}} \cdot \mathbf{D}_{ij} \cdot \bar{\mathbf{R}}^{-1} \cdot \bar{\mathbf{R}} \cdot \Delta^{ij}\mathbf{q} = \bar{\mathbf{R}} \cdot \mathbf{D}_{ij} \cdot \Delta^{ij}\mathbf{q}$$

and therefore:

$$\boldsymbol{\phi}'^{ROE}_{u,ij} = \bar{\mathbf{R}} \cdot \boldsymbol{\phi}^{ROE}_{u,ij} \tag{291}$$

As a result, the equality (288) immediately follows from (289) and (291), in view of the definitions (278) and (280).

5.1.3 Incorporation of the preconditioning strategy

It is possible to extend the preconditioning strategy introduced in sec. 3.4.3 so as to be incorporated into the proposed 3D Roe numerical flux (280)-(284). To the purpose, the preconditioner (194) is firstly recalled. Consistently with the sweep approximation introduced in the previous section, the representation of the preconditioner $\mathbf{P}'_{ij}$ -to be incorporated into the Roe flux $\boldsymbol{\phi}'^{ROE}_{ij}$ defined in (278)- can be derived from the expression (194) by replacing $(u_{ij}, \xi_{ij}, \eta_{ij})^T$ with the Roe averages (273), namely:

$$\mathbf{P}'_{ij} := \mathbf{I} + \left(\beta^2 - 1\right) \begin{pmatrix} 1 & \mathbf{0}^T \\ & \\ \mathbf{u}'_{ij} & \mathbf{O} \end{pmatrix} \tag{292}$$

The matrix (292) satisfies the following relation:

$$\mathbf{P}'_{ij} = \bar{\mathbf{R}} \cdot \mathbf{P}_{ij} \cdot \bar{\mathbf{R}}^{-1}$$

where, of course:

$$\mathbf{P}_{ij} := \mathbf{I} + \left(\beta^2 - 1\right) \begin{pmatrix} 1 & \mathbf{0}^T \\ & \\ \mathbf{u}_{ij} & \mathbf{O} \end{pmatrix} \tag{293}$$

with the Roe averages $\mathbf{u}_{ij}$ defined in (274). The preconditioner (293) must be associated with the Roe matrix $\tilde{\mathbf{J}}_{ij}$ defined in (284) and the resulting 3D preconditioned Roe numerical flux finally reads:

$$\boldsymbol{\phi}^{ROE,p}_{ij} := \boldsymbol{\phi}^{ROE}_{c,ij} + \boldsymbol{\phi}^{ROE,p}_{u,ij} \tag{294}$$

where the centred component $\boldsymbol{\phi}^{ROE}_{c,ij}$ is given by (281) while the upwind one reads:

$$\begin{aligned} \boldsymbol{\phi}^{ROE,p}_{u,ij} &:= \mathbf{D}^{p}_{ij} \cdot \Delta^{ij}\mathbf{q} \\ \mathbf{D}^{p}_{ij} &:= -\frac{1}{2} \left(\mathbf{P}_{ij}\right)^{-1} \cdot \left| \mathbf{P}_{ij} \cdot \tilde{\mathbf{J}}_{ij} \right| \end{aligned} \tag{295}$$

It may be worth remarking that, as for the starting augmented-1D case, the matrix $\mathbf{P}_{ij} \cdot \tilde{\mathbf{J}}_{ij}$ appearing in (295) is diagonalizable with real eigenvalues and therefore the operator $|\cdot|$ -defined in (6)- can be rightfully applied. The Roe numerical flux (294)-(295) can be considered as an instance of the 3D flux function (269); for later convenience, it is marked as follows:

$$\boldsymbol{\phi}^{p}\left(\mathbf{q}_i, \mathbf{q}_j, \hat{\boldsymbol{\nu}}_{ij}\right) := \boldsymbol{\phi}^{ROE,p}_{ij} \tag{296}$$

5.1.4 Discretization of the fluxes and semi-discrete formulation

The discretization of the surface integral appearing in the balance (8) is considered in the present section. More precisely, the convective flux across the boundary ∂C_i of the generic cell C_i is considered.

In general, the relevant integrand can be recast as follows:

$$\sum_{k=1}^{3} \hat{n}_k \, \mathbf{f}^{(k)} = \mathbf{f}^{(\hat{\mathbf{n}})}$$

where the definition of $\mathbf{f}^{(\hat{\mathbf{n}})}(\cdot)$ is trivially derived from (283), namely:

$$\mathbf{f}^{(\hat{\mathbf{n}})}(\mathbf{q}) := \left(\mathbf{u}^T \cdot \hat{\mathbf{n}}\right) \mathbf{q} + p \begin{pmatrix} 0 \\ \hat{\mathbf{n}} \end{pmatrix} \tag{297}$$

As far as the integration domain is concerned, it can be split into several parts, as described below. The cell C_i is adjacent to a certain number $s(i)$ of other cells C_j, clearly equal to the number of vertices $\mathbf{P}_j$ which are connected to $\mathbf{P}_i$ by an edge of the underlying tetrahedral lattice (see sec. 5.1.1). Consequently, the boundary ∂C_i of C_i can be decomposed as follows [58]:

$$\partial C_i = \left(\bigcup_{j \in s(i)} \partial C_i \cap \partial C_j \right) \cup \left(\partial C_i \cap \partial \mathcal{D}^{pol}\right) \tag{298}$$

where $\partial \mathcal{D}^{pol}$ denotes the boundary of the flow domain $\mathcal{D}^{pol}$. This boundary, in turn, is assumed to be split as follows:

$$\partial \mathcal{D}^{pol} = \partial \mathcal{D}^{I} \cup \partial \mathcal{D}^{O} \cup \partial \mathcal{D}^{B} \cup \partial \mathcal{D}^{C} \tag{299}$$

where $\partial \mathcal{D}^I$ and $\partial \mathcal{D}^O$ respectively denote the inflow and the outflow surfaces, $\partial \mathcal{D}^B$ represents the wall of a rigid body immersed within the flow (if any) and $\partial \mathcal{D}^C$ indicates a rigid wall encasing the flow. Consequently, once introduced the following definitions:

$$\mathcal{S}_{ij} := \partial C_i \cap \partial C_j \quad , \quad \boldsymbol{\varphi}_{ij} := \int_{\mathcal{S}_{ij}} \mathbf{f}^{(\hat{\mathbf{n}})} \, \mathrm{dS}$$

$$\mathcal{S}_{iX} := \partial C_i \cap \partial \mathcal{D}^X \quad , \quad \boldsymbol{\varphi}_{iX} := \int_{\mathcal{S}_{iX}} \mathbf{f}^{(\hat{\mathbf{n}})} \, \mathrm{dS} \quad , \quad X \in \{I, O, B, C\}$$

[58] A detailed characterization of the considered boundary can be found in [33].

the convective flux across the boundary ∂C_i in the balance (8) can be recast as follows:

$$\int_{\partial C_i} \mathbf{f}^{(\hat{\mathbf{n}})}\, \mathrm{dS} = \sum_{j \in s(i)} \boldsymbol{\varphi}_{ij} + \boldsymbol{\varphi}_{iI} + \boldsymbol{\varphi}_{iO} + \boldsymbol{\varphi}_{iB} + \boldsymbol{\varphi}_{iC} \tag{300}$$

The discretization of each flux appearing on the right-hand side of (300) is discussed below.

In order to define a numerical approximation $\tilde{\boldsymbol{\varphi}}_{ij}$ of $\boldsymbol{\varphi}_{ij}$, the following average direction $\hat{\boldsymbol{\nu}}_{ij}$ associated with $\mathcal{S}_{ij}$ is introduced:

$$\boldsymbol{\nu}_{ij} := \int_{\mathcal{S}_{ij}} \hat{\mathbf{n}}\ \mathrm{dS} \quad , \quad \hat{\boldsymbol{\nu}}_{ij} := \frac{\boldsymbol{\nu}_{ij}}{\|\boldsymbol{\nu}_{ij}\|} \tag{301}$$

and $\boldsymbol{\varphi}_{ij}$ is firstly approximated by a 3D numerical flux of the type of (269) crossing an "equivalent" planar surface having measure $\|\boldsymbol{\nu}_{ij}\|$ and normal $\hat{\boldsymbol{\nu}}_{ij}$, namely:

$$\boldsymbol{\varphi}_{ij} \approx \|\boldsymbol{\nu}_{ij}\|\ \boldsymbol{\phi}\left(\mathbf{q}_i, \mathbf{q}_j, \hat{\boldsymbol{\nu}}_{ij}\right)$$

In consideration of the fact that the convective flux along $\hat{\boldsymbol{\nu}}_{ij}$, as obtained by substituting $\hat{\mathbf{n}} = \hat{\boldsymbol{\nu}}_{ij}$ into (297), coincides with the expression (283), it is possible to choose the proposed 3D Roe flux $\boldsymbol{\phi}_{ij}^{ROE}$ defined in (280)-(284) for approximating $\boldsymbol{\varphi}_{ij}$. More in general, the 3D preconditioned Roe flux $\boldsymbol{\phi}_{ij}^{ROE,p}$ given in (294)-(295) can be considered and therefore, in view of (296), the following approximation is defined:

$$\boldsymbol{\varphi}_{ij} \approx \tilde{\boldsymbol{\varphi}}_{ij} := \|\boldsymbol{\nu}_{ij}\|\ \boldsymbol{\phi}^{p}\left(\mathbf{q}_i, \mathbf{q}_j, \hat{\boldsymbol{\nu}}_{ij}\right) \tag{302}$$

An approximation of the type of (302) is also adopted for the fluxes $\boldsymbol{\varphi}_{iI}$ and $\boldsymbol{\varphi}_{iO}$. More precisely, once introduced a fictitious inflow state vector $\mathbf{q}_i^I$, the following relation is introduced:

$$\boldsymbol{\varphi}_{iI} \approx \tilde{\boldsymbol{\varphi}}_{iI} := \|\boldsymbol{\nu}_{iI}\|\ \boldsymbol{\phi}^{p}\left(\mathbf{q}_i, \mathbf{q}_i^I, \hat{\boldsymbol{\nu}}_{iI}\right) \tag{303}$$

where $\hat{\boldsymbol{\nu}}_{iI}$ is defined in the spirit of (301). Similarly, the chosen approximation of the outflow flux reads:

$$\boldsymbol{\varphi}_{iO} \approx \tilde{\boldsymbol{\varphi}}_{iO} := \|\boldsymbol{\nu}_{iO}\|\ \boldsymbol{\phi}^{p}\left(\mathbf{q}_i, \mathbf{q}_i^O, \hat{\boldsymbol{\nu}}_{iO}\right) \tag{304}$$

where $\mathbf{q}_i^O$ represents a fictitious outflow state vector and $\hat{\boldsymbol{\nu}}_{iO}$ is defined in the spirit of (301). It should be noticed that the approximations (303) and (304), besides being consistent with the discretization of the inner fluxes (302), take into account the wave structure of the flow entering/exiting the computational domain by means of the upwinding component of the considered Roe numerical flux function.

At the walls $\partial\mathcal{D}^B$ and $\partial\mathcal{D}^C$ the classical slip condition [88]:

$$\mathbf{u}^T \cdot \hat{\mathbf{n}} = \mathbf{0} \tag{305}$$

is imposed, consistently with the adopted inviscid approximation (see sec. 2.2). The condition (305) can be introduced into (297), thus leading to the following approximations:

$$\boldsymbol{\varphi}_{iB} \approx \tilde{\boldsymbol{\varphi}}_{iB} := \|\boldsymbol{\nu}_{iB}\| \begin{pmatrix} 0 \\ p_i\, \hat{\boldsymbol{\nu}}_{iB} \end{pmatrix} \tag{306}$$

$$\boldsymbol{\varphi}_{iC} \approx \tilde{\boldsymbol{\varphi}}_{iC} := \|\boldsymbol{\nu}_{iC}\| \begin{pmatrix} 0 \\ p_i\, \hat{\boldsymbol{\nu}}_{iC} \end{pmatrix} \tag{307}$$

where $\hat{\boldsymbol{\nu}}_{iB}$ and $\hat{\boldsymbol{\nu}}_{iC}$ are clearly defined in the spirit of (301).

In consideration of the material introduced in the present section, the convective flux (300) is discretized as follows:

$$\int_{\partial C_i} \mathbf{f}^{(\hat{\mathbf{n}})}\, \mathrm{dS} \approx \tilde{\boldsymbol{\varphi}}_i := \sum_{j \in s(i)} \tilde{\boldsymbol{\varphi}}_{ij} + \tilde{\boldsymbol{\varphi}}_{iI} + \tilde{\boldsymbol{\varphi}}_{iO} + \tilde{\boldsymbol{\varphi}}_{iB} + \tilde{\boldsymbol{\varphi}}_{iC} \tag{308}$$

with $\tilde{\boldsymbol{\varphi}}_{ij}$, $\tilde{\boldsymbol{\varphi}}_{iI}$, $\tilde{\boldsymbol{\varphi}}_{iO}$, $\tilde{\boldsymbol{\varphi}}_{iB}$ and $\tilde{\boldsymbol{\varphi}}_{iC}$ respectively defined in (302), (303), (304), (306) and (307). The expression (308) can be formally introduced for all the finite volume cells; indeed, if $\mathcal{S}_{iX} = \{\emptyset\}$ $(X \in \{I, O, B, C\})$ then $\|\boldsymbol{\nu}_{iX}\| = 0$ and the term $\tilde{\boldsymbol{\varphi}}_{iX}$ correctly vanishes. As a result, by combining (268) and (308), the following semi-discrete formulation of the considered balance (8) is finally obtained:

$$\mu_i \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{q}_i + \tilde{\boldsymbol{\varphi}}_i = \mathbf{0} \quad , \quad i \in \mathcal{I} \tag{309}$$

5.1.5 Extension to rotating frames

Let $\mathcal{B}$ denote a rigid body immersed within the flow, which rotates with constant angular velocity $\boldsymbol{\omega}$ (e.g. an axial inducer of the type of those introduced in sec. 1). The representation of the governing equations with respect to a frame rotating with $\mathcal{B}$ (hereafter referred to as body-frame) is given in (17); the corresponding semi-discrete formulation is considered in the present section.

The external portion $\partial\mathcal{D}^{pol(ext)}$ of the boundary $\partial\mathcal{D}^{pol}$ in (299) is clearly given by:

$$\partial\mathcal{D}^{pol(ext)} := \partial\mathcal{D}^I \cup \partial\mathcal{D}^O \cup \partial\mathcal{D}^C$$

The surface $\partial\mathcal{D}^{pol(ext)}$ is here assumed to be symmetrical with respect to the rotation axis; in such a circumstance, it behaves like a fixed one in the body-frame and therefore it is possible to discretize the balance (17) without dealing with moving computational grids. While the previous assertion is clear as far as the inflow and outflow components are concerned [59], it may be useful to further discuss the term related to the external wall $\partial\mathcal{D}^C$. In the body frame, $\partial\mathcal{D}^C$ is a moving surface on which the slip condition is properly formulated as follows [88]:

$$\mathbf{u}^T \cdot \hat{\mathbf{n}} = (\boldsymbol{\omega} \wedge \mathbf{x})^T \cdot \hat{\mathbf{n}} \tag{310}$$

where the vector product on the right-hand side represents the dragging velocity associated with the point on $\partial\mathcal{D}^C$ which is identified by the position vector $\mathbf{x}$. However, by virtue of the assumed symmetry, the vectors $\boldsymbol{\omega}$, $\mathbf{x}$ and $\hat{\mathbf{n}}$ are necessarily coplanar and therefore the right-hand side of (310) is systematically equal to zero. As a result, the condition (310) reduces to its non-rotating counterpart (305) and $\partial\mathcal{D}^C$ behaves as a non-rotating boundary.

In view of the aforementioned considerations, it is possible to derive the sought semi-discrete formulation from the non-rotating one (309), as described below. Let $\mathbf{g}_i$ denote the centroid associated with C_i, namely:

$$\mathbf{g}_i := \frac{1}{\mu_i} \int_{C_i} \mathbf{x}\,\mathrm{dV} \tag{311}$$

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

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

where $\hat{\boldsymbol{\omega}}$ represents the versor associated with $\boldsymbol{\omega}$. Then, once introduced the following definition (derived from (18)):

$$\mathbf{s}_i := \|\boldsymbol{\omega}\| \begin{pmatrix} 0 \\ -\,2\,\hat{\boldsymbol{\omega}} \wedge \rho_i \mathbf{u}_i + \rho_i\,\|\boldsymbol{\omega}\|\,\mathbf{r}_i \end{pmatrix} \tag{312}$$

it is possible to approximate the right-hand side of the balance (17) -written for C_i- as follows:

$$\int_{C_i} \mathbf{s}\,\mathrm{dV} \approx \mu_i\,\mathbf{s}_i \tag{313}$$

Finally, by combining (309) and (313), it is straightforward to introduce the following semi-discrete formulation for the considered balance (17):

$$\mu_i \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{q}_i + \tilde{\boldsymbol{\varphi}}_i = \mu_i\,\mathbf{s}_i \quad , \quad i \in \mathcal{I} \tag{314}$$

[59] Of course, the rotation affects the representation of the fictitious state vectors $\mathbf{q}_i^I$ and $\mathbf{q}_i^O$ appearing in the approximations (303) and (304) which, however, can be formally kept.

5.2 Time discretization

A discrete scheme is presented, based on a generalization of the linearized implicit time-advancing proposed in sec. 3.5.

5.2.1 Linearization of the 3D Roe numerical flux

It turns out to be straightforward to extend the proposed linearization (229) of the preconditioned, augmented-1D Roe flux function to the 3D case. Indeed, as highlighted in sec. 5.1.2, the preconditioner $\mathbf{P}_{ij}$ defined in (293) and the Roe matrix $\tilde{\mathbf{J}}_{ij}$ defined in (284) respectively generalize their augmented-1D counterparts (namely $\mathbf{P}_{ij}^{(A)}$ and $s_{ij}\,\tilde{\mathbf{J}}_{ij}^{(A)}$) and therefore the linearization of the preconditioned Roe numerical flux (296) reads (δ^n being defined in (103)):

$$\delta^n \boldsymbol{\phi}^p\left(\mathbf{q}_i, \mathbf{q}_j, \hat{\boldsymbol{\nu}}_{ij}\right) \approx \mathbf{A}_{ij}^n \cdot \delta^n \mathbf{q}_i + \mathbf{B}_{ij}^n \cdot \delta^n \mathbf{q}_j \tag{315}$$

where:

$$\left\{\begin{array}{rcl} \mathbf{A}_{ij}^n & := & \mathbf{A}\left(\mathbf{q}_i^n, \mathbf{q}_j^n, \hat{\boldsymbol{\nu}}_{ij}\right) \\ \\ \mathbf{B}_{ij}^n & := & \mathbf{B}\left(\mathbf{q}_i^n, \mathbf{q}_j^n, \hat{\boldsymbol{\nu}}_{ij}\right) \end{array}\right. \tag{316}$$

$$\left\{\begin{array}{rcl} \mathbf{A}\left(\mathbf{q}_i, \mathbf{q}_j, \hat{\boldsymbol{\nu}}_{ij}\right) & := & \left(\mathbf{P}_{ij}\right)^{-1} \cdot \left(\mathbf{P}_{ij} \cdot \tilde{\mathbf{J}}_{ij}\right)^{+} \\ \\ \mathbf{B}\left(\mathbf{q}_i, \mathbf{q}_j, \hat{\boldsymbol{\nu}}_{ij}\right) & := & \left(\mathbf{P}_{ij}\right)^{-1} \cdot \left(\mathbf{P}_{ij} \cdot \tilde{\mathbf{J}}_{ij}\right)^{-} \end{array}\right. \tag{317}$$

5.2.2 Linearized implicit time-advancing

Starting from the semi-discrete formulation (314), a linearized implicit time-advancing strategy is defined, as described below:

- the time derivative term in (314) is approximated by a backward finite difference, namely:

$$\mu_i \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{q}_i \approx \frac{\mu_i}{\delta^n t}\, \delta^n \mathbf{q}_i = \frac{\mu_i}{\delta^n t} \mathbf{I} \cdot \delta^n \mathbf{q}_i \tag{318}$$

- from the relevant definition (308), the variation of the term $\tilde{\boldsymbol{\varphi}}_i$ in (314) reads:

$$\delta^n \tilde{\boldsymbol{\varphi}}_i = \sum_{j \in s(i)} \delta^n \tilde{\boldsymbol{\varphi}}_{ij} + \delta^n \tilde{\boldsymbol{\varphi}}_{iI} + \delta^n \tilde{\boldsymbol{\varphi}}_{iO} + \delta^n \tilde{\boldsymbol{\varphi}}_{iB} + \delta^n \tilde{\boldsymbol{\varphi}}_{iC} \tag{319}$$

Then, in consideration of the definitions (302)-(304) and by recalling the material discussed in sec. 5.2.1, it is possible to introduce the following approximations:

$$\begin{aligned}
\delta^n \tilde{\boldsymbol{\varphi}}_{ij} &\approx \|\boldsymbol{\nu}_{ij}\| \, \left(\mathbf{A}^n_{ij} \cdot \delta^n \mathbf{q}_i + \mathbf{B}^n_{ij} \cdot \delta^n \mathbf{q}_j\right) && (320) \\
\delta^n \tilde{\boldsymbol{\varphi}}_{iI} &\approx \|\boldsymbol{\nu}_{iI}\| \, \left(\mathbf{A}^n_{iI} \cdot \delta^n \mathbf{q}_i + \mathbf{B}^n_{iI} \cdot \delta^n \mathbf{q}^I_i\right) \\
\delta^n \tilde{\boldsymbol{\varphi}}_{iO} &\approx \|\boldsymbol{\nu}_{iO}\| \, \left(\mathbf{A}^n_{iO} \cdot \delta^n \mathbf{q}_i + \mathbf{B}^n_{iO} \cdot \delta^n \mathbf{q}^O_i\right) && (321)
\end{aligned}$$

where $\mathbf{A}^n_{ij}$ and $\mathbf{B}^n_{ij}$ are given by (316) and the remaining coefficients are defined in the spirit of (316), by suitably replacing $\mathbf{q}_j$ with the fictitious state vectors $\mathbf{q}^I_i$ and $\mathbf{q}^O_i$. Moreover, let $\mathbf{K}^{(\hat{\boldsymbol{\nu}})}(\mathbf{q})$ denote the following Jacobian:

$$\mathbf{K}^{(\hat{\boldsymbol{\nu}})}(\mathbf{q}) := \partial_{\mathbf{q}} \begin{pmatrix} 0 \\ p\,\hat{\boldsymbol{\nu}} \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{0}^T \\ a^2\,\hat{\boldsymbol{\nu}} & \mathbf{O} \end{pmatrix}$$

Then, by defining the following matrices:

$$\mathbf{K}^n_{iB} := \mathbf{K}^{(\hat{\boldsymbol{\nu}}_{iB})}(\mathbf{q}^n_i) \quad , \quad \mathbf{K}^n_{iC} := \mathbf{K}^{(\hat{\boldsymbol{\nu}}_{iC})}(\mathbf{q}^n_i)$$

it is possible to introduce the following linearization for the remaining numerical fluxes in (319):

$$\delta^n \tilde{\boldsymbol{\varphi}}_{iB} \approx \|\boldsymbol{\nu}_{iB}\| \, \mathbf{K}^n_{iB} \cdot \delta^n \mathbf{q}_i \qquad (322)$$

$$\delta^n \tilde{\boldsymbol{\varphi}}_{iC} \approx \|\boldsymbol{\nu}_{iC}\| \, \mathbf{K}^n_{iC} \cdot \delta^n \mathbf{q}_i \qquad (323)$$

By combining (320)-(321), (322) and (323) it is possible to recast (319) as follows:

$$\delta^n \tilde{\boldsymbol{\varphi}}_i \approx \mathbf{F}^n_{ii} \cdot \delta^n \mathbf{q}_i + \sum_{j \in s(i)} \mathbf{F}^n_{ij} \cdot \delta^n \mathbf{q}_j + \mathbf{F}^n_{iI} \cdot \delta^n \mathbf{q}^I_i + \mathbf{F}^n_{iO} \cdot \delta^n \mathbf{q}^O_i \qquad (324)$$

where:

$$\left\{ \begin{aligned}
\mathbf{F}^n_{ii} &:= \sum_{j \in s(i)} \|\boldsymbol{\nu}_{ij}\| \, \mathbf{A}^n_{ij} + \\
&\quad \|\boldsymbol{\nu}_{iI}\| \, \mathbf{A}^n_{iI} + \|\boldsymbol{\nu}_{iO}\| \, \mathbf{A}^n_{iO} + \|\boldsymbol{\nu}_{iB}\| \, \mathbf{K}^n_{iB} + \|\boldsymbol{\nu}_{iC}\| \, \mathbf{K}^n_{iC} \\
\mathbf{F}^n_{ij} &:= \|\boldsymbol{\nu}_{ij}\| \, \mathbf{B}^n_{ij} \\
\mathbf{F}^n_{iI} &:= \|\boldsymbol{\nu}_{iI}\| \, \mathbf{B}^n_{iI} \\
\mathbf{F}^n_{iO} &:= \|\boldsymbol{\nu}_{iO}\| \, \mathbf{B}^n_{iO}
\end{aligned} \right.$$

- let $\mathbf{S}_i$ denote the Jacobian of the term $\mathbf{s}_i$ introduced in (312), namely:

$$\mathbf{S}_i := \partial_{\mathbf{q}_i} \partial \mathbf{s}_i = \|\boldsymbol{\omega}\| \begin{pmatrix} 0 & \mathbf{0}^T \\ \\ \|\boldsymbol{\omega}\|\, \mathbf{r}_i & -2\,\boldsymbol{\Omega} \end{pmatrix} \tag{325}$$

with $\boldsymbol{\Omega}$ defined by the following relation:

$$\hat{\boldsymbol{\omega}} \wedge \mathbf{y} = \boldsymbol{\Omega} \cdot \mathbf{y}$$

where $\mathbf{y}$ is a generic vector in $\mathbb{R}^3$. Once noticed that the matrix (325) does not depend on the specific instance of the state vector $\mathbf{q}_i$ (and therefore on the time-level), it is possible to linearize the right-hand side of (314) as follows:

$$\mu_i\, \mathbf{s}_i^{n+1} = \mu_i\, \mathbf{s}_i^n + \mu_i\, \mathbf{S}_i \cdot \delta^n \mathbf{q}_i \tag{326}$$

By combining (318), (324) and (326), it is straightforward to introduce the following discrete scheme:

$$\mathbf{E}_i^n \cdot \delta^n \mathbf{q}_i + \sum_{j \in s(i)} \mathbf{F}_{ij}^n \cdot \delta^n \mathbf{q}_j + \mathbf{F}_{iI}^n \cdot \delta^n \mathbf{q}_i^I + \mathbf{F}_{iO}^n \cdot \delta^n \mathbf{q}_i^O = \mathbf{b}_i^n \quad , \quad i \in \mathcal{I} \tag{327}$$

where:

$$\begin{cases} \mathbf{E}_i^n & := & \dfrac{\mu_i}{\delta^n t}\, \mathbf{I} + \mathbf{F}_{ii}^n - \mu_i\, \mathbf{S}_i \\ \\ \mathbf{b}_i^n & := & \mu_i\, \mathbf{s}_i^n - \tilde{\boldsymbol{\varphi}}_i^n \end{cases}$$

Note 47 *The equation (327) clearly represents a sparse linear system which can be solved once the boundary terms $\delta^n \mathbf{q}_i^I$ and $\delta^n \mathbf{q}_i^O$ have been suitably associated to specific BCs. For instance, if a uniform inflow is assumed with respect to the non-rotating frame, associated with the state vector $\mathbf{q}_\infty(t)$, then $\mathbf{q}_i^I$ admits the following representation in the body frame:*

$$\mathbf{q}_i^I = \mathbf{q}_\infty(t) - \|\boldsymbol{\omega}\| \begin{pmatrix} 0 \\ \hat{\boldsymbol{\omega}} \wedge \mathbf{g}_i \end{pmatrix} \tag{328}$$

with $\mathbf{g}_i$ given by (311). In consideration of the fact that the corresponding variation:

$$\delta^n \mathbf{q}_i^I = \mathbf{q}_\infty(t^{n+1}) - \mathbf{q}_\infty(t^n)$$

does not involve any unknown, the term $\mathbf{F}_{iI}^n \cdot \delta^n \mathbf{q}_i^I$ in (327) must be formally incorporated into the known term $\mathbf{b}_i^n$. If, in addition, the following transmissive outflow BC is assumed:

$$\mathbf{q}_i^O = \mathbf{q}_i \tag{329}$$

then the corresponding variation, namely:

$$\delta^n \mathbf{q}_i^O = \delta^n \mathbf{q}_i$$

clearly implies that the coefficient $\mathbf{F}^n_{iO}$ *in (327) must be formally incorporated into* $\mathbf{E}^n_i$. *As a result, when adopting the BCs (328) and (329), the system (327) becomes:*

$$\bar{\mathbf{E}}^n_i \cdot \delta^n \mathbf{q}_i + \sum_{j \in s(i)} \mathbf{F}^n_{ij} \cdot \delta^n \mathbf{q}_j = \bar{\mathbf{b}}^n_i \tag{330}$$

with:

$$\begin{cases} \bar{\mathbf{E}}^n_i & := & \dfrac{\mu_i}{\delta^n t}\,\mathbf{I} + \mathbf{F}^n_{ii} + \mathbf{F}^n_{iO} - \mu_i\,\mathbf{S}_i \\ \\ \bar{\mathbf{b}}^n_i & := & \mu_i\,\mathbf{s}^n_i - \tilde{\boldsymbol{\varphi}}^n_i - \mathbf{F}^n_{iI} \cdot \delta^n \mathbf{q}^I_i \end{cases}$$

6 3D Applications

In the present section, the numerical method proposed in sec. 5 is applied to the liquid flow around a hydrofoil (sec. 6.1) as well as to the flow around an axial inducer (sec. 6.2). For both cases, suitable instances of the barotropic state law introduced in sec. 4.1 are adopted.

6.1 Simulation of the 3D flow around a hydrofoil

The water flow around a 3D NACA0015 hydrofoil having chord $c = 115$ mm and mounted within a water tunnel at 4° angle of attack is considered, as a validation benchmark for the linearized implicit scheme proposed in sec. 5. After introducing the problem in sec. 6.1.1, some issues regarding the numerical discretization as well as the used computational resource are presented in secs. 6.1.2 to 6.1.4. Non-cavitating as well as cavitating numerical simulations are respectively presented in secs. 6.1.5 and 6.1.6.

6.1.1 Problem description

The geometry of the test-chamber is sketched in Fig. 54 while the test-section, which is obtained by cutting the chamber along its symmetry plane, is sketched in Fig. 55.

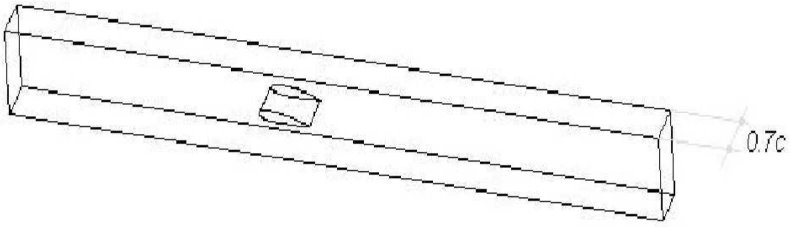

Figure 54: Sketch of the 3D test-chamber.

The considered temperature of the water is $T_L = 293.16$ K. Let the subscript ∞ denote the free-stream (unperturbed) conditions; experimental data are available for the conditions reported in Tab. 15. More precisely, mea-

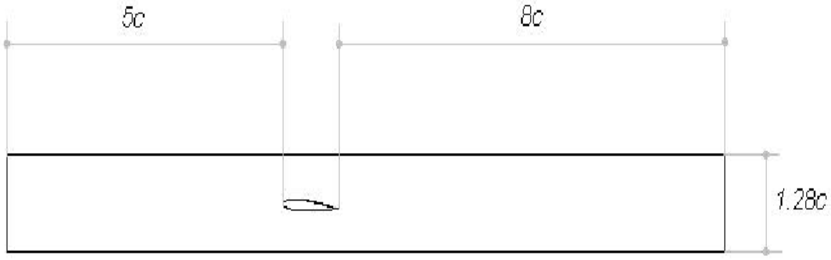

Figure 55: Sketch of the test-section.

Free-stream	p_∞ (Pa)	$\|\mathbf{u}_\infty\|$ (m/s)	M_∞	σ_∞
FS1	59050	3.115	$2.2 \cdot 10^{-3}$	11.7
FS2	12000	3.460	$2.4 \cdot 10^{-3}$	1.5

Table 15: Free-stream conditions of the available experiments.

surements of the pressure coefficient:

$$C_p := \frac{p - p_\infty}{\frac{1}{2}\,\rho_\infty\,\|\mathbf{u}_\infty\|^2}$$

are available [81] along the curve which is defined by intersecting the hydrofoil surface and the test-section. The velocity $\mathbf{u}_\infty$ is orthogonal to the inlet section. The symbols M_∞ and σ_∞ in Tab. 15 respectively denote the free-stream Mach number, defined as follows:

$$M_\infty := \frac{\|\mathbf{u}_\infty\|}{a_\infty} \tag{331}$$

and the cavitation number, defined as follows [9]:

$$\sigma_\infty := \frac{p_\infty - p_{sat}}{\frac{1}{2}\,\rho_\infty\,\|\mathbf{u}_\infty\|^2} \tag{332}$$

In view of the definition (332), it is clear that cavitation phenomena are likely to take place in correspondence of low cavitation numbers. The conditions

Grid	N_c	N_t
GR1	27220	137756
GR2	19322	88400

Table 16: Considered computational grids.

FS1 in Tab. 15, in particular, are associated with a non-cavitating flow, which can be considered as a (very) low Mach number validation benchmark for numerical solvers. Conversely, the conditions FS2 are associated with a cavitating flow. At the considered liquid temperature, the transition between non-cavitating and cavitating flow regions is extremely abrupt [10]; this behaviour is described by e.g. the complex state laws shown in Figs. 44 and 45, whose numerical treatment is particularly tough.

6.1.2 Computational grids

The domain sketched in Fig. 54 is discretized by means of a 3D tetrahedral unstructured grid. The considered grids are reported in Tab. 16, in which the symbols N_c and N_t (defined in sec. 5.1.1) respectively denote the number of cells (i.e. nodes) and elements (i.e. tetrahedra). It is worth mentioning that:

- both the grids GR1 and GR2 are 3D tetrahedral, unstructured grids. However, by construction they are symmetrical with respect to the test-section and therefore their imprint on the test-section appears as a 2D triangular, unstructured grid (see Fig. 56);
- neither GR1 nor GR2 is highly refined in order to contain the computational cost of the simulations while validating/developing the considered numerical schemes (examples of finer grids discretizing the domain under consideration can be found in [6]);
- while GR1 represents the whole test-chamber, GR2 only discretizes a "slice" of it (its span-wise width being 0.1 c instead of 0.7 c) and it is used for reducing the computational cost while validating/developing the considered numerical schemes.

The considered grids must be partitioned in order to be incorporated into the parallel numerical frame mentioned in the introduction to the present document (i.e. the AERO code). The grids GR1 and GR2, in particular,

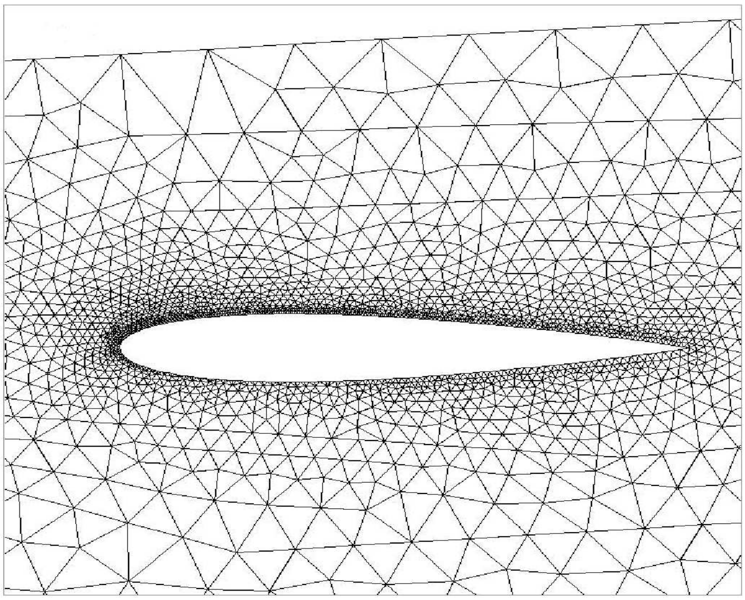

Figure 56: Imprint of the grid GR1 on the test-section (detail).

Computer	CPU	No. of CPUs	Total RAM
COMP1	Intel Pentium4, 2.66 GHz	1	512 MB
COMP2	Intel Pentium4 Xeon, 3.06 GHz	2	8 GB
COMP3	IBM POWER4, 1.3 GHz	512	1088 GB

Table 17: Available computers.

have been divided into 5 and 2 sub-domains, respectively. To the purpose, the proprietary software "TopDomDec" as well as the open source tool "Metis" (http://www-users.cs.umn.edu/ karypis/metis/metis) have been exploited.

6.1.3 Computational resources

The considered computational resources are reported in Tab. 17. Among them, COMP3 denotes the IBM SP4 computing platform available at CINECA (currently upgraded to SP5, see http://www.cineca.it) while COMP1 and COMP2 are common PCs.

6.1.4 Numerical discretization

The linearized implicit discrete scheme which is derived from (330) by setting $\boldsymbol{\omega} = \mathbf{0}$ is considered for both the non-cavitating and the cavitating simulations.

A variable time-step is adopted, defined as follows:

$$\delta^n t = c^{(CFL)\,n} \min_{h \in \mathcal{H}} \left(\frac{\lambda_h}{\tilde{s}_h^n} \right) \tag{333}$$

where:

- λ_h denotes the minimum among the four heights which are associated with the $h-$th tetrahedron T_h ($h \in \mathcal{H} := \{1, \ldots, N_t\}$);
- $\tilde{s}_h^n$ denotes the value at time-level n of an estimate of the maximum wave speed $\tilde{s}_h$ associated with T_h. More precisely, $\tilde{s}_h$ is chosen as the maximum among the wave speeds arising in the Roe-linearized RPs associated with the four vertices of T_h;
- $c^{(CFL)\,n}$ denotes the value at time-level n of a CFL-like coefficient, $c^{(CFL)}$, which can be modulated during the simulation (see below).

Since, as shown by the experiments, the flows associated with the considered free-streams in Tab. 15 turn out to be substantially steady (even the cavitating one, due to the low angle of attack), the numerical simulations are advanced in time up to a steady-state.

6.1.5 Non-cavitating simulations

The considered non-cavitating test-cases are summarized in Tab. 18. The following state vector:

$$\mathbf{q}_\infty = \begin{pmatrix} \rho_\infty \\ \rho_\infty \, \mathbf{u}_\infty \end{pmatrix} \tag{334}$$

is derived, in particular, from the chosen free-stream FS1 (see Tab. 15 above). The state vector (334), in turn, is introduced in (328) for defining the fictitious inflow state vectors $\mathbf{q}_i^I$; moreover, it is exploited for defining the adopted initial conditions. The free-stream Mach number M_∞ (see Tab. 15 above) is assumed to be the characteristic Mach number $M_\star$ to be used for preconditioning the Roe numerical flux; in particular, the constant β_{ref} in (189) is chosen equal to 1 [60]. Furthermore, the parameters "Liquid" and "T_L" in

Test-case	Free-stream	Liquid	T_L (K)	Grid
NONCAV1	FS1	water	293.16	GR1
NONCAV2	FS1	water	293.16	GR2

Table 18: Considered non-cavitating test-cases.

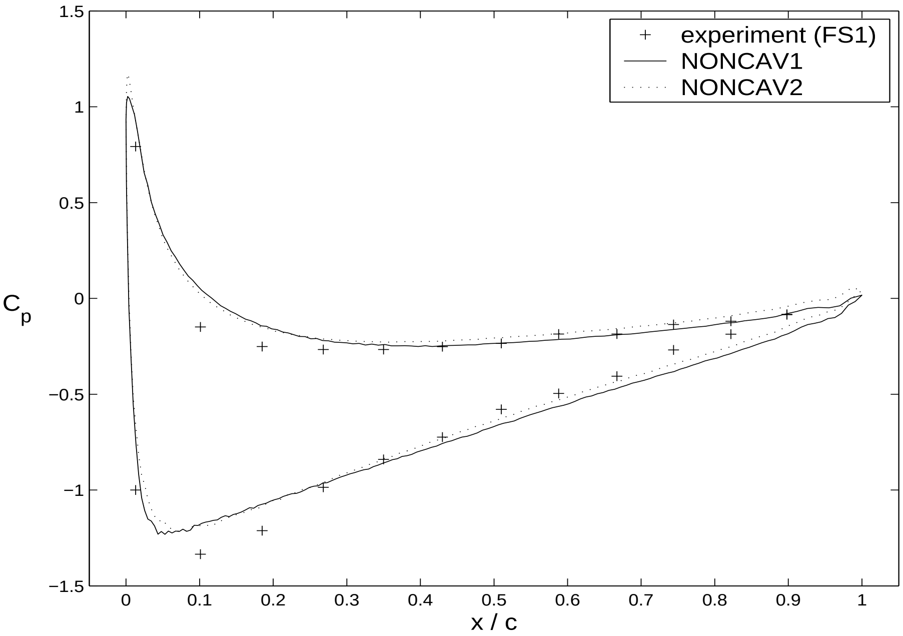

Figure 57: C_p distribution for the test-cases in Tab. 18.

Tab. 18 characterize the isentropic compressibility coefficient $\vartheta \approx 8.55 \cdot 10^5$ appearing in the chosen liquid model (247).

The resulting C_p distribution is reported in Fig. 57, against the relevant experimental data. It is worth noticing that:

- the numerical results respectively obtained by exploiting GR1 and GR2 are close to each other, thus showing that the obtained C_p distribution

[60] A sensitivity study has been performed in order to set β_{ref}, not reported here for the sake of conciseness. It has been observed that smaller time-steps must be adopted when decreasing β_{ref} from its upper bound, $\beta_{ref} = M_\star^{-1}$ (corresponding to the non-preconditioned case $\beta = 1$ in (189)) down to its "recommended" value, i.e. O(1). However, the resulting numerical solution turns out to be considerably inaccurate (in terms of C_p, against the experiments) when β_{ref} is distant from 1; vice versa, for β_{ref} in the neighbourhood of the unity, the most accurate result seems to be associated with $\beta_{ref} \approx 1$.

is almost independent of the grid. Moreover, no appreciable 3D effects take place along the span-wise direction (which is not surprising, by virtue of the assumed absence of viscosity effects);

- the agreement between the numerical results and the experimental data can be considered reasonably good, in view of the fact that the considered 3D numerical scheme is only first-order accurate and the used grids are relatively coarse. Both these issues seem to contribute, for instance, to underestimating the suction peak which is located near the leading edge of the hydrofoil [61];

- the coefficient $c^{(CFL)}$ in (333) has been increased, linearly with respect to n, during the first iterations for smoothly abandoning the initial flow field (which is, in general, a crude approximation of the final one). In particular, it has been increased up to more than 400 for both the considered test-cases, thus confirming the efficiency of the proposed linearized implicit schemes when dealing with smooth flows (see sec. 3.5.5). As far as the total CPU time is concerned, the test-case NONCAV1 requires 17 hours and 30 minutes on the computer COMP3 reported in Tab. 17 (a contained elapsed time is obtained, due to the parallelization strategy) while the test-case NONCAV2 requires 7 hours and 30 minutes on the computer COMP1 [62] reported in Tab. 17.

6.1.6 Cavitating simulations

The considered cavitating test-cases are reported in Tab. 19. A state vector of the type of (334) is introduced also for the present case, based on the free-stream FS3 defined in Tab. 20. More in detail, the considered state vector is obtained from that one associated with the free-stream FS1 in Tab. 15 (sec. 6.1.1), by decreasing the pressure p_∞; such a procedure has been actually performed for defining the considered inlet and initial conditions [63]. The free-stream Mach number M_∞ in Tab. 20 is assumed to be the characteristic Mach number of the liquid region. However, in consideration of the fact that

[61] A better result could be obtained by suitably refining the grid in the leading edge area and by increasing the order of spatial accuracy of the scheme (e.g. by a standard MUSCL technique [108]); these improvements are postponed to a subsequent research stage.

[62] The considered parallel code has been run on the mono-processor computer COMP1 by means of the "LAM" parallel environment (see http://www.lam-mpi.org), thus introducing a certain degree of communication overhead.

[63] A variable free-stream state vector $\mathbf{q}_\infty(t)$ is explicitly considered in (328). Moreover, a user-defined flow field (typically, the result of a previous simulation) can be read by the developed numerical solver for starting the simulation.

Test-case	Free-stream	Liquid	T_L (K)	ζ	Grid
CAV1	FS3	water	293.16	0.1	GR1
CAV2	FS3	water	293.16	0.1	GR2
CAV3	FS3	water	293.16	0.01	GR2

Table 19: Considered cavitating test-cases.

Free-stream	p_∞ (Pa)	$\|\mathbf{u}_\infty\|$ (m/s)	M_∞	σ_∞
FS3	7500	3.115	$2.2 \cdot 10^{-3}$	1.1

Table 20: Considered free-stream conditions.

no preconditioning is required within the cavitating region (where the flow can be easily hypersonic), a local preconditioning strategy is heuristically adopted. More precisely, a local preconditioning parameter β_{ij}^2, defined as follows (compare with (234)):

$$\beta_{ij}^2 := \begin{cases} M_\infty^2 & \text{if} \quad \min(\rho_i, \rho_j) \geq \rho_{Lsat} \\ 1 & \text{otherwise} \end{cases}$$

is introduced into the preconditioning matrix (293) in place of the original parameter β^2. As far as the state law is concerned, the parameters "Liquid", "T_L" and "ζ" in Tab. 19 characterize two instances of the barotropic model (247)-(255). The relevant model parameters are $\vartheta \approx 8.55 \cdot 10^5$, $\sigma_1 \approx 1.33 \cdot 10^3$, $\sigma_2 \approx -0.73$ and $\sigma_3 \approx 0.78$; the mixture branches of the considered laws are shown in Fig. 44.

It is possible to adopt the experimental data based on the free-stream FS2 in Tab. 15 (sec. 6.1.1) for validating the cavitating simulations at hand, since the corresponding cavitation number is similar to that one associated with the considered free-stream FS3. Hence, in Figs. 58 and 59, the C_p distribution -on the suction side of the considered hydrofoil- which is obtained from the considered simulations is compared with the aforementioned experimental points. It is worth noticing that:

- only small differences, located near the leading edge of the hydrofoil (i.e. where cavitation occurs), appear in the C_p distribution when adopting different grids (see Fig. 58);

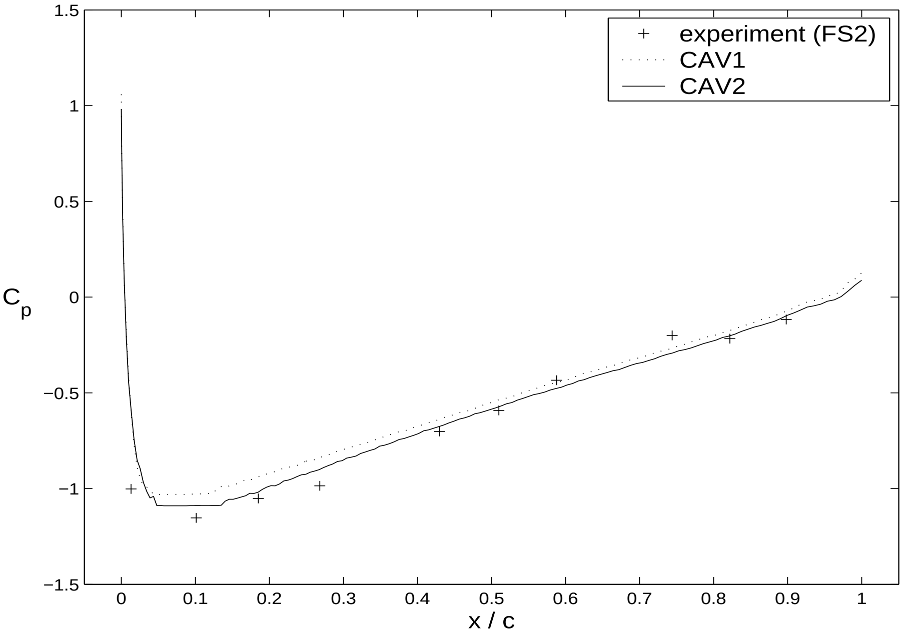

Figure 58: C_p distribution (suction side) for the test-cases CAV1 and CAV2 reported in Tab. 19.

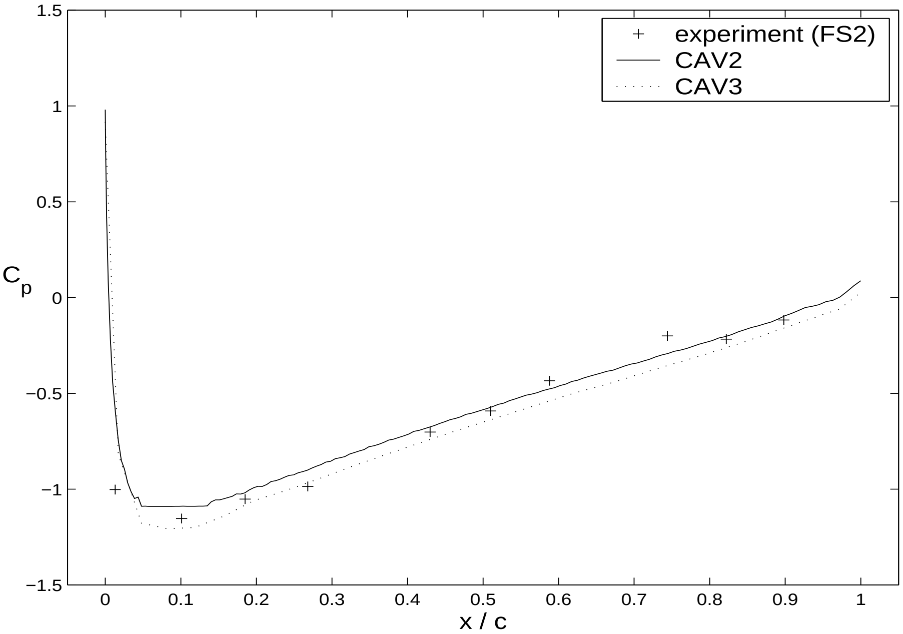

Figure 59: C_p distribution (suction side) for the test-cases CAV2 and CAV3 reported in Tab. 19.

- the agreement between the numerical results and the experimental data can be considered reasonably good, in view of the fact that it is very challenging to accurately simulate the cavitation phenomena at hand;

- on the basis of the numerical results in Fig. 59, it seems that the C_p distribution gradually varies with respect to ζ. Moreover, a lower value of ζ correctly leads to a less pronounced Mach number variation (indeed, as shown in Fig. 45, the minimum sound speed -in the cavitating region- increases when reducing ζ), as shown in Figs. 60 and 61. Furthermore, once defined a local cavitation number as follows (compare with (332)):
$$\sigma := \frac{p - p_{sat}}{\frac{1}{2}\,\rho_\infty \left\|\mathbf{u}_\infty\right\|^2}$$
it is possible to identify the cavity with the fluid sub-domain within which $\sigma < 0$. Then, as shown in Figs. 62 and 63, it is possible to see that a lower value of ζ results in a more extended cavity. Also this result seems to be correct, since the nearly constant pressure value in Fig. 44, which roughly provides a characteristic value of the cavity pressure, decreases when decreasing ζ. Nevertheless, a systematic investigation of the sensitivity of the numerical results to the free cavitation model parameter ζ is postponed to a further research stage;

- before the inception of cavitation, the coefficient $c^{(CFL)}$ introduced in (333) can be increased during the simulation up to $\mathrm{O}\left(10^2\right)$ for all the considered cavitating test-cases However, as soon as cavitation occurs, it must be reduced to $\mathrm{O}\left(10^{-2}\right)$ for all the considered simulations to remain stable. This point seems to confirm the hypothesis put forward in sec. 3.5.6 according to which the observed stability restriction can be caused by the presence/onset of discontinuities in the flow field (caused by the cavitation inception in the present case), which render it more difficult to exploit the proposed linearized scheme. As far as the total CPU time is concerned, the test-case CAV1 approximately requires 400 hours on the computer COMP3 (see Tab. 17) while the test-cases CAV2 and CAV3 approximately require 150 hours on the computer COMP2. In both cases, a contained elapsed time is obtained, thanks to the parallelization strategy.

Further investigation is definitely recommended in order to counteract the aforementioned efficiency problem. Moreover, according to the author, it would of interest to also assess the effects that the chosen local preconditioning strategy produces on the stability properties of the resulting numerical scheme. However, such a study is postponed to a subsequent research stage.

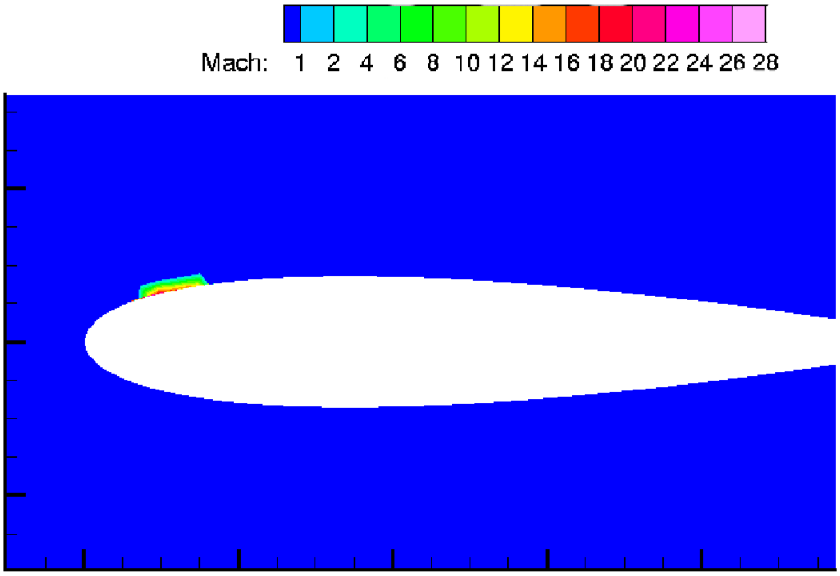

Figure 60: Contour plot of the local Mach number on the test-section (detail), for the test-case CAV2 reported in Tab. 19.

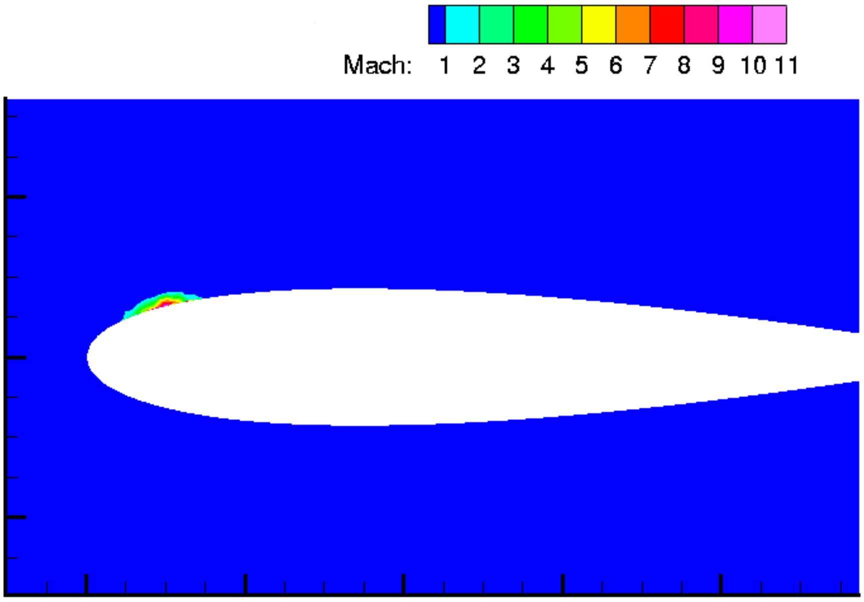

Figure 61: Contour plot of the local Mach number on the test-section (detail) for the test-case CAV3 reported in Tab. 19.

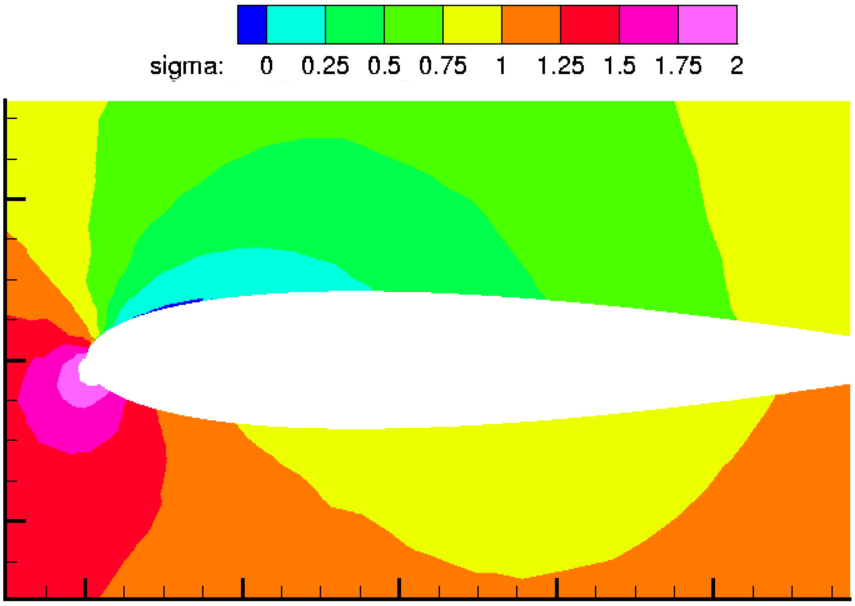

Figure 62: Contour plot of the local cavitation number (sigma) on the test-section (detail), for the test-case CAV2 reported in Tab. 19.

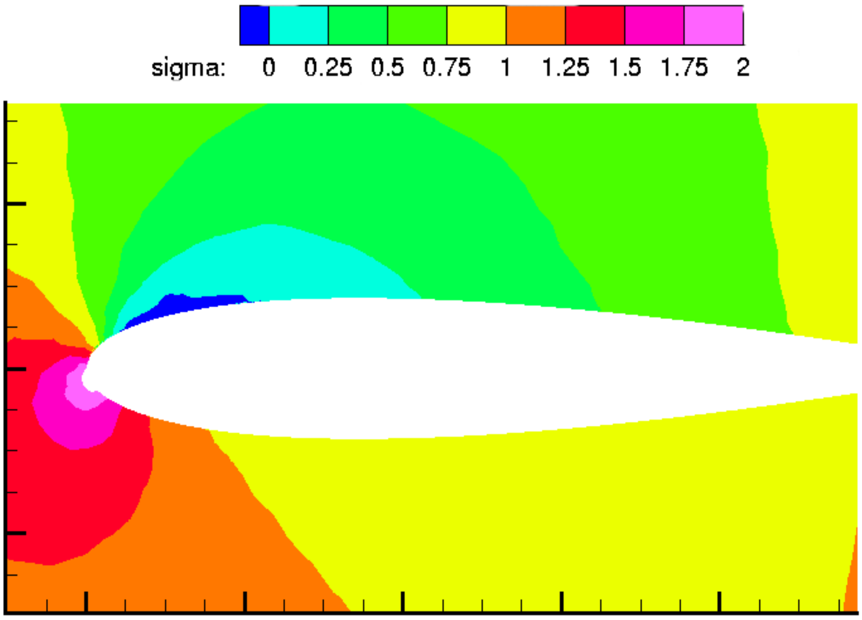

Figure 63: Contour plot of the local cavitation number (sigma) on the test-section (detail), for the test-case CAV3 reported in Tab. 19.

6.2 Simulation of the 3D flow around an axial inducer

The water flow around a turbo-pump inducer (see sec. 1.1) is considered in the present section, as a qualitative validation benchmark for the linearized implicit scheme proposed in sec. 5.

After introducing the problem in sec. 6.2.1, some issues regarding the numerical discretization as well as the used computational resource are presented in secs. 6.2.2 to 6.2.4. Some non-cavitating numerical results are finally presented in sec. 6.2.5. In consideration of the efficiency problems already discussed in secs. 3.5.6, 4.2 and 6.1.6, no cavitating simulations are considered for the inducer flow at hand. Indeed, the huge increase in computational cost, which is here amplified by the complexity of the considered geometry (see below), makes it practically impossible to advance the simulation unless exploiting specific supercomputing resources, that are not available within the scope of the present research project [64].

6.2.1 Problem description

The considered geometry is sketched in Fig. 64, where the inducer is denoted by "I". A nose "N" as well as an after-body "A" smoothly join "I"; in particular, the nose is part of an axisymmetrical ellipsoid while the after-body is a circular cylinder having a diameter equal to the base diameter of the inducer. The flow domain is bounded by a cylindrical case, whose diameter is equal to the maximum blade tip diameter D; hence, there is no tip clearance and a shrouded inducer (see sec. 1.1) is considered. The length L_{out} of the after-body, as well as the length L_{in} of the inflow section, are chosen equal to $1.5\,D$. The inducer angular speed is equal to 2000 rpm.

The chosen temperature of the water is $T_L = 296.16$ K. The considered free-stream conditions are reported in Tab. 21. Both M_∞ and σ_∞ in the aforementioned table, respectively defined in (331) and (332), are computed by exploiting the absolute velocity $\mathbf{u}_\infty$; the local Mach number and the local cavitation number are respectively higher and lower than those reported in Tab. 21, due to the dragging velocity appearing in the body frame.

6.2.2 Computational grids

The domain sketched in Fig. 64 is discretized by means of a 3D tetrahedral unstructured grid, whose main features are reported in Tab. 22 (N_c and N_t respectively denoting the number of nodes and elements). It is worth

[64] As mentioned in [93], a cavitating simulation has been stopped at the inception stage, due to the aforementioned efficiency problems.

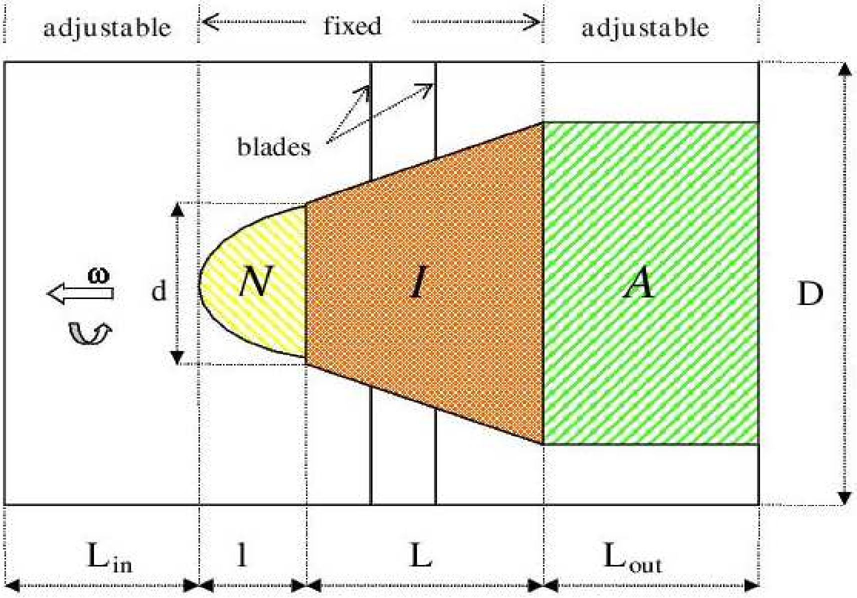

Figure 64: Schematic representation of the considered inducer geometry.

Free-stream	p_∞ (Pa)	$\|\mathbf{u}_\infty\|$ (m/s)	M_∞	σ_∞
FS4	115000	0.476	$3.4 \cdot 10^{-4}$	990

Table 21: Considered free-stream conditions.

Grid	N_c	N_t
GR3	549139	2588501

Table 22: Considered computational grid.

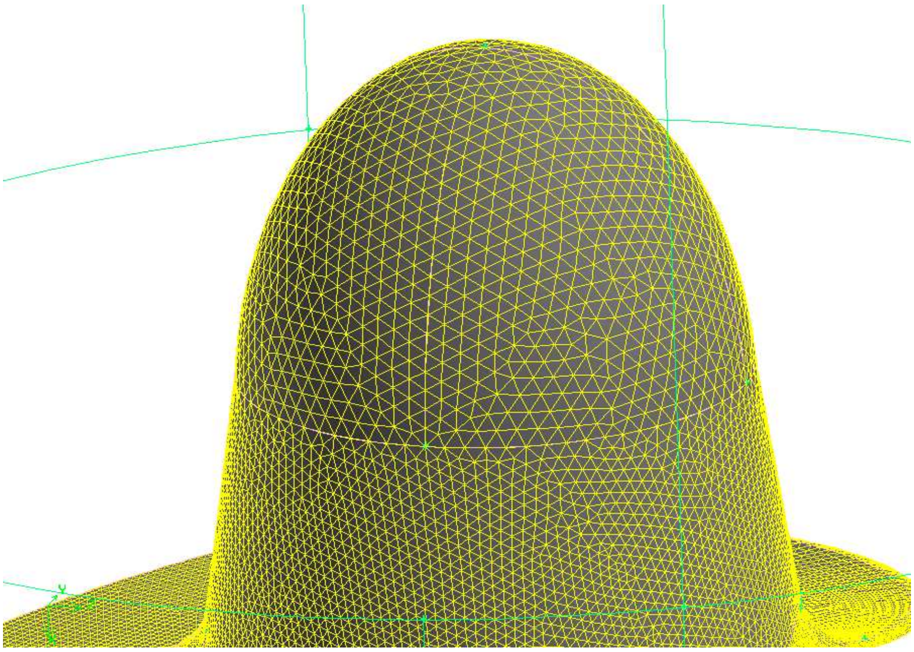

Figure 65: Detail of the grid GR3 at the nose-inducer junction.

mentioning that:

- the size of the grid elements smoothly transitions between different regions on the body surface (e.g. the nose-inducer junction shown in Fig. 65) and accurately follows the solid walls even within high-curvature regions (e.g. the hub-blade intersection shown in Fig. 66);
- as far as the external case is concerned, it is not possible to define a perfectly cylindrical wall due to the numerical errors (even if very small) related to the numerical format of the inducer geometry file. To counteract this problem, a kind of shell covering the inter-blade passages is modelled, whose external aspect is shown in Fig. 67.

The considered grid has been partitioned into 16 sub-domains in order to be incorporated into the parallel numerical frame mentioned in the introduction to the present document (i.e. the AERO code). To the purpose, the proprietary software "TopDomDec" has been exploited.

6.2.3 Computational resources

In consideration of the noticeable size of the grid at hand, the only supercomputer COMP3 reported in Tab. 17 (sec. 6.1.3) is considered.

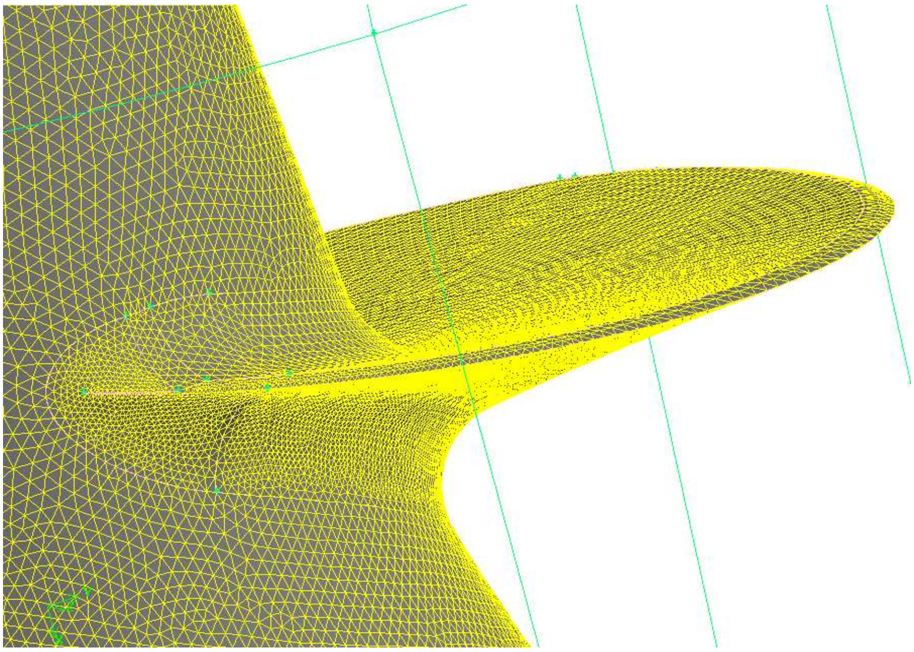

Figure 66: Detail of the hub-blade intersection for GR3.

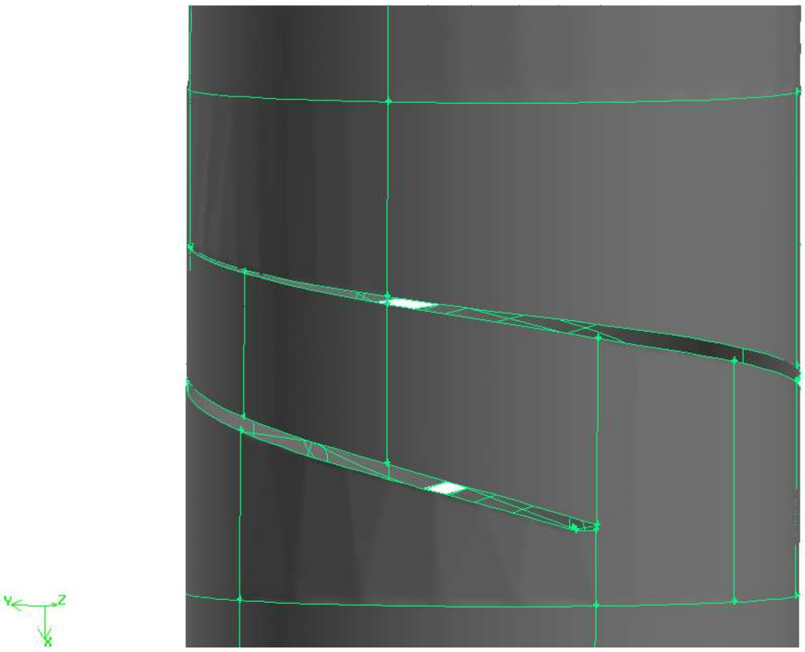

Figure 67: External view of the inter-blade covering created for GR3. The "cut" on the boundary surface represent the imprint of the inducer blade tip.

6.2.4 Numerical discretization

The linearized implicit discrete scheme (330) is considered, in which the terms related to the rotation are computed by exploiting the inducer (constant) angular velocity.

As far as the time-advancing is concerned, the variable time-step (333) is adopted.

6.2.5 Non-cavitating simulations

The considered non-cavitating test-case is reported in Tab. 23. A state vector of the type of (334), derived from the considered free-stream FS4 (see Tab. 21 in sec. 6.2.1), is introduced for the defining the fictitious inflow state vector $\mathbf{q}_i^I$ in (328). Moreover, a uniform initial flow field, determined by the free-stream conditions, is assumed; its representation in the body frame is therefore obtained (for the $i-$th finite volume cell) by evaluating the right-hand side of (328) in correspondence of the aforementioned free-stream state vector. In consideration of the fact that, with respect to the rotating frame, the local Mach number can undergo substantial variations along the radial direction due to the dragging velocity, a local preconditioning strategy is required (see the relevant paragraph in sec. 3.5.4). In particular, a local preconditioning parameter β_{ij}^2 (to be introduced into the preconditioner (293) in place of the original parameter β^2) is heuristically defined as follows (compare with (234)):

$$\beta_{ij}^2 := 1 - \exp\left(-\left(\hat{M}_{ij}\right)^2\right)$$

where:

$$\hat{M}_{ij} := \frac{\|\mathbf{u}_{ij}\|}{a_{ij}}$$

with $\mathbf{u}_{ij}$ and a_{ij} respectively defined in (274) and (154). Finally, as far as the state law is concerned, the parameters "Liquid" and "T_L" in Tab. 23 characterize the isentropic compressibility coefficient $\vartheta \approx 7.13 \cdot 10^5$ appearing in the chosen liquid model (247).

The pressure contours on the inducer surface, obtained after 27000 iterations, are reported in Fig. 68. It is worth noticing that:

- the behaviour of the flow field, as described by the numerical simulation, is in a good qualitative agreement with that one observed in a number of experimental works. Indeed, the working fluid gradually undergoes a pressure rise while flowing within the vanes between the

Test-case	Free-stream	Liquid	T_L (K)	Grid
NONCAV3	FS4	water	296.16	GR3

Table 23: Considered non-cavitating test-case.

rotating blades, as shown in Fig. 68. Moreover, according to this figure, the flow region which is most prone to cavitation is located near that portion of the blades where the volutes, detaching from the hub, firstly reach the external tip diameter D. This is in agreement with the experiments which, for similar flow conditions, observe the cavitation inception exactly in the flow region under consideration (see e.g. [14]);

- the considerable axial back-flow occurring near the blade tip where the diameter is less than D (i.e. where the volutes are not completely shrouded), which is well documented in a number of experimental works (e.g. [118]), is described by the numerical simulation as well, as shown in Fig. 69;

- the coefficient $c^{(CFL)}$ introduced in (333) has been increased during the simulation for smoothly abandoning the initial flow field (which is, in general, a crude approximation of the final one). In particular, it has been increased up to 350. As far as the total CPU time is concerned, the considered simulation approximately requires 1500 hours on the computer mentioned in sec. 6.2.3; the corresponding elapsed time can be contained by virtue of the parallelization strategy.

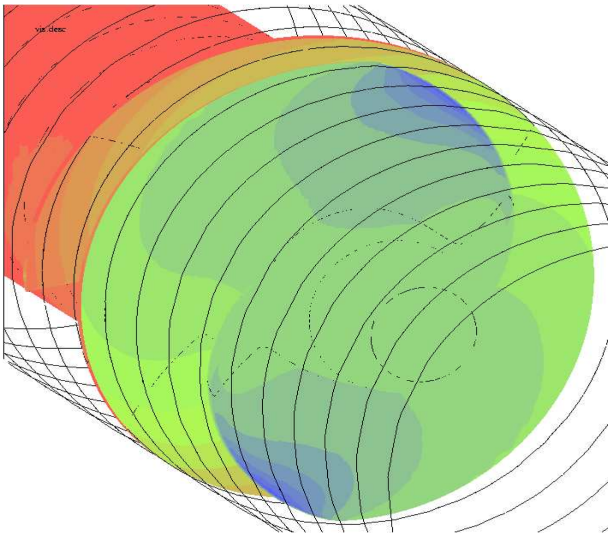

Figure 68: Pressure contours on the inducer surface for the test-case NONCAV3: max [red] 177700 (Pa), min [blue] 79700 (Pa), spacing 5000 (Pa).

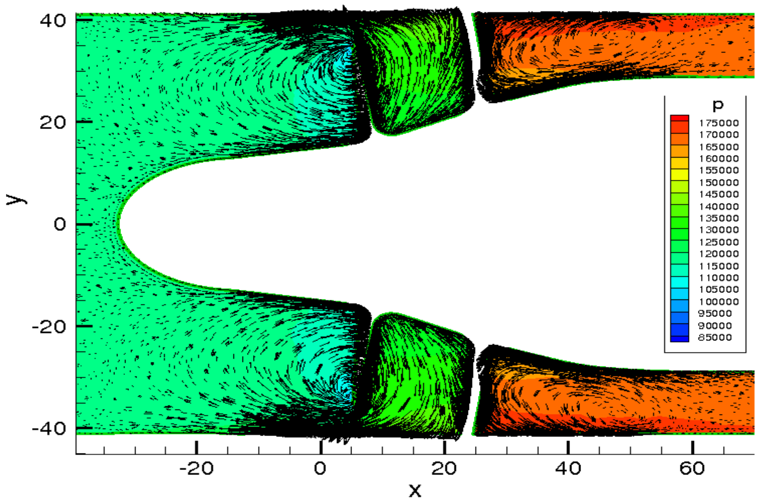

Figure 69: Velocity field (x: axial component, y: radial component) in a longitudinal cut plane of the flow domain for the test-case NONCAV3. Pressure contours are drawn in the background.

7 Concluding remarks

A numerical method for simulating 3D barotropic flows in complex, possibly rotating, geometries has been presented. The considered method can successfully cope with nearly-incompressible flows by *ad hoc* preconditioning and allows for an efficient linearized implicit time-advancing technique to be defined. All the proposed numerical ingredients were implemented within a parallel numerical framework; the resulting CFD solver was validated against 3D non-cavitating as well as cavitating liquid flows. The documented research activities were driven by an industrial program, funded by the Italian Space Agency (ASI), aimed at developing a numerical tool for simulating propellant flows around 3D rotating axial inducers belonging to the feed turbo-pump system of a liquid propellant rocket engine.

In view of the fact that, under typical operational conditions, cavitation phenomena can take place within the aforementioned turbo-machines, the choice of a suitable cavitation model was initially addressed. A literature review suggested considering an equivalent fluid cavitation model; a barotropic homogeneous flow model was adopted, in particular, which can take into account thermal cavitation effects and, possibly, the concentration of the active cavitation nuclei. This model was coupled with the mass and momentum balances of classical fluid dynamics; the effects of viscosity were neglected. In order to incorporate the chosen model into an existing numerical frame which was available to the research group, namely the AERO code described in the introduction, a density-based numerical approach was chosen. The AERO code was originally conceived for dealing with ideal gases and the specific expression and properties of the ideal gas state law deeply affected its implementation. In particular, both the definition of the Roe numerical flux function (characterizing the space discretization of the convective fluxes by a finite volume approach) and the linearized implicit time-advancing strategy (involving an approximate linearization of the aforementioned numerical flux) were based on the ideal gas state law. Moreover, also the preconditioning technique introduced for coping with low Mach number flows was affected by the specific form of the adopted state law. As a consequence, all these numerical issues needed to be replaced, if possible, with proper counterparts holding for a barotropic state law.

The definition of the new numerical ingredients was initially conceived in a 1D context; moreover, in order to keep a certain degree of generality, a generic barotropic state law was assumed. Once defined a Roe numerical flux applicable to generic barotropic fluids, the accuracy of the resulting semi-discrete formulation, as applied to nearly-incompressible flows, was addressed following [42]. This study showed that for low Mach number flows the accuracy of the proposed semi-discrete formulation degrades; the same result had already been found -and a suitable remedy (preconditioning) had been proposed- in [42] for the ideal gas case. The considered preconditioning strategy was successfully extended to the barotropic case; however, the introduction of the preconditioning narrowed the stability region of common explicit time-advancing schemes. To counteract this problem, a linearized implicit time-advancing strategy was proposed, only relying on the algebraic properties of the Roe flux function and therefore applicable to a variety of problems. In particular, differently from the linearization technique already implemented in AERO, the proposed one does not rely on the first-order homogeneity of the analytical flux function (since this properties, satisfied by the ideal gas state law, does not hold for the barotropic one). The implicit scheme was further extended so as to incorporate the aforementioned preconditioning strategy. All these ingredients were qualitatively validated in a 1D context, namely the water flow in a convergent-divergent nozzle, for both non-cavitating and cavitating conditions [91]. The proposed preconditioning technique turned out to effectively counteract the accuracy problem at low Mach numbers. Furthermore, the proposed linearized implicit scheme allowed for an efficient time-advancing to be performed when considering non-cavitating flows; as soon as cavitation occurred, however, significantly smaller time-steps had to be adopted. The proposed 1D numerical techniques were then extended to the 3D case, firstly to non-rotating and then to rotating frames. The generalization of the Roe numerical flux, in particular, was accomplished by exploiting the tensorial character of the considered governing equations while the extension to rotating frames was performed by introducing a suitable term in the aforementioned equations, accounting for non-inertial effects. The proposed 3D numerical method was firstly validated by considering the water flow around a NACA0015 hydrofoil, for which experimental data concerning the pressure coefficient distribution were available. In particular, water at 20° C was considered, possibly leading to the occurrence of “cold cavitation” phenomena whose numerical treatment is extremely challenging. All the issues highlighted in the 1D numerical experiments appeared in the 3D case as well; in particular, the proposed scheme proved out to efficiently compute non-cavitating flows but, as soon as cavitation takes place, its efficiency was significantly reduced. As far as the accuracy is concerned, the obtained results (which appeared to be independent of the grid) seemed reasonably good for both non-cavitating and cavitating conditions, in view of the fact that the considered 3D numerical scheme was only first-order accurate and the used grids were relatively coarse. A few thousand iterations of a non-cavitating simulation of the water flow around an axial turbo-pump inducer were carried out as well. The behaviour of the flow field, as described by the considered numerical simulation, turned out to be in a good qualitative agreement with that one observed in a number of experimental works. In particular, the numerical solution correctly described the pressure contours on the surface of the inducer blades as well as the considerable axial back-flow occurring where the inducer volutes are not completely shrouded. Moreover, also for the non-cavitating case under consideration, an efficient time-advancing could be performed.

A more systematic investigation of the aforementioned 1D numerical ingredients was then started. In this context, the exact solution of the 1D Riemann problem associated with a generic convex barotropic state law was addressed and a solution procedure was proposed (which was also exploited for defining exact benchmarks for the validation of the 1D numerical schemes considered in the present document). A Godunov numerical flux function based on the aforementioned exact solution was defined as well.

Clearly, the efficiency problem emerging when considering non-smooth flow fields, like those originating from cavitation inception when adopting realistic homogeneous flow models, deserves special attention. In view of the numerical results reported in the present document, this efficiency issue seems to be imputable to the approximate linearization of the Roe numerical flux in the implicit time-advancing (as briefly mentioned, the specifically adopted linearization does not seem to play a crucial role in this problem). Consequently, it could be of interest to also consider different (i.e. more robust, even if less refined) numerical flux functions as, for instance, the Rusanov flux, the HLL/HLLC flux, etc... [98] (in this spirit, the proposed solution to the 1D Riemann problem associated with convex barotropic state laws could be exploited for investigating further Godunov methods). The aforementioned point could be supported by the fact that the considered Roe flux function (as it stands, without fixes) may provide entropy-violating solutions within the transonic regime associated with cavitation inception. Furthermore, the fact that phase transition (and therefore cavitation) is a major reason in the lack of convexity of the considered state law [69] may add to the complexity of the problem, since the convexity may be important when seeking entropic solutions [3]. It is therefore evident that there is room for improvement while keeping the adopted numerical frame (i.e. homogeneous flow cavitation model, compressible -generally preconditioned- algorithms, finite volume space discretization, linearized implicit time-advancing); further investigation in this direction is definitely recommended. Simultaneously, it would be of interest to increase the order of accuracy of the proposed method (e.g. by developing the "Defect Correction" strategy briefly discussed in sec. 3.5.3) as well as to investigate additional/different numerical ingredients as, for instance, relaxation techniques (see e.g. [3], [22] and [23]) and dual time-stepping strategies (see e.g. [22], [23] and [58]), which seem to improve the convergence properties of the considered algorithms.

As a concluding remark, it may be worth emphasizing that the assumed generality of the considered barotropic state law permits to apply the proposed material to several problems (e.g. to shallow water flows, see Note 12 in sec. 2.5.1). This aspect, together with the fact that the proposed linearization of the Roe numerical flux function may be applied when considering an arbitrary state law (not necessarily a barotropic one), endow the present work with a certain degree of generality.

A Appendix: auxiliary material for sec. 3

A.1 Derivation of the expression (170)

By applying the standard non-dimensionalization procedure mentioned in the relevant paragraph of sec. 3.4.1 to the continuous system (23), the following expression is obtained:

$$\left\{ \begin{array}{lclcl} \partial_t\,(\rho) & = & & & \Psi_c^{(0)} \\ \\ \partial_t\,(\rho u) & = & M_\star^{-2}\,\Theta_c^{(-2)} & + & \Theta_c^{(0)} \end{array} \right. \tag{335}$$

with:

$$\left\{ \begin{array}{lcl} \Psi_c^{(0)} & := & -\,\partial_x\,(\rho u) \\ \\ \Theta_c^{(-2)} & := & -\,\partial_x p \\ \\ \Theta_c^{(0)} & := & -\,\partial_x\,(\rho u^2) \end{array} \right. \tag{336}$$

and:

$$M_\star := \frac{u_{ref}}{a_{ref}} \tag{337}$$

The expressions (335) and (337) are copied in sec. 3.4.1, respectively to (170) and (171), for ease of presentation.

A.2 Derivation of the expression (172)

By recalling the relevant definitions, the equation (164) can be recast as follows:

$$\begin{array}{rcl} 2\,\mu_i\,\dfrac{\mathrm{d}}{\mathrm{d}t}\,\mathbf{q}_i^{(x)} & = & \mathbf{f}_{i-1}^{(x)} - \mathbf{f}_{i+1}^{(x)} + \\ \\ & & \left|\,\tilde{\mathbf{J}}_{i(i+1)}^{(x)}\,\right| \cdot \Delta^{i(i+1)}\mathbf{q}^{(x)} - \\ \\ & & \left|\,\tilde{\mathbf{J}}_{(i-1)i}^{(x)}\,\right| \cdot \Delta^{(i-1)i}\mathbf{q}^{(x)} \end{array} \tag{338}$$

As a preliminary step, a suitable representation is sought for the generic term $\left|\,\tilde{\mathbf{J}}_{ij}^{(x)}\,\right| \cdot \Delta^{ij}\mathbf{q}^{(x)}$ appearing, in particular, in (338). To the purpose, once introduced the eigenvalue-eigenvector pairs of $\tilde{\mathbf{J}}_{ij}^{(x)}$, namely:

$$\left\{ \begin{array}{l} \lambda_{ij}^{(1)} = u_{ij} + a_{ij} \quad , \quad \mathbf{r}_{ij}^{(1)} = \left(1, \lambda_{ij}^{(1)}\right)^T \\ \\ \lambda_{ij}^{(2)} = u_{ij} - a_{ij} \quad , \quad \mathbf{r}_{ij}^{(2)} = \left(1, \lambda_{ij}^{(2)}\right)^T \end{array} \right. \tag{339}$$

it is possible to introduce the following equality:

$$\left| \tilde{\mathbf{J}}_{ij}^{(x)} \right| \cdot \Delta^{ij} \mathbf{q}^{(x)} = \sum_{k=1}^{2} c_{ij}^{(k)} \, |\lambda_{ij}^{(k)}| \, \mathbf{r}_{ij}^{(k)} \tag{340}$$

where $c_{ij}^{(k)}$ denotes the k-th coordinate of $\Delta^{ij}\mathbf{q}^{(x)}$ with respect to the basis formed by the eigenvectors introduced in (339). Then, by exploiting the following classical property (see e.g. [1] or [111]):

$$\Delta^{ij}(\rho u) = u_{ij} \, \Delta^{ij} \rho + \tilde{\rho}_{ij} \, \Delta^{ij} u \tag{341}$$

with:

$$\tilde{\rho}_{ij} := (\rho_i \rho_j)^{1/2} \tag{342}$$

the following expressions are obtained:

$$\left\{ \begin{array}{lcl} c_{ij}^{(1)} & = & \dfrac{1}{2a_{ij}} \left(\dfrac{\Delta^{ij} p}{a_{ij}} + \tilde{\rho}_{ij} \, \Delta^{ij} u \right) \\ \\ c_{ij}^{(2)} & = & \dfrac{1}{2a_{ij}} \left(\dfrac{\Delta^{ij} p}{a_{ij}} - \tilde{\rho}_{ij} \, \Delta^{ij} u \right) \end{array} \right.$$

and (340) can be recast as follows:

$$\left| \tilde{\mathbf{J}}_{ij}^{(x)} \right| \cdot \Delta^{ij} \mathbf{q}^{(x)} = \frac{1}{2a_{ij}} \, \tilde{\mathbf{U}}_{ij}^{(x)} \cdot \left(\begin{array}{c} \Delta^{ij} p \\ \\ \tilde{\rho}_{ij} \, \Delta^{ij} u \end{array} \right) \tag{343}$$

where the components of the matrix $\tilde{\mathbf{U}}_{ij}^{(x)}$ read:

$$\left\{ \begin{array}{lcl} \tilde{\mathbf{U}}_{ij}^{(x)}(1,1) & = & \dfrac{|\lambda_{ij}^{(1)}| + |\lambda_{ij}^{(2)}|}{a_{ij}} \\ \\ \tilde{\mathbf{U}}_{ij}^{(x)}(1,2) & = & |\lambda_{ij}^{(1)}| - |\lambda_{ij}^{(2)}| \\ \\ \tilde{\mathbf{U}}_{ij}^{(x)}(2,1) & = & \dfrac{\lambda_{ij}^{(1)}|\lambda_{ij}^{(1)}| + \lambda_{ij}^{(2)}|\lambda_{ij}^{(2)}|}{a_{ij}} \\ \\ \tilde{\mathbf{U}}_{ij}^{(x)}(2,2) & = & \lambda_{ij}^{(1)}|\lambda_{ij}^{(1)}| - \lambda_{ij}^{(2)}|\lambda_{ij}^{(2)}| \end{array} \right.$$

For nearly-incompressible flows $|u_{ij}| \ll a_{ij}$ [65] and, consequently, the representation of $\tilde{\mathbf{U}}^{(x)}_{ij}$ reduces to:

$$\tilde{\mathbf{U}}^{(x)}_{ij}\,(M_\star \to 0) \to 2\begin{pmatrix} 1 & u_{ij} \\ & \\ 2u_{ij} & u_{ij}^2 + a_{ij}^2 \end{pmatrix} \tag{344}$$

By substituting (344) into (343) and then back into the proper terms in (338), the following expression is obtained for the nearly-incompressible limit of the (dimensional) semi-discrete system at hand:

$$\left\{\begin{array}{lcrcrcrcr} 2\mu_i \dfrac{\mathrm{d}}{\mathrm{d}t}\,(\rho_i) & = & & & \Psi^{(-1)}_{sd} & + & \Psi^{(0)}_{sd} & + & \Psi^{(1)}_{sd} \\ & & & & & & & & \\ 2\mu_i \dfrac{\mathrm{d}}{\mathrm{d}t}\,(\rho_i u_i) & = & \Theta^{(-2)}_{sd} & + & \Theta^{(-1)}_{sd} & + & \Theta^{(0)}_{sd} & + & \Theta^{(1)}_{sd} \end{array}\right. \tag{345}$$

where:

$$\left\{\begin{array}{lcl} \Psi^{(-1)}_{sd} & := & \dfrac{\Delta^{i(i+1)}p}{a_{i(i+1)}} - \dfrac{\Delta^{(i-1)i}p}{a_{(i-1)i}} \\ \Psi^{(0)}_{sd} & := & \rho_{i-1}u_{i-1} - \rho_{i+1}u_{i+1} \\ \Psi^{(1)}_{sd} & := & M_{i(i+1)}\,\tilde{\rho}_{i(i+1)}\,\Delta^{i(i+1)}u - \\ & & M_{(i-1)i}\,\tilde{\rho}_{(i-1)i}\,\Delta^{(i-1)i}u \\ \Theta^{(-2)}_{sd} & := & p_{i-1} - p_{i+1} \\ \Theta^{(-1)}_{sd} & := & 2\left(M_{i(i+1)}\,\Delta^{i(i+1)}p - M_{(i-1)i}\,\Delta^{(i-1)i}p\right) + \\ & & a_{i(i+1)}\,\tilde{\rho}_{i(i+1)}\,\Delta^{i(i+1)}u - a_{(i-1)i}\,\tilde{\rho}_{(i-1)i}\,\Delta^{(i-1)i}u \\ \Theta^{(0)}_{sd} & := & \rho_{i-1}u_{i-1}^2 - \rho_{i+1}u_{i+1}^2 \\ \Theta^{(1)}_{sd} & := & M_{i(i+1)}\,u_{i(i+1)}\,\tilde{\rho}_{i(i+1)}\,\Delta^{i(i+1)}u - \\ & & M_{(i-1)i}\,u_{(i-1)i}\,\tilde{\rho}_{(i-1)i}\,\Delta^{(i-1)i}u \end{array}\right. \tag{346}$$

[65] Indeed, the density is practically constant and therefore a_{ij}, as given by (154), is of the order of the characteristic sound speed $a_\star$ of the flow. On the other hand, due to the relevant convex combination in (153), $u_{ij} \leq \max(u_i, u_j) = a_\star\,\mathrm{O}\,(M_\star)$.

and:

$$M_{ij} := \frac{u_{ij}}{a_{ij}} \tag{347}$$

Finally, by applying the standard non-dimensionalization procedure mentioned in the relevant paragraph of sec. 3.4.1 to the system (345), the following expression is obtained:

$$\begin{cases} 2\mu_i \dfrac{\mathrm{d}}{\mathrm{d}t}\left(\rho_i\right) & = & & & M_\star^{-1}\,\Psi_{sd}^{(-1)} & + & \hat{\Psi}_{sd}^{(0)} \\ & & & & & & \\ 2\mu_i \dfrac{\mathrm{d}}{\mathrm{d}t}\left(\rho_i u_i\right) & = & M_\star^{-2}\,\Theta_{sd}^{(-2)} & + & M_\star^{-1}\,\Theta_{sd}^{(-1)} & + & \hat{\Theta}_{sd}^{(0)} \end{cases} \tag{348}$$

where:

$$\begin{cases} \hat{\Psi}_{sd}^{(0)} & := & \Psi_{sd}^{(0)} + M_\star\,\Psi_{sd}^{(1)} \\ & & \\ \hat{\Theta}_{sd}^{(0)} & := & \Theta_{sd}^{(0)} + M_\star\,\Theta_{sd}^{(1)} \end{cases} \tag{349}$$

and the relevant coefficients are recalled from (346). The system (348) is copied in sec. 3.4.1, namely to (172), for ease of presentation.

A.3 Proof of the Proposition 7 (sec. 3.4.1)

By substituting the expansion of the continuous solution (173) into the relevant system (170), the following relations are obtained:

$$\begin{cases} \partial_t \rho_0 & = & & & & & \ddot{\Psi}_c^{(0)} \\ & & & & & & \\ \partial_t\left(\rho_0 u_0\right) & = & M_\star^{-2}\,\check{\Theta}_c^{(-2)} & + & M_\star^{-1}\,\check{\Theta}_c^{(-1)} & + & \ddot{\Theta}_c^{(0)} \end{cases} \tag{350}$$

where:

$$\begin{cases} \ddot{\Psi}_c^{(0)} & := & -\,\partial_x\left(\rho_0 u_0\right) + M_\star\left(\cdots\right) \\ & & \\ \check{\Theta}_c^{(-2)} & := & -\,\partial_x p_0 \\ & & \\ \check{\Theta}_c^{(-1)} & := & -\,\partial_x p_1 \\ & & \\ \ddot{\Theta}_c^{(0)} & := & -\,\partial_x\left(\rho_0 u_0^2 + p_2\right) + M_\star\left(\cdots\right) \\ & & \end{cases} \tag{351}$$

Clearly, it is possible to solve the system (350) for $M_\star \to 0$ only if:

$$\check{\Theta}_c^{(-2)} = 0 \quad , \quad \check{\Theta}_c^{(-1)} = 0 \tag{352}$$

The equations (352) imply that:

$$p_0(x,t) = \bar{p}_0(t) \quad , \quad p_1(x,t) = \bar{p}_1(t)$$

for suitable functions $\bar{p}_0$ and $\bar{p}_1$ and therefore the asymptotic expression (176) is obtained.

As far as the semi-discrete problem is concerned, by substituting the expansion (174) into the relevant system (172), the following relations are obtained:

$$\left\{ \begin{array}{lcrcrcr} 2\mu_i \dfrac{\mathrm{d}}{\mathrm{d}t}(\rho_{0i}) & = & & & M_\star^{-1}\,\check{\Psi}_{sd}^{(-1)} & + & \ddot{\Psi}_{sd}^{(0)} \\ & & & & & & \\ 2\mu_i \dfrac{\mathrm{d}}{\mathrm{d}t}(\rho_{0i}u_{0i}) & = & M_\star^{-2}\,\check{\Theta}_{sd}^{(-2)} & + & M_\star^{-1}\,\check{\Theta}_{sd}^{(-1)} & + & \ddot{\Theta}_{sd}^{(0)} \end{array} \right. \tag{353}$$

where the coefficients are suitably defined in terms of the entities introduced in (346) (the definitions are not reported here because inessential to the present purposes). As for the continuous case, the coefficients associated with the negative powers of $M_\star$ in (353) must be identically equal to zero in order to allow for the solution to be defined when $M_\star \to 0$. In particular, the following relations are obtained by respectively imposing $\check{\Psi}_{sd}^{(-1)} = 0$ and $\check{\Theta}_{sd}^{(-2)} = 0$:

$$\frac{\Delta^{i(i+1)}p_0}{a_{0i(i+1)}} - \frac{\Delta^{(i-1)i}p_0}{a_{0(i-1)i}} = 0 \tag{354}$$

$$\Delta^{i(i+1)}p_0 + \Delta^{(i-1)i}p_0 = 0 \tag{355}$$

The equations (354) and (355) above only admit the following solution (as usual, $i \in \mathcal{I}$ and $j \in \pi_i$):

$$\Delta^{ij}p_0 = 0 \tag{356}$$

Indeed, according to (354), $\Delta^{i(i+1)}p_0$ has the same sign as $\Delta^{(i-1)i}p_0$ (a_{0ij} being positive), in contrast with (355) unless (356) holds. The relation (356), in turn, implies that:

$$p_{0i}(t) = \tilde{p}_0(t)$$

for a certain function $\tilde{p}_0$. In consideration of (356), the condition $\check{\Theta}_{sd}^{(-1)} = 0$ leads to the following relation:

$$\tilde{\rho}_0\left(a_{0i(i+1)}\Delta^{i(i+1)}u_0 - a_{0(i-1)i}\Delta^{(i-1)i}u_0\right) + p_{1(i+1)} - p_{1(i-1)} = 0 \tag{357}$$

where $\tilde{\rho}_0$ is defined by inverting the first relation in (175) in correspondence of $p_0 = \tilde{p}_0$. The equation (357), in general, does not impose any specific constraint on $p_{1i}(t)$. As a result, the asymptotic expression (177) is recovered. This completes the proof. ■

A.4 Derivation of the expression (188)

By recalling the relevant definitions, the equation (187) can be recast as follows:

$$\begin{aligned}2\,\mu_i\,\frac{\mathrm{d}}{\mathrm{d}t}\,\mathbf{q}_i^{(x)} &= \mathbf{f}_{i-1}^{(x)} - \mathbf{f}_{i+1}^{(x)} + \\ &\left(\mathbf{P}_{i(i+1)}^{(x)}\right)^{-1} \cdot \left|\mathbf{P}_{i(i+1)}^{(x)} \cdot \tilde{\mathbf{J}}_{i(i+1)}^{(x)}\right| \cdot \Delta^{i(i+1)}\mathbf{q}^{(x)} - \\ &\left(\mathbf{P}_{(i-1)i}^{(x)}\right)^{-1} \cdot \left|\mathbf{P}_{(i-1)i}^{(x)} \cdot \tilde{\mathbf{J}}_{(i-1)i}^{(x)}\right| \cdot \Delta^{(i-1)i}\mathbf{q}^{(x)}\end{aligned} \tag{358}$$

As a preliminary step, a suitable representation is sought for the generic term $\left(\mathbf{P}_{ij}^{(x)}\right)^{-1} \cdot \left|\mathbf{P}_{ij}^{(x)} \cdot \tilde{\mathbf{J}}_{ij}^{(x)}\right| \cdot \Delta^{ij}\mathbf{q}^{(x)}$ appearing, in particular, in (358). To the purpose, once introduced the eigenvalue-eigenvector pairs of $\left(\mathbf{P}_{ij}^{(x)} \cdot \tilde{\mathbf{J}}_{ij}^{(x)}\right)$, namely:

$$\begin{cases} \lambda_{ij}^{(1,p)} = u_{ij}^{(p)} + a_{ij}^{(p)} \quad , \quad \mathbf{r}_{ij}^{(1,p)} = \left(1, \lambda_{ij}^{(1,p)}\right)^T \\ \\ \lambda_{ij}^{(2,p)} = u_{ij}^{(p)} - a_{ij}^{(p)} \quad , \quad \mathbf{r}_{ij}^{(2,p)} = \left(1, \lambda_{ij}^{(2,p)}\right)^T \end{cases} \tag{359}$$

with:

$$u_{ij}^{(p)} := \frac{1+\beta^2}{2}\,u_{ij}$$

and:

$$a_{ij}^{(p)} := \left(\left(\frac{1-\beta^2}{2}\,u_{ij}\right)^2 + (\beta\,a_{ij})^2\right)^{\frac{1}{2}} \tag{360}$$

it is possible to introduce the following equality:

$$\left(\mathbf{P}_{ij}^{(x)}\right)^{-1} \cdot \left|\mathbf{P}_{ij}^{(x)} \cdot \tilde{\mathbf{J}}_{ij}^{(x)}\right| \cdot \Delta^{ij}\mathbf{q}^{(x)} = \left(\mathbf{P}_{ij}^{(x)}\right)^{-1} \cdot \left(\sum_{k=1}^{2} c_{ij}^{(k,p)}\,|\lambda_{ij}^{(k,p)}|\,\mathbf{r}_{ij}^{(k,p)}\right) \tag{361}$$

where $c_{ij}^{(k,p)}$ denotes the k-th coordinate of $\Delta^{ij}\mathbf{q}^{(x)}$ with respect to the basis formed by the eigenvectors introduced in (359). Then, by exploiting the classical property (341), the following expressions are obtained:

$$\begin{cases} c_{ij}^{(1,p)} &= \dfrac{\beta^2}{2\,a_{ij}^{(p)}}\left(\dfrac{\Delta^{ij}p}{\sigma_{ij}} + \tilde{\rho}_{ij}\,\Delta^{ij}u\right) \\ \\ c_{ij}^{(2,p)} &= \dfrac{\beta^2}{2\,a_{ij}^{(p)}}\left(\dfrac{\Delta^{ij}p}{\tau_{ij}} - \tilde{\rho}_{ij}\,\Delta^{ij}u\right) \end{cases}$$

with:

$$\sigma_{ij} := a_{ij}^{(p)} + \frac{1-\beta^2}{2}\, u_{ij} \quad , \quad \tau_{ij} := a_{ij}^{(p)} - \frac{1-\beta^2}{2}\, u_{ij}$$

and (361) can be recast as follows:

$$\left(\mathbf{P}_{ij}^{(x)}\right)^{-1} \cdot \left|\mathbf{P}_{ij}^{(x)} \cdot \tilde{\mathbf{J}}_{ij}^{(x)}\right| \cdot \Delta^{ij}\mathbf{q}^{(x)} = \frac{1}{2\, a_{ij}^{(p)}}\, \tilde{\mathbf{U}}_{ij}^{(x)p} \cdot \begin{pmatrix} \Delta^{ij} p \\ \\ \tilde{\rho}_{ij}\, \Delta^{ij} u \end{pmatrix} \tag{362}$$

where the components of the matrix $\tilde{\mathbf{U}}_{ij}^{(x)p}$ read:

$$\begin{cases} \tilde{\mathbf{U}}_{ij}^{(x)p}(1,1) &= \dfrac{|\lambda_{ij}^{(1,p)}|}{\sigma_{ij}} + \dfrac{|\lambda_{ij}^{(2,p)}|}{\tau_{ij}} \\ \tilde{\mathbf{U}}_{ij}^{(x)p}(1,2) &= |\lambda_{ij}^{(1,p)}| - |\lambda_{ij}^{(2,p)}| \\ \tilde{\mathbf{U}}_{ij}^{(x)p}(2,1) &= \dfrac{\lambda_{ij}^{(1,p)}\, |\lambda_{ij}^{(1,p)}|}{\sigma_{ij}} + \dfrac{\lambda_{ij}^{(2,p)}\, |\lambda_{ij}^{(2,p)}|}{\tau_{ij}} + \left(1-\beta^2\right)\, u_{ij}\, \tilde{\mathbf{U}}_{ij}^{(x)p}(1,1) \\ \tilde{\mathbf{U}}_{ij}^{(x)p}(2,2) &= \lambda_{ij}^{(1,p)}\, |\lambda_{ij}^{(1,p)}| - \lambda_{ij}^{(2,p)}\, |\lambda_{ij}^{(2,p)}| + (1-\beta^2)\, u_{ij}\, \tilde{\mathbf{U}}_{ij}^{(x)p}(1,2) \end{cases}$$

For nearly-incompressible flows $|u_{ij}^{(p)}| \ll a_{ij}^{(p)}$; indeed:

$$\left(u_{ij}^{(p)}\right)^2 - \left(a_{ij}^{(p)}\right)^2 = \beta^2 \left(u_{ij}^2 - a_{ij}^2\right)$$

and therefore the considerations already introduced when discussing the non-preconditioned case in sec. A.2 can be applied. Consequently, the representation of the matrix $\tilde{\mathbf{U}}_{ij}^{(x)p}$ in (362) reduces to:

$$\tilde{\mathbf{U}}_{ij}^{(x)p}\,(M_\star \to 0) \to 2 \begin{pmatrix} 1 - \dfrac{1-\beta^2}{2} M_{ij}^2 & \dfrac{1+\beta^2}{2} u_{ij} \\ \left(\dfrac{3+\beta^2}{2} - \dfrac{1-\beta^2}{2} M_{ij}^2\right) u_{ij} & u_{ij}^2 + \beta^2\, a_{ij}^2 \end{pmatrix} \tag{363}$$

where M_{ij} is defined in (347). By substituting (363) into (362) and then back into the proper terms in (358), the following expression is obtained for the nearly-incompressible limit of the (dimensional) semi-discrete system at hand:

$$\begin{cases} 2\mu_i \dfrac{\mathrm{d}}{\mathrm{d}t}(\rho_i) &= & & \Psi_{sd,p}^{(-1)} & + & \Psi_{sd,p}^{(0)} & + & \Psi_{sd,p}^{(1)} \\ 2\mu_i \dfrac{\mathrm{d}}{\mathrm{d}t}(\rho_i u_i) &= \Theta_{sd,p}^{(-2)} & + & \Theta_{sd,p}^{(-1)} & + & \Theta_{sd,p}^{(0)} & + & \Theta_{sd,p}^{(1)} \end{cases} \tag{364}$$

where:

$$
\left\{
\begin{aligned}
\Psi^{(-1)}_{sd,p} &:= \frac{\Delta^{i(i+1)}p}{a^{(p)}_{i(i+1)}} - \frac{\Delta^{(i-1)i}p}{a^{(p)}_{(i-1)i}} \\
\Psi^{(0)}_{sd,p} &:= \rho_{i-1}u_{i-1} - \rho_{i+1}u_{i+1} \\
\Psi^{(1)}_{sd,p} &:= \frac{1+\beta^2}{2}\Big(M^{(p)}_{i(i+1)}\,\tilde{\rho}_{i(i+1)}\,\Delta^{i(i+1)}u - \\
&\quad M^{(p)}_{(i-1)i}\,\tilde{\rho}_{(i-1)i}\,\Delta^{(i-1)i}u\Big) - \\
&\quad \frac{1-\beta^2}{2}\left(\big(M_{i(i+1)}\big)^2\,\frac{\Delta^{i(i+1)}p}{a^{(p)}_{i(i+1)}} - \big(M_{(i-1)i}\big)^2\,\frac{\Delta^{(i-1)i}p}{a^{(p)}_{(i-1)i}}\right) \\
\Theta^{(-2)}_{sd,p} &:= p_{i-1} - p_{i+1} \\
\Theta^{(-1)}_{sd,p} &:= \frac{3+\beta^2}{2}\Big(M^{(p)}_{i(i+1)}\,\Delta^{i(i+1)}p - M^{(p)}_{(i-1)i}\,\Delta^{(i-1)i}p\Big) + \\
&\quad \beta^2\left(\frac{a^2_{i(i+1)}}{a^{(p)}_{i(i+1)}}\,\tilde{\rho}_{i(i+1)}\,\Delta^{i(i+1)}u - \frac{a^2_{(i-1)i}}{a^{(p)}_{(i-1)i}}\,\tilde{\rho}_{(i-1)i}\,\Delta^{(i-1)i}u\right) \\
\Theta^{(0)}_{sd,p} &:= \rho_{i-1}u^2_{i-1} - \rho_{i+1}u^2_{i+1} \\
\Theta^{(1)}_{sd,p} &:= M^{(p)}_{i(i+1)}\,u_{i(i+1)}\,\tilde{\rho}_{i(i+1)}\,\Delta^{i(i+1)}u - \\
&\quad M^{(p)}_{(i-1)i}\,u_{(i-1)i}\,\tilde{\rho}_{(i-1)i}\,\Delta^{(i-1)i}u - \\
&\quad \frac{1-\beta^2}{2}\Big(\big(M_{i(i+1)}\big)^2\,M^{(p)}_{i(i+1)}\,\Delta^{i(i+1)}p - \\
&\quad \big(M_{(i-1)i}\big)^2\,M^{(p)}_{(i-1)i}\,\Delta^{(i-1)i}p\Big)
\end{aligned}
\right.
\tag{365}
$$

and:

$$M^{(p)}_{ij} := \frac{u_{ij}}{a^{(p)}_{ij}}$$

The non-dimensional counterpart of the system (364) is obtained by applying the standard non-dimensionalization procedure mentioned in the relevant paragraph of sec. 3.4.1. It is worth remarking that, to the purpose, the reference sound speed a_{ref} introduced in (169) is exploited for non-dimensionalizing $a_{ij}^{(p)}$, in view of the fact that $a_{ij}^{(p)}\left(\beta^2 \to 1\right) \to a_{ij}$. Hence, the non-dimensional counterpart of (360) in particular reads [66]:

$$a_{ij}^{(p)} = \left(\left(\frac{1-\beta^2}{2} \, M_\star \, u_{ij} \right)^2 + \left(\beta \, a_{ij} \right)^2 \right)^{\frac{1}{2}} \tag{366}$$

Then, by assuming that the parameter β is formally of the order of the unity, the following non-dimensional expression is obtained for the system (364):

$$\left\{ \begin{array}{lclclcl} 2\mu_i \dfrac{\mathrm{d}}{\mathrm{d}t}\left(\rho_i\right) & = & & & M_\star^{-1}\,\Psi_{sd,p}^{(-1)} & + & \hat{\Psi}_{sd,p}^{(0)} \\ 2\mu_i \dfrac{\mathrm{d}}{\mathrm{d}t}\left(\rho_i u_i\right) & = & M_\star^{-2}\,\Theta_{sd,p}^{(-2)} & + & M_\star^{-1}\,\Theta_{sd,p}^{(-1)} & + & \hat{\Theta}_{sd,p}^{(0)} \end{array} \right. \tag{367}$$

where:

$$\left\{ \begin{array}{lcl} \hat{\Psi}_{sd,p}^{(0)} & := & \Psi_{sd,p}^{(0)} + M_\star\,\Psi_{sd,p}^{(1)} \\ \hat{\Theta}_{sd,p}^{(0)} & := & \Theta_{sd,p}^{(0)} + M_\star\,\Theta_{sd,p}^{(1)} \end{array} \right. \tag{368}$$

and the relevant coefficients (of course, here understood as non-dimensional) are recalled from (365). The system (367) is copied in sec. 3.4.1, namely to (188), for ease of presentation.

[66] The same symbol is used for corresponding dimensional and non-dimensional entities, as declared in Note 33 (sec. 3.4.1).

A.5 A remark on the expression (190)

Both the (non-dimensional) systems (188) and (190) derive from the (dimensional) system (364) by the non-dimensionalization procedure mentioned in the relevant paragraph of sec. 3.4.1. More in detail, in the former case β is assumed of the order of the unity while in the latter one it is assumed of the order of $M_\star$ by means of the position (189). Clearly, (189) directly affects the considered non-dimensional equations through the definitions of the relevant coefficients associated with the powers of $M_\star$. As an example, the non-dimensional form of $a_{ij}^{(p)}$, given by (366), can be considered. Once introduced the following definition:

$$\alpha_{ij} := \left(a_{ij}^{(p)}\right)^2$$

the expansion of $\left(a_{ij}^{(p)}\right)^{-1}$, which appears in (365) both directly and via $M_{ij}^{(p)}$, varies as shown below:

- without the position (189) the expansion of α_{ij} reads:

$$\alpha_{ij} = \alpha_{0ij} + M_\star\,\alpha_{1ij} + \cdots$$

where:

$$\left\{\begin{array}{lcl} \alpha_{0ij} & = & \beta^2\,a_{0ij}^2 \\ \\ \alpha_{1ij} & = & 2\,\beta^2\,a_{0ij}\,a_{1ij} \end{array}\right.$$

and therefore (α_{0ij} is positive, see Note 35 in sec. 3.4.1):

$$\left(a_{ij}^{(p)}\right)^{-1} = \alpha_{0ij}^{-\frac{1}{2}}\left(1 - M_\star\,\frac{\alpha_{1ij}}{2\,\alpha_{0ij}} + \cdots\right)$$

- with the position (189), the following expansion must be considered:

$$\alpha_{ij} = M_\star^2\,(\alpha_{2ij} + M_\star\,\alpha_{3ij} + \cdots)$$

where:

$$\left\{\begin{array}{lcl} \alpha_{2ij} & = & \beta_{ref}^2\,a_{0ij}^2 + \dfrac{1}{4}\,u_{0ij}^2 \\ \\ \alpha_{3ij} & = & 2\,\beta_{ref}^2\,a_{0ij}\,a_{1ij} + \dfrac{1}{2}\,u_{0ij}\,u_{1ij} \end{array}\right. \tag{369}$$

and thus (α_{2ij} is clearly positive, see Note 35 in sec. 3.4.1):

$$\left(a_{ij}^{(p)}\right)^{-1} = M_\star^{-1}\,\alpha_{2ij}^{-\frac{1}{2}}\left(1 - M_\star\,\frac{\alpha_{3ij}}{2\,\alpha_{2ij}} + \cdots\right)$$

A.6 Proof of the Proposition 8 (sec. 3.4.3)

The coefficients associated with the negative powers of $M_\star$ in (190) must be identically equal to zero in order to allow for the solution to be defined for $M_\star \to 0$. In particular, the following relations are derived by respectively imposing $\check{\Psi}^{(-2)}_{sd,p} = 0$ and $\check{\Theta}^{(-2)}_{sd,p} = 0$:

$$\frac{\Delta^{i(i+1)}p_0}{\alpha^{1/2}_{2i(i+1)}} - \frac{\Delta^{(i-1)i}p_0}{\alpha^{1/2}_{2(i-1)i}} = 0 \tag{370}$$

$$\begin{aligned} &\Delta^{i(i+1)}p_0 - \Delta^{(i-1)i}p_0 + \\ &\frac{3}{2}\left(\frac{u_{0i(i+1)}}{\alpha^{1/2}_{2i(i+1)}}\,\Delta^{i(i+1)}p_0 - \frac{u_{0(i-1)i}}{\alpha^{1/2}_{2(i-1)i}}\,\Delta^{(i-1)i}p_0\right) \quad = \quad 0 \end{aligned} \tag{371}$$

with α_{2ij} given by (369). The equations (370) and (371) constitute a system of two homogeneous difference equations for the two unknowns $\Delta^{(i-1)i}p_0$ and $\Delta^{i(i+1)}p_0$. Due to the arbitrariness of the coefficients (depending on the solution itself), it necessarily follows that (as usual, $i \in \mathcal{I}$ and $j \in \pi_i$):

$$\Delta^{ij}p_0 = 0 \tag{372}$$

and therefore p_{0i} does not depend on the spatial index i:

$$p_{0i}(t) = \hat{p}_0(t) \tag{373}$$

and therefore for a suitable function $\hat{p}_0$. In consideration of (372), the conditions $\check{\Psi}^{(-1)}_{sd,p} = 0$ and $\check{\Theta}^{(-1)}_{sd,p} = 0$ respectively read:

$$\frac{\Delta^{i(i+1)}p_1}{\alpha^{1/2}_{2i(i+1)}} - \frac{\Delta^{(i-1)i}p_1}{\alpha^{1/2}_{2(i-1)i}} = 0 \tag{374}$$

$$\begin{aligned} &\Delta^{i(i+1)}p_1 - \Delta^{(i-1)i}p_1 + \\ &\frac{3}{2}\left(\frac{u_{0i(i+1)}}{\alpha^{1/2}_{2i(i+1)}}\,\Delta^{i(i+1)}p_1 - \frac{u_{0(i-1)i}}{\alpha^{1/2}_{2(i-1)i}}\,\Delta^{(i-1)i}p_1\right) \quad = \quad 0 \end{aligned} \tag{375}$$

The equations (374) and (375) are identical to (370) and (371), respectively, once replaced p_1 with p_0. Hence, the following condition can be immediately drawn from (373):

$$p_{1i}(t) = \hat{p}_1(t)$$

where $\hat{p}_1$ denotes a suitable function. As a result, the asymptotic expression (191) is recovered. This completes the proof. ■

A.7 Proof of the Proposition 9 (sec. 3.5.1)

In order to simplify the notation, a generic function $\mathbf{g}(\mathbf{u}, \mathbf{v})$ is considered at a preliminary stage, $\mathbf{u}$ and $\mathbf{v}$ hereafter denoting generic vectors. Moreover, the following definitions, directly derived from (220), are introduced:

$$\begin{cases} \Delta_{(L)}\mathbf{g} & := & \mathbf{g}(\mathbf{u}, \mathbf{v}_0) & - & \mathbf{g}(\mathbf{u}_0, \mathbf{v}_0) \\ \Delta_{(R)}\mathbf{g} & := & \mathbf{g}(\mathbf{u}_0, \mathbf{v}) & - & \mathbf{g}(\mathbf{u}_0, \mathbf{v}_0) \\ \bar{\Delta}_{(L)}\mathbf{g} & := & \mathbf{g}(\mathbf{u}_0, \mathbf{v}) & - & \mathbf{g}(\mathbf{u}, \mathbf{v}) \\ \bar{\Delta}_{(R)}\mathbf{g} & := & \mathbf{g}(\mathbf{u}, \mathbf{v}_0) & - & \mathbf{g}(\mathbf{u}, \mathbf{v}) \end{cases}$$

where $\mathbf{u}_0$ and $\mathbf{v}_0$ represent specific instances of $\mathbf{u}$ and $\mathbf{v}$, respectively.

The proof under consideration exploits the algebraic relation described by the subsequent:

Lemma 1 *Let* $\mathbf{M}(\cdot, \cdot)$ *and* $\mathbf{N}(\cdot, \cdot)$ *denote suitable matrices and let* $\hat{\mathbf{r}}(\cdot, \cdot, \cdot, \cdot)$ *represent a suitable vector, such that:*

$$\begin{aligned} \Delta_{(L)}\mathbf{g} + \Delta_{(R)}\mathbf{g} \;=\; & \mathbf{M}(\mathbf{u}, \mathbf{v}_0) \cdot (\mathbf{u} - \mathbf{u}_0) \;+ \\ & \mathbf{N}(\mathbf{u}_0, \mathbf{v}) \cdot (\mathbf{v} - \mathbf{v}_0) \;+ \\ & \hat{\mathbf{r}}(\mathbf{u}_0, \mathbf{v}_0, \mathbf{u}, \mathbf{v}) \end{aligned} \tag{376}$$

for any value of $\mathbf{u}_0$*,* $\mathbf{v}_0$*,* $\mathbf{u}$ *and* $\mathbf{v}$*. Then, the following relation is satisfied:*

$$\begin{aligned} \mathbf{g}(\mathbf{u}, \mathbf{v}) - \mathbf{g}(\mathbf{u}_0, \mathbf{v}_0) \;=\; & \mathbf{M}(\mathbf{u}_0, \mathbf{v}_0) \cdot (\mathbf{u} - \mathbf{u}_0) \;+ \\ & \mathbf{N}(\mathbf{u}_0, \mathbf{v}_0) \cdot (\mathbf{v} - \mathbf{v}_0) \;+ \\ & \frac{1}{2}\;\mathbf{r}(\mathbf{u}_0, \mathbf{v}_0, \mathbf{u}, \mathbf{v}) \end{aligned} \tag{377}$$

with:

$$\begin{aligned} \mathbf{r}(\mathbf{u}_0, \mathbf{v}_0, \mathbf{u}, \mathbf{v}) := \; & \left(\Delta_{(L)}\mathbf{M} + \Delta_{(R)}\mathbf{M}\right) \cdot (\mathbf{u} - \mathbf{u}_0) \;+ \\ & \left(\Delta_{(L)}\mathbf{N} + \Delta_{(R)}\mathbf{N}\right) \cdot (\mathbf{v} - \mathbf{v}_0) \;- \\ & \Delta_{(LR)}\hat{\mathbf{r}}(\cdot, \cdot, \mathbf{u}_0, \mathbf{v}) \;- \\ & \Delta_{(LR)}\hat{\mathbf{r}}(\cdot, \cdot, \mathbf{u}, \mathbf{v}_0) \end{aligned}$$

and:

$$\Delta_{(LR)}\hat{\mathbf{r}}\left(\cdot,\cdot,\bar{\mathbf{u}},\bar{\mathbf{v}}\right) := \hat{\mathbf{r}}\left(\mathbf{u},\mathbf{v},\bar{\mathbf{u}},\bar{\mathbf{v}}\right) - \hat{\mathbf{r}}\left(\mathbf{u}_0,\mathbf{v}_0,\bar{\mathbf{u}},\bar{\mathbf{v}}\right)$$

Proof *By firstly choosing* $\mathbf{v} = \mathbf{v}_0$ *and then* $\mathbf{u} = \mathbf{u}_0$ *in (376), the following expressions are respectively obtained:*

$$\Delta_{(L)}\mathbf{g} = \mathbf{M}\left(\mathbf{u},\mathbf{v}_0\right)\cdot\left(\mathbf{u}-\mathbf{u}_0\right) + \hat{\mathbf{r}}\left(\mathbf{u}_0,\mathbf{v}_0,\mathbf{u},\mathbf{v}_0\right) \tag{378}$$

$$\Delta_{(R)}\mathbf{g} = \mathbf{N}\left(\mathbf{u}_0,\mathbf{v}\right)\cdot\left(\mathbf{v}-\mathbf{v}_0\right) + \hat{\mathbf{r}}\left(\mathbf{u}_0,\mathbf{v}_0,\mathbf{u}_0,\mathbf{v}\right) \tag{379}$$

Furthermore, by inverting the role of $(\mathbf{u}_0,\mathbf{v}_0)$ *and* $(\mathbf{u},\mathbf{v})$ *it follows that:*

$$\bar{\Delta}_{(L)}\mathbf{g} = \mathbf{M}\left(\mathbf{u}_0,\mathbf{v}\right)\cdot\left(\mathbf{u}_0-\mathbf{u}\right) + \hat{\mathbf{r}}\left(\mathbf{u},\mathbf{v},\mathbf{u}_0,\mathbf{v}\right) \tag{380}$$

$$\bar{\Delta}_{(R)}\mathbf{g} = \mathbf{N}\left(\mathbf{u},\mathbf{v}_0\right)\cdot\left(\mathbf{v}_0-\mathbf{v}\right) + \hat{\mathbf{r}}\left(\mathbf{u},\mathbf{v},\mathbf{u},\mathbf{v}_0\right) \tag{381}$$

Then, once noticed that:

$$2\;\left(\,\mathbf{g}\left(\mathbf{u},\mathbf{v}\right) - \mathbf{g}\left(\mathbf{u}_0,\mathbf{v}_0\right)\,\right) = \Delta_{(L)}\mathbf{g} + \Delta_{(R)}\mathbf{g} - \bar{\Delta}_{(L)}\mathbf{g} - \bar{\Delta}_{(R)}\mathbf{g} \tag{382}$$

the equality (377) is immediately obtained by substituting (378)-(381) into the corresponding terms on the right-hand side of (382). This completes the proof. ■

Let $\boldsymbol{\phi}^{ROE}\left(\mathbf{u},\mathbf{v}\right)$ denote, in the present context, the generic Roe numerical flux $\boldsymbol{\phi}_{LR}^{(g)ROE}$ considered in sec. 3.5.1. It is possible to recast the considered numerical flux as follows (from (135), by a trivial change of notation):

$$\begin{cases} \boldsymbol{\phi}^{ROE}\left(\mathbf{u},\mathbf{v}\right) = \mathbf{f}\left(\mathbf{u}\right) + \tilde{\mathbf{J}}^{-}\left(\mathbf{u},\mathbf{v}\right)\cdot\left(\mathbf{v}-\mathbf{u}\right) \\ \\ \boldsymbol{\phi}^{ROE}\left(\mathbf{u},\mathbf{v}\right) = \mathbf{f}\left(\mathbf{v}\right) - \tilde{\mathbf{J}}^{+}\left(\mathbf{u},\mathbf{v}\right)\cdot\left(\mathbf{v}-\mathbf{u}\right) \end{cases} \tag{383}$$

Then, from the first relation in (383) it follows that:

$$\Delta_{(R)}\boldsymbol{\phi}^{ROE} = \tilde{\mathbf{J}}^{-}\left(\mathbf{u}_0,\mathbf{v}\right)\cdot\left(\mathbf{v}-\mathbf{u}_0\right) - \tilde{\mathbf{J}}^{-}\left(\mathbf{u}_0,\mathbf{v}_0\right)\cdot\left(\mathbf{v}_0-\mathbf{u}_0\right) \tag{384}$$

while from the second one it follows that:

$$\Delta_{(L)}\boldsymbol{\phi}^{ROE} = \tilde{\mathbf{J}}^{+}\left(\mathbf{u}_0,\mathbf{v}_0\right)\cdot\left(\mathbf{v}_0-\mathbf{u}_0\right) - \tilde{\mathbf{J}}^{+}\left(\mathbf{u},\mathbf{v}_0\right)\cdot\left(\mathbf{v}_0-\mathbf{u}\right) \tag{385}$$

Moreover, by combining (384) and (385) the subsequent relation is obtained:

$$\begin{aligned} \Delta_{(L)}\boldsymbol{\phi}^{ROE} + \Delta_{(R)}\boldsymbol{\phi}^{ROE} \;\; = \;\; & \tilde{\mathbf{J}}^{+}\left(\mathbf{u},\mathbf{v}_0\right)\cdot\left(\mathbf{u}-\mathbf{u}_0\right) \;\; + \\ & \tilde{\mathbf{J}}^{-}\left(\mathbf{u}_0,\mathbf{v}\right)\cdot\left(\mathbf{v}-\mathbf{v}_0\right) \;\; + \\ & \hat{\mathbf{r}}^{ROE}\left(\mathbf{u}_0,\mathbf{v}_0,\mathbf{u},\mathbf{v}\right) \end{aligned} \tag{386}$$

with:

$$\hat{\mathbf{r}}^{ROE}\left(\mathbf{u}_0,\mathbf{v}_0,\mathbf{u},\mathbf{v}\right) := \left(\Delta_{(R)}\tilde{\mathbf{J}}^{-} - \Delta_{(L)}\tilde{\mathbf{J}}^{+}\right)\cdot\left(\mathbf{v}_0 - \mathbf{u}_0\right)$$

In consideration of the similarity between (376) and (386), it is possible to apply the Lemma 1 introduced above, thus obtaining the following relation:

$$\begin{aligned} \boldsymbol{\phi}^{ROE}\left(\mathbf{u},\mathbf{v}\right) - \boldsymbol{\phi}^{ROE}\left(\mathbf{u}_0,\mathbf{v}_0\right) \quad = \quad & \tilde{\mathbf{J}}^{+}\left(\mathbf{u}_0,\mathbf{v}_0\right)\cdot\left(\mathbf{u}-\mathbf{u}_0\right) \quad + \\ & \tilde{\mathbf{J}}^{-}\left(\mathbf{u}_0,\mathbf{v}_0\right)\cdot\left(\mathbf{v}-\mathbf{v}_0\right) \quad + \\ \frac{1}{2}\;\; & \mathbf{r}^{ROE}\left(\mathbf{u}_0,\mathbf{v}_0,\mathbf{u},\mathbf{v}\right) \end{aligned} \tag{387}$$

with:

$$\begin{aligned} \mathbf{r}^{ROE}\left(\mathbf{u}_0,\mathbf{v}_0,\mathbf{u},\mathbf{v}\right) := \quad & \left(\Delta_{(L)}\tilde{\mathbf{J}}^{+} + \Delta_{(R)}\tilde{\mathbf{J}}^{+}\right)\cdot\left(\mathbf{u}-\mathbf{u}_0\right) \quad + \\ & \left(\Delta_{(L)}\tilde{\mathbf{J}}^{-} + \Delta_{(R)}\tilde{\mathbf{J}}^{-}\right)\cdot\left(\mathbf{v}-\mathbf{v}_0\right) \quad + \\ & \left(\Delta_{(R)}\tilde{\mathbf{J}}^{-} - \Delta_{(L)}\tilde{\mathbf{J}}^{+}\right)\cdot\left(\mathbf{v}_0-\mathbf{u}_0\right) \quad + \\ & \left(\bar{\Delta}_{(L)}\tilde{\mathbf{J}}^{+} - \bar{\Delta}_{(R)}\tilde{\mathbf{J}}^{-}\right)\cdot\left(\mathbf{v}-\mathbf{u}\right) \end{aligned}$$

Finally, the equality (218) is directly obtained from (387) by means of a straightforward change of notation. This completes the proof. ■

B Appendix: efficient access to the table (256) for pressure-based algorithms

A typical access to the table (256) within a pressure-based algorithm is aimed at finding the density ρ and the sound speed a corresponding to a certain input value of the independent variable $p < p_{sat}$.

It is possible to define a fast look-up strategy by firstly noticing that the distribution along the y-axis in Fig. 44 of the pressure "nodes" p_i, as provided by an ordinary adaptive integration algorithm (e.g. a classical fourth-order Runge-Kutta scheme [79] with adaptive step-size control), is typically clustered around a node $p_{i_\star}$ corresponding to a density $\rho_{i_\star}$ such that $\rho_{i_\star} \approx 0.5 \cdot \rho_{Lsat}$. It is therefore possible to approximate the original pressure sequence p_i by a new one, say p''_k, obtained by juxtaposing two geometric sequences, $p_k^{(up)}$ and $p_k^{(down)}$, both starting from $p_{i_\star}$ and respectively marching towards p_0 and p_{n-1}. Let $\gamma_u > 1$ and $\gamma_d > 1$ denote the ratios of $p_k^{(up)}$ and $p_k^{(down)}$, respectively. Once defined the number of points in each sequence, say n_u and n_d respectively, the following representations are easily obtained:

$$p_k^{(up)} := p_{i_\star} + \frac{\gamma_u^{(n_u-1)-k} - 1}{\gamma_u - 1}\,\delta_u \quad , \quad k \in \{0, \ldots, (n_u - 1)\} \tag{388}$$

$$p_k^{(down)} := p_{i_\star} - \frac{\gamma_d^{k-(n_u-1)} - 1}{\gamma_d - 1}\,\delta_d \quad , \quad k \in \{(n_u - 1), \ldots, (n_u + n_d - 2)\} \tag{389}$$

where:

$$\delta_u := (p_0 - p_{i_\star})\,\frac{\gamma_u - 1}{\gamma_r^{(n_u-1)} - 1}$$

$$\delta_d := (p_{i_\star} - p_{n-1})\,\frac{\gamma_d - 1}{\gamma_d^{n_d-1} - 1}$$

and the new pressure sequence finally reads:

$$p''_k := \begin{cases} p_0 & , \quad k = 0 \\ p_k^{(up)} & , \quad k \in \{1, \ldots, (n_u - 2)\} \\ p_{i_\star} & , \quad k = (n_u - 1) \\ p_k^{(down)} & , \quad k \in \{n_u, \ldots, (n_u + n_d - 3)\} \\ p_{n-1} & , \quad k = (n_u + n_d - 2) \end{cases} \tag{390}$$

The new pressure sequence has a noticeable advantage over the old one: it permits to analytically identify the nodal span to which a given value of the pressure p belongs by inverting (388) and (389) as follows (the cases $p = p_0$, $p = p_{i_\star}$ and $p = p_{n-1}$ are neglected because trivial):

$$p \in \begin{cases} \left[p''_{\mu(p)}, p''_{\mu(p)-1}\right) & , \quad p_{i_\star} < p < p_0 \\ \\ \left(p''_{\nu(p)+1}, p''_{\nu(p)}\right] & , \quad p_{n-1} < p < p_{i_\star} \end{cases} \tag{391}$$

with:

$$\mu(p) := (n_u - 1) - \left\lfloor \frac{1}{\ln(\gamma_u)} \ln\left\{1 + (p - p_{i_\star}) \frac{\gamma_u - 1}{\delta_u}\right\}\right\rfloor \tag{392}$$

$$\nu(p) := (n_u - 1) + \left\lfloor \frac{1}{\ln(\gamma_d)} \ln\left\{1 + (p_{i_\star} - p) \frac{\gamma_d - 1}{\delta_d}\right\}\right\rfloor \tag{393}$$

where, of course, the symbol $\lfloor\cdot\rfloor$ denotes the floor function.

Once defined the new pressure sequence p''_k, a new table can be built either by solving the o.d.e. (250)once more, now in correspondence of the sequence p''_k, or by interpolating the original table. The latter strategy is considered here and the following new table, in particular, is built:

$$(\rho''_k\,, p''_k\,, a''_k) \quad , \quad k \in \{0, \ldots, (n_u + n_d - 2)\} \tag{394}$$

by linearly interpolating the original one (256) in correspondence of the new pressure sequence (390). Clearly, the original table can be discarded at this point, since it is never accessed by the considered algorithm. It may be worth noticing that, besides being attractive for its simplicity, a linear interpolation preserves the strict monotonicity of the p-ρ curve.

For suitable values of the relevant parameters, the new table very well approximates the original one: a fitting practically identical to that one shown in Fig. 46 is obtained, not reported here for brevity. It is therefore natural to define the following two-step access strategy based on table (394):

- given an input pressure p (the cases $p = p_0$, $p = p_{i_\star}$ and $p = p_{n-1}$ are not considered here because trivial), the corresponding span within (394) is identified by means of (391)-(393);

- the values of ρ and a corresponding to p are defined by linear interpolation within the identified span. Of course, this procedure can be extended to an arbitrary number of dependent variables (e.g. the function Ψ, defined in (69), to be used for solving RPs associated with convex state laws, see sec. 2.5.1).

Evidently, the aforementioned access strategy is more efficient than a crude look-up within the original table (256).

References

[1] R. Abgrall. An extension of Roe's upwind scheme to algebraic equilibrium real gas models. *Computers & Fluids*, 19:171–182, 1991.

[2] R. K. Avva, A. K. Singhal, and D. H. Gibson. An enthalpy based model of cavitation. *ASME-FED*, 226:63–70, 1995.

[3] T. Barberon and P. Helluy. Finite volume simulation of cavitating flows. *Computers & Fluids*, 34:832–858, 2005.

[4] A. Bermúdez, A. Dervieux, J. A. Désidéri, and M. E. Vázquez. Upwind schemes for the two-dimensional shallow water equations with variable depth using unstructured meshes. Rapport de recherche 2738, INRIA, 1995.

[5] F. Beux. *Conception optimale de formes aérodynamiques et méthodes d'approximations décentrées pour des écoulements incompressibles.* PhD thesis, University of Nice-Sophia-Antipolis, 1993.

[6] F. Beux, M.V. Salvetti, A. Ignatyev, D. Li, C. Merkle, and E. Sinibaldi. A numerical study of non-cavitating and cavitating liquid flow around a hydrofoil. *ESAIM - Mathematical Modelling and Numerical Analysis*, 39(3):577–590, 2005.

[7] M. Bilanceri. Studio dell'effetto della legge di stato nella simulazione di un flusso cavitante barotropico. Tesi di laurea in Ingegneria Aerospaziale, Università di Pisa, Pisa (Italy), a.a. 2005.

[8] P. Birken and A. Meister. Stability of preconditioned finite volume schemes at low Mach numbers. *BIT-Numerical Mathemathics*, 45(3):463–480, 2005.

[9] C. E. Brennen. *Hydrodynamics of Pumps.* Concepts ETI Inc. and Oxford University Press, 1994.

[10] C. E. Brennen. *Cavitation and Bubble Dynamics.* Oxford University Press, 1995.

[11] W. R. Briley and H. McDonald. An overview and generalization of implicit Navier-Stokes algorithms and approximate factorization. *Computers & Fluids*, 30:807–828, 2001.

[12] H. B. Callen. *Thermodynamics and an Introduction to Thermostatistics.* John Wiley & Sons, 1985.

[13] L. Castelletti, G. Quaranta, and L. Quartapelle. Solution of the Riemann problem for van der Waals gas. Technical Report DIA SR 03-04, Dipartimento di Ingegneria Aerospaziale, Politecnico di Milano, Milano (Italy), 2003.

[14] A. Cervone, L. Torre, C. Bramanti, E. Rapposelli, and L. d'Agostino. Experimental characterization of the cavitation instabilities in the AVIO FAST2 inducer. In *Proc. 41st AIAA/ASME/SAE/ASEE Joint Propulsion Conference*, Tucson (Arizona, USA), July 2005.

[15] Y. Chen and S. D. Heister. Modeling hydrodynamic nonequilibrium in cavitating flows. *Journal of Fluids Engineering*, 118:172–178, 1996.

[16] C.H. Choi, S.-S. Hong, B.J. Cha, and S. Yang. Study on the hydraulic performance of a turbopump inducer. In *Proc. ASME FEDSM'03 - 4th ASME/JSME Joint Fluids Engineering Conference*, Honolulu (Hawaii, USA), July 2003.

[17] Y.-H. Choi and C. L. Merkle. The application of preconditioning in viscous flows. *Journal of Computational Physics*, 105:207–223, 1993.

[18] A. J. Chorin. A numerical method for solving incompressible viscous flow problems. *Journal of Computational Physics*, 2:12–26, 1967.

[19] A. J. Chorin and J. E. Marsden. *A mathematical introduction to fluid mechanics.* Springer, 1993.

[20] R. Courant and K. O. Friedrichs. *Supersonic flow and shock waves.* Springer, 1985.

[21] O. Coutier-Delgosha and J.A. Astolfi. Numerical prediction of the cavitating flow on a two-dimensional symmetrical hydrofoil with a single fluid model. In *Proc. CAV2003 - Fifth International Symposium on Cavitation*, Osaka (Japan), November 2003.

[22] O. Coutier-Delgosha, R. Fortes-Patella, J. L. Reboud, N. Hakimi, and C. Hirsch. Numerical simulation of cavitating flow in 2D and 3D inducer geometries. *International Journal for Numerical Methods in Fluids*, 48:135–167, 2005.

[23] O. Coutier-Delgosha, R. Fortes-Patella, J-L. Reboud, N. Hakimi, and C. Hirsch. Stability of preconditioned Navier-Stokes equations associated with a cavitation model. *Computers & Fluids*, 34:319–349, 2005.

[24] O. Coutier-Delgosha, P. Morel, R. Fortes-Patella, and J. L. Reboud. Numerical simulation of turbopump inducer cavitating behavior. *International Journal of Rotating Machinery*, 2:135–142, 2005.

[25] O. Coutier-Delgosha, J-L. Reboud, and R. Fortes-Patella. Numerical study of the effect of the leading edge shape on cavitation around inducer blade sections. In *Proc. CAV2001 - Fourth International Symposium on Cavitation*, Pasadena (California, USA), June 2001.

[26] L. d'Agostino. Isenthalpic cavitation model. (Private communication).

[27] L. d'Agostino and E. Rapposelli. A modified bubbly isenthalpic model for numerical simulation of cavitating flows. *AIAA paper 2001-3402*, 2001.

[28] C. Debiez and A. Dervieux. Mixed element volume MUSCL methods with weak viscosity for steady and unsteady flow calculation. *Computers & Fluids*, 29:89–118, 2000.

[29] Y. Delannoy. *Modélisation d'écoulements instationnaires et cavitants.* PhD thesis, INPG, Grenoble (France), 1989.

[30] Y. Delannoy and J. L. Kueny. Cavity flow predictions based on the euler equations. *ASME Cavitation and Multiphase Flow Forum*, 109:153–158, 1990.

[31] A. I. Delis, C. P. Skeels, and S. C. Ryrie. Implicit high-resolution methods for modelling one-dimensional open channel flow. *Journal of Hydraulic Research*, 5:369–382, 2000.

[32] A. Dervieux. Steady Euler simulations using unstructured meshes. Von Karman Institute for Fluid Dynamics, Lecture series 1985-04, 1985.

[33] A. Dervieux and J. A. Désidéri. Compressible flow solvers using unstructured grids. Rapport de recherche 1732, INRIA, 1992.

[34] L. C. Evans. *Partial differential equations.* Number 19 in Graduate Studies in Mathematics. American Mathematical Society, 1998.

[35] C. Farhat. High performance simulation of coupled nonlinear transient aeroelastic problems. Special course on parallel computing in CFD. Technical Report R-807, NATO AGARD, October 1995.

[36] L. Fezoui and B. Stoufflet. A class of implicit upwind schemes for Euler simulations with unstructured meshes. *Journal of Computational Physics*, 84:174–206, 1989.

[37] P. Glaister. An approximate linearised Riemann solver for the Euler equations for real gases. *Journal of Computational Physics*, 74:382–408, 1988.

[38] P. Glaister. A Riemann solver for barotropic flow. *Journal of Computational Physics*, 93:477–480, 1991.

[39] E. Godlewski and P.A. Raviart. *Numerical approximation of hyperbolic systems of conservation laws.* Number 118 in Applied Mathematical Sciences. Springer, 1996.

[40] S. K. Godunov. A finite difference method for the computation of discontinuous solutions of the equations of fluid dynamics. *Matematicheskii Sbornik*, 47:357–393, 1959.

[41] A. Guardone and L. Vigevano. Roe linearization for the van der Waals gas. *Journal of Computational Physics*, 175:50–78, 2002.

[42] H. Guillard and C. Viozat. On the behaviour of upwind schemes in the low Mach number limit. *Computers & Fluids*, 28:63–86, 1999.

[43] A. Harten and J. M. Hyman. Self-adjusting grid methods for one-dimensional hyperbolic conservation laws. *Journal of Computational Physics*, 50:235–269, 1983.

[44] C. Hirsch. *Numerical Computation of Internal and External Flows*, volume 1 (Fundamentals of Numerical Discretizations). Wiley, 1988.

[45] H.W.M. Hoeijmakers, M.E. Janssens, and W. Kwan. Numerical simulation of sheet cavitation. In *Proc. Third International Symposium on Cavitation*, Grenoble (France), April 1998.

[46] A. Hosangadi, V. Ahuja, and R.J. Ungewitter. Simulations of cavitating flows in turbopumps. *AIAA paper 2003-1261*, 2003.

[47] A. Hosangadi, V. Ahuja, and R.J. Ungewitter. Simulations of cavitating cryogenic inducers. *AIAA paper 2004-4023*, 2004.

[48] T. Y. Hou and P. Le Floch. Why non-conservative schemes converge to the wrong solutions: error analysis. *Mathematics of Computation*, 62(206):497–530, 1994.

[49] U. Iben, F. Wrona, C.-D. Munz, and M. Beck. Cavitation in hydraulic tools based on thermodynamic properties of liquid and gas. *Journal of Fluids Engineering*, 124:1011–1016, 2002.

[50] M. Ishii. *Thermo-fluid Dynamic Theory of Two-Phase Flow.* Eyrolles, 1975.

[51] M. J. Ivings, D. M. Causon, and E. F. Toro. Riemann solvers for compressible water. In *Proc. Computational Fluid Dynamics '96 - ECCOMAS*, pages 944–949, 1996.

[52] A. Jameson. Time-dependent calculations using multigrid, with applications to unsteady flows past airfoils and wings. *AIAA paper 91-1596*, 1991.

[53] A. Jameson, W. Schmidt, and E. Turkel. Numerical solutions of the Euler equations by finite volume methods using Runge-Kutta time-stepping schemes. *AIAA paper 81-1259*, 1981.

[54] M.E. Janssens, S.J. Hulshoff, and H.W.M. Hoeijmakers. Calculation of unsteady attached cavitation. *AIAA paper 97-1936*, 1997.

[55] A. Jeffrey. *Quasilinear hyperbolic systems and waves.* Pitman, 1976.

[56] A. Jeffrey and T. Taniuti. *Non-linear wave propagation, with applications to physics and magnetohydrodynamics.* New York Academic Press, 1964.

[57] A. Kubota, H. Kato, and H. Yamaguchi. A new modelling of cavitating flows: a numerical study of unsteady cavitation on a hydrofoil section. *Journal of Fluid Mechanics*, 240:59–96, 1992.

[58] R. F. Kunz, D. A. Boger, D. R. Stinebring, T. S. Chyczewski, J. W. Lindau, H. J. Gibeling, S. Venkateswaran, and T. R. Govindan. A preconditioned Navier-Stokes method for two-phase flows application to cavitation prediction. *Computers & Fluids*, 29:849–875, 2000.

[59] B. Lakshminarayana. Fluid dynamics of inducers - a review. *Journal of Fluids Engineering*, 104:411–427, 1982.

[60] P. D. Lax. Development of singularities of solutions of nonlinear hyperbolic partial differential equations. *Journal of Mathematical Physics*, 5:611–613, 1964.

[61] P. D. Lax. *Hyperbolic systems of conservation laws and the mathematical theory of shock waves.* SIAM, 1973.

[62] P. D. Lax and B. Wendroff. Systems of conservation laws. *Communications On Pure & Applied Mathematics*, 13:217–237, 1960.

[63] R. J. Le Veque. *Numerical methods for conservation laws.* Birkhäuser, 1992.

[64] R. J. Le Veque. *Finite volume methods for hyperbolic problems.* Cambridge University Press, 2002.

[65] T. Levi-Civita. *Caratteristiche dei sistemi differenziali e propagazione ondosa.* Zanichelli, 1931.

[66] H. W. Liepmann and A. Roshko. *Elements of Gasdynamics.* John Wiley & Sons, 1957.

[67] R. Martin and H. Guillard. Second-order defect-correction scheme for unsteady problems. *Computers & Fluids*, 25(1):9–27, 1996.

[68] G. Mattei. *Lezioni di meccanica razionale.* SEU - Servizio Editoriale Universitario di Pisa, 1995.

[69] R. Menikoff and B. J. Plohr. The Riemann problem for fluid flow of real materials. *Reviews of Modern Physics*, 61(1):75–130, 1989.

[70] C. L. Merkle, J. Feng, and P. E. O. Buelow. Computational modeling of the dynamics of sheet cavitation. In *Proc. Third International Symposium on Cavitation*, Grenoble (France), April 1998.

[71] B. N'Konga and H. Guillard. Godunov type method on non-structured meshes for three dimensional moving boundary problems. *Computer Methods in Applied Mechanics and Engineering*, 113:183–204, 1994.

[72] O. Oleinik. Discontinuous solutions of nonlinear differential equations. *American Mathematical Society translations series*, 2(26):95–172, 1957.

[73] S. A. Pandya, S. Venkateswaran, and T. H. Pulliam. Implementation of preconditioned dual-time procedures in OVERFLOW. *AIAA paper 2003-0072*, 2003.

[74] S. V. Patankar. *Numerical heat transfer and fluid flow.* Hemisphere, 1980.

[75] R. Peyret and T. Taylor. *Computational methods for fluid flows.* Springer, 1983.

[76] B. Pouffary, R. Fortes-Patella, and J-L Reboud. Numerical simulation of cavitating flow around a 2D hydrofoil: a barotropic approach. In *Proc. CAV2003 - Fifth International Symposium on Cavitation*, Osaka (Japan), November 2003.

[77] A. Preston, T. Colonius, and C. E. Brennen. Toward efficient computation of heat and mass transfer effects in the continuum model for bubbly cavitating flows. In *Proc. CAV2001 - Fourth International Symposium on Cavitation*, Pasadena (California, USA), June 2001.

[78] Q. Qin, C. C. S. Song, and R.E.A. Arndt. A virtual single-phase natural cavitation model and its application to CAV2003 hydrofoil. In *Proc. CAV2003 - Fifth International Symposium on Cavitation*, Osaka (Japan), November 2003.

[79] A. Quarteroni, R. Sacco, and F. Saleri. *Matematica numerica.* Springer, 2000.

[80] A. Quarteroni and A. Valli. *Numerical approximation of partial differential equations.* Number 23 in SCM Series. Springer, 1994.

[81] E. Rapposelli, A. Cervone, C. Bramanti, and L. d'Agostino. Thermal cavitation experiments on a NACA0015 hydrofoil. In *Proc. ASME FEDSM'03 - 4th ASME/JSME Joint Fluids Engineering Conference*, Honolulu (Hawaii, USA), July 2003.

[82] J-L. Reboud, B. Stutz, and O. Coutier. Two-phase flow structure of cavitation: experiment and modelling of unsteady effects. In *Proc. Third International Symposium on Cavitation*, Grenoble (France), April 1998.

[83] W. C. Reynolds. *Thermodynamics properties in SI.* Dept. of Mechanical Engineering, Stanford University, 1979.

[84] P. L. Roe. Approximate Riemann solvers, parameter vectors, and difference schemes. *Journal of Computational Physics*, 43:357–372, 1981.

[85] P. Rostand. *Sur une méthode de volumes finis en maillage non structuré pour le calcul d'écoulements visqueux compressibles.* PhD thesis, Université da Paris VI, 1989.

[86] Y. Saito, I. Nakamori, and T. Ikohagi. Numerical analysis of unsteady vaporous cavitating flow around a hydrofoil. In *Proc. CAV2003 - Fifth International Symposium on Cavitation*, Osaka (Japan), November 2003.

[87] I. Senocak and W. Shyy. A pressure-based method for turbolent cavitating flow computations. *Journal of* C*omputational* P*hysics*, 176:363–383, 2002.

[88] J. Serrin. Mathematical principles of classical fluid mechanics. In *Handbuch der Physik*, volume VIII, pages 125–263. Springer, 1959.

[89] W. Shyy, J. Wu, and Y. Utturkar. Computational modeling of cavitation for liquid rocket applications. *AIAA paper 2004-3985*, 2004.

[90] E. Sinibaldi and F. Beux. A linearised implicit Roe scheme for inhomogeneous flux functions. In *Proc. VII Congresso SIMAI - Società Italiana di Matematica Applicata e Industriale (extended abstract)*, Venezia (Italy), September 2004.

[91] E. Sinibaldi, F. Beux, and M.V. Salvetti. A preconditioned implicit Roe scheme for barotropic flows: towards simulation of cavitation phenomena. Rapport de recherche 4891, INRIA, 2003.

[92] E. Sinibaldi, F. Beux, and M.V. Salvetti. A preconditioned compressible flow solver for numerical simulation of 3D cavitation phenomena. In *Proc. ECCOMAS2004*, Jyväskylä (Finland), July 2004.

[93] E. Sinibaldi, F. Beux, and M.V. Salvetti. A numerical method for 3D barotropic flows in turbomachinery. *Flow, Turbulence and Combustion*, 76, 2006.

[94] E. Sinibaldi, F. Beux, M.V. Salvetti, and L. d'Agostino. Numerical experiments with an homogeneous-flow model for thermal cavitation. In *Proc. CAV2003 - Fifth International Symposium on Cavitation*, Osaka (Japan), November 2003.

[95] G. A. Sod. A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws. *Journal of Computational Physics*, 27:1–31, 1978.

[96] C. C. S. 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.

[97] G. P. Sutton. *Rocket propulsion elements.* John Wiley & Sons, 2001.

[98] E. F. Toro. *Riemann solvers and numerical methods for fluid dynamics.* Springer, 1997.

[99] E. F. Toro. *Shock-capturing methods for free-surface shallow flows.* John Wiley & Sons, 2001.

[100] E. Turkel. Preconditioned methods for solving the incompressible and low speed compressible equations. *Journal of Computational Physics*, 72:277–298, 1987.

[101] E. Turkel. Preconditioning techniques in computational fluid dynamics. *Annual Reviews in Fluid Mechanics 1999*, 31:385–416, 1999.

[102] E. Turkel and V. N. Vatsa. Choice of variables and preconditioning for time dependent problems. *AIAA paper 2003-3692*, 2003.

[103] D. R. van der Heul. *A staggered scheme for nonconvex hyperbolic system of conservation laws.* PhD thesis, Delft University of Technology, 2000.

[104] D. R. van der Heul, C. Vuik, and P. Wesseling. A staggered scheme for hyperbolic conservation laws applied to unsteady sheet cavitation. *Computing and Visualization in Science*, 2:63–68, 1999.

[105] D. R. van der Heul, C. Vuik, and P. Wesseling. Efficient computation of flow with cavitation by compressible pressure correction. In *Proc. ECCOMAS 2000*, Barcelona (Spain), September 2000.

[106] B. van Leer. Towards the ultimate conservative difference scheme III: upstream-centered finite difference schemes for ideal compressible flow. *Journal of Computational Physics*, 23:263–275, 1977.

[107] B. van Leer. Towards the ultimate conservative difference scheme IV: a new approach to numerical convection. *Journal of Computational Physics*, 23:276–299, 1977.

[108] B. van Leer. Towards the ultimate conservative difference scheme V: a second-order sequel to Godunov's method. *Journal of Computational Physics*, 32:101–136, 1979.

[109] B. van Leer, T.-E. Lee, and P. Roe. Characteristic time-stepping or local preconditioning of the Euler equations. *AIAA paper 91-1552*, 1991.

[110] S. Venkateswaran, J. W. Lindau, R. F. Kunz, and C. L. Merkle. Preconditioning algorithms for the computation of multi-phase mixture flows. *AIAA paper 2001-0279*, 2001.

[111] M. Vinokur and J. L. Montagné. Generalized flux-vector splitting and Roe average for an equilibrium real gas. *Journal of Computational Physics*, 89:276–300, 1990.

[112] C. Viozat. Implicit upwind schemes for low Mach number compressible flows. Rapport de recherche 3084, INRIA, 1997.

[113] B. Wendroff. The Riemann problem for materials with nonconvex equations of state, I: Isentropic flow. *Journal of Mathematical Analysis and Applications*, 38:454–466, 1972.

[114] B. Wendroff. The Riemann problem for materials with nonconvex equations of state, II: General flow. *Journal of Mathematical Analysis and Applications*, 38:640–658, 1972.

[115] G. B. Whitham. *Linear and non-linear waves.* John Wiley & Sons, 1974.

[116] J. Wu, Y. Utturkar, and W. Shyy. Assessment of modeling strategies for cavitating flow around a hydrofoil. In *Proc. CAV2003 - Fifth International Symposium on Cavitation*, Osaka (Japan), November 2003.

[117] T. Y. Wu. Cavity and wake flows. *Annual Review of Fluid Mechanics*, 3:243–284, 1972.

[118] H. Yamada, S. Hasegawa, M. Watanabe, T. Hashimoto, T. Kimura, J. Takita, and I. Kubota. Observation of the inner flow in the inducer. In *Proc. 9th International Symposium on Transport Phenomena and Dynamics of Rotating Machinery*, Honolulu (Hawaii, USA), February 2002.