MATHEMATICS OF COMPUTATION
Volume 79, Number 271, July 2010, Pages 1493–1533
S 0025-5718(10)02339-2
Article electronically published on February 23, 2010

ENTROPY-SATISFYING RELAXATION METHOD WITH LARGE TIME-STEPS FOR EULER IBVPS

FRÉDÉRIC COQUEL, QUANG LONG NGUYEN, MARIE POSTEL, AND QUANG HUY TRAN

Abstract. This paper could have been given the title: "How to positively and implicitly solve Euler equations using only linear scalar advections." The new relaxation method we propose is able to solve Euler-like systems—as well as initial and boundary value problems—with real state laws at very low cost, using a hybrid explicit-implicit time integration associated with the Arbitrary Lagrangian-Eulerian formalism. Furthermore, it possesses many attractive properties, such as: (i) the preservation of positivity for densities; (ii) the guarantee of min-max principle for mass fractions; (iii) the satisfaction of entropy inequality, under an expressible bound on the CFL ratio. The main feature that will be emphasized is the design of this optimal time-step, which takes into account data not only from the inner domain but also from the boundary conditions.

1. Introduction

The numerical simulation of compressible fluid flows governed by Euler-like equations has been the subject of extensive studies for several decades [18, 19, 24, 33]. This contribution is concerned with 1-D initial and boundary value problems (IBVPs) within a hybrid explicit-implicit time integration. Although the present work primarily comes within the scope of multiphase flows in pipelines [27, 30], the numerical method we propose extends well beyond it.

In industrial applications, the use of large time-steps by means of an implicit time integration is an essential requirement to reduce the computational cost to an acceptable level. The price to be paid for the CPU saving is that we no longer have any theoretical guarantee for *positivity*, although the supposedly greater amount of numerical dissipation plays in our favor. In the area of implicit methods for Euler equations, it seems that the schemes available so far are either positive, entropic but costly [23], or efficient but more "risky" [8, 28, 36]. The aim of this paper is to show that we can simultaneously achieve low cost and preserve positivity, while maintaining some degree of accuracy on slow waves, at least for the flow regimes described below.

In the flow regimes under consideration, there co-exist two kinds of waves that are clearly separated by their characteristic speeds: fast acoustic waves and slow

Received by the editor December 31, 2007 and, in revised form, February 27, 2009.
2010 *Mathematics Subject Classification.* Primary 65M08; Secondary 35L04.
Key words and phrases. Euler equations, multiphase flow, initial boundary value problems, explicit-implicit, relaxation methods, Lagrange-projection, entropy-satisfying, positivity-preserving.

©2010 American Mathematical Society
Reverts to public domain 28 years from publication

kinematic waves. From the petroleum engineer's standpoint, however, only the kinematic waves are of interest since they represent mass transportation. Therefore, it is wise to find some way to make the time integration implicit with respect to fast waves (to keep the time-step reasonably large), while remaining explicit with respect to slow waves (to maintain accuracy). Such a *hybrid explicit-implicit* scheme in Eulerian coordinates was attempted by Masella et al. [26], followed by Faille and Heintzé [17], in the framework of VFRoe methods. The idea is to forcibly alter the "Roe-matrix" (or more exactly, its VFRoe version) by canceling its slow components. This approach is exact for linear systems, but for nonlinear systems, it is mere heurism, even if it works well in most cases. In any case, it was reused by Baudin et al. [4], as well as by Evje and Flåtten [16]. Unfortunately, little can be said regarding the positivity of such methods.

There is another way, nevertheless, to design a *selectively implicit* scheme. Surprisingly, this second way is based upon a theoretical tool that had been created for quite a different purpose. The Arbitrary Lagrangian-Eulerian (ALE) formalism was introduced [20] to allow for computations over a moving mesh. It consists of two steps: (i) the *Lagrange step*, in which we take into account all physical phenomena except for the displacement of particles; (ii) the *convection step*, during which the quantities are remapped accordingly. When applied to a motionless grid, the two steps most naturally split the waves into two families: fast acoustic waves for the Lagrange step, and slow kinematic waves for the convection step (also called *projection step* or *remap step*). Consequently, all we have to do is to compute the Lagrange step by an implicit scheme, while carrying out the convection step using an explicit scheme.

As a matter of fact, this alternative explicit-implicit approach has already been implemented for years in KIVA [1, 21], a code for 3-D reactive flows, but without the motivation related to the separation of waves. In KIVA, there is no way to ensure positivity either. The time-step for the Lagrange step is assessed by a rule of thumb, whereas in the convection step, the current time-step has to be divided into smaller sub-cycles in order to comply with the CFL condition associated with explicit transport.

Our claim is that, in the 1-D case, it is possible to recover all of the good properties via an *a priori* estimate of the time-step. This estimate is the outcome of a complete theory including existence, uniqueness, positivity, entropy for the IBVP at the continuous and discrete levels. The success of our approach relies on *relaxation* [22, 25, 29], the benefits of which are manifold. First, it is well-known [3, 7, 10, 11] that explicit relaxation schemes can be made positivity-preserving. Second, relaxation provides us with a PDE interpretation, from which a correct treatment for boundary conditions can be derived in the framework proposed by Dubois and LeFloch [15]. Finally, as will be shown in §3, it reduces the Lagrange step to a set of two symmetric scalar linear advection equations with interacting boundary conditions. For this two-advection system, we put forward a short-cut solution procedure and a quick and nearly optimal estimate for L^∞-bounds. Thus, considering that the remap step also boils down to several independent linear scalar advections, it is not unfair to say that we have managed to solve Euler equations by means of linear scalar advections only!

This paper is outlined as follows: We start, in §2, by investigating the two-advection set with coupling boundary conditions as a preliminary tool for the rest

of the paper. Then, we tackle the core of the subject in §3, where we elaborate on the relaxation strategy and the ALE formalism for a simple two-phase flow model. Most importantly, we highlight the connection between the two-advection problem of §2 and the Lagrange step. The details of the scheme, at the fully discrete level, are supplied in §4, along with statements about its properties. In §5, we show how to adapt the new scheme for two-phase flow to Euler's standard equations. Finally, numerical results are given in §6.

2. Symmetric advections with coupling boundary conditions

2.1. **The continuous problem.** Let $a > 0$ and $Z > 0$ be two real constants. Over the time-space domain $\mathbb{R}_+ \times [0, Z]$, we consider the following problem, called *symmetric advections*.

Problem (SA) Given

– the initial data $z \in [0, Z] \mapsto \overrightarrow{w}_\flat(z), \overleftarrow{w}_\flat(z) \in \mathbb{R}^2$,
– the boundary data $t \in \mathbb{R}_+ \mapsto \sigma_0(t), \sigma_Z(t) \in \mathbb{R}^2$,
– the coupling factors $t \in \mathbb{R}_+ \mapsto \theta_0(t), \theta_Z(t) \in \mathbb{R}^2$.

Find

$$t, z \in \mathbb{R}_+ \times [0, Z] \mapsto \overrightarrow{w}(t, z), \overleftarrow{w}(t, z) \in \mathbb{R}^2 \tag{2.1}$$

so as to satisfy the following conditions:

- for $(t, z) \in \mathbb{R}_+^* \times]0, Z[$, the interior advection equations

$$\partial_t \overrightarrow{w} + a\partial_z \overrightarrow{w} = 0, \tag{2.2a}$$

$$\partial_t \overleftarrow{w} - a\partial_z \overleftarrow{w} = 0; \tag{2.2b}$$

- for $z \in]0, Z[$, the initial Cauchy conditions

$$\overrightarrow{w}(t = 0, z) = \overrightarrow{w}_\flat(z), \tag{2.3a}$$

$$\overleftarrow{w}(t = 0, z) = \overleftarrow{w}_\flat(z); \tag{2.3b}$$

- for $t \in \mathbb{R}_+$, the boundary relationships

$$\overrightarrow{w}(t, z = 0) = \sigma_0(t) + \theta_0(t)\overleftarrow{w}(t, z = 0), \tag{2.4a}$$

$$\overleftarrow{w}(t, z = Z) = \sigma_Z(t) + \theta_Z(t)\overrightarrow{w}(t, z = Z). \tag{2.4b}$$

Despite its linearity, *Problem (SA)* will reveal itself to be a convenient building block for the numerical approximation of a class of nonlinear models for fluid flows. It can also be investigated *per se* from the theoretical point of view. This will be done in Appendix A. For the moment, we summarize the main results that will be needed later.

For any open subset $\mathscr{O}$ of $\mathbb{R}$ or $\mathbb{R}^2$ and any function $f \in L^\infty(\mathscr{O}; \mathbb{R})$, we denote by $\|f\|$ its norm, namely,

$$\|f\| = \inf\{M \text{ s.t. } |f(x)| \leq M \text{ for a.e. } x \in \mathscr{O}\}. \tag{2.5}$$

Of course, $\mathscr{O}$ may be the time domain $\mathbb{R}_+^*$ or the space domain $]0, Z[$ or the time-space domain $\mathbb{R}_+^* \times]0, Z[$.

Remark 2.1. The reason why we are using the L^∞-norm, instead of the L^2-norm traditionally associated with linear problems, is that this is the natural setting to express local stability, positivity and maximum principle results.

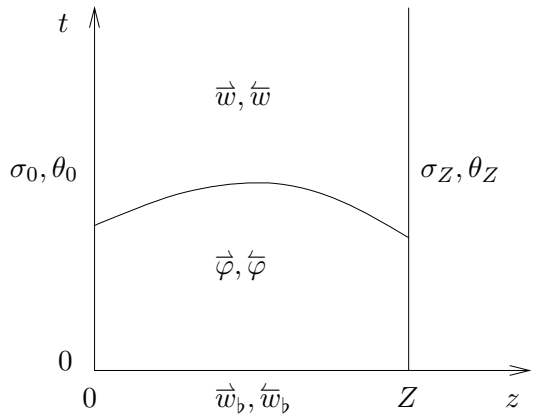

FIGURE 1. *Problem (SA).*

Theorem 2.1. *If $\|\theta_0\|\|\theta_Z\| < 1$, then Problem (SA) has a unique solution. This solution $(\vec{w}, \overleftarrow{w})$ depends continuously on the data $(\vec{w}_\flat, \overleftarrow{w}_\flat, \sigma_0, \sigma_Z)$, that is, there exists a constant $C = C(\|\theta_0\|, \|\theta_Z\|)$ so that*

$$\max\{\|\vec{w}\|, \|\overleftarrow{w}\|\} \le C \max\{\|\vec{w}_\flat\|, \|\overleftarrow{w}_\flat\|, \|\sigma_0\|, \|\sigma_Z\|\}. \tag{2.6}$$

Furthermore, the solution can be expressed by

$$\vec{w}(t,z) = \mathbf{1}_{\{at<z\}}\vec{w}_\flat(z-at) + \mathbf{1}_{\{at>z\}}\vec{w}_0(t-z/a), \tag{2.7a}$$

$$\overleftarrow{w}(t,z) = \mathbf{1}_{\{at<Z-z\}}\overleftarrow{w}_\flat(z+at) + \mathbf{1}_{\{at>Z-z\}}\overleftarrow{w}_Z(t-(Z-z)/a), \tag{2.7b}$$

where $\mathbf{1}_{\{.\}}$ is the characteristic function, and $(\vec{w}_0, \overleftarrow{w}_Z)$ are two auxiliary functions that can be defined in two equivalent manners, i.e.,

(1) *$(\vec{w}_0, \overleftarrow{w}_Z)$ is the unique solution to the coupled system*

$$\vec{w}_0(t) = \sigma_0(t) + \theta_0(t)[\mathbf{1}_{\{at<Z\}}\overleftarrow{w}_\flat(at) + \mathbf{1}_{\{at>Z\}}\overleftarrow{w}_Z(t-Z/a)], \tag{2.8a}$$

$$\overleftarrow{w}_Z(t) = \sigma_Z(t) + \theta_Z(t)[\mathbf{1}_{\{at<Z\}}\vec{w}_\flat(Z-at) + \mathbf{1}_{\{at>Z\}}\vec{w}_0(t-Z/a)]; \tag{2.8b}$$

(2) *$(\vec{w}_0, \overleftarrow{w}_Z)$ is the unique solution to the uncoupled system*

$$\vec{w}_0(t) - \theta_0(t)\theta_Z(t-Z/a)\mathbf{1}_{\{at>2Z\}}\vec{w}_0(t-2Z/a) = G_0(t), \tag{2.9a}$$

$$\overleftarrow{w}_Z(t) - \theta_Z(t)\theta_0(t-Z/a)\mathbf{1}_{\{at>2Z\}}\overleftarrow{w}_Z(t-2Z/a) = G_Z(t), \tag{2.9b}$$

where

$$\begin{aligned} G_0(t) = \sigma_0(t) &+ \theta_0(t)\mathbf{1}_{\{at<Z\}}\overleftarrow{w}_\flat(at) \\ &+ \theta_0(t)\mathbf{1}_{\{at>Z\}}\sigma_Z(t-Z/a) \\ &+ \theta_0(t)\mathbf{1}_{\{at>Z\}}\theta_Z(t-Z/a)\mathbf{1}_{\{at<2Z\}}\vec{w}_\flat(2Z-at), \\ G_Z(t) = \sigma_Z(t) &+ \theta_Z(t)\mathbf{1}_{\{at<Z\}}\vec{w}_\flat(Z-at) \\ &+ \theta_Z(t)\mathbf{1}_{\{at>Z\}}\sigma_0(t-Z/a) \\ &+ \theta_Z(t)\mathbf{1}_{\{at>Z\}}\theta_0(t-Z/a)\mathbf{1}_{\{at<2Z\}}\overleftarrow{w}_\flat(at-Z). \end{aligned} \tag{2.10}$$

Proof. See Appendix A. □

The auxiliary functions $\vec{w}_0$ and $\overleftarrow{w}_Z$ embody the incoming values $\vec{w}(t, z=0)$ and $\overleftarrow{w}(t, z=Z)$. As for

$$\mathbf{1}_{\{at<Z\}}\overleftarrow{w}_\flat(at) + \mathbf{1}_{\{at>Z\}}\overleftarrow{w}_Z(t-Z/a) \quad \text{and} \quad \mathbf{1}_{\{at<Z\}}\vec{w}_\flat(at) + \mathbf{1}_{\{at>Z\}}\vec{w}_0(t-Z/a)$$

in (2.8), in point of fact they represent the outgoing values $\overleftarrow{w}(t, z = 0)$ and $\overrightarrow{w}(t, z = Z)$.

Another result is the min-max principle below, that can be considered as a refined version of the estimate (2.6). Its purpose is to compare solutions at two close time values t and $t + \Delta t$.

Proposition 2.1. *If $0 < \Delta t < Z/a$, then*

(1) *The functions $(\overrightarrow{w}_0, \overleftarrow{w}_Z)$ introduced in Theorem 2.1 and representing incoming boundary values are enclosed by*

$$\overrightarrow{m}_0(t;\Delta t) \le \overrightarrow{w}_0(t') \le \overrightarrow{M}_0(t;\Delta t), \tag{2.11a}$$

$$\overleftarrow{m}_Z(t;\Delta t) \le \overleftarrow{w}_Z(t') \le \overleftarrow{M}_Z(t;\Delta t), \tag{2.11b}$$

for $t' \in [t, t + \Delta t]$, where

$$\begin{aligned}
\overrightarrow{M}_0(t;\Delta t) &= \max_{t_1\in[t,t+\Delta t]} \sigma_0(t_1) + \theta_0(t_1)\overleftarrow{w}(t, a(t_1 - t)),\\
\overrightarrow{m}_0(t;\Delta t) &= \min_{t_1\in[t,t+\Delta t]} \sigma_0(t_1) + \theta_0(t_1)\overleftarrow{w}(t, a(t_1 - t)),\\
\overleftarrow{M}_Z(t;\Delta t) &= \max_{t_1\in[t,t+\Delta t]} \sigma_Z(t_1) + \theta_Z(t_1)\overrightarrow{w}(t, Z - a(t_1 - t)),\\
\overleftarrow{m}_Z(t;\Delta t) &= \min_{t_1\in[t,t+\Delta t]} \sigma_Z(t_1) + \theta_Z(t_1)\overrightarrow{w}(t, Z - a(t_1 - t)).
\end{aligned} \tag{2.12}$$

(2) *The solution functions $(\overrightarrow{w}, \overleftarrow{w})$ at time $t + \Delta t$ are enclosed by*

$$\overrightarrow{m}^{\Delta t}(t, z) \le \overrightarrow{w}(t + \Delta t, z) \le \overrightarrow{M}^{\Delta t}(t, z), \tag{2.13a}$$

$$\overleftarrow{m}^{\Delta t}(t, z) \le \overleftarrow{w}(t + \Delta t, z) \le \overleftarrow{M}^{\Delta t}(t, z), \tag{2.13b}$$

where

$$\begin{aligned}
\overrightarrow{M}^{\Delta t}(t, z) &= \max\{\overrightarrow{M}_0(t;\Delta t), \langle\overrightarrow{M}\rangle(t, z)\} & \overleftarrow{M}^{\Delta t}(t, z) &= \max\{\overleftarrow{M}_Z(t;\Delta t), \langle\overleftarrow{M}\rangle(t, z)\},\\
\overrightarrow{m}^{\Delta t}(t, z) &= \min\{\overrightarrow{m}_0(t;\Delta t), \langle\overrightarrow{m}\rangle(t, z)\} & \overleftarrow{m}^{\Delta t}(t, z) &= \min\{\overleftarrow{m}_Z(t;\Delta t), \langle\overleftarrow{m}\rangle(t, z)\}
\end{aligned}$$

with

$$\begin{aligned}
\langle\overrightarrow{M}\rangle(t, z) &= \max_{z'\in[0,z]} \overrightarrow{w}(t, z') & \langle\overleftarrow{M}\rangle(t, z) &= \max_{z'\in[z,Z]} \overleftarrow{w}(t, z'),\\
\langle\overrightarrow{m}\rangle(t, z) &= \min_{z'\in[0,z]} \overrightarrow{w}(t, z') & \langle\overleftarrow{m}\rangle(t, z) &= \min_{z'\in[z,Z]} \overleftarrow{w}(t, z').
\end{aligned} \tag{2.14}$$

Proof. The first part is a consequence of (2.8), where we have replaced $(\overrightarrow{w}_\flat(.), \overleftarrow{w}_\flat(.))$ by $(\overrightarrow{w}(t, .), \overleftarrow{w}(t, .))$ and t by Δt in the brackets. This is the same as considering the solution at time t as initial data and looking ahead for a small time interval Δt. Since $a\Delta t < Z$, the terms containing $\mathbf{1}_{\{a\Delta t>Z\}}$ disappear and we get (2.12) easily.

To prove the second part, we go along the same lines to deduce (2.14) from (2.7), but this time the initial data and the boundary data have been taken into account. □

The bounds for $\overrightarrow{w}(t+\Delta t, z)$ depend only on what lies on the left of z, while those for $\overleftarrow{w}(t + \Delta t, z)$ depend only on what lies on the right of z. The coupling between $\overrightarrow{w}$ and $\overleftarrow{w}$ is achieved, in reality, via the bounds on $\overrightarrow{w}_0$ and $\overleftarrow{w}_Z$, as demonstrated by (2.12). In (2.14), it would have been sharper to restrict the dependence domains to $[z - a\Delta, z]$ and $[z, z + a\Delta t]$, but our aim is to prepare the ground for a parallel comparison between the continuous and the discrete problems.

2.2. **The discrete problem.** The space domain $[0, Z]$ is divided into N cells of variable lengths Δz_i so that $\sum_{i=1}^{N} \Delta z_i = Z$. In each cell $]z_{i-1/2}, z_{i+1/2}[$ we consider ψ_i representing an approximation of $\psi(z_i)$. To this grid we add two fictitious points, located at $i = 0$ and $i = N+1$ in order to deal with boundary conditions. However, the discrete norm

$$\|\psi\| = \max_{1 \le i \le N} |\psi_i| \tag{2.15}$$

is taken over inner points. Let $\Delta t > 0$ be a time-step. The superscript n will denote the time level t^n, while $n\sharp$ will denote the time level $t^{n\sharp} = t^n + \Delta t$. The problem below is meant to be a discrete version of the continuous *Problem (SA)*.

Problem $(SA)_N^n$ Given, for $0 \le i \le N+1$,

$$\overrightarrow{w}_i^n, \overleftarrow{w}_i^n \in \mathbb{R} \times \mathbb{R}, \quad \sigma_0^n, \sigma_Z^n \in \mathbb{R} \times \mathbb{R}, \quad \theta_0^n, \theta_Z^n \in \mathbb{R} \times \mathbb{R}. \tag{2.16}$$

Find

$$\overrightarrow{w}_i^{n\sharp}, \overleftarrow{w}_i^{n\sharp} \in \mathbb{R} \times \mathbb{R} \quad \text{so as to satisfy} \tag{2.17}$$

- the implicit scheme for interior points $1 \le i \le N$, i.e.,

$$\frac{\overrightarrow{w}_i^{n\sharp} - \overrightarrow{w}_i^n}{\Delta t} + a \frac{\overrightarrow{w}_i^{n\sharp} - \overrightarrow{w}_{i-1}^{n\sharp}}{\Delta z_i} = 0, \tag{2.18a}$$

$$\frac{\overleftarrow{w}_i^{n\sharp} - \overleftarrow{w}_i^n}{\Delta t} - a \frac{\overleftarrow{w}_{i+1}^{n\sharp} - \overleftarrow{w}_i^{n\sharp}}{\Delta z_i} = 0; \tag{2.18b}$$

- the boundary relationships for the two fictitious points, i.e.,

$$\overrightarrow{w}_0^{n\sharp} = \sigma_0^n + \theta_0^n \overleftarrow{w}_0^{n\sharp}, \tag{2.19a}$$

$$\overleftarrow{w}_{N+1}^{n\sharp} = \sigma_Z^n + \theta_Z^n \overrightarrow{w}_{N+1}^{n\sharp}; \tag{2.19b}$$

- the Neumann relationships for outgoing waves, i.e.,

$$\overleftarrow{w}_0^{n\sharp} = \overleftarrow{w}_1^{n\sharp}, \tag{2.20a}$$

$$\overrightarrow{w}_{N+1}^{n\sharp} = \overrightarrow{w}_N^{n\sharp}. \tag{2.20b}$$

We have already mentioned the reason why we chose to work with an implicit scheme such as (2.18): in applications, $(\overrightarrow{w}, \overleftarrow{w})$ will correspond to fast acoustic waves. In (2.19), which is a discrete version of (2.4), the data $(\sigma_0, \sigma_Z, \theta_0, \theta_Z)$ have been frozen to time n to make the presentation easier. Note that the conditions are imposed at the centers of the fictitious cells, not at the edges of the physical domain. The Neumann relationships (2.20) correspond to a wave-cancellation strategy adapted from Dubois and LeFloch [15].

Definition 2.1. Let us introduce

- the local acoustic CFL ratios

$$\mu_i = \frac{a \Delta t}{\Delta z_i}, \tag{2.21}$$

- the local apparent propagation factor

$$e_i = \frac{\mu_i}{1 + \mu_i}, \tag{2.22}$$

- the global cumulated propagation factors

$$E_k^\ell = \begin{cases} \Pi_{j=k}^\ell e_j & \text{if } k \le \ell, \\ 1 & \text{if } k > \ell. \end{cases} \tag{2.23}$$

Although μ_i can be larger than 1, e_i and E_k^ℓ can never exceed or be equal to 1. The name "apparent propagation factor" comes from the following observation. Rewriting the inner equations (2.18) under the form

$$(1+\mu_i)\overrightarrow{w}_i^{n\sharp} - \mu_i \overrightarrow{w}_{i-1}^{n\sharp} = \overrightarrow{w}_i^n, \tag{2.24a}$$

$$(1+\mu_i)\overleftarrow{w}_i^{n\sharp} - \mu_i \overleftarrow{w}_{i+1}^{n\sharp} = \overleftarrow{w}_i^n, \tag{2.24b}$$

we can deduce that

$$\overrightarrow{w}_i^{n\sharp} = e_i \overrightarrow{w}_{i-1}^{n\sharp} + (1-e_i)\overrightarrow{w}_i^n, \tag{2.25a}$$

$$\overleftarrow{w}_i^{n\sharp} = e_i \overleftarrow{w}_{i+1}^{n\sharp} + (1-e_i)\overleftarrow{w}_i^n. \tag{2.25b}$$

In the above convex combinations, the factor e_i accounts for the influence of the upwind cell (i.e., $i-1$ for $\overrightarrow{w}_i$ and $i+1$ for $\overleftarrow{w}_i$) in the updated values at i.

Theorem 2.2. *If $\theta_0^n \theta_Z^n < 1$, then Problem $(SA)_N^n$ is well-posed, in the sense that it has a unique solution. This solution $(\overrightarrow{w}_i^{n\sharp}, \overleftarrow{w}_i^{n\sharp})$ depends continuously on the initial data $(\overrightarrow{w}_i^n, \overleftarrow{w}_i^n, \sigma_0^n, \sigma_Z^n)$, i.e., there is a constant $C = C(\theta_0^n, \theta_Z^n)$, independent of Δt, so that*

$$\max\{\|\overrightarrow{w}^{n\sharp}\|, \|\overleftarrow{w}^{n\sharp}\|\} \le C \max\{\|\overrightarrow{w}^n\|, \|\overleftarrow{w}^n\|, |\sigma_0^n|, |\sigma_Z^n|\}. \tag{2.26}$$

Furthermore, the solution can be given by

$$\overrightarrow{w}_i^{n\sharp} = \textstyle\sum_{k=1}^i (E_{k+1}^i - E_k^i)\overrightarrow{w}_k^n + E_1^i \overrightarrow{w}_0^{n\sharp}, \tag{2.27a}$$

$$\overleftarrow{w}_j^{n\sharp} = \textstyle\sum_{\ell=j}^N (E_j^{\ell-1} - E_j^\ell)\overleftarrow{w}_\ell^n + E_j^N \overleftarrow{w}_{N+1}^{n\sharp} \tag{2.27b}$$

for $0 \le i \le N$, $1 \le j \le N+1$, where the boundary values $(\overrightarrow{w}_0^{n\sharp}, \overleftarrow{w}_{N+1}^{n\sharp})$ can be defined in two equivalent ways, i.e.,

(1) *$(\overrightarrow{w}_0^{n\sharp}, \overleftarrow{w}_{N+1}^{n\sharp})$ is the unique solution to the coupled system*

$$\overrightarrow{w}_0^{n\sharp} = \sigma_0^n + \theta_0^n [\textstyle\sum_{\ell=1}^N \ (E_1^{\ell-1} - E_1^\ell)\overleftarrow{w}_\ell^n + E_1^N \overleftarrow{w}_{N+1}^{n\sharp}], \tag{2.28a}$$

$$\overleftarrow{w}_{N+1}^{n\sharp} = \sigma_Z^n + \theta_Z^n [\textstyle\sum_{k=1}^N (E_{k+1}^N - E_k^N)\overrightarrow{w}_k^n + E_1^N \overrightarrow{w}_0^{n\sharp}]; \tag{2.28b}$$

(2) *$(\overrightarrow{w}_0^{n\sharp}, \overleftarrow{w}_{N+1}^{n\sharp})$ is the unique solution to the uncoupled system*

$$\begin{aligned}[1-\theta_0^n\theta_Z^n (E_1^N)^2]\overrightarrow{w}_0^{n\sharp} \quad &= \sigma_0^n + \theta_0^n \textstyle\sum_{\ell=1}^N (E_1^{\ell-1} - E_1^\ell)\overleftarrow{w}_\ell^n \\ &\quad + \theta_0^n E_1^N [\sigma_Z^n + \theta_Z^n \textstyle\sum_{k=1}^N (E_{k+1}^N - E_k^N)\overrightarrow{w}_k^n],\end{aligned} \tag{2.29a}$$

$$\begin{aligned}[1-\theta_0^n\theta_Z^n (E_1^N)^2]\overleftarrow{w}_{N+1}^{n\sharp} &= \sigma_Z^n + \theta_Z^n \textstyle\sum_{k=1}^N (E_{k+1}^N - E_k^N)\overrightarrow{w}_k^n, \\ &\quad + \theta_Z^n E_1^N [\sigma_0^n + \theta_0^n \textstyle\sum_{\ell=1}^N (E_1^{\ell-1} - E_1^\ell)\overleftarrow{w}_\ell^n].\end{aligned} \tag{2.29b}$$

The formal analogy between this theorem and Theorem 2.1, as reflected by contemplating (2.27)–(2.29) vs. (2.7)–(2.9), is worth mentioning. Note that the assumption $\theta_0^n \theta_Z^n < 1$ at the discrete level is weaker than the condition $\|\theta_0\|\|\theta_Z\| < 1$

at the continuous level. As in the continuous case, there is a refined min-max estimate. For $\theta \in \mathbb{R}$ and $(w_k)_{1\le k\le N}$, we define the upper-bound

$$M(\theta, w) = \begin{cases} \theta \max\limits_{1\le k\le N} w_k & \text{if } \theta \ge 0, \\ -\theta \min\limits_{1\le k\le N} w_k & \text{if } \theta < 0, \end{cases} \tag{2.30}$$

and the lower-bound

$$m(\theta, w) = \begin{cases} \theta \min\limits_{1\le k\le N} w_k & \text{if } \theta \ge 0, \\ -\theta \max\limits_{1\le k\le N} w_k & \text{if } \theta < 0. \end{cases} \tag{2.31}$$

Proposition 2.2. *If $\theta_0^n \theta_Z^n < 1$, then for all $\Delta t > 0$:*

(1) *The values of fictitious points $(\overrightarrow{w}_0, \overleftarrow{w}_Z)$ introduced in Theorem 2.1 are enclosed by*

$$\overrightarrow{m}_0^{n\sharp} \le \overrightarrow{w}_0^{n\sharp} \le \overrightarrow{M}_0^{n\sharp} \qquad \text{and} \qquad \overleftarrow{m}_{N+1}^{n\sharp} \le \overleftarrow{w}_{N+1}^{n\sharp} \le \overleftarrow{M}_{N+1}^{n\sharp}, \tag{2.32}$$

$$\begin{aligned} \overrightarrow{m}_0^{n\sharp} &= \min_{\xi\in[0,1]} \frac{\sigma_0^n + \theta_0^n \sigma_Z^n \xi + m(\theta_0^n\theta_Z^n, \overrightarrow{w}^n)\xi(1-\xi) + m(\theta_0^n, \overleftarrow{w}^n)(1-\xi)}{1-\theta_0^n\theta_Z^n\xi^2}, \\ \overrightarrow{M}_0^{n\sharp} &= \max_{\xi\in[0,1]} \frac{\sigma_0^n + \theta_0^n \sigma_Z^n \xi + M(\theta_0^n\theta_Z^n, \overrightarrow{w}^n)\xi(1-\xi) + M(\theta_0^n, \overleftarrow{w}^n)(1-\xi)}{1-\theta_0^n\theta_Z^n\xi^2}, \\ \overleftarrow{m}_{N+1}^{n\sharp} &= \min_{\xi\in[0,1]} \frac{\sigma_Z^n + \theta_Z^n \sigma_0^n \xi + m(\theta_0^n\theta_Z^n, \overleftarrow{w}^n)\xi(1-\xi) + m(\theta_Z^n, \overleftarrow{w}^n)(1-\xi)}{1-\theta_0^n\theta_Z^n\xi^2}, \\ \overleftarrow{M}_{N+1}^{n\sharp} &= \max_{\xi\in[0,1]} \frac{\sigma_Z^n + \theta_Z^n \sigma_0^n \xi + M(\theta_0^n\theta_Z^n, \overleftarrow{w}^n)\xi(1-\xi) + M(\theta_Z^n, \overleftarrow{w}^n)(1-\xi)}{1-\theta_0^n\theta_Z^n\xi^2}. \end{aligned} \tag{2.33}$$

(2) *The values of inner points $1 \le i, j \le N$ are enclosed by*

$$\overrightarrow{m}_i^{n\sharp} \le \overrightarrow{w}_i^{n\sharp} \le \overrightarrow{M}_i^{n\sharp} \qquad \text{and} \qquad \overleftarrow{m}_j^{n\sharp} \le \overleftarrow{w}_j^{n\sharp} \le \overleftarrow{M}_j^{n\sharp}, \tag{2.34}$$

with

$$\begin{aligned} \overrightarrow{M}_i^{n\sharp} &= \max\{\, \overrightarrow{M}_0^{n\sharp}, \langle \overrightarrow{M} \rangle_i^n \,\} & \overleftarrow{M}_j^{n\sharp} &= \max\{\, \overleftarrow{M}_{N+1}^{n\sharp}, \langle \overleftarrow{M} \rangle_j^n \,\}, \\ \overrightarrow{m}_i^{n\sharp} &= \min\{\, \overrightarrow{m}_0^{n\sharp}, \langle \overrightarrow{m} \rangle_i^n \,\} & \overleftarrow{m}_j^{n\sharp} &= \min\{\, \overleftarrow{m}_{N+1}^{n\sharp}, \langle \overleftarrow{m} \rangle_j^n \,\} \end{aligned}$$

and

$$\begin{aligned} \langle \overrightarrow{M} \rangle_i^n &= \max_{1\le k\le i} \overrightarrow{w}_k^n & \langle \overleftarrow{M} \rangle_j^n &= \max_{j\le \ell\le N} \overleftarrow{w}_\ell^n, \\ \langle \overrightarrow{m} \rangle_i^n &= \min_{1\le k\le i} \overrightarrow{w}_k^n & \langle \overleftarrow{m} \rangle_j^n &= \min_{j\le \ell\le N} \overleftarrow{w}_\ell^n. \end{aligned} \tag{2.35}$$

Formal connections could be made between this proposition and Proposition 2.1, by comparing (2.32)–(2.35) to (2.11)–(2.14). The bounds supplied by (2.32)–(2.35) also have a practical purpose: they will be used for the numerical computation of some optimal CFL ratios in the upcoming Euler problems.

Continuous dependence of $(\overrightarrow{w}^{n\sharp}, \overleftarrow{w}^{n\sharp})$ with respect to $(\sigma_0^n, \sigma_Z^n, \overrightarrow{w}^n, \overleftarrow{w}^n)$, as stated in Theorem 2.2 and improved in Proposition 2.2, can be interpreted as a property of stability. In the case of *Problem* $(SA)_N^n$, however, there is an additional stability property via energy inequalities.

Theorem 2.3. *For any strictly convex function $(\overrightarrow{w}, \overleftarrow{w}) \in \mathbb{R}^2 \mapsto \mathscr{S}(\overrightarrow{w}, \overleftarrow{w}) \in \mathbb{R}$ that is of the form*

$$\mathscr{S}(\overrightarrow{w}, \overleftarrow{w}) = S(\overrightarrow{w}) + S(\overleftarrow{w}), \tag{2.36}$$

in which $w \in \mathbb{R} \mapsto S(w) \in \mathbb{R}$ is strictly a convex function, the implicit scheme (2.18) *of Problem* $(SA)_N^n$ *satisfies the implicit local energy dissipation inequality*

$$\frac{\mathscr{S}(\vec{w}_i^{n\sharp}, \overleftarrow{w}_i^{n\sharp}) - \mathscr{S}(\vec{w}_i^{n}, \overleftarrow{w}_i^{n})}{\Delta t} + \frac{\mathscr{H}(\vec{w}_i^{n\sharp}, \overleftarrow{w}_{i+1}^{n\sharp}) - \mathscr{H}(\vec{w}_{i-1}^{n\sharp}, \overleftarrow{w}_i^{n\sharp})}{\Delta z_i} \leq 0, \tag{2.37}$$

for $1 \leq i \leq N$, *where* $\mathscr{H}(\vec{w}, \overleftarrow{w}) = a[S(\vec{w}) - S(\overleftarrow{w})]$ *is the consistent energy-flux.*

We recall that for smooth solutions of the continuous *Problem (SA)*, combining (2.2a) and (2.2b) leads to

$$\partial_t \mathscr{S}(\vec{w}, \overleftarrow{w}) + \partial_z \mathscr{H}(\vec{w}, \overleftarrow{w}) = 0, \tag{2.38}$$

provided that $\mathscr{S}$ is smooth itself. The fact that this additional conservation law has an inequality counterpart at the discrete level is a major asset for the stability of a scheme.

Remark 2.2. More general energies can be considered for $\mathscr{S}$, but the form (2.36) will be enough for our future purpose.

Proof of Theorem 2.2. *Uniqueness and existence.* Suppose $\vec{w}_0^{n\sharp}$ is known. Then, by (2.25a), we have $\vec{w}_1^{n\sharp} = e_1 \vec{w}_0^{n\sharp} + (1 - e_1)\vec{w}_1^{n}$. By induction on $1 \leq i \leq N$, we carry out a left-to-right sweeping

$$\vec{w}_i^{n\sharp} = E_1^i \vec{w}_0^{n\sharp} + \textstyle\sum_{k=1}^{i} (E_{k+1}^i - E_k^i)\vec{w}_k^n. \tag{2.39}$$

Specifying $i = N$ in (2.39), combining with (2.20b) and using (2.19b), we have

$$\overleftarrow{w}_{N+1}^{n\sharp} = \sigma_Z^n + \theta_Z^n [\textstyle\sum_{\ell=1}^{N} (E_{k+1}^N - E_k^N)\overleftarrow{w}_\ell^n + E_1^N \overleftarrow{w}_0^{n\sharp}]. \tag{2.40}$$

In a similar fashion, if $\vec{w}_{N+1}^{n\sharp}$ is known, we can derive

$$\overleftarrow{w}_j^{n\sharp} = E_j^N \overleftarrow{w}_{N+1}^{n\sharp} + \textstyle\sum_{\ell=j}^{N} (E_j^{\ell-1} - E_j^\ell)\overleftarrow{w}_\ell^n \tag{2.41}$$

for $1 \leq j \leq N$, then

$$\vec{w}_0^{n\sharp} = \sigma_0^n + \theta_0^n [\textstyle\sum_{\ell=1}^{N} (E_1^{\ell-1} - E_1^\ell)\overleftarrow{w}_\ell^n + E_1^N \overleftarrow{w}_{N+1}^{n\sharp}]. \tag{2.42}$$

The system (2.39), (2.41) coincides exactly with (2.27), while the system (2.42), (2.40) is none other than (2.28). A little more algebra shows the equivalence between (2.28) and (2.29). Note that if $\theta_0^n \theta_Z^n < 1$, since $E_1^N < 1$, the bracket $1 - \theta_0^n \theta_Z^n (E_1^N)^2$ always remains positive.

Continuous dependence. Equation (2.29a) gives rise to the abrupt upper-bound

$$|\vec{w}_0^{n\sharp}| \leq C_0 \max\{|\sigma_0^n|, |\sigma_Z^n|, \|\vec{w}^n\|, \|\overleftarrow{w}^n\|\}, \tag{2.43}$$

with

$$\begin{aligned}[1 - \theta_0^n \theta_Z^n (E_1^N)^2] C_0 = 1 &+ |\theta_0^n| \textstyle\sum_{\ell=1}^{N} (E_1^{\ell-1} - E_1^\ell) \\ &+ |\theta_0^n| E_1^N + |\theta_0^n||\theta_Z^n| E_1^N \textstyle\sum_{\ell=1}^{N} (E_{k+1}^i - E_k^i).\end{aligned} \tag{2.44}$$

Note that on one hand $E_1^{\ell-1} - E_1^\ell \geq 0$ and $E_{k+1}^i - E_k^i \geq 0$. On the other hand, the sums involved are telescoping sums, i.e.,

$$\textstyle\sum_{\ell=1}^{N} (E_1^{\ell-1} - E_1^\ell) = \sum_{k=1}^{N} (E_{k+1}^N - E_k^N) = 1 - E_1^N. \tag{2.45}$$

As a result,

$$C_0 = C_0(E_1^N) = \frac{1 + |\theta_0^n|(1 - E_1^N) + |\theta_0^n| E_1^N + |\theta_0^n||\theta_Z^n| E_1^N (1 - E_1^N)}{1 - \theta_0^n \theta_Z^n (E_1^N)^2}. \tag{2.46}$$

To get rid of Δt (through E_1^N) in C_0, we can upper-bound it by

$$\|C_0\| = \max_{\xi\in[0,1]} C_0(\xi),$$

which is finite because $C_0(\xi)$ is a continuous function of $\xi \in [0,1]$. In a similar way, we can show that

$$|\overleftarrow{w}_{N+1}^{n\sharp}| \le \|C_{N+1}\| \max\{|\sigma_0^n|, |\sigma_Z^n|, \|\overrightarrow{w}^n\|, \|\overleftarrow{w}^n\|\} \tag{2.47}$$

for a constant $\|C_{N+1}\|$, which depends on (θ_0^n, θ_Z^n) but not on Δt. This enables us to write

$$\max\{|\overrightarrow{w}_0^{n\sharp}|, |\overleftarrow{w}_{N+1}^{n\sharp}|\} \le C(\theta_0^n, \theta_Z^n) \max\{|\sigma_0^n|, |\sigma_Z^n|, \|\overrightarrow{w}^n\|, \|\overleftarrow{w}^n\|\}, \tag{2.48}$$

with $C(\theta_0^n, \theta_Z^n) = \max(\|C_0\|, \|C_{N+1}\|)$. As for points inside the domain, from the first equation of (2.27), we have

$$|\overrightarrow{w}_i^{n\sharp}| \le [E_1^i + \textstyle\sum_{k=1}^i (E_{k+1}^i - E_k^i)] \max\{|\overrightarrow{w}_0^{n\sharp}|, \|\overrightarrow{w}^n\|\} = \max\{|\overrightarrow{w}_0^{n\sharp}|, \|\overrightarrow{w}^n\|\} \tag{2.49}$$

for all $1 \le i \le N$, so that $\|\overrightarrow{w}^{n\sharp}\|$ is also bounded by the right-hand side of (2.48). The same conclusion holds true for $\|\overleftarrow{w}^{n\sharp}\|$. □

Proof of Proposition 2.2. The first part is a direct consequence of formulae (2.29). The second part is based on (2.27). □

Proof of Theorem 2.3. From the convex combinations (2.25), we infer that

$$S(\overrightarrow{w}_i^{n\sharp}) \le e_i S(\overrightarrow{w}_{i-1}^{n\sharp}) + (1-e_i) S(\overrightarrow{w}_i^n), \tag{2.50a}$$

$$S(\overleftarrow{w}_i^{n\sharp}) \le e_i S(\overleftarrow{w}_{i+1}^{n\sharp}) + (1-e_i) S(\overleftarrow{w}_i^n), \tag{2.50b}$$

insofar as S is a convex function. Using the definition (2.22) of e_i, we cast (2.50) into

$$(1+\mu_i) S(\overrightarrow{w}_i^{n\sharp}) - \mu_i S(\overrightarrow{w}_{i-1}^{n\sharp}) \le S(\overrightarrow{w}_i^n), \tag{2.51a}$$

$$(1+\mu_i) S(\overleftarrow{w}_i^{n\sharp}) - \mu_i S(\overleftarrow{w}_{i+1}^{n\sharp}) \le S(\overleftarrow{w}_i^n). \tag{2.51b}$$

Using the definition (2.21) of μ_i, we go back to the discretized form

$$\frac{S(\overrightarrow{w}_i^{n\sharp}) - S(\overrightarrow{w}_i^n)}{\Delta t} + a \frac{S(\overrightarrow{w}_i^{n\sharp}) - S(\overrightarrow{w}_{i-1}^{n\sharp})}{\Delta z_i} \le 0, \tag{2.52a}$$

$$\frac{S(\overleftarrow{w}_i^{n\sharp}) - S(\overleftarrow{w}_i^n)}{\Delta t} - a \frac{S(\overleftarrow{w}_{i+1}^{n\sharp}) - S(\overleftarrow{w}_i^{n\sharp})}{\Delta z_i} \le 0. \tag{2.52b}$$

To complete the proof, we add (2.52a) and (2.52b). □

2.3. **Practical procedures.** For the sake of computational efficiency, we recommend the following solution procedure, to be implemented in place of explicit formulae (2.27)–(2.29). The basic idea rests on the following observation.

Lemma 2.1. *The mapping $\overrightarrow{w}_0^{n\sharp} \mapsto F(\overrightarrow{w}_0^{n\sharp})$ defined by the diagram*

$$\begin{array}{ccccccccccc}
\boxed{F(\overrightarrow{w}_0^{n\sharp}) \,|\, \overrightarrow{w}_0^{n\sharp}} & \xrightarrow{(2.25a)} & \overrightarrow{w}_1^{n\sharp} & \xrightarrow{(2.25a)} & \dots & \overrightarrow{w}_i^{n\sharp} & \dots & \xrightarrow{(2.25a)} & \overrightarrow{w}_N^{n\sharp} & \xrightarrow{(2.20b)} & \overrightarrow{w}_{N+1}^{n\sharp} \\
\Big\uparrow{\scriptstyle (2.19a)} & & & & & & & & & & \Big\downarrow{\scriptstyle (2.19b)} \\
\overleftarrow{w}_0^{n\sharp} & \xleftarrow[(2.20a)]{} & \overleftarrow{w}_1^{n\sharp} & \xleftarrow[(2.25b)]{} & \dots & \overleftarrow{w}_i^{n\sharp} & \dots & \xleftarrow[(2.25b)]{} & \overleftarrow{w}_N^{n\sharp} & \xleftarrow[(2.25b)]{} & \overleftarrow{w}_{N+1}^{n\sharp}
\end{array}$$

is an affine function, whose derivative is equal to $\theta_0^n \theta_Z^n (E_1^N)^2$.

Proof. Each elementary step of the diagram is an affine operation, therefore, the overall process is an affine function. In the left-to-right propagation using (2.25a), the cumulated factor by which the variable $\vec{w}_0^{n\sharp}$ is multiplied is E_1^N. The outlet condition (2.19b) multiplies it by θ_Z^n. In the right-to-left propagation using (2.25b), the cumulated factor is also E_1^N. The inlet condition (2.19a) multiplies the variable by θ_0^n. □

Of course, the expected value for $\vec{w}_0^{n\sharp}$ is a fixed point of F. The fact that $F(\vec{w}_0) = \theta_0^n \theta_Z^n (E_1^N)^2 \vec{w}_0 + \beta$ naturally suggests a two-step procedure:

(1) First, we set $\vec{w}_0^{n\sharp} = 0$ and apply the sweep process described in the diagram in order to compute $\beta = F(0)$.
(2) Second, we deduce the correct value for $\vec{w}_0^{n\sharp}$ by

$$\vec{w}_0^{n\sharp} = \frac{\beta}{1 - \theta_0^n \theta_Z^n (E_1^N)^2}. \tag{2.53}$$

Once this value is known, a second sweep loop is performed in order to assign the correct values to every other point in the computational domain.

Note that, because of linearity, the existence of a unique fixed point for F only requires $\theta_0^n \theta_Z^n (E_1^N)^2 \neq 1$, which is implied by $\theta_0^n \theta_Z^n < 1$, instead of the contracting property $|\theta_0^n \theta_Z^n (E_1^N)^2| < 1$. In the latter case, we would have to impose $|\theta_0^n||\theta_Z^n| < 1$.

Finally, we wish to point out a cost-effective routine for the computation of the bounds (2.33). The following result is valid only when $\theta = \theta_0^n \theta_Z^n < 0$, but this will be sufficient for our applications.

Lemma 2.2. *For a given $\theta < 0$, the extremal values of the function*

$$f(\xi) = \frac{A\xi^2 + B\xi + C}{1 - \theta\xi^2}, \qquad \xi \in [0,1] \tag{2.54}$$

are given by

$$\min_{\xi\in[0,1]} f(\xi) = \mathbf{1}_{\{B\geq 0\}} f(0) + \mathbf{1}_{\{B<0\}} f(\min(\xi^\star, 1)), \tag{2.55a}$$

$$\max_{\xi\in[0,1]} f(\xi) = \mathbf{1}_{\{B\geq 0\}} f(1) + \mathbf{1}_{\{B<0\}} \max\{f(0), f(1)\} \tag{2.55b}$$

where $\xi^\star$ only needs to be defined for when $B < 0$ by

$$\xi^\star = \frac{-(A + \theta C) + \sqrt{(A + \theta C)^2 - \theta B^2}}{\theta B}. \tag{2.56}$$

Proof. The proof is based on a discussion about the roots of the derivative

$$f'(\xi) = \frac{\theta B \xi^2 + 2(A + \theta C)\xi + B}{(1 - \theta\xi^2)^2}. \tag{2.57}$$

We leave it to the readers. □

3. Two-phase flow model: the continuous problem

3.1. **The original problem.** In this section, we deal with a hydrodynamic model built from an *internal energy* function $(\tau, Y) \in \mathbb{R}_+^* \times [0,1] \mapsto \varepsilon(\tau, Y) \in \mathbb{R}_+$. This

function ε must be smooth enough and have the following properties, in accordance with the framework for compressible fluids proposed by Weyl [35]:

$$\begin{array}{lll} \text{(a)} \quad \varepsilon > 0; & \text{(b)} \quad \varepsilon_\tau < 0; & \text{(c)} \ \ \varepsilon_{\tau\tau} > 0; \\ \text{(d)} \ \ \varepsilon_{\tau\tau\tau} < 0; & \text{(e)} \ \ \varepsilon_{\tau\tau}\varepsilon_{YY} > (\varepsilon_{\tau Y})^2. & \end{array} \tag{3.1}$$

From the internal energy ε, we define

$$\text{the pressure} \quad P(\tau, Y) = -\varepsilon_\tau(\tau, Y); \tag{3.2a}$$

$$\text{the sound speed} \quad c(\tau, Y) = \tau\sqrt{\varepsilon_{\tau\tau}(\tau, Y)} = \tau\sqrt{-P_\tau(\tau, Y)}. \tag{3.2b}$$

Conditions (3.1c), (3.1e) express the fact that ε is strictly convex with respect to (τ, Y). From the standpoint of physics, τ is a specific volume, that is, the inverse of some density ρ, while Y is a mass-fraction. For a prescribed internal energy ε, we state the following IBVP for a fluid model within the phase space

$$\Omega_{\mathbf{U}} = \{\mathbf{U} = (\rho Y, \rho, \rho u) \in \mathbb{R}^3 \ \mid \ \rho > 0, \, u \in \mathbb{R} \ \text{ and } \ Y \in [0,1]\}, \tag{3.3}$$

where u denotes the velocity.

Problem (TP) Given

– the initial data $x \in [0, X] \mapsto \mathbf{U}_\flat(x) \in \Omega_{\mathbf{U}}$,
– the inlet data $t \in \mathbb{R}_+ \mapsto q_0(t), \, g_0(t) \in \mathbb{R}_+^2$,
– the outlet data $t \in \mathbb{R}_+ \mapsto p_X(t), \, Y_X(t) \in \mathbb{R}_+ \times [0,1]$.

Find

$$\mathbf{U} : \, (t, x) \in \mathbb{R}^+ \times \, [0, X] \mapsto \mathbf{U}(t, x) \in \Omega_{\mathbf{U}} \tag{3.4}$$

so as to satisfy (in the usual sense of distributions) the following conditions:

- for $(t, x) \in \mathbb{R}_+^* \times]0, X[$, the system of conservation laws

$$\partial_t(\rho Y) + \partial_x(\rho Y u) \quad = 0, \tag{3.5a}$$

$$\partial_t(\rho) \quad + \partial_x(\rho u) \quad = 0, \tag{3.5b}$$

$$\partial_t(\rho u) \ + \partial_x(\rho u^2 + p) = 0, \tag{3.5c}$$

  with $p = P\left(\dfrac{1}{\rho}, \dfrac{\rho Y}{\rho}\right)$, where P is the pressure defined in (3.2a);

- for $(t, x) \in \mathbb{R}_+^* \times]0, X[$, the energy inequality

$$\partial_t\{\rho\mathfrak{E}\}(\mathbf{U}) + \partial_x\{\rho\mathfrak{E}u + pu\}(\mathbf{U}) \leq 0, \tag{3.6}$$

  with

$$\{\rho\mathfrak{E}\}(\mathbf{U}) = \frac{1}{2}\frac{(\rho u)^2}{\rho} + \rho\varepsilon\left(\frac{1}{\rho}, \frac{\rho Y}{\rho}\right); \tag{3.7}$$

- for $x \in]0, X[$, the initial Cauchy conditions

$$\rho(t = 0, x) = \rho_\flat(x), \quad u(t = 0, x) = u_\flat(x), \quad Y(t = 0, x) = Y_\flat(x); \tag{3.8}$$

- for $t \in \mathbb{R}_+$, the boundary relationships

$$\rho u(t, x = 0) = q_0(t) \qquad \text{if } \ u(t, 0) \ > -c(\rho^{-1}(t, 0), \, Y(t, 0)), \tag{3.9a}$$

$$\rho Y u(t, x = 0) = g_0(t) \qquad \text{if } \ u(t, 0) \ > 0, \tag{3.9b}$$

$$p(t, x = X) = p_X(t) \qquad \text{if } \ u(t, X) < \ \ c(\rho^{-1}(t, X), \, Y(t, X)), \tag{3.9c}$$

$$Y(t, x = X) = Y_X(t) \qquad \text{if } \ u(t, X) < 0, \tag{3.9d}$$

  where c is the sound speed defined in (3.2b).

This problem, called *(TP)* for *two-phase*, is one of the simplest models for flows in pipelines. It is a particular yet prominent case of a more sophisticated model used in the industrial code TACITE [27, 30]. Here, ρ denotes the total density, ρY is the gas density, so that the liquid density can be computed as $\rho(1-Y)$. Both gas and liquid phases move at the same velocity u. The PDE part of this model consists of two mass-balances (3.5b), (3.5a) and one total momentum-balance (3.5c). Then, it is well known [18] that the formula-definition (3.2a) of the pressure law gives rise to a further conservation law

$$\partial_t\{\rho\mathfrak{E}\}(\mathbf{U}) + \partial_x\{\rho\mathfrak{E}u + pu\}(\mathbf{U}) = 0 \tag{3.10}$$

for the smooth solutions of (3.5). In addition, assumptions (3.1c)–(3.1e) ensure that the mapping $\mathbf{U} \in \Omega_{\mathbf{U}} \to \{\rho\mathfrak{E}\}(\mathbf{U}) \in \mathbb{R}_+$ is strictly convex. Hence, $(\rho\mathfrak{E},\, \rho\mathfrak{E}u + pu)$ may serve as an entropy pair for selecting the physical weak solution of (3.5) via the energy inequality (3.6).

The boundary conditions (3.9) represent the operating modes available to the pipeline monitors. At the inlet $x = 0$, we would like to impose the flow rates (3.9a), (3.9b) whenever the physics of waves allows us to do so. At the outlet $x = X$, we would like to impose the pressure (3.9c) whenever the physics is in agreement with our wishes; should the flow direction happen to be reverted at the outlet, we would also like to prescribe the incoming gas fraction (3.9d). In practice, since the flows considered are always subsonic, the first three conditions (3.9a)–(3.9c) are systematically active, while (3.9d) depends on the test case at hand. In order to express the above boundary conditions, we adopt the theory developed by Dubois-LeFloch [15] based on the notion of half-Riemann problems.

For conciseness in the notation, the PDE model (3.5) is written in the condensed form

$$\partial_t\mathbf{U} + \partial_x\mathbf{F}(\mathbf{U}) = 0. \tag{3.11}$$

The following proposition collects the classical properties of (3.5) that we will use later.

Proposition 3.1. *The system* (3.5) *is hyperbolic over* $\Omega_{\mathbf{U}}$*, i.e., for any state* $\mathbf{U} \in \Omega_{\mathbf{U}}$*, the Jacobian matrix* $\nabla_{\mathbf{U}}\mathbf{F}(\mathbf{U})$ *has real eigenvalues*

$$u - c(\mathbf{U}) \,<\, u \,<\, u + c(\mathbf{U}) \tag{3.12}$$

and is $\mathbb{R}$*-diagonalizable. The two extreme fields are genuinely nonlinear while the intermediate one is linearly degenerate.*

Furthermore, the mapping $\mathbf{U} \in \Omega_{\mathbf{U}} \to \{\rho\mathfrak{E}\}(\mathbf{U}) \in \mathbb{R}_+$ *is strictly convex.*

Proof. The calculations can be found in [18], for instance. Hyperbolicity is due to (3.1c), genuine nonlinearity of the extreme fields is due to (3.1d), the additional law (3.10) follows from (3.2a) and strict convexity of $\rho\mathfrak{E}$ is due to (3.1c), (3.1e). □

3.2. **The relaxation problem.** As explained in [5, 10, 22] (see also [2, 3, 4]), it is judicious to approximate the entropic weak solutions of the original problem by those of a relaxation model: this helps us cope more easily with the nonlinearities in the closure laws.

It is well known [18] that the (strict) convexity $\partial_{\tau\tau}P(\tau, Y) > 0$ stated in (3.1d) is responsible for the (genuine) nonlinearities in the two extreme fields. Following

the strategy developed in [10] (see also [3, 6, 7]), we propose to modify the reported nonlinearities by approximating the exact pressure law by

$$\Pi(\tau,\zeta,Y) = P(\zeta,Y) + a^2(\zeta-\tau), \tag{3.13}$$

for some given positive constant $a>0$. The new unknown ζ is intended to coincide with τ in the limit of an infinite relaxation parameter so as to restore the original nonlinearities. The estimates

$$\partial_\tau \Pi(\tau,\zeta,Y) < 0 \quad \text{and} \quad \partial_{\tau\tau}\Pi(\tau,\zeta,Y) = 0,$$

to be compared with (3.1c), (3.1d), actually ensure that the relaxation PDE model is hyperbolic but with only linearly degenerate fields [6, 7].

Over the phase space,

$$\Omega_{\mathbb{U}} = \{\mathbb{U} = (\rho Y, \rho, \rho u, \rho\zeta) \in \mathbb{R}^4 \mid \rho>0,\ \zeta>0 \text{ and } Y\in[0,1]\}, \tag{3.14}$$

and for a fixed relaxation parameter $\lambda>0$, we introduce the relaxation approximation *(TP-R)*$^a_\lambda$ of the original problem *(TP)* in the following way.

Problem (TP-R)$^a_\lambda$ Given

- the initial data $x\in[0,X] \mapsto \mathbb{U}_\flat(x)\in\Omega_{\mathbb{U}}$,
- the inlet boundary data $t\in\mathbb{R}_+ \mapsto q_0(t),\, g_0(t)\in\mathbb{R}_+^2$,
- the outlet boundary data $t\in\mathbb{R}_+ \mapsto p_X(t),\, Y_X(t)\in\mathbb{R}_+\times[0,1]$.

Find

$$\mathbb{U}^\lambda : (t,x)\in\mathbb{R}_+\times[0,X] \mapsto \mathbb{U}^\lambda(t,x)\in\Omega_{\mathbb{U}} \tag{3.15}$$

so as to satisfy in the usual weak sense (for clarity the superscripts λ for the components of $\mathbb{U}^\lambda$ are omitted):

- for $(t,x)\in\mathbb{R}_+^*\times]0,X[$, the system of conservation laws

$$\partial_t(\rho Y) + \partial_x(\rho Y u) = 0, \tag{3.16a}$$
$$\partial_t(\rho) + \partial_x(\rho u) = 0, \tag{3.16b}$$
$$\partial_t(\rho u) + \partial_x(\rho u^2 + \Pi(\tau,\zeta,Y)) = 0, \tag{3.16c}$$
$$\partial_t(\rho\zeta) + \partial_x(\rho\zeta u) = \lambda\rho[\tau-\zeta]; \tag{3.16d}$$

- for $(t,x)\in\mathbb{R}_+^*\times]0,X[$, the energy inequality

$$\partial_t\{\rho\mathscr{E}\}(\mathbb{U}^\lambda) + \partial_x\{\rho\mathscr{E}u + \Pi u\}(\mathbb{U}^\lambda) \le 0, \tag{3.17}$$

with

$$\{\rho\mathscr{E}\}(\mathbb{U}^\lambda) = \frac{1}{2}\frac{(\rho u)^2}{\rho} + \rho\varepsilon\left(\frac{\rho\zeta}{\rho},\frac{\rho Y}{\rho}\right) + \frac{\rho}{2a^2}\left[\Pi^2 - P^2\left(\frac{\rho\zeta}{\rho},\frac{\rho Y}{\rho}\right)\right]; \tag{3.18}$$

- for $x\in]0,X[$, the initial Cauchy conditions

$$\rho(t=0,x) = \rho_\flat(x), \qquad u(t=0,x) = u_\flat(x), \tag{3.19a}$$
$$Y(t=0,x) = Y_\flat(x), \qquad \zeta(t=0,x) = \zeta_\flat(x); \tag{3.19b}$$

- for $t \in \mathbb{R}_+$, the boundary relationships

$$\rho u(t, x = 0) = q_0(t) \qquad \text{if } u(t,0) > -a\tau(t,0), \tag{3.20a}$$

$$\rho Y u(t, x = 0) = g_0(t) \qquad \text{if } u(t,0) > 0, \tag{3.20b}$$

$$\Pi(t, x = X) = p_X(t) \qquad \text{if } u(t,X) < a\tau(t,X), \tag{3.20c}$$

$$Y(t, x = X) = Y_X(t) \qquad \text{if } u(t,X) < 0. \tag{3.20d}$$

Clearly, the limit $\lambda \to +\infty$ in (3.16) formally gives $\zeta = \tau$ and thus restores $\Pi = P(\tau, Y)$ and $\mathscr{E} = \mathfrak{E}(\tau, Y, u)$. In other words, the original equations (3.5) together with the entropy diminishing condition (3.6) are formally recovered in the limit of an infinite relaxation parameter. However, to prevent the relaxation approximation from instabilities in the asymptotic regime $\lambda \to +\infty$, the relaxation system (3.16) is required to be uniformly compatible with the privileged entropy $\mathscr{E}$, according to the work by Liu [25] and Chen et al. [9]. The relaxation entropy inequality (3.17) accounts for this stability requirement. Its detailed form reads [7]

$$\partial_t\{\rho\mathscr{E}\}(\mathbb{U}^\lambda) + \partial_x\{\rho\mathscr{E}u + \Pi u\}(\mathbb{U}^\lambda) = -\lambda\rho[a^2 + P_\tau(\zeta, Y)](\tau - \zeta)^2 \leq 0. \tag{3.21}$$

For this inequality to be valid for all $\lambda > 0$, the positive constant a entering the definition of the relaxation pressure law (3.13) must be chosen in order to obey the subcharacteristic condition [6, 9, 25]

$$a^2 > -P_\tau(\zeta, Y) \tag{3.22}$$

for all (ζ, Y) under consideration. We also refer to (3.22) as the Whitham condition.

For simplicity in the notation, the relaxation system (3.16) is rewritten in the condensed form

$$\partial_t \mathbb{U}^\lambda + \partial_x \mathbb{F}(\mathbb{U}^\lambda) = \lambda \mathrm{R}(\mathbb{U}^\lambda). \tag{3.23}$$

Let us summarize the main properties of (3.16) that will soon be of interest.

Proposition 3.2. *The first order system in* (3.16) *is hyperbolic over* $\Omega_{\mathbb{U}}$*, i.e., for any state* $\mathbb{U} \in \Omega_{\mathbb{U}}$*, the Jacobian matrix* $\nabla_{\mathbb{U}}\mathbb{F}(\mathbb{U})$ *has real eigenvalues*

$$u - a\tau < u = u < u + a\tau \tag{3.24}$$

and is $\mathbb{R}$*-diagonalizable. The eigenvalues all correspond to linearly degenerate fields and are associated with the strong Riemann invariants*

$$\overleftarrow{w} = \Pi - au, \quad Y, \quad \mathscr{I} = \Pi + a^2\tau, \quad \overrightarrow{w} = \Pi + au. \tag{3.25}$$

Furthermore, the solutions of (3.16) *satisfy the additional conservation law*

$$\partial_t\{\rho\Pi\}(\mathbb{U}^\lambda) + \partial_x\{\rho\Pi u + a^2 u\}(\mathbb{U}^\lambda) = \lambda\rho[1 + a^{-2}P_\tau(\zeta, Y)](P(\zeta, Y) - \Pi). \tag{3.26}$$

Proof. The calculations are easily adapted from [3, 4]. Because of the linear degeneracy of all fields, the additional law (3.26) holds with equality in the sense of distributions for the discontinuous solutions of (3.16). □

Let Δt be the time-step. As explained in [3, 4] and illustrated in (3.27) below, the relaxation strategy consists of two steps. First, starting from the data $\mathbb{U}_\flat = \mathbb{U}^n = \mathbb{U}(t^n, .)$ at equilibrium, that is, with $\zeta^n = \tau^n$, we solve *Problem (TP-R)*$^a_{\lambda=0}$ from t^n until $t^{n\text{X}} = t^n + \Delta t$. Since the relaxation parameter is set at $\lambda = 0$, the outcome $\mathbb{U}^{n\text{X}}$ will be out of equilibrium, i.e., $(\rho\zeta)^{n\text{X}} \neq 1$. Second, we project it onto the equilibrium manifold by setting $\zeta^{n+1} = \tau^{n\text{X}}$, while keeping the remaining components:

(3.27)

$$\begin{array}{ccccc} \mathbb{U}^n = \mathbb{U}_\flat & \xrightarrow[\text{by some method,}]{\text{solve } (TP\text{-}R)^a_0} & \mathbb{U}^{n\text{X}} & \xrightarrow[\text{equilibrium}]{\text{return to}} & \mathbb{U}^{n+1} \\ \| & & \| & & \| \\ (\mathbf{U}^n, (\rho\zeta)^n = 1) & \xrightarrow[\text{projection}]{\text{e.g., Lagrange-}} & (\mathbf{U}^{n\text{X}}, (\rho\zeta)^{n\text{X}}) & \longrightarrow & (\mathbf{U}^{n+1} = \mathbf{U}^{n\text{X}}, 1) \end{array}$$

The question remains as to how we can find a good scheme for the first step. In [3, 4], *Problem (TP-R)*$^a_0$ was solved by a direct Eulerian approach. In this paper, we propose an indirect but much more advantageous approach, based on the Lagrange-Euler decomposition of the relaxation system (3.16).

3.3. **The relaxation problem in ALE coordinates.** Let us introduce a new referential frame, in which the coordinates are denoted by χ. This frame is neither the material (Lagrangian) configuration $\mathscr{X}$ nor the laboratory (Eulerian) configuration x. Instead, it moves at the imposed speed $u - v$ with respect to the laboratory. Then, the velocity of the particles with respect to the moving frame, as seen from the laboratory, is equal to v.

Let $x = x(\chi, t)$ be the correspondence between the moving frame and the laboratory frame, and let $J = \partial_\chi x_{|t}$ be the dilatation rate. Then, from the calculations presented in [13, 14, 20], it is a classical exercise to prove that system (3.16) is equivalent to

$$\partial_t(J) + \partial_\chi(v) - \partial_\chi(u) = 0, \tag{3.28a}$$

$$\partial_t(\rho Y J) + \partial_\chi(\rho Y v) = 0, \tag{3.28b}$$

$$\partial_t(\rho J) + \partial_\chi(\rho v) = 0, \tag{3.28c}$$

$$\partial_t(\rho u J) + \partial_\chi(\rho u v) + \partial_\chi(\Pi) = 0, \tag{3.28d}$$

$$\partial_t(\rho\zeta J) + \underbrace{\partial_\chi(\rho\zeta v)}_{\text{projection}} = \underbrace{\lambda\rho J(\tau - \zeta)}_{\text{Lagrange}}. \tag{3.28e}$$

The formulation (3.28) most naturally separates fast acoustic waves from slow kinematic waves. Therefore, the basic idea of Arbitrary Lagrangian-Eulerian (ALE) approaches is to perform a splitting of (3.28), as indicated above, within a time-step Δt. The Lagrange-projection method that we are going to detail is a special case of ALE, in which v is chosen so as to come back to Eulerian coordinates after the two steps, namely, to secure $J^{n\text{X}} = 1$:

$$(J^n = 1, \mathbb{U}^n) \xrightarrow{\text{Lagrange}} (J^{n\sharp}, \mathbb{U}^{n\sharp}) \xrightarrow{\text{projection}} (J^{n\text{X}} = 1, \mathbb{U}^{n\text{X}}). \tag{3.29}$$

3.3.1. *Lagrange step.* In the Lagrange step, which takes into account only acoustic effects due to the pressure, the PDE system to be solved reads

$$\partial_t(J) \quad - \partial_\chi(u) = 0, \tag{3.30a}$$
$$\partial_t(\rho Y J) \qquad = 0, \tag{3.30b}$$
$$\partial_t(\rho J) \qquad = 0, \tag{3.30c}$$
$$\partial_t(\rho u J) + \partial_\chi(\Pi) = 0, \tag{3.30d}$$
$$\partial_t(\rho\zeta J) \qquad = \lambda\rho J(\tau - \zeta). \tag{3.30e}$$

This system is equipped with the initial data $(J_\flat, \mathbb{U}_\flat) = (J^n = 1, \mathbb{U}^n)$ and a suitably modified version of the boundary conditions (3.20), namely,

$$\begin{array}{ll} \text{(a)} \quad \rho u(t, \chi = 0) = q_0(t); & \text{(c)} \quad \Pi(t, \chi = X) = p_X(t); \\ \text{(b)} \quad \rho Y u(t, \chi = 0) = g_0(t); & \text{(d)} \quad Y(t, \chi = X) = Y_X(t). \end{array} \tag{3.31}$$

It is important to note that when $\lambda = 0$, it is possible to solve (3.30)–(3.31) by means of *Problem (SA)*. In other words, we can reduce the Lagrange step to the problem of two symmetric advections with coupling boundary conditions.

Theorem 3.1. *Let $m = \rho_\flat = \rho^n > 0$ be the initial density. Define*

$$z = \int_0^\chi m(\varkappa)\, d\varkappa \qquad \text{and} \qquad Z = \int_0^X m(\varkappa)\, d\varkappa. \tag{3.32}$$

Then, the Lagrange step (3.30)–(3.31) *with $\lambda = 0$ is equivalent to*

- *the PDE system*

$$\begin{array}{ll} \text{(a)} \quad \partial_t Y = 0; & \text{(c)} \quad \partial_t \overrightarrow{w} + a\partial_z \overrightarrow{w} = 0; \\ \text{(b)} \quad \partial_t \mathscr{I} = 0; & \text{(d)} \quad \partial_t \overleftarrow{w} - a\partial_z \overleftarrow{w} = 0, \end{array} \tag{3.33}$$

  where Y and $(\overrightarrow{w}, \overleftarrow{w}, \mathscr{I}) = (\Pi + au, \Pi - au, \Pi + a^2\tau)$, already introduced in (3.25)*, are to be considered as functions of $(t, z) \in [t^n, t^{n+1}] \times [0, Z]$;*
- *and the boundary conditions*

$$\begin{array}{ll} \text{(a)} \ Y(t, z = 0) = g_0(t)/q_0(t); & \text{(b)} \ \overrightarrow{w}(t, z = 0) = \sigma_0(t) + \theta_0(t)\overleftarrow{w}(t, z = 0); \\ \text{(c)} \ Y(t, z = Z) = Y_X(t); & \text{(d)} \ \overleftarrow{w}(t, z = Z) = \sigma_Z(t) + \theta_Z(t)\overrightarrow{w}(t, z = Z), \end{array} \tag{3.34}$$

  where

$$\sigma_0(t) = \frac{2\mathscr{I}_\flat(0) q_0(t)}{1 + q_0(t)/a}, \qquad \theta_0(t) = \frac{1 - q_0(t)/a}{1 + q_0(t)/a}, \tag{3.35a}$$
$$\sigma_Z(t) = 2p_X(t), \qquad \theta_Z(t) = -1. \tag{3.35b}$$

Proof. Equation (3.30c) implies that $m = \rho J$ is a function of χ alone, and it coincides with its initial value, i.e., $m = \rho_\flat J_\flat = \rho^n$. Factoring m out of the time derivatives in the remaining equations of (3.30), dividing each equation by $m > 0$, and using $dz = m(\chi)d\chi$, we end up with

$$\begin{array}{ll} \text{(a)} \quad \partial_t Y = 0; & \text{(c)} \quad \partial_t \tau - \partial_z u = 0; \\ \text{(b)} \quad \partial_t \zeta = 0; & \text{(d)} \quad \partial_t u + \partial_z \Pi = 0. \end{array} \tag{3.36}$$

where we recall that $\Pi = P(\zeta, Y) + a^2(\zeta - \tau)$. This system, in which z appears as the Lagrangian mass-coordinate [34], can be shown to be hyperbolic with eigenvalues $\pm a$ and 0 (double), all of them being linearly degenerate fields. Hence, it is

equivalent to

$$\partial_t \tau - \partial_z u = 0, \tag{3.37a}$$

$$\partial_t Y = 0, \tag{3.37b}$$

$$\partial_t u + \partial_z \Pi = 0, \tag{3.37c}$$

$$\partial_t \Pi + a^2 \partial_z u = 0, \tag{3.37d}$$

from which (3.33) follows. On the other hand, the boundary conditions (3.31) can be rewritten as

(3.38)

(a) $Y(t, z = 0) = g_0(t)/q_0(t)$; (b) $u(t, z = 0) = q_0(t)\tau(t, z = 0)$;
(c) $Y(t, z = Z) = Y_X(t)$; (d) $\Pi(t, z = Z) = p_X(t)$.

Substituting the inverse transformation

$$\Pi = \tfrac{1}{2}(\overrightarrow{w} + \overleftarrow{w}), \quad u = \tfrac{1}{2a}(\overrightarrow{w} - \overleftarrow{w}), \quad \tau = \tfrac{1}{2a^2}[2\mathscr{I} - (\overrightarrow{w} + \overleftarrow{w})] \tag{3.39}$$

into (3.38) and invoking $\mathscr{I}(t, z = 0) = \mathscr{I}_\flat(0)$ yield (3.34)–(3.35). □

3.3.2. *Projection step.* The outcome of the fast Lagrange step, denoted by $(J^{n\sharp}, \mathbb{U}^{n\sharp})$ in (3.29), is now the input data for the slow projection step. The latter amounts to solving

$$\partial_t(J) + \partial_\chi(v) = 0, \tag{3.40a}$$

$$\partial_t(\rho Y J) + \partial_\chi(\rho Y v) = 0, \tag{3.40b}$$

$$\partial_t(\rho J) + \partial_\chi(\rho v) = 0, \tag{3.40c}$$

$$\partial_t(\rho u J) + \partial_\chi(\rho u v) = 0, \tag{3.40d}$$

$$\partial_t(\rho \zeta J) + \partial_\chi(\rho \zeta v) = 0, \tag{3.40e}$$

where v is a given velocity field. Comparing the evolution equations (3.40a) and (3.30a) for J, we see that in order for J to go back to its initial value 1, we have to take $v = u$. For the moment, it is not obvious as to how we can achieve this, but things will become clearer at the fully discrete level. Taking $v = u$ for granted and writing the system (3.40) under the condensed form

$$\partial_t(J) + \partial_\chi(u) = 0, \tag{3.41a}$$

$$\partial_t(\mathbb{U} J) + \partial_\chi(\mathbb{U} u) = 0, \tag{3.41b}$$

we can combine the equations to obtain the componentwise advection equation

$$\partial_t \mathbb{U} + \frac{u}{J} \partial_x \mathbb{U} = 0. \tag{3.42}$$

Thus, the projection step is merely a remap of the variables contained in $\mathbb{U}$.

4. Two-phase flow model: the numerical scheme

The connection made by Theorem 3.1 between the Lagrange step and *Problem (SA)* opens up the possibility of us applying the scheme considered in *Problem* $(SA)^n_N$.

4.1. **Updating formulae.** We divide the domain $[0, X]$ into N cells $[x_{j-1/2}, x_{j+1/2}]$ of size $\Delta x = X/N$. The inner cells are numbered from 1 to N. We also define two ghost cells labeled 0 and $N + 1$.

4.1.1. *For the Lagrange step.* At the beginning of each time step $n \to n\sharp$, the variables χ and x coincide with each, so that we can identify them. Since all data, including ρ^n, are assumed to be piecewise constant, the local step-size of the mass-coordinate z is

$$\Delta z_i = \rho_i^n \Delta x. \tag{4.1}$$

To update $(\overrightarrow{w}, \overleftarrow{w})$ in (3.33)–(3.34), we use formulae (2.18)–(2.20). Updating $(Y, \mathscr{I})$ inside the domain is easy, since $\partial_t Y = \partial_t \mathscr{I} = 0$. As for $(Y, \mathscr{I})$ at the boundaries, we need to specify two more conditions at each ghost cell, as indicated in (4.3a) and (4.4a) below. Note that

- the "wave-cancellation" conditions for $\mathscr{I}$ are justified by the fact that the $\mathscr{I}$-wave, artificially created by the relaxation model, has no real physical meaning;
- the "mass-fraction" conditions for Y do not conflict with the evolution equation $\partial_t Y = 0$, since the latter is valid only for inner points.

To summarize, the comprehensive set of equations for the Lagrange step is:

- For $1 \le i \le N$,

$$\frac{Y_i^{n\sharp} - Y_i^n}{\Delta t} = 0, \tag{4.2a}$$

$$\frac{\mathscr{I}_i^{n\sharp} - \mathscr{I}_i^n}{\Delta t} = 0, \tag{4.2b}$$

$$\frac{\overrightarrow{w}_i^{n\sharp} - \overrightarrow{w}_i^n}{\Delta t} + a\frac{\overrightarrow{w}_i^{n\sharp} - \overrightarrow{w}_{i-1}^{n\sharp}}{\Delta z_i} = 0, \tag{4.2c}$$

$$\frac{\overleftarrow{w}_i^{n\sharp} - \overleftarrow{w}_i^n}{\Delta t} - a\frac{\overleftarrow{w}_{i+1}^{n\sharp} - \overleftarrow{w}_i^{n\sharp}}{\Delta z_i} = 0. \tag{4.2d}$$

- For $i = 0$,

$$Y_0^{n\sharp} = g_0^n/q_0^n, \qquad \mathscr{I}_0^{n\sharp} = \mathscr{I}_1^{n\sharp}, \tag{4.3a}$$

$$\overleftarrow{w}_0^{n\sharp} = \overleftarrow{w}_1^{n\sharp}, \qquad \overrightarrow{w}_0^{n\sharp} = \sigma_0^n + \theta_0^n \overleftarrow{w}_0^{n\sharp}, \tag{4.3b}$$

with $\sigma_0^n = \dfrac{2q_0^n/a}{1 + q_0^n/a}\mathscr{I}_1^n$ and $\theta_0^n = \dfrac{1 - q_0^n/a}{1 + q_0^n/a}$.

- For $i = N + 1$,

$$Y_{N+1}^{n\sharp} = Y_X^n, \qquad \mathscr{I}_{N+1}^{n\sharp} = \mathscr{I}_N^{n\sharp}, \tag{4.4a}$$

$$\overrightarrow{w}_{N+1}^{n\sharp} = \overrightarrow{w}_N^{n\sharp}, \qquad \overleftarrow{w}_{N+1}^{n\sharp} = \sigma_Z^n + \theta_Z^n \overrightarrow{w}_{N+1}^{n\sharp}, \tag{4.4b}$$

with $\sigma_Z^n = 2p_X^n$ and $\theta_Z^n = -1$.

To gain more insight into this scheme, it is helpful to rewrite it in terms of the original variables. After some algebra, we see that (4.2) is equivalent to

$$\rho_i^n \frac{Y_i^{n\sharp} - Y_i^n}{\Delta t} = 0, \tag{4.5a}$$

$$\rho_i^n \frac{\tau_i^{n\sharp} - \tau_i^n}{\Delta t} - \frac{\widetilde{u}_{i+1/2}^{n\sharp} - \widetilde{u}_{i-1/2}^{n\sharp}}{\Delta x} = 0, \tag{4.5b}$$

$$\rho_i^n \frac{u_i^{n\sharp} - u_i^n}{\Delta t} + \frac{\widetilde{\Pi}_{i+1/2}^{n\sharp} - \widetilde{\Pi}_{i-1/2}^{n\sharp}}{\Delta x} = 0, \tag{4.5c}$$

$$\rho_i^n \frac{\zeta_i^{n\sharp} - \zeta_i^n}{\Delta t} = 0, \tag{4.5d}$$

where

$$\widetilde{\Pi}_{i+1/2}^{n\sharp} = \tfrac{1}{2}(\Pi_j^{n\sharp} + \Pi_{j+1}^{n\sharp}) - \tfrac{a}{2}(u_{j+1}^{n\sharp} - u_j^{n\sharp}), \tag{4.6a}$$

$$\widetilde{u}_{i+1/2}^{n\sharp} = \tfrac{1}{2}(u_j^{n\sharp} + u_{j+1}^{n\sharp}) - \tfrac{1}{2a}(\Pi_{j+1}^{n\sharp} - \Pi_j^{n\sharp}) \tag{4.6b}$$

appear to be the pressure and the velocity of the solution to the Riemann problem associated with (3.37) at the interface $i+1/2$. A straightforward calculation shows that we can replace (4.5d) by

$$\rho_i^n \frac{\Pi_j^{n\sharp} - \Pi_j^n}{\Delta t} + a^2 \frac{\widetilde{u}_{i+1/2}^{n\sharp} - \widetilde{u}_{i-1/2}^{n\sharp}}{\Delta x} = 0 \tag{4.7}$$

so as to be able to work with Π as a full-fledged variable. Since $J_i^n = 1$, equation (4.5b) can still be interpreted as

$$\frac{J_i^{n\sharp} - J_i^n}{\Delta t} - \frac{\widetilde{u}_{i+1/2}^{n\sharp} - \widetilde{u}_{i-1/2}^{n\sharp}}{\Delta x} = 0, \tag{4.8}$$

which is the discrete version of (3.30a). Following the widely adopted terminology in continuum mechanics (see [12] for a mathematical presentation), we shall refer to (4.8) as Piola's identity. If, in (4.5), we replace (4.5b) with $(\rho J)_i^{n\sharp} = \rho_i^n$, then the new system can be condensed under the conservative form

$$\frac{(\mathbb{U}J)_i^{n\sharp} - (\mathbb{U}J)_i^n}{\Delta t} + \frac{\mathbb{A}_{i+1/2}^{n\sharp} - \mathbb{A}_{i-1/2}^{n\sharp}}{\Delta x} = 0, \tag{4.9}$$

where $\mathbb{A}_{i+1/2}^{n\sharp} = (0, 0, \widetilde{\Pi}_{i+1/2}^{n\sharp}, 0)$ denotes the acoustic part of the flux. For later use, we write $\mathbf{A}_{i+1/2}^{n\sharp} = (0, 0, \widetilde{\Pi}_{i+1/2}^{n\sharp})$.

Remark 4.1. In the pure Eulerian setting of [4] and within the frame of an implicit time integration, Baudin et al. strongly recommended handling the discrete version of the relaxation equation (3.16d) in the limit $\lambda \to +\infty$. In contrast, the solution procedure proposed here seems to rely on the choice $\lambda = 0$ as advocated by formulae (3.33). Let us stress, however, that no contradiction arises with [4]. Had we discretized the last equation (3.30e) by the consistent approximation

$$\rho_i^n \frac{\zeta_i^{n\sharp} - \zeta_i^n}{\Delta t} = \lambda \rho_i^n (\tau_i^n - \zeta_i^{n\sharp}), \tag{4.10}$$

then for any $\lambda \geq 0$, we would have obtained the expected value (4.5d)

$$\zeta_i^{n\sharp} = \tau_i^n \tag{4.11}$$

because at time n, the variable ζ is at equilibrium, i.e., $\zeta_i^n = \tau_i^n$. This algebraic miracle occurs solely in Lagrangian coordinates.

4.1.2. *For the projection step.* Piola's identity (4.8) clearly shows that, at the discrete level, we have to use the velocity field $v_{i+1/2} = \widetilde{u}_{i+1/2}^{n\sharp}$, defined at the interfaces, to remap the variables. More concretely, we have to discretize (3.41) by

$$\frac{J_i^{n\mathsf{X}} - J_i^{n\sharp}}{\Delta t} + \frac{\widetilde{u}_{i+1/2}^{n\sharp} - \widetilde{u}_{i-1/2}^{n\sharp}}{\Delta x} = 0, \tag{4.12a}$$

$$\frac{(\mathbb{U}J)_i^{n\mathsf{X}} - (\mathbb{U}J)_i^{n\sharp}}{\Delta t} + \frac{(\mathbb{U}\widetilde{u})_{i+1/2}^{n\sharp} - (\mathbb{U}\widetilde{u})_{i-1/2}^{n\sharp}}{\Delta x} = 0, \tag{4.12b}$$

the product $(\mathbb{U}\widetilde{u})_{i+1/2}^{n\sharp}$ being upwinded as

$$(\mathbb{U}\widetilde{u})_{i+1/2}^{n\sharp} = \mathbb{U}_i^{n\sharp}(\widetilde{u}_{i+1/2}^{n\sharp})^+ + \mathbb{U}_{i+1}^{n\sharp}(\widetilde{u}_{i+1/2}^{n\sharp})^-, \tag{4.13}$$

where u^+ (respectively u^-) stands for the positive (resp. negative) part of u. Note that $\widetilde{\Pi}_{i+1/2}^{n\sharp}$ and $\widetilde{u}_{i+1/2}^{n\sharp}$ are byproducts of the Lagrange step and can be computed as

$$\widetilde{\Pi}_{i+1/2}^{n\sharp} = \tfrac{1}{2}(\overrightarrow{w}_i^{n\sharp} + \overleftarrow{w}_{i+1}^{n\sharp}), \qquad \widetilde{u}_{i+1/2}^{n\sharp} = \tfrac{1}{2a}(\overrightarrow{w}_i^{n\sharp} - \overleftarrow{w}_{i+1}^{n\sharp}). \tag{4.14}$$

To better understand this projection step, let us multiply (4.8) by $\mathbb{U}_i^{n\sharp}$ and add it to (4.12b). Arguing that $J^{n\mathsf{X}} = 1$, according to (4.12a), we have

$$\frac{\mathbb{U}_i^{n\mathsf{X}} - \mathbb{U}_i^{n\sharp}}{\Delta t} + (\widetilde{u}_{i-1/2}^{n\sharp})^+ \frac{\mathbb{U}_i^{n\sharp} - \mathbb{U}_{i-1}^{n\sharp}}{\Delta x} + (\widetilde{u}_{i+1/2}^{n\sharp})^- \frac{\mathbb{U}_{i+1}^{n\sharp} - \mathbb{U}_i^{n\sharp}}{\Delta x} = 0 \tag{4.15}$$

after some cancellations. Undoubtedly, this is a first-order explicit discretization of (3.42), where J has been "implicit" to $J^{n\mathsf{X}}$. Let us introduce the algebraic CFL ratios

$$\lambda_{i+1/2} = \frac{\widetilde{u}_{i+1/2}^{n\sharp}\Delta t}{\Delta x} \tag{4.16}$$

based on the transport velocities. Then, equation (4.15) becomes

$$\mathbb{U}_i^{n\mathsf{X}} = \lambda_{i-1/2}^+ \mathbb{U}_{i-1}^{n\sharp} + (1 - \lambda_{i-1/2}^+ + \lambda_{i+1/2}^-)\mathbb{U}_i^{n\sharp} - \lambda_{i+1/2}^- \mathbb{U}_{i+1}^{n\sharp}, \tag{4.17}$$

and we see that a CFL-like condition should be imposed on Δt so that the right-hand side of (4.17) is a convex combination. This is the objective of the next subsection.

4.2. **Positivity, stability and energy properties.** The novelty we wish to put forward lies in the guarantee of positivity, stability and energy dissipation, as stated in the following theorem.

Theorem 4.1. *The overall scheme* (3.27), (3.29) *has the following properties:*

(1) *It can be expressed as the locally conservative form*

$$\frac{\mathbf{U}_i^{n+1} - \mathbf{U}_i^n}{\Delta t} + \frac{\mathbf{F}_{i+1/2}^{n\sharp} - \mathbf{F}_{i-1/2}^{n\sharp}}{\Delta x} = 0 \tag{4.18}$$

with $\mathbf{F}_{i+1/2}^{n\sharp} = \mathbf{U}_i^{n\sharp}(\widetilde{u}_{i+1/2}^{n\sharp})^+ + \mathbf{U}_{i+1}^{n\sharp}(\widetilde{u}_{i+1/2}^{n\sharp})^- + \mathbf{A}_{i+1/2}^{n\sharp}$.

(2) *Under the CFL constraint*

$$\frac{\Delta t}{\Delta x} < \frac{2a}{\max\limits_{1\le i\le N}\left\{(\overleftarrow{M}_i^{n\sharp} - \overrightarrow{m}_{i+1}^{n\sharp})^+ - (\overleftarrow{m}_i^{n\sharp} - \overrightarrow{M}_{i+1}^{n\sharp})^-\right\}}, \tag{4.19}$$

where the various $\overrightarrow{M}, \overleftarrow{M}, \overrightarrow{m}, \overleftarrow{m}$*'s, defined by* (2.33)–(2.34) *of Proposition* 2.2*, are explicitly computable from data at time* n *and do not depend on* Δt*, we have*

$$\rho_i^{n+1} > 0 \qquad \text{and} \qquad Y_i^{n+1} \in [0,1]. \tag{4.20}$$

(3) *Under the CFL restriction* (4.19)*, there is the min-max principle*

$$\min\{Y_{i-1}^n,\, Y_i^n,\, Y_{i+1}^n\} \le Y_i^{n+1} \le \max\{Y_{i-1}^n,\, Y_i^n,\, Y_{i+1}^n\}. \tag{4.21}$$

(4) *Under the CFL restriction* (4.19) *and the subcharacteristic condition*

$$a^2 > \max_{i\in\{1,\dots,N\}} \max_{\sigma\in[0,1]} \{-P_\tau(\sigma\tau_i^n + (1-\sigma)\tau_i^{n\sharp}, Y_i^n)\}, \tag{4.22}$$

we have

$$\frac{\{\rho\mathfrak{E}\}(\mathbf{U}_i^{n+1}) - \{\rho\mathfrak{E}\}(\mathbf{U}_i^n)}{\Delta t} + \frac{(\rho\mathfrak{E}\widetilde{u} + \widetilde{\Pi}\widetilde{u})_{i+1/2}^{n\sharp} - (\rho\mathfrak{E}\widetilde{u} + \widetilde{\Pi}\widetilde{u})_{i-1/2}^{n\sharp}}{\Delta x} \le 0. \tag{4.23}$$

This discrete energy inequality is consistent with (3.6).

(5) *Stationary contact discontinuities are preserved exactly.*

To our knowledge, the stability results mentioned above seem to be new for a time implicit approximation of the solutions of the Euler's IBVP. This is why Theorem 4.1 deserves our attention. Before proving this theorem, we wish to make two comments.

First, the CFL restriction (4.19) results from enforcing the validity of the estimate

$$\frac{\Delta t}{\Delta x}[(\widetilde{u}_{i-1/2}^{n\sharp})^+ - (\widetilde{u}_{i+1/2}^{n\sharp})^-] < 1, \tag{4.24}$$

which is nothing but a CFL condition based on the intermediate wave velocity u. Such a condition is expected, precisely because the proposed scheme is time-explicit with respect to this wave. Numerical benchmarks testify that the estimate (4.19) actually provides a sharp lower-bound of the time step Δt dictated by the "exact" condition (4.24).

Second, the subcharacteristic condition (4.22) reads the same as that for a fully time explicit setting [6]. In this respect, the sharp version (4.22) of the Whitham condition (3.22) is quite natural.

Now, let us turn to the proof. The derivation of the energy inequality (4.23) relies on the following preliminary result.

Lemma 4.1. *Assume the subcharacteristic condition* (4.22) *is met. Then, the solution of the Lagrange step satisfies the energy inequality*

$$\rho_i^n \frac{\mathfrak{E}(\mathbf{U}_i^{n\sharp}) - \mathfrak{E}(\mathbf{U}_i^n)}{\Delta t} + \frac{(\widetilde{\Pi}\widetilde{u})_{i+1/2}^{n\sharp} - (\widetilde{\Pi}\widetilde{u})_{i-1/2}^{n\sharp}}{\Delta x} \le 0, \tag{4.25}$$

where $\mathfrak{E}$ *is defined in* (3.7)*, and* $(\widetilde{\Pi}_{i+1/2}^{n\sharp}, \widetilde{u}_{i+1/2}^{n\sharp})$ *by* (4.14).

Observe that the proposed discrete inequality is nothing but a consistent approximation of the energy inequality (3.6) expressed in Lagrangian coordinates

$$\partial_t(\rho\mathfrak{E}J) + \partial_\chi(Pu) \le 0. \tag{4.26}$$

Proof of Theorem 4.1. *Locally conservative form.* Adding (4.9) and (4.12b), we get

$$\frac{\mathbb{U}_i^{n\chi} - \mathbb{U}_i^n}{\Delta t} + \frac{\mathbb{F}_{i+1/2}^{n\sharp} - \mathbb{F}_{i-1/2}^{n\sharp}}{\Delta x} = 0, \tag{4.27}$$

with $\mathbb{F}_{i+1/2}^{n\sharp} = \mathbb{U}_i^{n\sharp}(\widetilde{u}_{i+1/2}^{n\sharp})^+ + \mathbb{U}_{i+1}^{n\sharp}(\widetilde{u}_{i+1/2}^{n\sharp})^- + \mathbb{A}_{i+1/2}^{n\sharp}$. Extract the first three components of (4.27) to have (4.18).

Positivity for density and gas mass-fraction. Since $(\rho J)_i^{n\sharp} = \rho_i^n$, we have $\rho_i^{n\sharp} > 0$ as soon as $J_i^{n\sharp} > 0$. By virtue of Piola's identity (4.8), we must ask for

$$\frac{\Delta t}{\Delta x}[\widetilde{u}_{i-1/2}^{n\sharp} - \widetilde{u}_{i+1/2}^{n\sharp}] < 1. \tag{4.28}$$

From (4.17), we see that the estimate $\rho_i^{n\sharp} > 0$ implies $\rho_i^{n\chi} > 0$ as soon as the combination in the right-hand side is convex. It suffices that $1-\lambda_{i-1/2}^+ + \lambda_{i+1/2}^- > 0$, that is,

$$\frac{\Delta t}{\Delta x}[(\widetilde{u}_{i-1/2}^{n\sharp})^+ - (\widetilde{u}_{i+1/2}^{n\sharp})^-] < 1. \tag{4.29}$$

Obviously, (4.29) is stronger than (4.28), therefore we just have to focus on (4.29). Thanks to (4.14) and to Proposition 2.2, we have

$$\tfrac{1}{2a}(\overleftarrow{m}_j^{n\sharp} - \overrightarrow{M}_{j+1}^{n\sharp}) \le \widetilde{u}_{i+1/2}^{n\sharp} \le \tfrac{1}{2a}(\overleftarrow{M}_j^{n\sharp} - \overrightarrow{m}_{j+1}^{n\sharp}). \tag{4.30}$$

Consequently,

$$(\widetilde{u}_{i-1/2}^{n\sharp})^+ - (\widetilde{u}_{i+1/2}^{n\sharp})^- \le \tfrac{1}{2a}[(\overleftarrow{M}_{j-1}^{n\sharp} - \overrightarrow{m}_j^{n\sharp})^+ - (\overleftarrow{m}_j^{n\sharp} - \overrightarrow{M}_{j+1}^{n\sharp})^-], \tag{4.31}$$

hence the sufficient condition (4.19) to ensure $\rho_i^{n\chi} = \rho_i^{n+1} > 0$.

Min-max principle. In (4.17), we subtract the second equation, multiplied by any constant A, to the first equation to obtain

$$\begin{aligned}\rho_j^{n\chi}(Y_j^{n\chi} - A) = {}& \lambda_{i-1/2}^+\rho_{j-1}^{n\sharp}(Y_{j-1}^{n\sharp} - A) - \lambda_{i+1/2}^-\rho_{j+1}^{n\sharp}(Y_{j+1}^{n\sharp} - A)\\ &+ [1 - \lambda_{i-1/2}^+ + \lambda_{i+1/2}^-]\rho_j^{n\sharp}(Y_j^{n\sharp} - A).\end{aligned} \tag{4.32}$$

Again, $Y^{n\sharp} = Y^n$. Now by selecting $A = \max\{Y_{i-1}^n, Y_i^n, Y_{i+1}^n\}$, then $A = \min\{Y_{i-1}^n, Y_i^n, Y_{i+1}^n\}$, and discussing the signs, we obtain (4.21). This implies $Y_i^{n+1} \in [0,1]$.

Energy inequality. Leaving out the last component of (4.17), we may write

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^{n\chi} = \lambda_{i-1/2}^+\mathbf{U}_{i-1}^{n\sharp} + (1 - \lambda_{i-1/2}^+ + \lambda_{i+1/2}^-)\mathbf{U}_i^{n\sharp} - \lambda_{i+1/2}^-\mathbf{U}_{i+1}^{n\sharp}, \tag{4.33}$$

which is a convex combination under constraint (4.19). By Jensen's inequality, applied to the convex function $\mathbf{U} \mapsto \{\rho\mathfrak{E}\}(\mathbf{U})$, we infer

$$(\rho\mathfrak{E})_i^{n+1} \le \lambda_{i-1/2}^+(\rho\mathfrak{E})_{i-1}^{n\sharp} + [1 - \lambda_{i-1/2}^+ + \lambda_{i+1/2}^-](\rho\mathfrak{E})_i^{n\sharp} - \lambda_{i+1/2}^-(\rho\mathfrak{E})_{i+1}^{n\sharp}. \tag{4.34}$$

However, by construction

$$1 - \lambda_{i-1/2}^+ + \lambda_{i+1/2}^- = J_i^{n\sharp} + (\lambda_{i-1/2}^- - \lambda_{i+1/2}^+). \tag{4.35}$$

As a result, the inequality (4.34) becomes

$$(\rho\mathfrak{E})_i^{n+1} \leq \rho_i^n \mathfrak{E}_i^{n\sharp} - \frac{\Delta t}{\Delta x}[(\rho\mathfrak{E}\widetilde{u})_{i+1/2}^{n\sharp} - (\rho\mathfrak{E}\widetilde{u})_{i-1/2}^{n\sharp}], \tag{4.36}$$

again with the notation

$$(\rho\mathfrak{E}\widetilde{u})_{i+1/2}^{n\sharp} = (\rho\mathfrak{E})_i^{n\sharp}(\widetilde{u}_{i+1/2}^{n\sharp})^+ + (\rho\mathfrak{E})_{i+1}^{n\sharp}(\widetilde{u}_{i+1/2}^{n\sharp})^- \tag{4.37}$$

for the upwinded product. According to Lemma 4.1,

$$\rho_i^n \mathfrak{E}_i^{n\sharp} \leq (\rho\mathfrak{E})_i^n - \frac{\Delta t}{\Delta x}[(\widetilde{\Pi}\widetilde{u})_{i+1/2}^{n\sharp} - (\widetilde{\Pi}\widetilde{u})_{i-1/2}^{n\sharp}]. \tag{4.38}$$

Inserting (4.38) into the right-hand sides of (4.36) leads to (4.23).

Preservation of steady contact discontinuities. It is clear that the equivalent form (4.5)–(4.6) comes from a stationary contact discontinuity (say, at time n)

$$\rho_i^n, \quad u_i^n = 0, \quad P_i^n = P^\star, \quad i \in \{1, ..., N\},$$

the Lagrangian updated values

$$\rho_i^{n\sharp} = \rho_i^n, \quad u_i^{n\sharp} = 0, \quad P_i^{n\sharp} = P^\star, \quad i \in \{1, ..., N\},$$

namely $\widetilde{u}_{i+1/2}^{n\sharp} = 0$ and $\widetilde{\Pi}_{i+1/2}^{n\sharp} = P^\star$ so that the Eulerian projection step ends up with

$$(\rho Y)_i^{n+1} = (\rho Y)_i^n, \quad \rho_i^{n+1} = \rho_i^n, \quad (\rho u)_i^{n+1} = 0. \tag{4.39}$$

In other words, steady contact discontinuities are preserved exactly. □

Proof of Lemma 4.1. Let us use Theorem 2.3 with $S(w) = \dfrac{w^2}{4a^2}$ in order to get

$$\frac{\mathscr{S}_i^{n\sharp} - \mathscr{S}_i^n}{\Delta t} + \frac{\mathscr{H}_{i+1/2}^{n\sharp} - \mathscr{H}_{i-1/2}^{n\sharp}}{\rho_i^n \Delta x} \leq 0, \tag{4.40}$$

with

$$\mathscr{S}_i^n = \tfrac{1}{2}[u^2 + (\Pi/a)^2]_i^n, \quad \mathscr{S}_i^{n\sharp} = \tfrac{1}{2}[u^2 + (\Pi/a)^2]_i^{n\sharp}, \quad \mathscr{H}_{i+1/2}^{n\sharp} = (\widetilde{\Pi}\widetilde{u})_{i+1/2}^{n\sharp}. \tag{4.41}$$

Since $\mathscr{S} = \mathfrak{E} - \varepsilon + \frac{1}{2}(\Pi/a)^2$, equation (4.40) can be cast into

$$\rho_i^n \frac{\mathfrak{E}_i^{n\sharp} - \mathfrak{E}_i^n}{\Delta t} + \frac{(\widetilde{\Pi}\widetilde{u})_{i+1/2}^{n\sharp} - (\widetilde{\Pi}\widetilde{u})_{i-1/2}^{n\sharp}}{\Delta x} \leq \rho_i^n \mathscr{R}_i^{n\sharp}, \tag{4.42}$$

where

$$\mathscr{R}_i^{n\sharp} = \varepsilon_i^{n\sharp} - \varepsilon_i^n - \frac{1}{2a^2}[(\Pi_j^{n\sharp})^2 - (\Pi_j^n)^2] \tag{4.43}$$

and $\varepsilon_i^{n\sharp} = \varepsilon(\tau_i^{n\sharp}, Y_i^{n\sharp}) = \varepsilon(\tau_i^{n\sharp}, Y_i^n)$. Since the relaxation system is brought back to equilibrium at each time step, we have $\Pi_j^n = P(\tau_j^n, Y_j^n) = P_j^n$. Therefore, we can rewrite the previous equation as

$$\mathscr{R}_i^{n\sharp} = \varepsilon_i^{n\sharp} - \varepsilon_i^n - \frac{1}{a^2} P_i^n (\Pi_j^{n\sharp} - P_i^n) - \frac{1}{2a^2}(\Pi_i^{n\sharp} - P_i^n)^2. \tag{4.44}$$

Because of $\zeta_i^{n\sharp} = \tau_i^n$, as shown in (4.11), we have

$$\Pi_j^{n\sharp} - P_j^n = -a^2(\tau_j^{n\sharp} - \tau_j^n). \tag{4.45}$$

Consequently, (4.44) becomes

$$\mathscr{R}_i^{n\sharp} = \varepsilon_i^{n\sharp} - \varepsilon_i^n + P_i^n(\tau_i^{n\sharp} - \tau_i^n) - \tfrac{1}{2}a^2(\tau_i^{n\sharp} - \tau_i^n)^2. \tag{4.46}$$

Resorting to the Taylor expansion with integral remainder

(4.47)
$$\varepsilon(\tau_i^{n\sharp},Y_i^n)-\varepsilon(\tau_i^n,Y_i^n)-\partial_\tau\varepsilon(\tau_i^n,Y_i^n)(\tau_i^{n\sharp}-\tau_i^n)=\int_{\tau_j^n}^{\tau_j^{n\sharp}}\partial_{\tau\tau}\varepsilon(\varsigma,Y_i^n)(\tau_i^{n\sharp}-\varsigma)\,d\varsigma,$$

we can easily derive

$$\mathscr{R}_i^{n\sharp}=(\tau_i^{n\sharp}-\tau_i^n)^2\int_0^1[\partial_{\tau\tau}\varepsilon(\sigma\tau_i^n+(1-\sigma)\tau_i^{n\sharp},Y_i^n)-a^2](1-\sigma)\,d\sigma. \tag{4.48}$$

This quantity is negative if a^2 is chosen large enough, in compliance with the subcharacteristic condition (4.22). □

5. Euler's standard single-phase model: an easy extension

5.1. **The continuous problem.** This section deals with the Euler equations for real compressible materials governed by an *internal energy* $(\tau,s)\in\mathbb{R}_+^*\times\mathbb{R}_+\mapsto\varepsilon(\tau,s)\in\mathbb{R}_+$. This function ε is assumed to be smooth enough and to satisfy Weyl's conditions [35]

$$\begin{array}{lll}\text{(a)}\quad \varepsilon>0; & \text{(b)}\quad \varepsilon_\tau<0; & \text{(c)}\quad \varepsilon_{\tau\tau}>0;\\ \text{(d)}\quad \varepsilon_{\tau\tau\tau}<0; & \text{(e)}\quad \varepsilon_{\tau\tau}\varepsilon_{ss}>(\varepsilon_{\tau s})^2; & \text{(f)}\quad \varepsilon_s<0.\end{array} \tag{5.1}$$

From the internal energy ε, we define

$$\text{the pressure}\quad P(\tau,s)=-\varepsilon_\tau(\tau,s), \tag{5.2a}$$

$$\text{the sound speed}\quad c(\tau,s)=\tau\sqrt{\varepsilon_{\tau\tau}(\tau,s)}=\tau\sqrt{-P_\tau(\tau,s)}, \tag{5.2b}$$

$$\text{the temperature}\quad \Theta(\tau,s)=-\varepsilon_s(\tau,s). \tag{5.2c}$$

Here, τ still denotes the specific volume while s stands for the specific entropy. Comparing (5.1) to (3.1) using the formal identification $s\equiv Y$, we see that the conditions for *Problem (EU)* are more stringent than those for *Problem (TP)*: here, we have to require the temperature to be positive. The strict monotonicity property (5.1f) also enables us to define

$$(\tau,\varepsilon)\mapsto s(\tau,\varepsilon)\text{ as the inverse function of }(\tau,s)\mapsto\varepsilon(\tau,s). \tag{5.3}$$

The fact of paramount importance is that this inverse function s decreases with respect to ε, as

$$s_\varepsilon=1/\varepsilon_s=-1/\Theta<0. \tag{5.4}$$

To use notations from thermodynamics, we have $\mathrm{d}\varepsilon=-P\mathrm{d}\tau-\Theta\mathrm{d}s$. Once the internal energy ε has been selected, we state the following IBVP over the natural phase space

$$\Omega_{\mathbf{V}}=\{\mathbf{V}=(\rho\mathfrak{E},\rho,\rho u)\in\mathbb{R}^3\ \mid\ \rho>0,\ u\in\mathbb{R},\ \varepsilon=\mathfrak{E}-\tfrac12u^2>0\}. \tag{5.5}$$

Problem (EU) Given

– the initial data $x\in[0,X]\mapsto \mathbf{V}_\flat(x)\in\Omega_{\mathbf{V}}$,
– the inlet data $t\in\mathbb{R}_+\mapsto q_0(t),\Theta_0(t)\in\mathbb{R}_+^2$,
– the outlet data $t\in\mathbb{R}_+\mapsto p_X(t)\in\mathbb{R}_+$.

Find

$$\mathbf{V} : (t,x) \in \mathbb{R}_+ \times [0,X] \mapsto \mathbf{V}(t,x) \in \Omega_{\mathbf{V}} \tag{5.6}$$

so as to satisfy (in the weak sense) the following conditions:

- for $(t,x) \in \mathbb{R}_+^* \times]0,X[$, the system of conservation laws

$$\partial_t(\rho\mathfrak{E}) + \partial_x(\rho\mathfrak{E}u + pu) = 0, \tag{5.7a}$$

$$\partial_t(\rho) \quad + \partial_x(\rho u) \qquad = 0, \tag{5.7b}$$

$$\partial_t(\rho u) + \partial_x(\rho u^2 + p) \quad = 0, \tag{5.7c}$$

  with $p = P(\rho^{-1}, s)$, where P is the pressure defined in (5.2a), and s is computed by (5.3), using $\varepsilon = \mathfrak{E} - \frac{1}{2}u^2$. We have intentionally put the energy balance (5.7a) in the first row in order to compare (5.7) with (3.5);
- for $(t,x) \in \mathbb{R}_+^* \times]0,X[$, the entropy inequality

$$\partial_t\{\rho s\}(\mathbf{V}) + \partial_x\{\rho s u\}(\mathbf{V}) \leq 0; \tag{5.8}$$

- for $x \in]0,X[$, the initial Cauchy conditions

$$\rho(t=0,x) = \rho_\flat(x), \quad u(t=0,x) = u_\flat(x), \quad \mathfrak{E}(t=0,x) = \mathfrak{E}_\flat(x); \tag{5.9}$$

- for $t \in \mathbb{R}_+$, the boundary relationships

$$\rho u(t, x=0) = q_0(t) \qquad \text{if } u(t,0) > -c(\rho^{-1}(t,0),\, s(t,0)), \tag{5.10a}$$

$$\Theta(t, x=0) = \Theta_0(t) \qquad \text{if } u(t,0) > 0, \tag{5.10b}$$

$$p(t, x=X) = p_X(t) \qquad \text{if } u(t,X) < c(\rho^{-1}(t,X),\, s(t,X)), \tag{5.10c}$$

  where c is the sound speed defined in (5.2b).

This problem is the usual Euler model for single-phase flows. It is well known [18] that smooth solutions of (5.7) obey the additional conservation law

$$\partial_t\{\rho s\}(\mathbf{V}) + \partial_x\{\rho s u\}(\mathbf{V}) = 0, \tag{5.11}$$

while discontinuous solutions of (5.7) are selected according to the entropy inequality (5.8). As far as the boundary conditions (5.10) are concerned, they are based on real-life operating modes. To shorten the notation, the PDE system (5.7) is given the clear condensed form

$$\partial_t \mathbf{V} + \partial_x \mathbf{G}(\mathbf{V}) = 0. \tag{5.12}$$

Now let us recapitulate the main properties that will be used later.

Proposition 5.1. *The system* (5.7) *is hyperbolic over* $\Omega_{\mathbf{V}}$*, i.e., for any state* $\mathbf{V} \in \Omega_{\mathbf{V}}$*, the Jacobian matrix* $\nabla_{\mathbf{V}}\mathbf{G}(\mathbf{V})$ *has real eigenvalues*

$$u - c(\mathbf{V}) < u < u + c(\mathbf{V}), \tag{5.13}$$

and is $\mathbb{R}$*-diagonalizable. The two extreme fields are genuinely nonlinear, while the intermediate field is linearly degenerate.*

Furthermore, the mapping $\mathbf{V} \in \Omega_{\mathbf{V}} \to \{\rho s\}(\mathbf{V}) \in \mathbb{R}$ *is strictly convex.*

Proof. See [18] for the details. □

We are going to design the evolution strategy in two steps, based on an idea introduced in [10]. First, we replace the energy-balance equation (5.7a) in the

PDE system by the entropy-balance equation (5.8), with the equal sign, namely, we consider weak solutions of the following auxiliary hyperbolic system

$$\partial_t(\rho s) + \partial_x(\rho s u) = 0, \tag{5.14a}$$

$$\partial_t(\rho) + \partial_x(\rho u) = 0, \tag{5.14b}$$

$$\partial_t(\rho u) + \partial_x(\rho u^2 + p) = 0, \tag{5.14c}$$

selected according to the natural energy inequality

$$\partial_t\{\rho\mathfrak{E}\}(\rho, \rho u, \rho s) + \partial_x\{(\rho\mathfrak{E} + p)u\}(\rho, \rho u, \rho s) \le 0. \tag{5.15}$$

Classical considerations [18] do prove the strict convexity of the mapping $(\rho, \rho u, \rho s) \to \{\rho\mathfrak{E}\}(\rho, \rho u, \rho s)$ from assumptions (5.1a), (5.1c), (5.1e). In other words, this mapping naturally yields an entropy for discriminating the physically relevant discontinuous solutions of (5.14).

The formal identification $Y \equiv s$ brings us back to *Problem (TP)*. After solving the new *Problem (TP)* (5.7) thanks to the relaxation/Lagrange-projection method proposed earlier, we obtain $\mathbf{U}^{n\ddagger} = (\rho s, \rho, \rho u)^{n\ddagger}$ with a discrete analog of the energy inequality (5.15), which we rewrite in a semi-discrete form to shorten the notation

$$\{\rho\mathfrak{E}\}(\rho^{n\ddagger}, (\rho u)^{n\ddagger}, (\rho s)^{n\ddagger}) \le (\rho\mathfrak{E})^n - \Delta t\partial_x(\rho\mathfrak{E}\widetilde{u} + \widetilde{\Pi}\widetilde{u})^{n\sharp}. \tag{5.16}$$

In order to enforce the validity of the conservation of the total energy at time $(n+1)$, we set $(\rho\mathfrak{E})^{n+1} = (\rho\mathfrak{E})^n - \Delta t\partial_x(\rho\mathfrak{E}\widetilde{u} + \widetilde{\Pi}\widetilde{u})^{n\sharp}$, while keeping the updated values of density and momentum unchanged. In other words, we choose $\rho^{n+1} = \rho^{n\ddagger}$ and $(\rho u)^{n+1} = (\rho u)^{n\ddagger}$. The procedure is shown in the following diagram. The reason why this process truly guarantees the entropy decay

$$(\rho s)^{n+1} \equiv \{\rho s\}(\rho^{n+1}, (\rho u)^{n+1}, (\rho\mathfrak{E})^{n+1}) \le (\rho s)^{n\ddagger} = (\rho s)^n - \Delta t\partial_x(\rho s\widetilde{u})^{n\sharp},$$

and hence the consistency with the expected entropy inequality (5.8) will be elaborated in the next subsection.

(5.17)

$$\begin{array}{ccccc}
\mathbf{V}^n = (\rho\mathfrak{E}, \rho, \rho u)^n & & \mathbf{V}^{n+1} = (\rho\mathfrak{E}, \rho, \rho u)^{n+1} & \Rightarrow (\rho s)^{n+1} \le (\rho s)^{n\ddagger} \\
\big\downarrow & & \mathfrak{E},s \,\big\uparrow \text{swap} & \\
\mathbf{U}^n = (\rho s, \rho, \rho u)^n & \xrightarrow[\text{with } Y\equiv s]{\text{Pb. } (TP)} & \mathbf{U}^{n\ddagger} = (\rho s, \rho, \rho u)^{n\ddagger} & \Rightarrow\ (\rho\mathfrak{E})^{n\ddagger} \le (\rho\mathfrak{E})^{n+1}
\end{array}$$

Remark 5.1. In comparison with *Problem (TP)*, there is a subtle difference regarding boundary conditions. In §3, we imposed the gas flow rate $\rho Y(t, x = 0)$ at the inlet, from which we deduced the incoming fraction $Y(t, x = 0)$. Here, we impose the temperature $\Theta(t, x = 0)$. By inverting the mapping $s \mapsto \Theta(\tau, s)$ at fixed τ (made possible thanks to $\Theta_s = -\varepsilon_{ss} < 0$), we obtain $s(t, x = 0)$ as a function of $\tau(t, x = 0)$ and $\Theta_0(t)$. Since τ is decoupled from s in the scheme, this enables us to proceed as if $s(t, x = 0)$ were known. At the outlet, we choose to leave $s(t, x = X)$ unspecified because in the applications, the velocity is expected to keep the constant sign $u(t, X) > 0$.

5.2. **The discrete scheme.** Step $\mathbf{U}^n \to \mathbf{U}^{n\ddagger}$ in (5.17) is of course performed via the scheme (3.27), (3.29), which consists of the two steps

$$\mathbf{U}^n \xrightarrow{\text{Lagrange}} \mathbf{U}^{n\sharp} \xrightarrow{\text{projection}} \mathbf{U}^{n\ddagger}. \tag{5.18}$$

Invoking once again the formal identification $Y \equiv s$, we can derive the formulae for the Lagrange step *mutatis-mutandis* from (4.2)–(4.4), which reads:

- For $1 \le i \le N$,

$$\frac{s_i^{n\sharp} - s_i^n}{\Delta t} = 0, \tag{5.19a}$$

$$\frac{\mathscr{I}_i^{n\sharp} - \mathscr{I}_i^n}{\Delta t} = 0, \tag{5.19b}$$

$$\frac{\overrightarrow{w}_i^{n\sharp} - \overrightarrow{w}_i^n}{\Delta t} + a\frac{\overrightarrow{w}_i^{n\sharp} - \overrightarrow{w}_{i-1}^{n\sharp}}{\Delta z_i} = 0, \tag{5.19c}$$

$$\frac{\overleftarrow{w}_i^{n\sharp} - \overleftarrow{w}_i^n}{\Delta t} - a\frac{\overleftarrow{w}_{i+1}^{n\sharp} - \overleftarrow{w}_i^{n\sharp}}{\Delta z_i} = 0. \tag{5.19d}$$

- For $i = 0$,

$$s_0^{n\sharp} = S(\tau_0^{n\sharp}, \Theta_0^n), \qquad \mathscr{I}_0^{n\sharp} = \mathscr{I}_1^{n\sharp}, \tag{5.20a}$$

$$\overleftarrow{w}_0^{n\sharp} = \overleftarrow{w}_1^{n\sharp}, \qquad \overrightarrow{w}_0^{n\sharp} = \sigma_0^n + \theta_0^n \overleftarrow{w}_0^{n\sharp}, \tag{5.20b}$$

  with $\sigma_0^n = \dfrac{2q_0^n/a}{1+q_0^n/a}\mathscr{I}_1^n$ and $\theta_0^n = \dfrac{1-q_0^n/a}{1+q_0^n/a}$.

- For $i = N+1$,

$$s_{N+1}^{n\sharp} = s_X^n, \qquad \mathscr{I}_{N+1}^{n\sharp} = \mathscr{I}_N^{n\sharp}, \tag{5.21a}$$

$$\overrightarrow{w}_{N+1}^{n\sharp} = \overrightarrow{w}_N^{n\sharp}, \qquad \overleftarrow{w}_{N+1}^{n\sharp} = \sigma_Z^n + \theta_Z^n \overrightarrow{w}_{N+1}^{n\sharp}, \tag{5.21b}$$

  with $\sigma_Z^n = 2p_X^n$ and $\theta_Z^n = -1$.

In (5.20), the function $\Theta \mapsto S(\tau, \Theta)$ is the inverse of the temperature function $s \mapsto \Theta(\tau, s)$ with respect to s, at fixed τ. The first variable in S is taken either at time n or at time $n\sharp$. This is not a difficulty in itself, since in view of the structure of the equations, the specific volume $\tau_0^{n\sharp}$ can be obtained before and independently of $s_0^{n\sharp}$. In this Lagrange step, the relaxation parameter a is assumed to satisfy a Whitham condition similar to (3.22), that is,

$$a^2 > -P_\tau(\zeta, s) \tag{5.22}$$

for all (ζ, s) under consideration.

Once $(s, \mathscr{I}, \overrightarrow{w}, \overleftarrow{w})_i^{n\sharp}$ has been converted to $(\mathbf{U}, \rho\zeta)_i^{n\sharp} = (\rho s, \rho, \rho u, \rho\zeta)_i^{n\sharp}$, the projection step is applied according to (4.12)–(4.14). Because we are interested only in the first three components, we are going to rewrite this step as

$$\frac{(\mathbf{U}J)_i^{n\ddagger} - (\mathbf{U}J)_i^{n\sharp}}{\Delta t} + \frac{(\mathbf{U}\widetilde{u})_{i+1/2}^{n\sharp} - (\mathbf{U}\widetilde{u})_{i-1/2}^{n\sharp}}{\Delta x} = 0, \tag{5.23}$$

the product $(\mathbf{U}\widetilde{u})_{i+1/2}^{n\sharp}$ being upwinded as

$$(\mathbf{U}\widetilde{u})_{i+1/2}^{n\sharp} = \mathbf{U}_i^{n\sharp}(\widetilde{u}_{i+1/2}^{n\sharp})^+ + \mathbf{U}_{i+1}^{n\sharp}(\widetilde{u}_{i+1/2}^{n\sharp})^-,$$

with $\widetilde{u}_{i+1/2}^{n\sharp} = \frac{1}{2a}(\overrightarrow{w}_i^{n\sharp} - \overleftarrow{w}_{i+1}^{n\sharp})$. So far, we have

$$(\rho s)_i^{n\ddagger} = (\rho s)_i^n - \frac{\Delta t}{\Delta x}[(\rho s\widetilde{u})_{i+1/2}^{n\sharp} - (\rho s\widetilde{u})_{i-1/2}^{n\sharp}], \tag{5.24a}$$

$$(\rho\mathfrak{E})_i^{n\ddagger} \le (\rho\mathfrak{E})_i^n - \frac{\Delta t}{\Delta x}[(\rho\mathfrak{E}\widetilde{u} + \widetilde{\Pi}\widetilde{u})_{i+1/2}^{n\sharp} - (\rho\mathfrak{E}\widetilde{u} + \widetilde{\Pi}\widetilde{u})_{i-1/2}^{n\sharp}] \tag{5.24b}$$

for all $1 \le i \le N$. By construction, the first equation gives the updated value for $(\rho s)_i^{n\sharp}$, whereas the second equation reflects the energy property of Theorem 4.1. Now, the "swap" step consists of ruling that

$$(\rho\mathfrak{E})_i^{n+1} = (\rho\mathfrak{E})_i^n - \frac{\Delta t}{\Delta x}[(\rho\mathfrak{E}\widetilde{u} + \widetilde{\Pi}\widetilde{u})_{i+1/2}^{n\sharp} - (\rho\mathfrak{E}\widetilde{u} + \widetilde{\Pi}\widetilde{u})_{i-1/2}^{n\sharp}], \tag{5.25a}$$

$$(\rho)_i^{n+1} = (\rho)_i^{n\sharp}, \tag{5.25b}$$

$$(\rho u)_i^{n+1} = (\rho u)_i^{n\sharp}, \tag{5.25c}$$

from which we deduce

$$(\rho s)_i^{n+1} = \rho_i^{n+1} s(\tau_i^{n+1}, \varepsilon_i^{n+1}) = \rho_i^{n+1} s(\tau_i^{n+1}, \mathfrak{E}_i^{n+1} - \tfrac{1}{2}(u_i^{n+1})^2). \tag{5.26}$$

Theorem 5.1. *The overall scheme* (5.17), (5.18) *has the following properties:*

(1) *It can be put under the locally conservative form*

$$\frac{\mathbf{V}_i^{n+1} - \mathbf{V}_i^n}{\Delta t} + \frac{\mathbf{G}_{i+1/2}^{n\sharp} - \mathbf{G}_{i-1/2}^{n\sharp}}{\Delta x} = 0 \tag{5.27}$$

with

$$\mathbf{G}_{i+1/2}^{n\sharp} = \mathbf{V}_i^{n\sharp}(\widetilde{u}_{i+1/2}^{n\sharp})^+ + \mathbf{V}_{i+1}^{n\sharp}(\widetilde{u}_{i+1/2}^{n\sharp})^- + \mathbf{B}_{i+1/2}^{n\sharp} \tag{5.28}$$

and $\mathbf{B}_{i+1/2}^{n\sharp} = (\widetilde{\Pi}_{i+1/2}^{n\sharp}\widetilde{u}_{i+1/2}^{n\sharp}, 0, \widetilde{\Pi}_{i+1/2}^{n\sharp})$.

(2) *Under the CFL constraint*

$$\frac{\Delta t}{\Delta x} < \frac{2a}{\max\limits_{1\le i\le N}\left\{(\overleftarrow{M}_i^{n\sharp} - \overrightarrow{m}_{i+1}^{n\sharp})^+ - (\overleftarrow{m}_i^{n\sharp} - \overrightarrow{M}_{i+1}^{n\sharp})^-\right\}}, \tag{5.29}$$

where $\overrightarrow{M}, \overleftarrow{M}, \overrightarrow{m}, \overleftarrow{m}$ *are defined by* (2.33)–(2.34) *of Proposition* 2.2, *we have*

$$\rho_i^{n+1} > 0 \qquad \text{and} \qquad \varepsilon_i^{n+1} > 0. \tag{5.30}$$

(3) *Under the CFL condition* (5.29), *there is the max principle*

$$s_i^{n+1} \le \max\{s_{i-1}^n,\, s_i^n,\, s_{i+1}^n\}. \tag{5.31}$$

(4) *Under the CFL condition* (5.29), *and the subcharacteristic condition*

$$a^2 > \max_{i\in\{1,\dots,N\}} \max_{\sigma\in[0,1]} \{-P_\tau(\sigma\tau_i^n + (1-\sigma)\tau_i^{n\sharp}, s_i^n)\}, \tag{5.32}$$

we have the entropy inequality

$$\frac{\{\rho s\}(\mathbf{V}_i^{n+1}) - \{\rho s\}(\mathbf{V}_i^n)}{\Delta t} + \frac{(\rho s\widetilde{u})_{i+1/2}^{n\sharp} - (\rho s\widetilde{u})_{i-1/2}^{n\sharp}}{\Delta x} \le 0, \tag{5.33}$$

which is consistent with (5.8).

(5) *Stationary contact discontinuities are preserved exactly.*

Proof. The locally conservative form is straightforward. The positivity of the density follows exactly the same steps as those developed in the previous section and gives rise to the CFL restriction (5.29). If we are able to prove that

$$\varepsilon_i^{n+1} \ge \varepsilon_i^{n\sharp} \qquad \text{and} \qquad s_i^{n+1} \le s_i^{n\sharp}, \tag{5.34}$$

then the remaining claims will follow suit, because

- by assumption (5.1a), $\varepsilon_i^{n\sharp} = \varepsilon(\tau_i^{n\sharp}, s_i^{n\sharp}) > 0$,
- by the min-max principle (4.21), $s_i^{n\sharp} \le \max\{s_{i-1}^n, s_i^n, s_{i+1}^n\}$,

- by definition, $(\rho s)_i^{n\ddagger} = (\rho s)_i^n - \frac{\Delta t}{\Delta x}[(\rho s\widetilde{u})_{i+1/2}^{n\sharp} - (\rho s\widetilde{u})_{i-1/2}^{n\sharp}]$.

To get (5.34), we first note that from (5.24a) and (5.25b), we have $(\rho\mathfrak{E})_i^{n+1} \geq (\rho\mathfrak{E})_i^{n\ddagger}$. Therefore,

$$\mathfrak{E}_i^{n+1} \geq \mathfrak{E}_i^{n\ddagger} \tag{5.35}$$

because $\rho_i^{n+1} = \rho_i^{n\ddagger}$. Since $\mathfrak{E} = \frac{1}{2}u^2 + \varepsilon$ and $u_i^{n+1} = u_i^{n\ddagger}$, we infer that $\varepsilon_i^{n+1} \geq \varepsilon_i^{n\ddagger}$. Now, as already shown in (5.4), s is decreasing with respect to ε at fixed τ. As a consequence,

$$s_i^{n+1} = s(\tau_i^{n+1}, \varepsilon_i^{n+1}) \leq s(\tau_i^{n+1}, \varepsilon_i^{n\ddagger}) = s(\tau_i^{n\ddagger}, \varepsilon_i^{n\ddagger}) = s_i^{n\ddagger}, \tag{5.36}$$

which completes the proof. □

6. Numerical results

The relaxation scheme using the Lagrange-projection formalism presented in §4 is applied to two test cases inspired from real operating situations encountered by pipeline monitors. The results are compared with those produced by the semi-implicit relaxation scheme in Eulerian coordinates, formerly introduced by Baudin et al. [4]. In both cases, we use the same pressure law as in Baudin et al. [3, 4], namely,

$$P(\tau, Y) = \frac{\alpha_G^2 Y}{\tau - \tau_L^\bullet(1-Y)}, \tag{6.1}$$

where

$$\alpha_G^2 = 10^5\,\mathrm{m}^2/\mathrm{s}^2, \qquad \tau_L^\bullet = 10^{-3}\,\mathrm{m}^3/\mathrm{kg}. \tag{6.2}$$

This amounts to assuming an ideal gas and an incompressible liquid. For simplicity, the constant a is chosen at each time-step according to

$$a^2 = \max_{1\leq i\leq N} -P_\tau(\tau_i^n, Y_i^n), \tag{6.3}$$

which is a rough version of the subcharacteristic condition (3.22). We also take for granted that, at the initial time $t = 0$, the pipeline is in the stationary state corresponding to the boundary values $q_0(t=0)$, $g_0(t=0)$ and $p_X(t=0)$.

6.1. **A simple scenario.** A mixture of gas and oil is injected into a pipeline of length $X = 4000\,\mathrm{m}$. The flow rates at the inlet are given by

$$g_0(t) = \ 10 + 0.2(t-100)\cdot \mathbf{1}_{\{100<t<200\}} + 20\cdot \mathbf{1}_{\{t>200\}}, \tag{6.4a}$$

$$q_0(t) = 990 + g_0(t). \tag{6.4b}$$

This means that within 100 seconds, we increase the gas flow rate linearly from 10 to $30\,\mathrm{kg/m^2s}$, while maintaining that of the liquid at the constant value $990\,\mathrm{kg/m^2s}$. At the outlet, the data

$$p_X(t) = 10^5\,\mathrm{Pa}, \qquad Y_X(t) = 1 \tag{6.5}$$

are also kept constant throughout the experiment.

Figure 2 displays the solutions computed at the end time $T = 300\mathrm{s}$ for the mesh size $\Delta x = 10\,\mathrm{m}$ and the CFL ratio 0.5, based on the slow wave. In terms of the density ρ and the mass-fraction Y, there is a good agreement between the two schemes. In terms of the velocity u and the pressure P, the discrepancy is more

visible but still small. This is due to the fact that both schemes are implicit—therefore less accurate—with respect to acoustic waves.

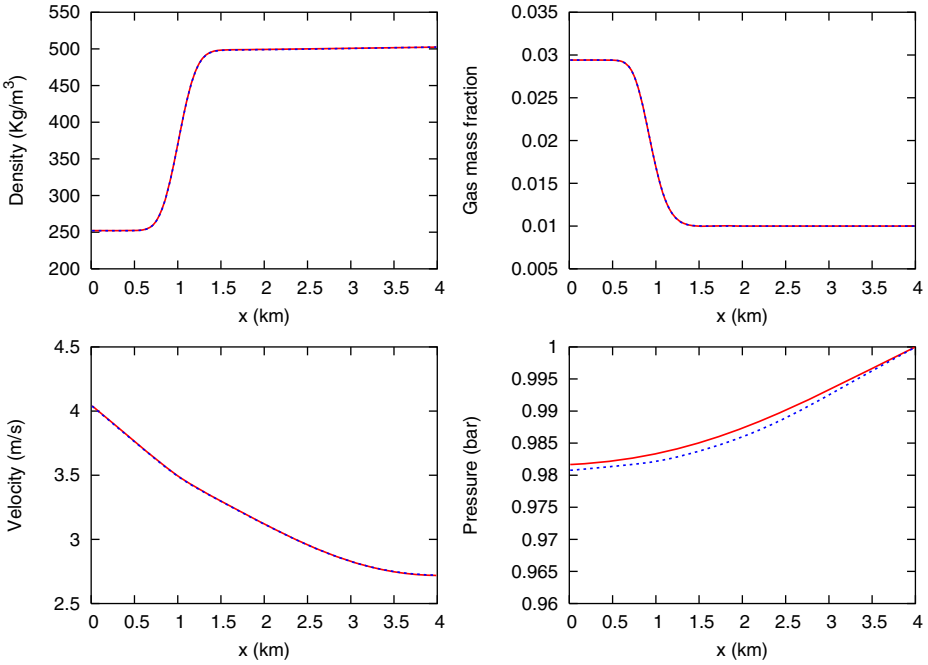

FIGURE 2. Numerical solutions obtained with the Lagrange-projection relaxation scheme (solid) and Eulerian relaxation scheme (dotted).

The Lagrange-projection relaxation scheme is about two times faster than the Eulerian relaxation scheme. This speed-up stems from the practical procedure of §2.3 for solving the linear system. Such a short-cut procedure is not possible in the Eulerian relaxation scheme. We remind the readers that the semi-implicit Eulerian relaxation scheme [4] is itself about 10 times faster than its fully explicit version [3].

We carried out a study of convergence for the two schemes. In Figure 3, we show the L^1-relative error of total density ρ versus mesh size $\Delta x = 80, 40, 20, 10\,\mathrm{m}$ in the log-log scale. This error is computed between the current solution and a reference solution, obtained with a very fine mesh ($\Delta x = 2.5\,\mathrm{m}$). It can be seen clearly that both schemes converge. From the sequence of errors we infer the rates of convergence by linear regression. The numerical orders of convergence are

$$\begin{aligned} \text{Lagrange-projection relaxation scheme} \quad & 0.82390, \\ \text{Eulerian relaxation scheme} \quad & 0.83675. \end{aligned}$$

These values are quite typical of first-order schemes with nonsmooth data.

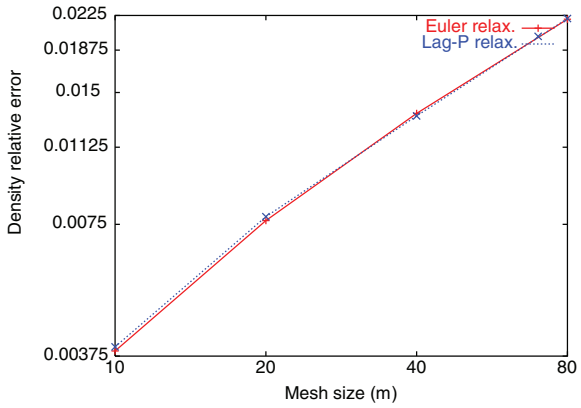

FIGURE 3. Convergence of the two relaxation schemes with respect to the mesh size Δx.

6.2. **A complex scenario.** In the second test case, the flow rates at the inlet are given by

$$g_0(t) = \ 10 + 0.2(t-100)\cdot \mathbf{1}_{\{100<t<200\}} + \ 20\cdot \mathbf{1}_{\{t>200\}} \ \text{kg/m}^2\text{s}, \tag{6.6a}$$

$$q_0(t) = 990 - 9.9(t-100)\cdot \mathbf{1}_{\{100<t<200\}} - 990\cdot \mathbf{1}_{\{t>200\}} \ \text{kg/m}^2\text{s}. \tag{6.6b}$$

This means that within 100 seconds, not only the gas flow rate is increased from 10 to 30 kg/m^{2}s, but also the liquid flow rate is decreased from 990 kg/m^{2}s to 0. As a consequence, the gas mass-fraction Y rises from 0.1 to its upper-bound 1, which is the main interest of this complex scenario. At the outlet, the pressure is increased by 100% according to

$$p_X(t) = 10^5 + 10^3(t-100)\cdot \mathbf{1}_{\{100<t<200\}} + 10^5\cdot \mathbf{1}_{\{t>200\}} \ \text{Pa}, \tag{6.7}$$

so as to allow the gas to return into the pipeline, thus activating the boundary condition $Y_X(t) = 1$.

Figure 4 displays the solutions computed at the final time $T = 300$ s for the mesh size $\Delta x = 10$ m and the CFL ratio 0.5, based on the slow wave. Again, in terms of density ρ and mass-fraction Y, there is a good agreement between the two schemes. In terms of velocity u and pressure P, the discrepancy is even more noticeable than in the previous test case. Let us repeat, however, that u and P are more closely associated with fast acoustic waves, in which engineers are not interested. The only wave worthy of their attention is the slow kinematic wave by which Y is transported.

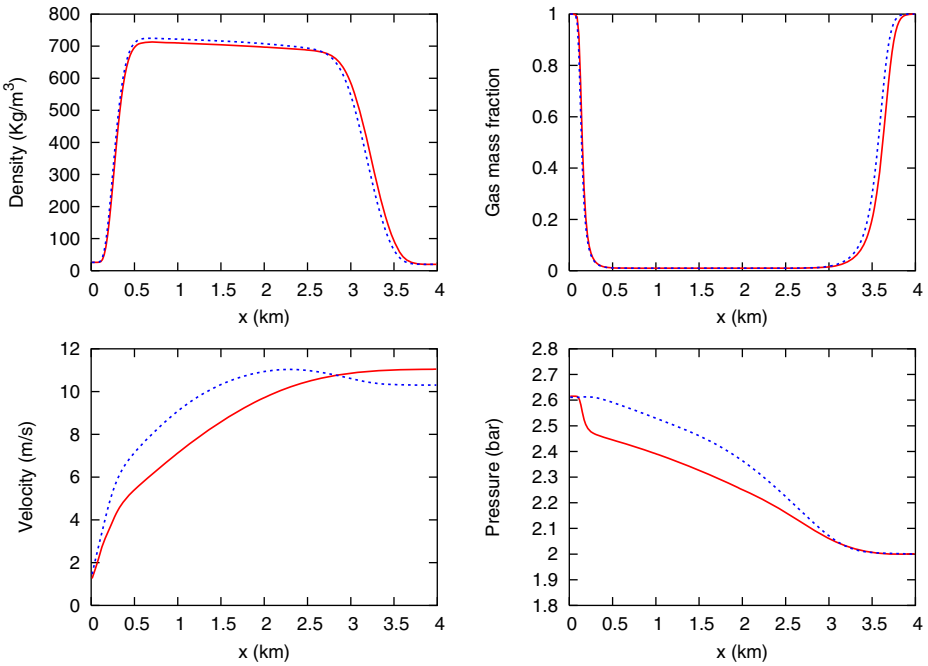

FIGURE 4. Numerical solutions obtained with the Lagrange-projection relaxation scheme (solid) and Eulerian relaxation scheme (dotted).

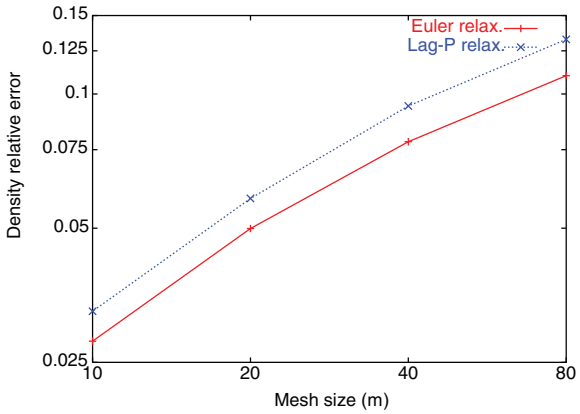

FIGURE 5. Convergence of the two relaxation schemes with respect to the mesh size Δx.

Since there is no guarantee of positivity for the Eulerian relaxation scheme, we resorted to the following device in order to maintain ρ and Y in the proper ranges. Whenever necessary, that is, as soon as those variables are found to get out of range, we re-do the current time-step after dividing Δt by 2. Keeping this in mind, we proceed to a study of convergence along the same line as in the previous case. The results are displayed in Figure 5. The numerical orders of convergence are

Lagrange-projection relaxation scheme 0.67695,
Eulerian relaxation scheme 0.65819.

Similarly to the first test case, the Lagrange-projection relaxation scheme is about two times faster than the Eulerian relaxation scheme.

7. Concluding remarks

Throughout this paper, we have opted for an axiomatic layout to introduce the various problems considered. This presentation allows us to put the boundary conditions on an equal footing with the PDEs and the initial data.

For the sake of clarity, the transformation of the Lagrange step into two symmetric advection equations was carried out using the invariants $\overrightarrow{w} = \Pi + av$ and $\overleftarrow{w} = \Pi - av$ involving the main variables. Actually, at the discrete level, there is an alternative formulation that makes use of the time variations

$$\delta(.) = (.)^{n\sharp} - (.)^n. \tag{7.1}$$

This incremental formulation is more convenient when we want to discretize the boundary conditions (2.4) on the basis of the values of (σ_0, σ_Z) at time $n\sharp$ instead of time n. It is also of great help when we wish to extend the new explicit-implicit method to a quasi second-order approximation. In this case, it is still possible to apply the same philosophy in order to find an optimal time-step that preserves positivity, even though the entropy inequality cannot be ascertained.

We are currently working on the extension of the method to more general and realistic two-phase flow systems, in which the gas mass balance reads

$$\partial_t(\rho Y) + \partial_x(\rho Y u - \sigma) = 0, \tag{7.2}$$

where $\sigma = \sigma(\rho, Y, u)$ represents a hydrodynamic closure law [3, 4].

Appendix A. Functional framework for Problem (SA)

At first sight, *Problem (SA)* seems to be somewhat of a "classic", and one might suspect that it has already been investigated. However, a more careful look reveals that the coupling of boundary conditions through (2.4) could be something new, essentially because (θ_0, θ_Z) depend on time t. The only reference we have been able to find that contains a similar two-advection system is a review by Russell [31], in which only a subcase of *Problem (SA)* is considered. On the other hand, the functional framework usually associated with linear problems involves L^2-spaces, as is the case in Russell's paper. However, in view of the application of *Problem (SA)* to the approximation of *Problem (TP)* and *Problem (EU)*, what we really need are L^∞-norms, as already explained in §2.1.

We use the L^∞-norm defined by (2.5) and the notation of §2.1. We will also write $L^\infty \cap C^1(\overline{\mathscr{O}}; \mathbb{R})$ for $L^\infty(\overline{\mathscr{O}}; \mathbb{R}) \cap C^1(\overline{\mathscr{O}}; \mathbb{R})$, equipped with the same norm. Finally, for short-hand convenience, we write $\mathbb{R}^+_Z = \mathbb{R}^+ \times [0, Z]$. The set of C^1-functions $\varphi(t, z)$ whose supports are compact and included in $\mathbb{R}^+_Z$ is denoted by $\mathscr{C}^1_0(\mathbb{R}^+_Z)$.

We are going to work out a weak formulation for *Problem (SA)*.

Definition A.1. Given

(A.1a) – the initial data $\overrightarrow{w}_\flat, \overleftarrow{w}_\flat \in L^\infty([0, Z]; \mathbb{R}) \times L^\infty([0, Z]; \mathbb{R})$,

(A.1b) – the boundary data $\sigma_0, \sigma_Z \in L^\infty(\mathbb{R}^+; \mathbb{R}) \times L^\infty(\mathbb{R}^+; \mathbb{R})$,

(A.1c) – the coupling factors $\theta_0, \theta_Z \in L^\infty \cap C^1(\mathbb{R}^+; \mathbb{R}) \times L^\infty \cap C^1(\mathbb{R}^+; \mathbb{R})$,

the pair

$$\vec{w}, \overleftarrow{w} \in L^\infty(\mathbb{R}^+_Z;\mathbb{R}) \times L^\infty(\mathbb{R}^+_Z;\mathbb{R}) \tag{A.2}$$

is said to be a weak solution to *Problem (SA)* if, for any pair of test functions $(\vec{\varphi}, \overleftarrow{\varphi}) \in \mathscr{C}^1_0(\mathbb{R}^+_Z) \times \mathscr{C}^1_0(\mathbb{R}^+_Z)$ subject to

$$\overleftarrow{\varphi}(t,0) = \theta_0(t)\vec{\varphi}(t,0) \quad \text{and} \quad \vec{\varphi}(t,Z) = \theta_Z(t)\overleftarrow{\varphi}(t,Z) \quad \text{for all } t>0, \tag{A.3}$$

we have

$$\begin{aligned} 0 = &\iint_{\mathbb{R}^+_Z} \vec{w}(\partial_t\vec{\varphi} + a\partial_z\vec{\varphi})\,dt\,dz + \iint_{\mathbb{R}^+_Z} \overleftarrow{w}(\partial_t\overleftarrow{\varphi} - a\partial_z\overleftarrow{\varphi})\,dt\,dz \\ &+ \int_{[0,Z]} \vec{w}_\flat(z)\vec{\varphi}(0,z)\,dz \qquad + \int_{[0,Z]} \overleftarrow{w}_\flat(z)\overleftarrow{\varphi}(0,z)\,dz \\ &+ \int_{\mathbb{R}^+} a\sigma_0(t)\vec{\varphi}(t,0)\,dt \qquad + \int_{\mathbb{R}^+} a\sigma_Z(t)\overleftarrow{\varphi}(t,Z)\,dt. \end{aligned} \tag{A.4}$$

It can be verified that under the assumptions (A.1)–(A.2), all integrals involved in (A.4) are well defined. This weak formulation comes from standard techniques [18]. We first suppose $(\vec{w}, \overleftarrow{w})$ to be a classical solution. Multiplying (2.2a) by $\vec{\varphi}$, (2.2b) by $\overleftarrow{\varphi}$, integrating by parts, then adding them together, we replace the initial data by (2.3) and make use of (2.4) to get rid of the boundary terms. The constraints (A.3) on test functions $(\vec{\varphi}, \overleftarrow{\varphi})$ reflect the fact that $(\vec{w}, \overleftarrow{w})$ influence each other through boundary conditions. The subset of $\mathscr{C}^1_0(\mathbb{R}^+_Z) \times \mathscr{C}^1_0(\mathbb{R}^+_Z)$ containing pairs of test functions satisfying (A.3) is not empty.

This weak formulation allows us to clarify Theorem 2.1 and to prove it.

Theorem A.1. *If $\|\theta_0\|\|\theta_Z\| < 1$, then Problem (SA) is well posed, in the sense that it has a unique weak solution depending continuously on the data. All other statements of Theorem* 2.1 *hold true. Furthermore, the auxiliary functions $(\vec{w}_0, \overleftarrow{w}_Z)$ both belong to $L^\infty(\mathbb{R}^+;\mathbb{R})$.*

In order to prove Theorem A.1, we need three preliminary results. The first two are technical devices for existence, while the last one is the keystone for uniqueness.

Lemma A.1. *Let $T>0$ and let α, g be two functions in $L^\infty(\mathbb{R}^+;\mathbb{R})$. If $\|\alpha\| < 1$, then the functional equation*

$$w(t) - \alpha(t)\mathbf{1}_{\{t>T\}}w(t-T) = g(t) \tag{A.5}$$

admits a unique solution $w \in L^\infty(\mathbb{R}^+;\mathbb{R})$. This solution depends continuously on the data g, and we have

$$\|w\| \le \frac{1}{1-\|\alpha\|}\,\|g\|. \tag{A.6}$$

Proof. Let us first assume existence and try to find out a formula for w. For $t<T$, it is clear that $w(t) = g(t)$. For $t>T$ let $d = \lfloor t/T \rfloor$, so that $0 \le t - dT < T$, and $w(t-dT) = g(t-dT)$. Then, combining the equalities

$$\begin{aligned} w(t) &= \alpha(t)\,w(t-T) &&+ g(t), \\ w(t-T) &= \alpha(t-T)\,w(t-2T) &&+ g(t-T), \\ w(t-2T) &= \alpha(t-2T)\,w(t-3T) &&+ g(t-2T), \\ &\dots = \dots \\ w(t-(d-1)T) &= \alpha(t-(d-1)T)\,w(t-dT) &&+ g(t-(d-1)T), \end{aligned} \tag{A.7}$$

we have

$$w(t) = g(t) + \sum_{r=1}^{d} \alpha(t)\alpha(t-T)\dots\alpha(t-(r-1)T)g(t-rT). \tag{A.8}$$

This ensures uniqueness. Conversely, we can readily check that

$$w(t) = g(t) + \sum_{r=1}^{\lfloor t/T \rfloor} \alpha(t)\alpha(t-T)\dots\alpha(t-(r-1)T)g(t-rT) \tag{A.9}$$

is indeed a solution to (A.5), the latter formula being valid for any $t \geq 0$. From (A.9), it follows that

$$|w(t)| \leq \|g\|(1+\sum_{r=1}^{\lfloor t/T \rfloor} |\alpha(t)\alpha(t-T)\dots\alpha(t-(r-1)T)|) \leq \|g\|(1+\sum_{r=1}^{\lfloor t/T \rfloor} \|\alpha\|^r) \tag{A.10}$$

so that if $\|\alpha\| < 1$, we have $w \in L^\infty(\mathbb{R}^+;\mathbb{R})$ and the desired estimate (A.6). □

Lemma A.2. *The two systems* (2.8) *and* (2.9) *are equivalent. If* $\|\theta_0\|\|\theta_Z\| < 1$, *then they have the same unique solution* $(\vec{w}_0, \overleftarrow{w}_Z)$, *which depends continuously on the data.*

Proof. We first prove that (2.9) is well posed. This is done by applying Lemma A.1 twice. The first time, with

$$T = 2Z/a,\ w(t) = \vec{w}_0(t),\ \alpha(t) = \theta_0(t)\theta_Z(t-Z/a),\ g(t) = G_0(t), \tag{A.11}$$

we get existence, uniqueness and continuous dependence for $\vec{w}_0$. In this case,

$$\|\vec{w}_0\| \leq \frac{1}{1-\|\theta_0\|\|\theta_Z\|}(\|\sigma_0\| + \|\theta_0\| \max\{\|\overleftarrow{w}_\flat\|, \|\sigma_Z\| + \|\theta_Z\|\|\vec{w}_\flat\|\}). \tag{A.12}$$

The second time, with

$$T = 2Z/a,\ w(t) = \overleftarrow{w}_Z(t),\ \alpha(t) = \theta_Z(t)\theta_0(t-Z/a),\ g(t) = G_Z(t), \tag{A.13}$$

we get existence, uniqueness and continuous dependence for $\overleftarrow{w}_Z$. In this case,

$$\|\overleftarrow{w}_Z\| \leq \frac{1}{1-\|\theta_0\|\|\theta_Z\|}(\|\sigma_Z\| + \|\theta_Z\| \max\{\|\vec{w}_\flat\|, \|\sigma_0\| + \|\theta_0\|\|\overleftarrow{w}_\flat\|\}). \tag{A.14}$$

Let us write (2.8b) at time $t - Z/a$, and plug the expression for $\overleftarrow{w}_Z(t-Z/a)$ into (2.8a). We then obtain (2.9a). A similar elimination enables us to deduce (2.9b) from (2.8). Thus, (2.8) ⇒ (2.9).

We can put (2.9)–(2.10) under the form

$$\vec{w}_0(t) = \sigma_0(t) + \theta_0(t)[\mathbf{1}_{\{at<Z\}}\overleftarrow{w}_\flat(at) + \mathbf{1}_{\{at>Z\}}H_Z(t-Z/a)], \tag{A.15a}$$

$$\overleftarrow{w}_Z(t) = \sigma_Z(t) + \theta_Z(t)[\mathbf{1}_{\{at<Z\}}\vec{w}_\flat(Z-at) + \mathbf{1}_{\{at>Z\}}H_0(t-Z/a)] \tag{A.15b}$$

with

$$H_0(t) = \sigma_0(t) + \theta_0(t)[\mathbf{1}_{\{at<Z\}}\overleftarrow{w}_\flat(at) + \mathbf{1}_{\{at>Z\}}\overleftarrow{w}_Z(t-Z/a)], \tag{A.16a}$$

$$H_Z(t) = \sigma_Z(t) + \theta_Z(t)[\mathbf{1}_{\{at<Z\}}\vec{w}_\flat(Z-at) + \mathbf{1}_{\{at>Z\}}\vec{w}_0(t-Z/a)]. \tag{A.16b}$$

Equations (A.15a), (A.16b) testify to the fact that $(\vec{w}_0, H_Z)$ solves (2.8). Likewise, $(H_0, \overleftarrow{w}_Z)$ also solves (2.8). Since (2.8) ⇒ (2.9), the two pairs $(\vec{w}_0, H_Z)$ and $(H_0, \overleftarrow{w}_Z)$ are solutions to (2.9). By virtue of uniqueness for (2.9), we infer that $H_0 = \vec{w}_0$ and $H_Z = \overleftarrow{w}_Z$. Hence, (2.9) ⇒ (2.8). □

Lemma A.3. *Every pair of functions* $(\vec{\kappa},\overleftarrow{\kappa}) \in \mathscr{C}^1_0(\mathbb{R}^+_Z) \times \mathscr{C}^1_0(\mathbb{R}^+_Z)$ *can be expressed as*

$$\vec{\kappa} = \partial_t \vec{\varphi} + a\partial_z \vec{\varphi}, \tag{A.17a}$$

$$\overleftarrow{\kappa} = \partial_t \overleftarrow{\varphi} - a\partial_z \overleftarrow{\varphi}, \tag{A.17b}$$

where the functions $(\vec{\varphi},\overleftarrow{\varphi}) \in \mathscr{C}^1_0(\mathbb{R}^+_Z) \times \mathscr{C}^1_0(\mathbb{R}^+_Z)$ *obey, for all* $t \geq 0$,

$$\overleftarrow{\varphi}(t,0) = \theta_0(t)\vec{\varphi}(t,0) \quad \textit{and} \quad \vec{\varphi}(t,Z) = \theta_Z(t)\overleftarrow{\varphi}(t,Z). \tag{A.18}$$

Proof. Suppose that the supports of $(\vec{\kappa},\overleftarrow{\kappa})$ are included in $[0,T_\natural[\times[0,Z]$, that is, $\vec{\kappa}(t,z) = \overleftarrow{\kappa}(t,z) = 0$ for $t \geq T_\natural$. For safety, we take $T = T_\natural + Z/2a$. We describe how to construct $(\vec{\varphi},\overleftarrow{\varphi})$ from $(\vec{\kappa},\overleftarrow{\kappa})$ with the additional requirement that the supports of $(\vec{\varphi},\overleftarrow{\varphi})$ be included in $[0,T[\times[0,Z]$.

From a fixed point $M = (t,z) \in [0,T[\times[0,Z]$, we draw characteristic lines forward, starting with the a slope, then alternating with $-a$ every time a boundary is met, and stopping at the upper boundary $t = T$. As illustrated in Figure 6, this gives rise to a path, for which we now provide an analytical definition in order to write accurate formulae later. Let

$$N = \left\lfloor \frac{z + a(T-t)}{Z} \right\rfloor \tag{A.19}$$

and introduce the points $A_n(t_n,z_n)$ defined by:

- for $n = 0$, $(t_0,z_0) = (t,z)$, which means that $A_0 = M$;
- for $1 \leq n \leq N$,

$$(t_n,z_n) = \left(t + \frac{nZ - z}{a}, \frac{1-(-1)^n}{2}Z\right), \tag{A.20}$$

  which means that A_n belongs to the right boundary if n is odd, and to the left boundary if n is even;
- for $n = N+1$, $(t_{N+1},z_{n+1}) = (T, z_N + (-1)^N a(T - t_N))$.

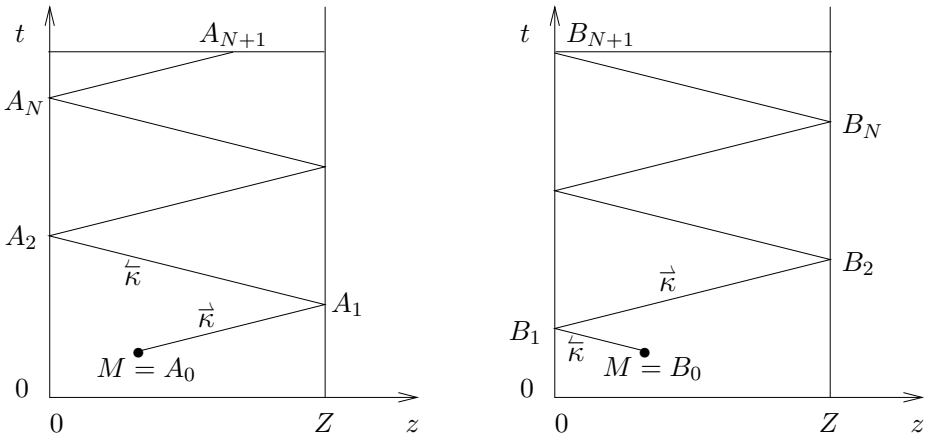

FIGURE 6. Characteristics for the solution to the adjoint problem.

If $(\vec{\varphi}, \overleftarrow{\varphi})$ is a solution pair, then in the first segment $[A_0A_1]$, we should have

$$\vec{\varphi}(A_0) = \vec{\varphi}(A_1) - \int_{A_0}^{A_1} \vec{\kappa}(t, z(t))\, dt = \theta_Z(t_1)\overleftarrow{\varphi}(A_1) - \int_{A_0}^{A_1} \vec{\kappa}(t, z(t))\, dt, \tag{A.21}$$

where $\int_{A_0}^{A_1}$ denotes integration along $[A_0A_1]$. More generally, we should have

$$\begin{aligned} \vec{\varphi}(A_{2k}) &= \vec{\varphi}(A_{2k+1}) - \int_{A_{2k}}^{A_{2k+1}} \vec{\kappa}\, dt = \theta_Z(t_{2k+1})\overleftarrow{\varphi}(A_{2k+1}) - \int_{A_{2k}}^{A_{2k+1}} \vec{\kappa}\, dt, \\ \vec{\varphi}(A_{2k+1}) &= \vec{\varphi}(A_{2k+2}) - \int_{A_{2k+1}}^{A_{2k+2}} \overleftarrow{\kappa}\, dt = \theta_0(t_{2k+2})\overleftarrow{\varphi}(A_{2k+2}) - \int_{A_{2k+1}}^{A_{2k+2}} \overleftarrow{\kappa}\, dt \end{aligned} \tag{A.22}$$

for k within an acceptable range. Combining the equalities (A.22) with the final data $\vec{\varphi}(A_{N+1}) = \overleftarrow{\varphi}(A_{N+1}) = 0$, we end up with

$$\vec{\varphi}(M) = -\sum_{k\geq 0} \Theta_{2k} \int_{A_{2k}}^{A_{2k+1}} \vec{\kappa}\, dt - \sum_{k\geq 1} \Theta_{2k-1} \int_{A_{2k-1}}^{A_{2k}} \overleftarrow{\kappa}\, dt, \tag{A.23}$$

where the sums automatically stop beyond A_{N+1}, and

$$\Theta_{2k} = \prod_{\ell=1}^{k} \theta_0(t_{2\ell})\theta_Z(t_{2\ell-1}), \qquad \Theta_{2k-1} = \theta_Z(t_{2k-1}) \prod_{\ell=1}^{k-1} \theta_0(t_{2\ell})\theta_Z(t_{2\ell-1}). \tag{A.24}$$

Consider the function $\vec{\varphi}$ defined by (A.23)–(A.24) for $t < T$ and by $\vec{\kappa}(t, z) = 0$ for all $t \geq T$. Since (t_n, z_n) are C^1-functions of (t, z), it is a straightforward matter to check that $\vec{\varphi}$ is C^1 with respect to (t, z) if (θ_0, θ_Z) are C^1-functions of t. On the other hand, it can be verified to be compact-supported. Therefore, $\vec{\varphi} \in \mathscr{C}_0^1(\mathbb{R}_Z^+)$.

Starting from M with the slope $-a$, we derive a similar construction for $\overleftarrow{\varphi} \in \mathscr{C}_0^1(\mathbb{R}_Z^+)$. It is easy to check that $(\vec{\varphi}, \overleftarrow{\varphi})$ is indeed a solution to (A.17)–(A.18). □

Proof of Theorem A.1. *Uniqueness.* Suppose there are two pairs $(\vec{w}_1, \overleftarrow{w}_1)$ and $(\vec{w}_2, \overleftarrow{w}_2)$ both satisfying the weak formulation (A.4). We are going to show that $(\vec{w}_1, \overleftarrow{w}_1) = (\vec{w}_2, \overleftarrow{w}_2)$ in the sense that

$$\iint_{\mathbb{R}_Z^+} (\vec{w}_2 - \vec{w}_1)\kappa\, dt\, dz = 0 \qquad \text{and} \qquad \iint_{\mathbb{R}_Z^+} (\overleftarrow{w}_2 - \overleftarrow{w}_1)\kappa\, dt\, dz = 0 \tag{A.25}$$

for all $\kappa \in \mathscr{C}_0^1(\mathbb{R}_Z^+)$. We make use of a nonlinear version of Holmgren's technique [32] and consider the adjoint problem (A.17)–(A.18) for a given pair $(\vec{\kappa}, \overleftarrow{\kappa}) \in \mathscr{C}_0^1(\mathbb{R}_Z^+) \times \mathscr{C}_0^1(\mathbb{R}_Z^+)$. According to Lemma A.3, this problem has a solution $(\vec{\varphi}, \overleftarrow{\varphi}) \in \mathscr{C}_0^1(\mathbb{R}_Z^+) \times \mathscr{C}_0^1(\mathbb{R}_Z^+)$. Specifying this solution pair as test functions and writing the weak formulation (A.4) for $(\vec{w}_1, \overleftarrow{w}_1)$ and $(\vec{w}_2, \overleftarrow{w}_2)$, we get

$$0 = \iint_{\mathbb{R}_Z^+} (\vec{w}_2 - \vec{w}_1)(\partial_t\vec{\varphi} + a\partial_z\vec{\varphi})\, dt\, dz + \iint_{\mathbb{R}_Z^+} (\overleftarrow{w}_2 - \vec{w}_1)(\partial_t\overleftarrow{\varphi} - a\partial_z\overleftarrow{\varphi})\, dt\, dz, \tag{A.26}$$

which implies

$$0 = \iint_{\mathbb{R}_Z^+} (\vec{w}_2 - \vec{w}_1)\vec{\kappa}\, dt\, dz + \iint_{\mathbb{R}_Z^+} (\overleftarrow{w}_2 - \vec{w}_1)\overleftarrow{\kappa}\, dt\, dz. \tag{A.27}$$

To reach claim (A.25), we set $(\vec{\kappa}, \overleftarrow{\kappa}) = (\kappa, 0)$, then $(\vec{\kappa}, \overleftarrow{\kappa}) = (0, \kappa)$.

Existence and continuous dependence. Our strategy is to insert the candidate functions (2.7)–(2.8) into the right-hand side of the weak formulation (A.4) and to check that it vanishes. The calculations are somewhat heavy, because of the many

changes of variables to be carried out for the double integrals. We just sketch out the intermediate steps, leaving the details to the reader.

First, using (2.7) and cutting the integration domain into subdomains, we have

$$\begin{aligned}
&\iint_{\mathbb{R}^+_Z} \vec{w}(\partial_t \vec{\varphi} + a\partial_z \vec{\varphi})\,dt\,dz + \int_0^Z \vec{w}_\flat(z)\vec{\varphi}(0,z)\,dz + \int_{\mathbb{R}^+} a\vec{w}_0(t)\vec{\varphi}(t,0)\,dt \\
&\qquad = \int_0^Z \vec{w}_\flat(z)\vec{\varphi}(\tfrac{Z-z}{a}, Z)\,dz + \textstyle\int_{\mathbb{R}^+} a\vec{w}_0(t)\vec{\varphi}(t+\frac{Z}{a}, Z)\,dt, \\
&\iint_{\mathbb{R}^+_Z} \overleftarrow{w}(\partial_t \overleftarrow{\varphi} - a\partial_z \overleftarrow{\varphi})\,dt\,dz + \int_0^Z \overleftarrow{w}_\flat(z)\overleftarrow{\varphi}(0,z)\,dz + \int_{\mathbb{R}^+} a\overleftarrow{w}_Z(t)\overleftarrow{\varphi}(t,Z)\,dt \\
&\qquad = \int_0^Z \overleftarrow{w}_\flat(z)\overleftarrow{\varphi}(\tfrac{z}{a}, 0)\,dz + \textstyle\int_{\mathbb{R}^+} a\overleftarrow{w}_Z(t)\overleftarrow{\varphi}(t+\frac{Z}{a}, 0)\,dt,
\end{aligned} \tag{A.28}$$

Invoking (2.8), taking advantage of (A.3), making appropriate changes of variables and invoking (2.8), once again leads us to the conclusion that the right-hand side of (A.4) is equal to 0.

It remains to check that $(\vec{w}, \overleftarrow{w})$ are L^∞-functions. From (2.7) and from the L^∞-assumptions made on $(\vec{w}_\flat, \overleftarrow{w}_\flat)$, it is obvious that we simply need to check that the auxiliary functions $(\vec{w}_0, \overleftarrow{w}_Z)$ are L^∞-functions. From (2.9)–(2.10) and by Lemma A.1, it can be seen that this is ensured as soon as $\|\theta_0\|\|\theta_Z\| < 1$. Finally, from the estimates (A.12)–(A.13), we infer that

$$\max\{\|\vec{w}_0\|, \|\overleftarrow{w}_Z\|\} \le \frac{(1+\|\theta_0\|)(1+\|\theta_Z\|)}{1-\|\theta_0\|\|\theta_Z\|} \max\{\|\vec{w}_\flat\|, \|\overleftarrow{w}_\flat\|, \|\sigma_0\|, \|\sigma_Z\|\}. \tag{A.29}$$

Recalling (2.7) again and arguing that the constant in (A.29) is greater than 1, we arrive at (2.6). □

Acknowledgments

This work was supported by the Ministère de la Recherche under grant ERT-22052274: *Simulation avancée du transport des hydrocarbures* and by the Institut Français du Pétrole.

References

1. A. A. Amsden, P. J. O'Rourke, and T. D. Butler, *KIVA-II: A computer program for chemically reactive flows with sprays*, Report LA-11560-MS, Los Alamos National Laboratory, 1989.
2. N. Andrianov, F. Coquel, M. Postel, and Q. H. Tran, *A relaxation multiresolution scheme for accelerating realistic two-phase flows calculations in pipelines*, Int. J. Numer. Meth. Fluids **54** (2007), 207–236. MR2313539 (2008e:76119)
3. M. Baudin, C. Berthon, F. Coquel, R. Masson, and Q. H. Tran, *A relaxation method for two-phase flow models with hydrodynamic closure law*, Numer. Math. **99** (2005), 411–440. MR2117734 (2005h:76079)
4. M. Baudin, F. Coquel, and Q. H. Tran, *A semi-implicit relaxation scheme for modeling two-phase flow in a pipeline*, SIAM J. Sci. Comput. **27** (2005), no. 3, 914–936. MR2199914 (2006k:76092)
5. François Bouchut, *Entropy satisfying flux vector splittings and kinetic BGK models*, Numer. Math. **94** (2003), 623–672. MR1990588 (2005e:65129)
6. ———, *Nonlinear stability of finite volume methods for hyperbolic conservation laws, and well-balanced schemes for sources*, Frontiers in Mathematics, Birkhäuser, 2004. MR2128209 (2005m:65002)

7. C. Chalons and F. Coquel, *Navier-Stokes equations with several independent pressure laws and explicit predictor-corrector schemes*, Numer. Math. **101** (2005), no. 3, 451–478. MR2194824 (2006m:76119)
8. Jean Jacques Chattot and Sylvie Mallet, *A "box-scheme" for the Euler equations*, Nonlinear Hyperbolic Problems (Berlin) (C. Carasso, P. A. Raviart, and D. Serre, eds.), Lecture Notes in Mathematics, vol. 1270, Springer-Verlag, 1987, pp. 82–102. MR0910106 (88h:76002)
9. Gui Qiang Chen, C. David Levermore, and Tai Ping Liu, *Hyperbolic conservation laws with stiff relaxation terms and entropy*, Comm. Pure Appl. Math. **47** (1994), no. 6, 787–830. MR1280989 (95h:35133)
10. Frédéric Coquel, Edwige Godlewski, Benoît Perthame, Arun In, and P. Rascle, *Some new Godunov and relaxation methods for two-phase flow problems*, Godunov methods: Theory and Applications (New York) (E. Toro, ed.), Proceedings of the International Conference on Godunov methods in Oxford, 1999, Kluwer Academic/Plenum Publishers, 2001, pp. 179–188.
11. F. Coquel and B. Perthame, *Relaxation of energy and approximate Riemann solvers for general pressure laws in fluid dynamics*, SIAM J. Numer. Anal. **35** (1998), no. 6, 2223–2249. MR1655844 (2000a:76129)
12. C.M. Dafermos, *Hyperbolic conservation laws in continuum physics*, Grundlehren der mathematischen Wissenschaften, vol. 325, Springer-Verlag, Berlin, 2000. MR1763936 (2001m:35212)
13. B. Després and C. Mazeran, *Lagrangian gas dynamics in two-dimensions and Lagrangian systems*, Arch. Ration. Mech. Anal. **178** (2005), no. 3, 327–372. MR2196496 (2006j:76135)
14. J. Donea, A. Huerta, J. P. Ponthot, and A. Rodríguez-Ferran, *Arbitrary Lagragian-Eulerian methods*, Encyclopedia of Computational Mechanics (E. Stein, R. de Borst, and T. Hughes, eds.), vol. 1, John Wiley & Sons, 2004, pp. 413–437.
15. F. Dubois and P. LeFloch, *Boundary conditions for nonlinear hyperbolic systems of conservation laws*, J. Diff. Eq. **71** (1988), no. 1, 93–122. MR922200 (89c:35099)
16. Steinar Evje and Tore Flåtten, *Weakly implicit numerical schemes for a two-fluid model*, SIAM J. Sci. Comput. **26** (2005), no. 5, 1449–1484. MR2142581 (2005k:76133)
17. I. Faille and E. Heintzé, *A rough finite volume scheme for modeling two phase flow in a pipeline*, Computers and Fluids **28** (1999), 213–241.
18. Edwige Godlewski and Pierre Arnaud Raviart, *Numerical approximation of hyperbolic systems of conservation laws*, Applied Mathematical Sciences, vol. 118, Springer-Verlag, New York, 1996. MR1410987 (98d:65109)
19. Charles Hirsch, *Numerical computation of internal and external flows, vol. I & II*, Wiley Series in Numerical Methods in Engineering, John Wiley and Sons, New York, 1990.
20. C. W. Hirt, A. A. Amsden, and J. L. Cook, *An arbitrary Lagrangian-Eulerian computing method for all flow speeds*, J. Comput. Phys. **14** (1974), 227–253.
21. M. J. Holst, *Notes on the KIVA-II software and chemically reactive fluid mechanics*, Report UCRL-ID-112019, Lawrence Livermore National Laboratory, California, 1992.
22. S. Jin and Z. P. Xin, *The relaxation schemes for systems of conservation laws in arbitrary space dimension*, Comm. Pure Appl. Math. **48** (1995), no. 3, 235–276. MR1322811 (96c:65134)
23. C. Johnson and A. Szepessy, *On the convergence of a finite element method for a nonlinear hyperbolic conservation law*, Math. Comp. **49** (1987), no. 180, 427–444. MR906180 (88h:65164)
24. R. J. LeVeque, *Numerical methods for conservation laws*, Lectures in Mathematics, ETH Zürich, Birkhäuser Verlag, Berlin, 1992. MR1153252 (92m:65106)
25. Tai Ping Liu, *Hyperbolic conservation laws with relaxation*, Comm. Math. Phys. **108** (1987), no. 1, 153–175. MR872145 (88f:35092)
26. Jean Marie Masella, Isabelle Faille, and Thierry Gallouët, *On an approximate Godunov scheme*, Int. J. Comput. Fluid Dynam. **12** (1999), no. 2, 133–149. MR1729206 (2000h:65122)
27. J. M. Masella, Q. H. Tran, D. Ferré, and C. Pauchon, *Transient simulation of two-phase flows in pipes*, Int. J. Multiph. Flow **24** (1998), no. 5, 739–755.
28. W. A. Mulder and B. van Leer, *Experiments with implicit upwind methods for the Euler equations*, J. Comput. Phys. **59** (1985), 232–246. MR796608
29. Roberto Natalini, *Convergence to equilibrium for the relaxation approximations of conservation laws*, Comm. Pure Appl. Math. **49** (1996), 1–30. MR1391756 (97c:35131)
30. Christian Pauchon, Henri Dhulesia, G. Binh-Cirlot, and Jean Fabre, *TACITE: A transient tool for multiphase pipeline and well simulation*, SPE Annual Technical Conference and Exhibition, New Orleans, September 1994, 1994, SPE Paper 28545.

31. David L. Russell, *Controllability and stabilizability theory for linear partial differential equations: recent progress and open questions*, SIAM Review **20** (1978), no. 4, 639–739. MR508380 (80c:93032)
32. Joel Smoller, *Shock waves and reaction-diffusion equations*, Grundlehren der mathematischen Wissenschaften, vol. 258, Springer-Verlag, New York, 1994. MR1301779 (95g:35002)
33. E. F. Toro, *Riemann solvers and numerical methods for fluid dynamics: A practical introduction*, Springer-Verlag, Berlin, 1997. MR1474503 (98h:76099)
34. D. H. Wagner, *Equivalence of the Euler and Lagrangian equations of gas dynamics for weak solutions*, J. Diff. Eq. **68** (1987), 118–136. MR885816 (88i:35100)
35. H. Weyl, *Shock waves in arbitrary fluids*, Comm. Pure Appl. Math **2** (1949), 103–122. MR0034677 (11:626a)
36. H. C. Yee, R. F. Warming, and A. Harten, *Implicit Total Variation Diminishing schemes for steady-state calculations*, J. Comput. Phys. **57** (1985), no. 3, 327–360. MR782986 (86h:65134)

UPMC Univ Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France

CNRS, UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France

Département Mathématiques Appliquées, Institut Français du Pétrole, 1 et 4 avenue de Bois-Préau, 92852 Rueil-Malmaison Cedex, France

Département Mathématiques Appliquées, Institut Français du Pétrole, 1 et 4 avenue de Bois-Préau, 92852 Rueil-Malmaison Cedex, France