PREPRINT 422

MAY 2015

On the Mach-Uniformity of Lagrange-Projection Scheme

Hamed Zakerzadeh

Institut für
Geometrie und Praktische Mathematik
Templergraben 55, 52056 Aachen, Germany

2000 Mathematics Subject Classification. Primary 35L81, 65M12; Secondary 35L65, 65M08.
Key words and phrases. All-Mach number scheme, Lagrange–projection scheme, asymptotic preserving scheme, stability analysis.
The author was supported by Richtlinien zur Förderung des wissenschaftlichen Nachwuchses (RFwN) Stipendium.

Mathematical Modelling and Numerical Analysis Will be set by the publisher
Modélisation Mathématique et Analyse Numérique

ON THE MACH-UNIFORMITY OF THE LAGRANGE–PROJECTION SCHEME *

Hamed Zakerzadeh[1]

Abstract. In the present work, we show that the Implicit-Explicit Lagrange–projection scheme applied to the barotropic Euler equations, presented in Coquel et al.'s paper (Math. of Comp. **79**.271 (2010): 1493–1533), is *asymptotic preserving* regarding the Mach number, i.e., it is asymptotically stable in ℓ_∞-norm with unrestrictive CFL condition for all-Mach flows, and asymptotically consistent which means that it gives a consistent discretization to the incompressible Euler equations in the limit, e.g., it preserves the incompressible limit as to satisfy the *div*-free condition or the analogues of continuous-level asymptotic expansion for the density. This consistency analysis has been done formally as well as rigorously. Moreover, we prove that the scheme is positivity-preserving and entropy-admissible under some Mach-uniform restrictions. The analysis is similar to what has been presented in the original paper, but with the emphasis on the uniformity regarding the Mach number, which is crucial for a scheme to be useful in the low-Mach regime. We then extend the modified (but similar) analysis to the shallow water equations with topography and get similar stability and consistency results.

Résumé. ...

1991 Mathematics Subject Classification. Primary 35L81, 65M12; Secondary 35L65, 65M08.

March 31, 2016.

1. Introduction

Singular limits of conservation laws (or more generally PDEs), may present severe difficulties to be treated either in analysis or numerics. The main issue is that the type of the equations changes in the limit [40], e.g., when the Mach number approaches zero for the Euler equations. This limit is singular, since the sound speed (the characteristic speed) goes to the infinity and the PDE changes to be hyperbolic-elliptic, in the so-called incompressible limit. So, there are difficulties to show the convergence of the solution of compressible Euler equations to the incompressible one (see [34,40]). Tackling this problem numerically is more complicated, since as the eigenvalues of the flux Jacobian blow up, the time step should tend to zero due to Courant–Friedrichs–Lewy (CFL) condition, which leads to very small time steps and thus huge computational cost. Also it has been shown that in the general case, the usual numerical schemes lose their accuracy in the limit for under-resolved mesh sizes; see [19,21,25,26,44–46].

Keywords and phrases: All-Mach number scheme, Lagrange–projection scheme, asymptotic preserving scheme, stability analysis

* *The author was supported by the scholarship of RWTH Aachen university through Graduiertenförderung nach Richtlinien zur Förderung des wissenschaftlichen Nachwuchses (RFwN)*

[1] Institut für Geometrie und Praktische Mathematik, RWTH Aachen
Templergraben 55, 52056, Aachen, Germany e-mail: `h.zakerzadeh@igpm.rwth-aachen.de`

$$\begin{tikzcd}
\mathcal{M}^\epsilon_\Delta \arrow[r, "\epsilon \to 0"] \arrow[d, "\Delta \to 0"'] & \mathcal{M}^0_\Delta \arrow[d, "\Delta \to 0"] \\
\mathcal{M}^\epsilon \arrow[r, "\epsilon \to 0"'] & \mathcal{M}^0
\end{tikzcd}$$

Figure 1. Illustration of Asymptotic Preserving schemes.

Throughout this paper, we assume that at least at the continuous level, the *solution* of compressible flow equations with the Mach number ϵ, converges to the solution of the limit equation, as $\epsilon \to 0$, and try to show that the counterpart of such a convergence also exists at the discrete level. This is in fact the idea of Asymptotic Preserving (AP) schemes, which has been introduced by Jin in [30, 31] for relaxation systems; see also [32] for a general review and [36] for older works. Figure 1 illustrates this definition; $\mathcal{M}^\epsilon$ stands for a continuous physical model with the (singular) perturbation parameter $\epsilon \in (0, 1]$, and $\mathcal{M}^\epsilon_\Delta$ is a discrete-level model which provides a consistent discretization of $\mathcal{M}^\epsilon$. As in [32], if $\mathcal{M}^0_\Delta$ is a *suitable* and *efficient* scheme for $\mathcal{M}^0$, then the scheme is called to be AP.

However, note that there are several definitions of AP schemes in different contexts owing to different interpretations of a *suitable* and *efficient* scheme. So we define an AP scheme for the framework of this article more precisely.

Definition 1.1. [AP schemes] A scheme is called to be AP, provided that it fulfills the following conditions.

(i) It gives a consistent discretization of $\mathcal{M}^\epsilon$ for all ϵ, in particular for the limit problem $\mathcal{M}^0$.
(ii) It is efficient uniformly in ϵ, e.g., the CFL condition should be uniform in ϵ and the implicit step should be solved efficiently for all ϵ.
(iii) It is stable in some suitable sense, uniformly in ϵ.

For brevity, we call these properties respectively Asymptotic Consistency (AC), Asymptotic Efficiency (AEf), and Asymptotic Stability (AS).

The AP property has been studied widely and several AP (in our terminology AC and AEf) schemes have been developed for the Euler or shallow water equations; see [4, 15, 18, 27, 42]. The bottom line of these schemes is a mixed implicit-explicit (IMEX) approach, stems from the more general *operator splitting* method; to split the flux (or its Jacobian) into two parts and treat one part explicitly in time and the other one implicitly in time. This approach is definitely necessary to find schemes with ϵ-uniform CFL conditions. But as mentioned in [17], it is not sufficient at all to claim for asymptotic stability; see for example [1] whence it is shown that for an explicit-explicit splitting with Lax–Wendroff scheme, even if both split parts are stable in terms of CFL condition, the resulting scheme is unconditionally unstable in L_2-norm using von Neumann stability analysis. On the other hand, it is shown in [27] that IMEX schemes are L_2-stable, as long as each step is L_2-stable. So, there is a huge gap between these two cases. Note that using IMEX splitting schemes makes the analysis more delicate compared to explicit splittings; see [8, 9] for some results on the Lagrange-projection scheme (which is also the topic of this paper) and [47] for a motivating study on linear systems.

Recently, inspired by the Arbitrary Lagrangian-Eulerian (ALE) approach, there have been some works devoted to the so-called Lagrange–projection scheme, which have resulted in some rigorous stability results; see [8, 9, 11, 14] for example. ALE nowadays is a classic approach in mechanics, trying to benefit from the advantages of Eulerian and Lagrangian formulations simultaneously; see [22] for a nice introduction, and this is the heart of the Lagrange–projection scheme, as we will see later on in Section 2. However, most of these works do not take the asymptotic limit into account. For example in [14], a rigorous numerical analysis has been done for a two-phase model including positivity of the density and entropy stability, but with no concern about the incompressible limit. Another study was carried over later for balance laws by Chalons et. al [8], in particular for the Euler equations with friction. Moreover, in [9], the Lagrange–projection scheme has been analyzed for the two-dimensional Euler equations, to construct an all-Mach scheme (the scheme with Mach-uniform consistency). So the focus is on the accuracy problems in the low-Mach regime, which one expects to see for Godunov-type schemes (of which Lagrange–projection scheme is a member), and to cure them by a careful look at the truncation error. In fact, it has been shown in [9] that the truncation error of the two-dimensional Lagrange-projection scheme blows up in the low Mach regime, i.e., it behaves as $\mathcal{O}(\frac{\Delta x}{Ma})$ where Ma stands for the Mach number. The authors of [9] could show that the truncation error can be made uniform regarding the Mach number for a particular modification of the scheme, namely by multiplying the dissipation involved in the discretization of the pressure terms by an $\mathcal{O}(Ma)$ term . Although this is a promising step, it is not clear if this uniform accuracy in terms of the truncation error is equivalent to the asymptotic consistency, due to the lack of convergence analysis of the scheme. Very recently in [10], the authors have extended the Lagrange–projection framework to the one-dimensional shallow water equations with particular attention to the well-balancing and the validity of the entropy inequality.

On the other hand, it is well-known that Godunov-type schemes show no accuracy problem for low-Mach one-dimensional problems as long as the initial condition is *well-prepared* (see Definition 3.2). The reader can consult with [9, 19, 21, 45, 46] for more details. This accuracy of Godunov-type schemes motivates the present paper, whose goal is to investigate the results of [14] regarding the Mach number. In this paper, we study the issue of consistency and stability of one-dimensional IMEX Lagrange-projection scheme, or the so-called LP-IMEX scheme as been proposed in [8], in the incompressible limit. In particular, we show that the stability conditions in [14] are uniform in the Mach number provided that the initial condition is well-prepared. So, all the stability properties in [14] hold without any restriction regarding the Mach number. Also we show that the solution is asymptotically consistent for well-prepared initial data (see Theorem 3.3). Indeed, these estimates imply convergence of a sub-sequence for fixed grids, as ϵ tends to zero (see Appendix A). The study has also been extended to the one-dimensional shallow water equations with topography. The source term presents an additional difficulty, and in order to prove asymptotic consistency, we had to use a more dissipative relaxation parameter (see Remark 4.9). Also note that in Sections 3.2 and 4.3 we prove ϵ-uniform bounds for the (implicit) solution, which justifies the asymptotic expansions used throughout the paper.

The paper is organized as follows. In Section 2 we introduce the splitting after a brief introduction to the ALE formalism and relaxation schemes. Then, in Section 3, we introduce the IMEX Lagrange–projection scheme with a specific relaxation approximation and then discuss the numerical analysis of the scheme. We prove the formal asymptotic consistency, positivity preserving, stability and entropy stability, all under a non-restrictive CFL condition. Then we show that the formal asymptotic consistency is in fact rigorous. In Section 4 we show similar results for the shallow water equations with a different and more diffusive relaxation approximation in the case of flat or non-flat bottom functions. Appendix A provides some results about the consequences of entropy stability on the stability of the solution in the limit $\epsilon \to 0$. We then conclude the discussion with some possible extensions and future works.

2. Lagrange–projection scheme: Continuous PDE level

In this section, we introduce the splitting to be used, let us call it *ALE splitting*, which is inspired by the classical Lagrange–projection scheme (see [23]). For the barotropic Euler equations, one natural way to split the waves, is to split them into acoustic and transport waves. The high speed acoustic waves are formulated in the Lagrangian framework and slow transport waves in Eulerian one. The idea is similar as for the Lagrange–projection scheme, which consists of solving Riemann problems for the acoustic system in the Lagrangian formulation and then projecting the computed solution onto the fixed Eulerian grid (which is equivalent to the transport system). In this way, the scheme handles Riemann problems in the Lagrangian coordinates which is much easier than the Eulerian one, and takes advantage of using a fixed grid; see [23, Chapter III, Section 2.5]. It is in this regard that the ALE splitting (as well as the Lagrange–projection scheme) can be understood (see [14]) in the framework of ALE, to write the equations in the *referential* coordinates χ which are necessarily neither spatial (Eulerian) x nor material (Lagrangian) X. The referential frame has a relative velocity v seen from the spatial frame, which is arbitrarily chosen. Note that the Lagrange–projection scheme is a special case of ALE, in which the velocity v is chosen such that after completing each step, the domain is the same as the fixed Eulerian one. We refer the reader to [14, Section 3.3] for more details.

Now, consider the system of barotropic Euler equations:

$$\partial_t \rho + \partial_x(\rho u) = 0, \tag{2.1}$$
$$\partial_t(\rho u) + \partial_x\left(\rho u^2 + p(\rho)\right) = 0, \tag{2.2}$$

when $p(\rho) = \kappa\rho^\gamma$ with $\kappa > 0$ and $\gamma > 1$ is the barotropic pressure law. For the rest of this paper, we set $k = \frac{1}{2}$ and $\gamma = 2$. As an entropy function, we choose the total energy of the solution ρE which can be shown to be strictly convex with respect to the conservative variables. The total energy density is written as $E = \mathcal{E} + \frac{u^2}{2}$ where $\mathcal{E}(\rho) := \frac{\kappa}{\gamma-1}\rho^{\gamma-1}$ is the internal energy density (see [38]).

The ALE splitting splits the original system, (2.1)-(2.2) into the following acoustic and transport sub-systems

$$\partial_t \rho + \rho\partial_x u = 0, \tag{2.3}$$
$$\partial_t(\rho u) + \rho u \partial_x u + \partial_x p = 0, \tag{2.4}$$

and

$$\partial_t \rho + u\partial_x \rho = 0, \tag{2.5}$$
$$\partial_t(\rho u) + u\partial_x(\rho u) = 0, \tag{2.6}$$

and solve them successively. Simply by using Taylor expansion it can be seen that this splitting is in general (globally) first-order accurate in time. We refer the reader to [28] for more details about the operator splitting methods. Note that the transport part is simply a transport of conservative variables $(\rho, \rho u)$ with the velocity field u.

2.1. **Lagrange step**

In the Lagrangian coordinates, the frame moves with the velocity field. So, what the observer sees is the acoustic part, (2.3)-(2.4). It is not difficult to show that they can also be written as

$$\partial_t \tau - \partial_z u = 0, \tag{2.7}$$
$$\partial_t u + \partial_z p = 0, \tag{2.8}$$

where τ is specific volume (the reciprocal of ρ) and $z := \rho \mathrm{d}x$ is the mass coordinate. This is exactly the classical form of the barotropic Euler equations in the Lagrangian framework. To obtain the *non-dimensionalized* equations, we set

$$\widehat{\tau} := \tau\rho_\circ, \quad \widehat{u} := \frac{u}{u_\circ}, \quad \widehat{p} := \frac{p}{p_\circ}, \quad p_\circ := \rho_\circ c_\circ^2,$$

where $c_\circ$ is the characteristic sound speed, defined as $c_\circ := \sqrt{\kappa\gamma\rho_\circ^{\gamma-1}}$, and $\rho_\circ$ and $u_\circ$ are characteristic density and velocity. Also we denote the Mach number as the ratio of characteristic speed to the sound speed, $Ma := \frac{u_\circ}{c_\circ}$. Thus, after suppressing hats, the equations become

$$\partial_t \tau - \partial_z u = 0, \tag{2.9}$$
$$\partial_t u + \frac{1}{\epsilon^2}\partial_z p_z = 0, \tag{2.10}$$

where $\epsilon := \sqrt{\gamma}Ma$. From now on and for simplicity, we call ϵ the Mach number, though it is different from Ma by the factor $\sqrt{\gamma}$.

Assuming a one-dimensional torus as the spatial domain $\Omega := \mathbb{T}$, which means the periodic boundaries, we define the domain of solutions as $\Omega_T := \mathbb{T} \times \mathbb{R}_+$ in space and time. So, we are left with a Cauchy initial value problem which needs solution of the Riemann problems. To ease the situation, we relax the system so that all characteristic fields would be linearly degenerate, which is easy to solve the Riemann problem for. We actually substitute the source of genuine nonlinearity $p(\rho)$ with some variable π, called relaxation pressure and add another equation for π. This is the heart of so-called *relaxation schemes*; we refer the reader to [5,12,33,39] for more details.

In the non-dimensionalized form, the Suliciu relaxation system [5,13] reads as

$$\partial_t \tau - \partial_z u = 0, \tag{2.11}$$

$$\partial_t u + \partial_z \Pi = 0, \tag{2.12}$$

$$\partial_t \Pi + \alpha^2 \partial_z u = \Lambda(p - \pi), \tag{2.13}$$

with the definitions

$$\Pi_\epsilon := \Pi := \frac{\pi}{\epsilon^2}, \qquad \alpha_\epsilon := \alpha := \frac{a}{\epsilon}, \qquad \Lambda_\epsilon := \Lambda := \frac{\lambda}{\epsilon^2},$$

where a is a constant to be specified and λ is the relaxation parameter.

Remark 2.1. One can choose α differently, for example as $\frac{a}{\epsilon^2}$, which is more diffusive compared to the current α. For now, we stick to $\frac{a}{\epsilon}$, and later on in Section 4 for the shallow water equations, we use the diffusive relaxation system.

At least formally, one can observe that in the asymptotic regime $\lambda \to \infty$, π tends to p and the original system would be recovered. Now, one can easily check that the relaxation system only has linearly-degenerate characteristic fields. To use the feature of linear degeneracy, at first we solve the problem out of equilibrium, setting $\lambda = 0$, and then we project the out-of-equilibrium solution to the equilibrium manifold, cf. [14].

In order to prevent the instabilities to happen for this relaxation system, or in other words to enforce the dissipativity of Chapman–Enskog expansion (see [12,39]), the parameter a must be chosen sufficiently large according to the so-called *sub-characteristic* or *Whitham stability* condition

$$\alpha^2 > \frac{\max(-p_\tau)}{\epsilon^2}, \tag{2.14}$$

see [7] for the proof.

Since the relaxation system with $\lambda = 0$ is strictly hyperbolic with eigenvalues given by $0, \pm a$—compared to exact eigenvalues $0, \pm c$ for the original system—the sub-characteristic condition means that information propagates faster in the relaxation model. Also linear degeneracy of the fields allows us to analytically solve the Riemann problem when $\lambda = 0$. This property justifies by itself the introduction of the proposed relaxation model and its simplicity [9].

So, for $\lambda = 0$, one can simply put the relaxation system (2.11)-(2.13) into an equivalent diagonal form like [8, eq. (12)]

$$\tau_t - u_z = 0, \tag{2.15}$$

$$\overrightarrow{w}_t + \alpha \overrightarrow{w}_z = 0, \quad \overrightarrow{w} := \Pi + \alpha u = \frac{\pi}{\epsilon^2} + \frac{a}{\epsilon} u, \tag{2.16}$$

$$\overleftarrow{w}_t - \alpha \overleftarrow{w}_z = 0, \quad \overleftarrow{w} := \Pi - \alpha u = \frac{\pi}{\epsilon^2} - \frac{a}{\epsilon} u. \tag{2.17}$$

Note that $\overrightarrow{w}$ and $\overleftarrow{w}$ are two of Riemann invariants of the relaxation system; the third one is $\mathscr{I} := \Pi + \alpha^2 \tau$. So, instead of the (2.15) one can use $\mathscr{I}_t = 0$.

Remark 2.2. Note that the naturally-split systems (2.3)-(2.4) and (2.5)-(2.6) are not conservative if they are written in the Eulerian coordinates, so to circumvent the complications coming with non-conservative products (see [16] for example) and also for solving Riemann problems with more ease and efficiency, we have changed the coordinates to Lagrangian, which provides a conservative formulation.

2.2. **Projection step**

Like the acoustic part, the transport part (2.5)-(2.6) can be written in the Lagrangian coordinates which provides a conservative form. In fact, since we remap the values onto the Eulerian grid, at the end of each step, the referential and spatial (Eulerian) coordinate should coincide. So following the notation in [14], the projection step can be summarized as

$$\partial_t \mathbb{U} + u\partial_x \mathbb{U} = 0, \tag{2.18}$$

where $\mathbb{U} := (\rho, \rho u)^T$ stands for the conservative variables. For further details on the derivation of the split systems, the reader can consult with [14]

3. Lagrange–projection scheme: Discrete numerical level

As mentioned above, for linearly-degenerate systems, it is straightforward to solve the Riemann problem. Moreover in this case, after writing the equations in terms of Riemann invariants, it would be in fact trivial since along each characteristic line, one of the Riemann invariants remains constant. In this way, there are just a set of two symmetric scalar linear advection equations to be solved for $\overrightarrow{w}$ and $\overleftarrow{w}$, while $\mathscr{I}$ does not change at all.

At the beginning of the Lagrange (acoustic) step from n to $n\dagger$, the Eulerian and Lagrangian coordinates coincide with each other, also $p_j^n = \pi_j^n$. The implicit Lagrange step reads as

$$\tau_j^{n\dagger} = \tau_j^n + \frac{\Delta t}{\Delta z_j^n}\left(\tilde{u}_{j+1/2}^{n\dagger} - \tilde{u}_{j-1/2}^{n\dagger}\right). \tag{3.1}$$

$$\overrightarrow{w}_j^{n\dagger} = \overrightarrow{w}_j^n - \frac{a\Delta t}{\epsilon\Delta z_j^n}\left(\overrightarrow{w}_j^{n\dagger} - \overrightarrow{w}_{j-1}^{n\dagger}\right), \tag{3.2}$$

$$\overleftarrow{w}_j^{n\dagger} = \overleftarrow{w}_j^n + \frac{a\Delta t}{\epsilon\Delta z_j^n}\left(\overleftarrow{w}_{j+1}^{n\dagger} - \overleftarrow{w}_j^{n\dagger}\right), \tag{3.3}$$

where $\Delta z^n := \rho_j^n \Delta x$ and $j \in \mathbb{S}$ denotes cell indices when $\mathbb{S}$ is a periodic set (the discretization of Ω). Also $\tilde{u}^{n\dagger}$ comes from solving a simple Riemann problem for the relaxation system with characteristic $0, \pm\frac{a}{\epsilon}$ (see [14]), and it is

$$\tilde{u}^{n\dagger} := \frac{1}{2a\epsilon}\left(a\epsilon(u_L + u_R) - (\pi_R - \pi_L)\right).$$

So, the interface velocity is defined as

$$\tilde{u}_{j+1/2}^{n\dagger} := \frac{u_j^{n\dagger} + u_{j+1}^{n\dagger}}{2} - \frac{1}{2a\epsilon}\left(\pi_{j+1}^{n\dagger} - \pi_j^{n\dagger}\right). \tag{3.4}$$

Note that there are several (equivalent) variants of the scheme (3.1)-(3.3), in different coordinates or with/without using the Riemann invariants; see [14] for further details.

In the next step, the explicit projection step from $n\dagger$ to $n+1$, we map updated values onto the fixed Eulerian grid. There are 4 cases based on the upwind direction [8, eq. (34)], which can be summarized as

$$\mathbb{U}_j^{n+1} = \mathbb{U}_j^{n\dagger} + \frac{\Delta t}{\Delta x}\left[(\tilde{u}_{j-1/2}^{n\dagger})^+\mathbb{U}_{j-1}^{n\dagger} + \left((\tilde{u}_{j+1/2}^{n\dagger})^- - (\tilde{u}_{j-1/2}^{n\dagger})^+\right)\mathbb{U}_j^{n\dagger} - (\tilde{u}_{j+1/2}^{n\dagger})^-\mathbb{U}_{j+1}^{n\dagger}\right], \tag{3.5}$$

with the definitions $\cdot^+ := \frac{\cdot + |\cdot|}{2}$ and $\cdot^- := \frac{\cdot - |\cdot|}{2}$.

This step, projects the values on the updated grid onto the fixed Eulerian one. Adding these two steps to each other is what we call LP-IMEX scheme.

3.1. Numerical analysis of LP-IMEX scheme

Considering the LP-IMEX scheme introduced in the previous section, one can obtain the stability results, gathered in Theorem 3.3. But at first, let us define the limit of barotropic Euler equations and the so-called well-prepared initial datum.

Definition 3.1. The formal limit of barotropic Euler equations is defined to be

$$\begin{aligned}
&\rho_{(0)} = \text{const.}, \qquad \rho_{(1)} = \text{const.},\\
&\operatorname{div} u_{(0)} = 0,\\
&\partial_t u_{(0)} + \partial_x \left(u_{(0)}^2 + p_{(2)} \right) = 0,
\end{aligned}$$

where the asymptotic expansion is

$$\begin{aligned}
\rho(x,t) &= \rho_{(0)} + \epsilon \rho_{(1)} + \epsilon^2 \rho_{(2)},\\
u(x,t) &= u_{(0)} + \epsilon u_{(1)} + \epsilon^2 u_{(2)}.
\end{aligned}$$

Definition 3.2. For the barotropic Euler equations (2.1)–(2.2), the well-prepared initial datum is defined as (see [21, 34, 35, 41])

$$\pi_{WP}^{(0)}(x) = p_{WP}^{(0)}(x) = p_{(0)} + \epsilon^2 p_{(2)}(x), \tag{3.6}$$

$$u_{WP}^{(0)}(x) = u_{(0)} + \epsilon u_{(1)}(x), \tag{3.7}$$

with constant p_0 and u_0.

Theorem 3.3. *The Lagrange–projection scheme* (3.1)–(3.3) *and* (3.5)*, with well-prepared initial datum, satisfies the following properties.*

(i) It can be expressed in the locally conservative form.

(ii) The scheme is AC, which means that it gives a consistent discretization of incompressible Euler equations in terms of Definition 3.1.

(iii) Under some non-ϵ-restrictive CFL constraint (3.20)*, the scheme is positivity preserving, i.e., $\rho_j^n > 0$ provided that $\rho_j^n > 0$ for all $j \in \mathbb{S}$. Moreover the density is bounded away from zero, i.e., there exists some $\rho_{LB}^n > 0$ such that $\rho_j^n \geq \rho_{LB}^n$ for all $j \in \mathbb{S}$.*

(iv) Under (3.20) *and sub-characteristic condition* (3.32)*, the solution fulfills the local (cell) entropy (energy) inequality, i.e.,*

$$\frac{(\rho E)_j^{n+1} - (\rho E)_j^n}{\Delta t} + \frac{\left(\rho E\tilde{u} + \dfrac{\tilde{\pi}\tilde{u}}{\epsilon^2}\right)_{j+1/2}^{n\dagger} - \left(\rho E\tilde{u} + \dfrac{\tilde{\pi}\tilde{u}}{\epsilon^2}\right)_{j-1/2}^{n\dagger}}{\Delta x} \leq 0, \tag{3.8}$$

which is consistent with

$$\partial_t (\rho E) + \partial_x \left((\rho E + \frac{p}{\epsilon^2}) u \right) \leq 0 \quad \text{in } \mathcal{D}'(\Omega_T).$$

(v) Under (3.20) *and sub-characteristic condition* (3.32)*, the computed density, momentum and velocity are stable, i.e., bounded in ℓ_∞-norm, uniformly in ϵ.*

We analyze the properties of this scheme in the subsequent subsections. Note that the locally conservative form of the scheme is proved in [14].

Remark 3.4. Throughout this section and the subsequent one, it is very natural to ask about the order of magnitudes of quantities (in terms of ϵ). For now, we only do the analysis formally; i.e., we only take the explicit ϵ into account and assume that all other quantities are $\mathcal{O}(1)$. But in a separate section, Section 3.2, we justify this assumption.

3.1.1. *Proof of asymptotic consistency (ii)*

At first, we show that the solution is close to the incompressible limit in the sense of Definition 3.1, i.e, constant density up to the second order of asymptotic expansions, and a divergence free (solenoidal) zeroth-order velocity component. Then, using these results, we prove that the scheme provides a consistent discretization of the PDE in the limit $\epsilon \to 0$. Thus, the asymptotic consistency in the sense of Definition 1.1 holds.

Considering Definitions 3.1 and 3.2, we consider a well-prepared solution at the step n, i.e.,

$$\rho_j^n = \rho_{0c}^n + \epsilon^2 \rho_{(2)j}^n, \tag{3.9}$$

$$p_j^n = \pi_{0c}^n + \epsilon^2 \pi_{(2)j}^n, \tag{3.10}$$

$$u_j^n = u_{0c}^n + \epsilon u_{(1)j}^n, \tag{3.11}$$

where ρ_{0c}^n, π_{0c}^n and u_{0c}^n are constant values. Here, we want to show that the scheme (3.1)–(3.3) preserves the well-preparedness of the solution at the step n to the intermediate step $n\dagger$ and then to the next time step $n+1$.

For the Lagrange step, we start with the $\mathcal{O}(1/\epsilon)$ terms in the mass equation, which yield

$$\pi_{(0)j+1}^{n\dagger} - 2\pi_{(0)j}^{n\dagger} + \pi_{(0)j-1}^{n\dagger} = 0,$$

thus $\pi_{(0)j}^{n\dagger}$ is a linear sequence over $j \in \mathbb{S}$ and due to periodic B.C. $\pi_{(0)j}^{n\dagger} = \pi_{(0)}^{n\dagger}$ which is constant in space.

Since π and ρ are two independent variables at this level, we cannot conclude immediately that the same is true for the density. But one can find their relation by combining (3.1)–(3.3) to find the update for the relaxation pressure as

$$\rho_j^n \frac{\pi_j^{n\dagger} - \pi_j^n}{\Delta t} + \frac{a^2}{\Delta x}\left(\tilde{u}_{j+1/2}^{n\dagger} - \tilde{u}_{j-1/2}^{n\dagger}\right) = 0. \tag{3.12}$$

Then using continuity (3.1), it yields

$$a^2\left(\tau_j^{n\dagger} - \tau_j^n\right) + \left(\pi_j^{n\dagger} - \pi_j^n\right) = 0. \tag{3.13}$$

From (3.13) it is clear that

$$a^2\big(\rho_j^{n\dagger} - \rho_j^n\big) = \rho_j^n \rho_j^{n\dagger}\big(\pi_j^{n\dagger} - \pi_j^n\big). \tag{3.14}$$

So, its $\mathcal{O}(1)$ part is

$$a^2\big(\rho_{(0)j}^{n\dagger} - \rho_{(0)j}^n\big) = \rho_{(0)j}^n \rho_{(0)j}^{n\dagger}\big(\pi_{(0)j}^{n\dagger} - \pi_{(0)j}^n\big),$$

which gives that

$$\rho_{(0)j}^{n\dagger}\left(a^2 - \rho_{0c}^n\big(\pi_{(0)}^{n\dagger} - \pi_{0c}^n\big)\right) = a^2 \rho_{0c}^n \Longrightarrow \rho_{(0)j}^{n\dagger} = \rho_{(0)}^{n\dagger} \text{ const. in space}$$

Then, due to periodic B.C. and by a spatial summation on (3.1), it can be found out that $\rho^{n\dagger}_{(0)j}$ is constant in time as well, i.e., $\rho^{n\dagger}_{(0)j} = \rho^{n}_{0c}$. Also from the update of relaxation pressure π, (3.12), and again periodic B.C. and spatial summation, the numerical fluxes cancel out with each other and it turns out that $\pi^{n\dagger}_{(0)j} = \pi^{n}_{0c}$, constant in both time and space.

Next, we continue with momentum equation (there is no difference between $\overrightarrow{w}_j$ and $\overleftarrow{w}_j$ in this regard).

$$\rho^n_j\big(\frac{\pi^{n\dagger}_j}{\epsilon^2} + \frac{a}{\epsilon}u^{n\dagger}_j\big) = \rho^n_j\big(\frac{\pi^{n}_j}{\epsilon^2} + \frac{a}{\epsilon}u^{n}_j\big) - \frac{a\Delta t}{\epsilon^2\Delta x}\Big(\frac{\pi^{n\dagger}_j - \pi^{n\dagger}_{j-1}}{\epsilon} + a\big(u^{n\dagger}_j - u^{n\dagger}_{j-1}\big)\Big).$$

So if one balances $\mathcal{O}(1/\epsilon^2)$ terms, one obtains

$$\rho^n_{0c}\pi^{n\dagger}_{(0)j} = \rho^n_{0c}\pi^n_{0c} - \frac{a\Delta t}{\Delta x}\Big(\pi^{n\dagger}_{(1)j} - \pi^{n\dagger}_{(1)j-1} + a\big(u^{n\dagger}_{(0)j} - u^{n\dagger}_{(0)j-1}\big)\Big),$$

which yields

$$\pi^{n\dagger}_{(1)j} - \pi^{n\dagger}_{(1)j-1} + a\big(u^{n\dagger}_{(0)j} - u^{n\dagger}_{(0)j-1}\big) = 0. \tag{3.15}$$

So, there is the possibility that both $\pi^{n\dagger}_{(1)j}$ and $u^{n\dagger}_{(0)j}$ be constant in space. To show it, note that from $\mathcal{O}(1)$ terms in continuity equation, one gets

$$\begin{aligned}\rho^n_{(0)j} =& \rho^{n\dagger}_{(0)j}\bigg(1 + \frac{\Delta t}{2a\Delta x}\Big(a\big(u^{n\dagger}_{(0)j+1} - u^{n\dagger}_{(0)j-1}\big) - \big(\pi^{n\dagger}_{(1)j-1} - 2\pi^{n\dagger}_{(1)j} + \pi^{n\dagger}_{(1)j+1}\big)\Big)\bigg)\\ &- \frac{\Delta t}{2a\Delta x}\rho^{n\dagger}_{(1)j}\big(\pi^{n\dagger}_{(0)j-1} - 2\pi^{n\dagger}_{(0)j} + \pi^{n\dagger}_{(0)j+1}\big).\end{aligned}$$

So,

$$a\big(u^{n\dagger}_{(0)j+1} - u^{n\dagger}_{(0)j-1}\big) - \big(\pi^{n\dagger}_{(1)j-1} - 2\pi^{n\dagger}_{(1)j} + \pi^{n\dagger}_{(1)j+1}\big) = 0 \tag{3.16}$$

Combining (3.16) and (3.15) yields that $\pi^{n\dagger}_{(1)j} = \pi^{n\dagger}_{(1)}$ and $u^{n\dagger}_{(0)j} = u^{n\dagger}_{(0)}$. So, $\operatorname{div} u^{n\dagger}_{(0)} = 0$.

Again, similar to the zeroth order, one can show that $\rho^{n\dagger}_{(1)j}$ is constant in space and even $\pi^{n\dagger}_{(1)j}$ and $\rho^{n\dagger}_{(1)j}$ are constant in time, i.e., $\pi^{n\dagger}_{(1)j} = \pi^n_{1c}$ and $\rho^{n\dagger}_{(1)j} = \rho^n_{1c}$. Hence, the solution of the Lagrange step is close to the incompressible limit.

For the projection step, we show asymptotic consistency for the first case, $\tilde{u}^{n\dagger}_{j-1/2} < 0$ and $\tilde{u}^{n\dagger}_{j+1/2} < 0$. The other cases can be done in a very similar way. In this case, it can be seen that

$$\rho^{n+1}_j = \rho^{n\dagger}_j - \frac{\Delta t}{2a\Delta x}\big(\rho^{n\dagger}_{j+1} - \rho^{n\dagger}_j\big)\Big(-\frac{\pi^{n\dagger}_{j+1} - \pi^{n\dagger}_j}{\epsilon} + a\big(u^{n\dagger}_{j+1} + u^{n\dagger}_j\big)\Big).$$

So $\mathcal{O}(1)$ terms give,

$$\begin{aligned}\rho^{n+1}_{(0)j} = \rho^{n\dagger}_{(0)j} - \frac{\Delta t}{2a\Delta x}\Big[&-\big(\rho^{n\dagger}_{(0)j+1} - \rho^{n\dagger}_{(0)j}\big)\big(\pi^{n\dagger}_{(1)j+1} - \pi^{n\dagger}_{(1)j}\big)\\ &-\big(\rho^{n\dagger}_{(1)j+1} - \rho^{n\dagger}_{(1)j}\big)\big(\pi^{n\dagger}_{(0)j+1} - \pi^{n\dagger}_{(0)j}\big)\\ &+a\big(\rho^{n\dagger}_{(0)j+1} - \rho^{n\dagger}_{(0)j}\big)\big(u^{n\dagger}_{(0)j+1} + u^{n\dagger}_{(0)j}\big)\Big],\end{aligned}$$

thus $\rho^{n+1}_{(0)j} = \rho^{n\dagger}_{(0)j} = \rho^n_{0c}$ and as a result $p^{n+1}_{(0)j} = p^n_{0c}$; it is constant as well. Similarly, one can find that the first order components are also constant in time and space; if they do not exist in the initial condition, so at the time t_{n+1} there is no $\mathcal{O}(\epsilon)$ pressure fluctuation:

$$\begin{aligned}\rho^{n+1}_{(1)j} = \rho^{n\dagger}_{(1)j} - \frac{\Delta t}{2a\Delta x}\Big[&- \big(\rho^{n\dagger}_{(0)j+1} - \rho^{n\dagger}_{(0)j}\big)\big(\pi^{n\dagger}_{(2)j+1} - \pi^{n\dagger}_{(2)j}\big) \\ &- \big(\rho^{n\dagger}_{(1)j+1} - \rho^{n\dagger}_{(1)j}\big)\big(\pi^{n\dagger}_{(1)j+1} - \pi^{n\dagger}_{(1)j}\big) \\ &- \big(\rho^{n\dagger}_{(2)j+1} - \rho^{n\dagger}_{(2)j}\big)\big(\pi^{n\dagger}_{(0)j+1} - \pi^{n\dagger}_{(0)j}\big) \\ &+ a\big(\rho^{n\dagger}_{(0)j+1} - \rho^{n\dagger}_{(0)j}\big)\big(u^{n\dagger}_{(1)j+1} + u^{n\dagger}_{(1)j}\big) \\ &+ a\big(\rho^{n\dagger}_{(1)j+1} - \rho^{n\dagger}_{(1)j}\big)\big(u^{n\dagger}_{(0)j+1} + u^{n\dagger}_{(0)j}\big)\Big],\end{aligned}$$

and $\rho^{n+1}_{(1)j} = \rho^{n\dagger}_{(1)j} = \rho^n_{0c} = 0$.

To show the div-free condition, one can consider $\mathcal{O}(1)$ terms of the momentum equation:

$$\begin{aligned}\rho^{n+1}_{(0)j}u^{n+1}_{(0)j} = \rho^{n\dagger}_{(0)j}u^{n\dagger}_{(0)j} - \frac{\Delta t}{2a\Delta x}\Big[&- \big(\rho^{n\dagger}_{(0)j+1}u^{n\dagger}_{(0)j+1} - \rho^{n\dagger}_{(0)j}u^{n\dagger}_{(0)j}\big)\big(\pi^{n\dagger}_{(1)j+1} - \pi^{n\dagger}_{(1)j}\big) \\ &- \big(\rho^{n\dagger}_{(1)j+1}u^{n\dagger}_{(0)j+1} - \rho^{n\dagger}_{(1)j}u^{n\dagger}_{(0)j}\big)\big(\pi^{n\dagger}_{(0)j+1} - \pi^{n\dagger}_{(0)j}\big) \\ &- \big(\rho^{n\dagger}_{(0)j+1}u^{n\dagger}_{(1)j+1} - \rho^{n\dagger}_{(0)j}u^{n\dagger}_{(1)j}\big)\big(\pi^{n\dagger}_{(0)j+1} - \pi^{n\dagger}_{(0)j}\big) \\ &+ a\big(\rho^{n\dagger}_{(0)j+1}u^{n\dagger}_{(0)j+1} - \rho^{n\dagger}_{(0)j}u^{n\dagger}_{(0)j}\big)\big(u^{n\dagger}_{(0)j+1} + u^{n\dagger}_{(0)j}\big)\Big].\end{aligned}$$

Thus $u^{n+1}_{(0)j} = u^{n\dagger}_{(0)j} = u^n_{0c}$, and the zeroth order component of velocity filed is solenoidal. Hence, combining the results for the Lagrange and projection steps together, it is obvious that the limit properties is satisfied.

To prove asymptotic consistency in the sense of Definition 1.1, it remains to show the consistency of the discretization in the limit. For the Lagrange step the consistency holds if the velocity update

$$\frac{u^{n\dagger}_j - u^n_j}{\Delta t} + \frac{1}{2\Delta z^n_j \epsilon^2}\left(\pi^{n\dagger}_{j+1} - \pi^{n\dagger}_{j-1}\right) - \frac{a/\epsilon}{2\Delta z^n_j}\left(u^{n\dagger}_{j+1} - 2u^{n\dagger}_j + u^{n\dagger}_{j-1}\right) = 0, \tag{3.17}$$

is a consistent discretization of $\partial_t u + \frac{1}{\epsilon^2}\partial_z \pi = 0$ in the limit, when (3.17) gives

$$\frac{u^{n\dagger}_{(0)j} - u^n_{(0)j}}{\Delta t} + \frac{1}{2\Delta z^n_{(0)j}}\left(\pi^{n\dagger}_{(2)j+1} - \pi^{n\dagger}_{(2)j-1}\right) - \frac{a}{2\Delta z^n_{(0)j}}\left(u^{n\dagger}_{(1)j+1} - 2u^{n\dagger}_{(1)j} + u^{n\dagger}_{(1)j-1}\right) = 0. \tag{3.18}$$

It is clear that (3.18) is the Rusanov scheme applied to $\partial_t u_{(0)} + \partial_z \pi_{(2)} = 0$, so the Lagrange step is AC.

To show the consistency of the discretization in the limit also for the projection step, comparing (2.18) and (3.5), it is sufficient to confirm that $\tilde{u}^{n\dagger}_{(0)j+1/2}$ is consistent with $u_{(0)}$. This is in fact the case, due to the definition of $\tilde{u}^{n\dagger}_{j+1/2}$ in (3.4) and the asymptotic behavior of $u^{n\dagger}_{(0)j}$ and $\pi^{n\dagger}_{(1)j}$, namely that $u^{n\dagger}_{(0)j}$ and $\pi^{n\dagger}_{(1)j}$ are constant in space. So, the projection step (3.5) is a consistent discretization of (2.18). Hence the scheme is AC in the sense of Definition 1.1.

3.1.2. *Proof of density positivity (iii)*

In this section, we show that the density is positive due to a time restriction which is not restrictive for small ϵ. From [14, eq. (2.25a)] we first define *local acoustic CFL ratios* μ_j, and *local apparent propagation factor* e_j as

$$\mu_j := \frac{a\Delta t}{\Delta z^n_j}, \qquad e_j := \frac{\mu_j/\epsilon}{1 + \mu_j/\epsilon}.$$

Then, one can write (3.2) as

$$\overrightarrow{w}_j^{n\dagger} = e_j \, \overrightarrow{w}_{j-1}^{n\dagger} + (1-e_j) \, \overrightarrow{w}_j^{n} \, .$$

Since $0 < e_j < 1$ (which can be satisfied for all ϵ uniformly), the updates for $\overrightarrow{w}_j$ and $\overleftarrow{w}_j$ are monotone, i.e., no new extremum can be generated. To show it for $\overrightarrow{w}^{n\dagger}$, assume that i is the index of maximum value of $\overrightarrow{w}_j^{n\dagger}$, that is $\overrightarrow{w}_i^{n\dagger} \geq \overrightarrow{w}_j^{n\dagger}$, $\forall j \in \mathbb{S}$. So,

$$\overrightarrow{w}_i^{n\dagger} \leq e_i \, \overrightarrow{w}_i^{n\dagger} + (1-e_i) \, \overrightarrow{w}_i^{n} \, .$$

Thus $\overrightarrow{w}_i^{n\dagger} \leq \overrightarrow{w}_i^{n}$ and then $\max_j \overrightarrow{w}_j^{n\dagger} \leq \max_j \overrightarrow{w}_j^{n}$. So, it is bounded from above. The proofs for the lower-bound and $\overleftarrow{w}^{n\dagger}$ are likewise. Hence,

$$\begin{aligned} \overrightarrow{m}^n \leq \overrightarrow{w}_j^{n\dagger} \leq \overrightarrow{M}^n, \\ \overleftarrow{m}^n \leq \overleftarrow{w}_j^{n\dagger} \leq \overleftarrow{M}^n \, . \end{aligned} \tag{3.19}$$

Having (3.19), one can show the following theorem.

Theorem 3.5. *For some Δt satisfying*

$$\frac{\Delta t}{\Delta x} \leq \frac{2a/\epsilon}{\left(\overrightarrow{M}^n - \overleftarrow{m}^n\right)^+ - \left(\overrightarrow{m}^n - \overleftarrow{M}^n\right)^-}, \tag{3.20}$$

the LP-IMEX scheme preserves the positivity of density provided that $\rho_j^0 > 0$ for all $j \in \mathbb{S}$.

Proof. In lines of [14], for the Lagrange step to satisfy positivity, one gets from τ-update that

$$\frac{\Delta t}{\Delta x} \left(\tilde{u}_{j-1/2}^{n\dagger} - \tilde{u}_{j+1/2}^{n\dagger}\right) < 1, \tag{3.21}$$

ensures $\rho_j^{n\dagger} > 0$ for all $j \in \mathbb{S}$. But on the other hand, Δt should be such that the projection step is a convex combination, thus

$$\frac{\Delta t}{\Delta x} \left(\left(\tilde{u}_{j-1/2}^{n\dagger}\right)^+ - \left(\tilde{u}_{j+1/2}^{n\dagger}\right)^-\right) < 1, \tag{3.22}$$

Between (3.21) and (3.22) the stronger condition should be chosen, which is (3.22). Then, based on the definition of $\tilde{u}^{n\dagger}$, we express Δt in terms of $\overrightarrow{M}$, $\overleftarrow{M}$, $\overrightarrow{m}$ and $\overleftarrow{m}$ and then the proof holds. ☐

The next goal is to show that this bound for the time step is uniform in ϵ, i.e., it does not vanish as the Mach number goes to zero. One can pose the following corollary.

Corollary 3.6. *For well-prepared initial data, the time step restriction* (3.20) *is uniform in ϵ.*

Proof. Recall that asymptotic consistency implies that with well-prepared initial datum and for $\epsilon \ll 1$, the density (and thus the pressure) has a constant part as the zeroth-order term in the asymptotic expansion. So the differences $\overrightarrow{M}^n - \overleftarrow{m}^n$ and $\overrightarrow{m}^n - \overleftarrow{M}^n$ are not of $\mathcal{O}(1/\epsilon^2)$ but $\mathcal{O}(1/\epsilon)$ and thus (3.20) is uniform in ϵ. In other words, the solution can be expanded w.r.t. ϵ, so

$$\begin{aligned} \overrightarrow{M}^n &\leq \frac{p_0^n}{\epsilon^2} + \max_j p_{2,j}^n + \frac{a}{\epsilon}\left(u_0^n + \epsilon \max_j(u_{1,j}^n)\right), & \overrightarrow{m}^n &\geq \frac{p_0^n}{\epsilon^2} + \min_j p_{2,j}^n + \frac{a}{\epsilon}\left(u_0^n + \epsilon \min_j(u_{1,j}^n)\right), \\ \overleftarrow{M}^n &\leq \frac{p_0^n}{\epsilon^2} + \max_j p_{2,j}^n - \frac{a}{\epsilon}\left(u_0^n + \epsilon \max_j(u_{1,j}^n)\right), & \overleftarrow{m}^n &\geq \frac{p_0^n}{\epsilon^2} + \min_j p_{2,j}^n - \frac{a}{\epsilon}\left(u_0^n + \epsilon \min_j(u_{1,j}^n)\right). \end{aligned}$$

Thus,

$$\begin{aligned}\overrightarrow{M}^n - \overleftarrow{m}^n &\leq \frac{a}{\epsilon}\left(2u_0^n + \epsilon\left(\max_j(u_{1,j}^n) + \min_j(u_{1,j}^n)\right)\right) + \left(\max_j p_{2,j}^n - \min_j p_{2,j}^n\right),\\ \overrightarrow{m}^n - \overleftarrow{M}^n &\leq \frac{a}{\epsilon}\left(2u_0^n - \epsilon\left(\max_j(u_{1,j}^n) + \min_j(u_{1,j}^n)\right)\right) - \left(\max_j p_{2,j}^n - \min_j p_{2,j}^n\right),\end{aligned}$$

and one gets

$$\lim_{\epsilon\to 0}\left[\frac{2a/\epsilon}{\left(\overrightarrow{M}^n - \overleftarrow{m}^n\right)^+ - \left(\overrightarrow{m}^n - \overleftarrow{M}^n\right)^-}\right] \geq \frac{2a/\epsilon}{\mathcal{O}(\frac{1}{\epsilon}) + \mathcal{O}(1)} \geq C. \tag{3.23}$$

Hence, there is an $\mathcal{O}(1)$ constant which bounds (3.20) from below, i.e., the condition (3.20) is uniform in ϵ. □

Now, to show that for a finite time the density is bounded from below, one should prove the following lemma.

Lemma 3.7. *Under the condition* (3.20)*, the computed density* ρ_j^{n+1} *is bounded away from zero in a finite time, where the lower-bound is given by*

$$\rho_{LB}^{n+1} := \frac{\min_j \rho_j^n}{1 + \dfrac{\Delta t\epsilon}{2a\Delta x}\left[\left(\overrightarrow{M}^n + \overleftarrow{M}^n\right) - \left(\overrightarrow{m}^n + \overleftarrow{m}^n\right)\right]} > 0. \tag{3.24}$$

Proof. From the τ-update (3.1) and $\tilde{u}_{j+1/2}^{n\dagger} = \frac{\epsilon}{2a}\left(\overrightarrow{w}_j^{n\dagger} - \overleftarrow{w}_{j+1}^{n\dagger}\right)$, one can get

$$\rho_j^n = \rho_j^{n\dagger}\left(1 + \frac{\epsilon\Delta t}{2a\Delta x}\left(\overrightarrow{w}_j^{n\dagger} - \overleftarrow{w}_{j+1}^{n\dagger} - \overrightarrow{w}_{j-1}^{n\dagger} + \overleftarrow{w}_j^{n\dagger}\right)\right).$$

So, to find the minimum value of the computed density, one should find the maximum value of the right-hand side. Due to (3.19) and under the condition (3.21), it can be seen that

$$\rho^{n\dagger} \geq \frac{\rho_j^n}{1 + \frac{\Delta t\epsilon}{2a\Delta x}\left[\left(\overrightarrow{M}^n + \overleftarrow{M}^n\right) - \left(\overrightarrow{m}^n + \overleftarrow{m}^n\right)\right]}.$$

Thus, since the projection step is a convex combination under (3.20), the lower-bound is obtained as (3.24). □

3.1.3. *Proof of local energy inequality (iv)*

We show that the solution of the scheme satisfies the energy inequality under an ϵ-independent time restriction. For the Lagrange step, based on [14, Theorem 2.3], we define the entropy function for symmetric advections problem, (2.16)-(2.17), as

$$\eta(\overrightarrow{w}, \overleftarrow{w}) := s(\overrightarrow{w}) + s(\overleftarrow{w}), \quad s(w) := \frac{\epsilon^2 w^2}{4a^2}.$$

So, it can be rewritten as

$$\eta(\overrightarrow{w}, \overleftarrow{w}) = \frac{1}{2}\left(u^2 + \frac{\pi^2}{\epsilon^2 a^2}\right) = E - \frac{\mathcal{E}}{\epsilon^2} + \frac{\pi^2}{2a^2\epsilon^2}, \tag{3.25}$$

since after non-dimensionalization, one gets $E = \frac{\mathcal{E}}{\epsilon^2} + \frac{u^2}{2}$ where $\mathcal{E}(\rho) = \frac{\kappa}{\gamma-1}\rho^{\gamma-1}$. For later use, we should mention that such a definition of internal energy fulfills the Weyl's assumptions as defined below.

Definition 3.8. The Weyl's assumption for the internal energy function are (see [14, 49])

$$\mathcal{E} > 0, \quad \mathcal{E}_\tau = -p < 0, \quad \mathcal{E}_{\tau\tau} > 0, \quad \mathcal{E}_{\tau\tau\tau} < 0.$$

We also define an entropy flux function $\psi(\overrightarrow{w}, \overleftarrow{w})$ as

$$\psi(\overrightarrow{w}, \overleftarrow{w}) := \frac{a}{\epsilon^2}\left(s(\overrightarrow{w}) - s(\overleftarrow{w})\right) = \frac{\pi u}{\epsilon^2}. \tag{3.26}$$

Then, the cell entropy inequality reads as

$$\frac{\eta_j^{n\dagger} - \eta_j^n}{\Delta t} + \frac{\psi_{j+1/2}^{n\dagger} - \psi_{j-1/2}^{n\dagger}}{\Delta z_j^n} \leq 0. \tag{3.27}$$

Substituting (3.25) and (3.26), one can relate the entropy inequality for symmetric advections problem, to the energy inequality for the barotropic Euler equations, i.e.,

$$\rho_j^n \frac{E_j^{n\dagger} - E_j^n}{\Delta t} + \frac{\left(\frac{\pi u}{\epsilon^2}\right)_{j+1/2}^{n\dagger} - \left(\frac{\pi u}{\epsilon^2}\right)_{j-1/2}^{n\dagger}}{\Delta x} \leq \frac{\rho_j^n}{\epsilon^2} \underbrace{\left[\mathcal{E}_j^{n\dagger} - \mathcal{E}_j^n - \frac{\left(\pi_j^{n\dagger}\right)^2 - \left(\pi_j^n\right)^2}{2a^2}\right]}_{=:\mathcal{R}_j^{n\dagger}}. \tag{3.28}$$

Then, to prove entropy stability of the scheme, one should show that the entropy residual $\mathcal{R}_j^{n\dagger}$ is non-positive. Considering $\pi_j^n = p_j^n$, we rewrite $\mathcal{R}_j^{n\dagger}$ as

$$\begin{aligned}\mathcal{R}_j^{n\dagger} &:= \mathcal{E}_j^{n\dagger} - \mathcal{E}_j^n - \frac{p_j^n}{2a^2}\left(\pi_j^{n\dagger} - p_j^n\right) - \frac{\left(\pi_j^{n\dagger} - p_j^n\right)^2}{2a^2} \\ \text{(due to (4.12))} &= \mathcal{E}_j^{n\dagger} - \mathcal{E}_j^n + p_j^n\left(\tau_j^{n\dagger} - \tau_j^n\right) - \frac{a^2}{2}\left(\tau_j^{n\dagger} - \tau_j^n\right)^2.\end{aligned}$$

On the other hand, from Taylor expansion with integral remainder, one gets

$$\mathcal{E}_j^{n\dagger} = \mathcal{E}_j^n + \mathcal{E}_\tau|_{x_j, t_n}\left(\tau_j^{n\dagger} - \tau_j^n\right) + \int_{\tau_j^n}^{\tau_j^{n\dagger}} \mathcal{E}_{\tau\tau}(\xi)\left(\tau_j^{n\dagger} - \xi\right) \mathrm{d}\xi.$$

Then, Weyl's assumptions and change of variables in the integral (re-parameterization) yield that

$$\mathcal{E}_j^{n\dagger} = \mathcal{E}_j^n - p_j^n\left(\tau_j^{n\dagger} - \tau_j^n\right) + \left(\tau_j^{n\dagger} - \tau_j^n\right)^2 \int_0^1 \mathcal{E}_{\tau\tau}(\tau_j^{n+1/2})(1-\zeta)\mathrm{d}\zeta, \tag{3.29}$$

where $\tau_j^{n+1/2} := \zeta\tau_j^{n\dagger} + (1-\zeta)\tau_j^n$. So, for the entropy residual to be non-positive, one gets

$$\mathcal{R}_j^{n\dagger} = \left(\tau_j^{n\dagger} - \tau_j^n\right)^2 \int_0^1 \left(\mathcal{E}_{\tau\tau}(\tau_j^{n+1/2}) - a^2\right)(1-\zeta)\mathrm{d}\zeta \tag{3.30}$$

$$= \left(\tau_j^{n\dagger} - \tau_j^n\right)^2 \int_0^1 \left(-p_\tau(\tau_j^{n+1/2}) - a^2\right)(1-\zeta)\mathrm{d}\zeta \leq 0, \tag{3.31}$$

and a sufficient condition would be to set the integrand to be negative. Since $p_\tau = \kappa\gamma\rho^{1+\gamma}$ it yields

$$a^2 \geq \kappa\gamma \max_j \max_\zeta \left(\left(\rho_j^{n+1/2}\right)^{\gamma+1} \right) = \max_j \left(\left(\rho_j^{n\dagger}\right)^3, \left(\rho_j^n\right)^3 \right) = \max\left(\|\rho^{n\dagger}\|_{\ell_\infty}^3, \|\rho^n\|_{\ell_\infty}^3 \right), \tag{3.32}$$

which satisfies the sub-characteristic condition as well as the energy inequality for the Lagrange step.

For the projection step, it is clear that due to Jensen's inequality the energy inequality holds as

$$(\rho E)_j^{n+1} \leq \rho_j^n E_j^{n\dagger} - \frac{\Delta t}{\Delta x} \left((\rho E \tilde{u})_{j+1/2}^{n\dagger} - (\rho E \tilde{u})_{j-1/2}^{n\dagger} \right). \tag{3.33}$$

Combining (3.28) and (3.33) we get the energy inequality (3.8) under an ϵ-uniform time restrictions (3.20) and sub-characteristic condition (3.32).

3.1.4. *Proof of ℓ_∞-stability (v)*

In this section, we prove the stability of LP-IMEX scheme in ℓ_∞-norm.

Lemma 3.9. *For the well-prepared initial data, the computed density, momentum and velocity are stable in ℓ_∞-norm uniformly in ϵ.*

Proof. As we have shown in the Appendix A for fixed ϵ, the entropy stability is enough to conclude the ℓ_∞ stability provided that the density is shown to be positive. Thus, the density, velocity and so the momentum are stable. For the proof of ϵ-uniformity of these result, we refer the reader to Appendix A. □

3.2. Rigorous analysis of asymptotic consistency

The proofs of asymptotic consistency for numerical schemes are often based on the formal asymptotic expansion as we presented in Section 3.1, i.e., using a well-prepared initial datum one analyzes the update as $\epsilon \to 0$. The analysis is rather formal; one usually does not show how the variables change in terms of ϵ, but inserts an asymptotic expansion into the scheme and does the formal asymptotic analysis, assuming implicitly that all the variables are $\mathcal{O}(1)$ in terms of ϵ. In this part, we show that it is possible for LP-IMEX to go further and show asymptotic consistency more rigorously.

In this approach, the main point is to study the implicit step to check how the updated solution behaves as $\epsilon \to 0$. Once we show that the *unique* updated solution does not blow up for the limit, one can combine it with the update and show asymptotic consistency, e.g., by studying the kernel of the solution operator of the implicit step. Such an analysis can also be done by combining the formal analysis in Section 3.1 with Theorem 3.10 below. The approach we present here to justify the formal analysis is akin to what has been used by Bispen in [3], in the context of the Finite Volume Evolution Galerkin (FVEG) scheme [4].

Note that for the scheme written in the form of (3.1)-(3.3) , $\overrightarrow{w}^{n\dagger}$ and $\overleftarrow{w}^{n\dagger}$ should be computed implicitly, and then $\tau^{n\dagger}$ is obtained explicitly. Now, let us define $\mathcal{N} := |\mathbb{S}|$ and consider the matrix $\overrightarrow{J}_\mu$ as the $\mathcal{N} \times \mathcal{N}$ coefficient matrix for the implicit update of $\overrightarrow{w}$, i.e.

$$\overrightarrow{J}_\mu \, \underline{\overrightarrow{w}}^{n\dagger} = \underline{\overrightarrow{y}}, \tag{3.34}$$

with

$$\overrightarrow{J}_\mu := \frac{a\Delta t}{\epsilon^\mu} \begin{bmatrix} 1 + \frac{\epsilon^\mu \Delta z_1^n}{a\Delta t} & 0 & \cdots & -1 \\ -1 & 1 + \frac{\epsilon^\mu \Delta z_2^n}{a\Delta t} & 0 & \cdots \\ 0 & \ddots & \ddots & 0 \\ 0 & \ddots & -1 & 1 + \frac{\epsilon^\mu \Delta z_{\mathcal{N}}^n}{a\Delta t} \end{bmatrix}, \tag{3.35}$$

where $\mu \in \{1,2\}$ corresponds to the relaxation approximation as $\alpha = \frac{a}{\epsilon^\mu}$, and $\underline{\overrightarrow{y}}$ is a vector specified in the case of barotropic Euler equations by (3.2), as $\underline{\overrightarrow{y}} := \underline{\Delta z}^n \odot \underline{\overrightarrow{w}}^n$, where $\odot$ denotes the entry-wise or Hadamard product. Notice that for the scheme (3.1)-(3.3), $\mu = 1$, and the reason to generalize the μ is the extension of the scheme for the shallow water equations with $\mu = 2$ that we will present later on in Section 4. Likewise we denote $\overleftarrow{J}_\mu$ and $\underline{\overleftarrow{y}}$; but, we will only present the proofs in this section and Section 4 for $\overrightarrow{w}$ since they are very similar for $\overleftarrow{w}$. With these definitions, one can pose the following theorem.

Theorem 3.10. *Consider the matrix $\overrightarrow{J}_\mu$ as defined in* (3.35)*. Then*

(i) $\overrightarrow{J}_\mu$ is non-singular for all $\epsilon > 0$;
(ii) $\lim_{\epsilon\to 0} \| \overrightarrow{J}_\mu^{-1} \|$ is bounded for any natural matrix norm.

Proof. To prove (i), we show that $\overrightarrow{J}_\mu$ has non-zero eigenvalues with positive real part. It is clear that matrix $\overrightarrow{J}_\mu$ is strictly diagonally dominant (SDD), and it is a classical result that SDD matrices are non-singular; see [29, Thm. 6.1.10]. This is enough to show the non-singularity of $\overrightarrow{J}_\mu$, and so to conclude that the solution of the implicit step, $\underline{\overrightarrow{w}}^{n\dagger}$, is unique. However it is worth to remark that using the Gerschgorin's circle theorem [29, Chap. 6], it is also possible to show that the eigenvalues have positive real parts.

For the point (ii), from [48], the infinity-norm of the inverse of matrix $M \in \mathbb{R}^{n\times n}$, provided that M is SDD, can be bounded as

$$\|M^{-1}\|_\infty \leq \max_{1\leq i\leq n} \frac{1}{\Delta_i(M)}, \qquad \Delta_i(M) := |M_{ii}| - \sum_{j\neq i} |M_{ij}|. \tag{3.36}$$

For $\overrightarrow{J}_\mu$, one can find $\Delta_i(\overrightarrow{J}_\mu) = \Delta z_i^n > 0$. So there is an ϵ-uniform bound for the infinity-norm of matrix inverse. Since $\mathcal{N}$ is fixed, all matrix norms are equivalent which implies the point (ii). ☐

From Theorem 3.10 one can immediately conclude that a unique solution $\|\underline{\overrightarrow{w}}^{n\dagger}\|$ (thus a unique numerical solution) exists, which has the same order as $\|\underline{\overrightarrow{w}}^n\|$ in terms of ϵ. Also by Theorem 3.10 and (3.34), one can see that $\overrightarrow{w}_j^{n\dagger} - \overrightarrow{w}_{j-1}^{n\dagger} = \mathcal{O}(\epsilon^2)$, thus (due to similar results for $\overleftarrow{w}$) $\pi_{(0)}^{n\dagger}$ and $u_{(0)}^{n\dagger}$ are constant in space. This is exactly the argument we made in the asymptotic consistency analysis in Section 3.1 but it is more rigorous since Theorem 3.10 has shown the behavior of $\overrightarrow{w}$ and $\overleftarrow{w}$ in terms of ϵ, rigorously. Since the projection step is explicit, its asymptotic consistency can be simply studied as in Section 3.1.

Remark 3.11*.* This approach proves asymptotic consistency rigorously, i.e., the solution at any time step moves to the limit as $\epsilon \to 0$. This is the result that makes the uniformity proofs of the previous sections rigorous, as been mentioned earlier in Remark 3.4.

Remark 3.12*.* To summarize, the scheme is AC and AS, and since the $\overrightarrow{J}_\mu$ and $\overleftarrow{J}_\mu$ can be inverted simply due to their structure and that the time step is ϵ-uniform, it is also AEf. Thus the scheme is AP in the sense of Definition 1.1.

4. Shallow water equations with topography

In this section, we show that similar stability arguments work for the LP-IMEX scheme applied to the non-dimensionalized shallow water equations with non-flat bottom:

$$\partial_t h + \partial_x(hu) = 0, \tag{4.1}$$

$$\partial_t(hu) + \partial_x\left(hu^2 + \frac{p(h)}{\epsilon^2}\right) = -\frac{hb_x}{\epsilon^2}, \tag{4.2}$$

where h stands for the water height and ϵ denotes the Froude number, defined as the ratio of characteristic speed to the speed of gravity waves. Also $b(x)$ is the bottom function, and the pressure function is chosen as before $p(h) = \frac{h^2}{2}$. Note that for this shallow water model to be valid, the bottom slope b_x should be small enough such that $\tan\theta \approx \theta$ where $\tan\theta$ is the bottom slope; see [6] for details.

As one can see later on in Section 4.2.1, the asymptotic consistency of the Lagrange–projection scheme for shallow water equations with non-flat bottom cannot be shown, if one uses the same $\alpha = \frac{a}{\epsilon}$ as in Section 3. So, here we propose a more diffusive relaxation system by using $\alpha = \frac{a}{\epsilon^2}$. We omit the details of splitting and numerical scheme, and refer the reader to [8] and [11]. We should only emphasize that we treat the source term implicitly in the Lagrange step, so the projection step is exactly like the homogeneous case (the barotropic Euler equations).

The Lagrange step of the scheme can be written as

$$\tau_j^{n\dagger} = \tau_j^n + \frac{\Delta t}{\Delta z_j^n}\left(\tilde{u}_{j+1/2}^{n\dagger} - \tilde{u}_{j-1/2}^{n\dagger}\right). \tag{4.3}$$

$$\overrightarrow{w}_j^{n\dagger} = \overrightarrow{w}_j^n - \frac{a\Delta t}{\epsilon^2 \Delta z_j^n}\left(\overrightarrow{w}_j^{n\dagger} - \overrightarrow{w}_{j-1}^{n\dagger}\right) - \frac{\Delta t a}{\epsilon^2}\frac{\Delta z_{j-1/2}^n}{\Delta z_j^n}\frac{b_{x,j-1/2}}{\epsilon^2}, \tag{4.4}$$

$$\overleftarrow{w}_j^{n\dagger} = \overleftarrow{w}_j^n + \frac{a\Delta t}{\epsilon^2 \Delta z_j^n}\left(\overleftarrow{w}_{j+1}^{n\dagger} - \overleftarrow{w}_j^{n\dagger}\right) + \frac{\Delta t a}{\epsilon^2}\frac{\Delta z_{j+1/2}^n}{\Delta z_j^n}\frac{b_{x,j+1/2}}{\epsilon^2}, \tag{4.5}$$

where $b_{x,j+1/2} := \frac{b_{j+1}-b_j}{\Delta x}$ is the one-sided discretization of the bottom function, and the interface velocity is defined as

$$\tilde{u}_{j+1/2}^{n\dagger} := \frac{u_j^{n\dagger} + u_{j+1}^{n\dagger}}{2} - \frac{1}{2a}\left(\pi_{j+1}^{n\dagger} - \pi_j^{n\dagger}\right) - \frac{1}{2a}\left(\Delta z_{j+1/2}^n b_{x,j+1/2}\right). \tag{4.6}$$

Notice that this choice of $b_{x,j+1/2}$ provides C-property (well-balancing for the lake at rest [5]) as we will see later on. The projection step is like (3.5).

Let us define the limit solution and the well-prepared initial datum for the shallow water equations.

Definition 4.1. The formal limit of shallow water equations (4.1)–(4.2) is defined to be

$$\begin{aligned}
&\eta_{(0)} = h_{(0)} + b = \text{const.}, \qquad h_{(1)} = \text{const.},\\
&\operatorname{div} m_{(0)} = 0,\\
&\partial_t m_{(0)} + \partial_x\left(\frac{m_{(0)}^2}{h_{(0)}} + p_{(2)}\right) = -h_{(2)} b_x,
\end{aligned}$$

where $\eta := h + b$ is the water surface, and the asymptotic expansion is written as

$$\begin{aligned}
h(x,t) &= h_{(0)} + \epsilon h_{(1)} + \epsilon^2 h_{(2)},\\
m(x,t) &= m_{(0)} + \epsilon m_{(1)} + \epsilon^2 m_{(2)}.
\end{aligned}$$

Definition 4.2. For the shallow water equations (4.1)–(4.2), the well-prepared initial datum is defined as

$$\begin{aligned}
h_{WP}^{(0)}(x) &= h_{(0)}(x) + \epsilon^2 h_{(2)}(x),\\
(hu)_{WP}^{(0)}(x) &= m_{(0)} + \epsilon m_{(1)}(x),
\end{aligned} \tag{4.7}$$

where $h_{(0)}(x) = \eta_{(0)} - b(x)$ with constant $\eta_{(0)}$ and $m_{(0)}$.

Before we discuss the stability results of shallow water equations with non-flat bottom, very briefly we review the results for the flat bottom case as it has the same system of equations as barotropic Euler equations, but this time the numerical approach uses different relaxation for the Lagrange step.

4.1. **Numerical analysis of LP-IMEX scheme for flat bottom**

The results of this section are similar to Section 3.1, though with some differences. Also they will be used later on for the non-flat bottom case. The following theorem summarizes the results.

Theorem 4.3. *The LP-IMEX scheme* (4.3)–(4.5) *and* (3.5), *with well-prepared initial datum and flat bottom assumption* $b_x \equiv 0$, *satisfies the following properties.*

(i) It can be expressed in locally conservative form.

(ii) The scheme is AC, which means that it gives a consistent discretization of the zero-Froude limit of shallow water equations in terms of Definition 4.1.

(iii) Under some non-ϵ-restrictive CFL constraint (4.13), *the scheme is positivity preserving, i.e.,* $h_j^n > 0$ *provided that* $h_j^0 > 0$ *for all* $j \in \mathbb{S}$. *Moreover the height is bounded away from zero, i.e., there exists some* $h_{LB}^n > 0$ *such that* $h_j^n \geq h_{LB}^n$ *for all* $j \in \mathbb{S}$.

(iv) Under (4.13) *and sub-characteristic condition* (4.17), *the solution fulfills the local (cell) entropy (energy) inequality* (3.8).

(v) Under (4.13) *and sub-characteristic condition* (4.17), *the computed height, momentum and velocity are stable, i.e., bounded in ℓ_∞-norm, uniformly in ϵ.*

We analyze the properties of this scheme in the subsequent subsections. Like Section 3.1, the proof of locally conservative form is skipped.

4.1.1. *Proof of asymptotic consistency (ii)*

To show that the solution approaches the limit manifold, we use the well-prepared data at time step n and prove that the computed solution is also well-prepared. The initial datum reads as

$$h_j^n = h_{0c}^n + \epsilon^2 h_{(2)j}^n, \tag{4.8}$$

$$p_j^n = \pi_{0c}^n + \epsilon^2 \pi_{(2)j}^n, \tag{4.9}$$

$$u_j^n = u_{0c}^n + \epsilon u_{(1)j}^n, \tag{4.10}$$

where h_{0c}^n, π_{0c}^n and u_{0c}^n are constant values.

For the Lagrange step, we start by combining (4.3)–(4.5) to find the update for the relaxation pressure as

$$h_j^n \frac{\pi_j^{n\dagger} - \pi_j^n}{\Delta t} + \frac{a^2}{\epsilon^2 \Delta x}\left(\tilde{u}_{j+1/2}^{n\dagger} - \tilde{u}_{j-1/2}^{n\dagger}\right) = 0. \tag{4.11}$$

If one combines it with the continuity equation (4.3), it yields

$$\frac{a^2}{\epsilon^2}\left(\tau_j^{n\dagger} - \tau_j^n\right) + \left(\pi_j^{n\dagger} - \pi_j^n\right) = 0. \tag{4.12}$$

So its $\mathcal{O}(1/\epsilon^2)$ and $\mathcal{O}(1/\epsilon)$ components give

$$\frac{a^2}{\epsilon^2}\left(h_{(k)j}^{n\dagger} - h_{(k)j}^n\right) = h_{(k)j}^n h_{(k)j}^{n\dagger}\left(\pi_{(k)j}^{n\dagger} - \pi_{(k)j}^n\right), \quad \text{for } k = 0, 1$$

which yields

$$h_{(0)j}^{n\dagger} = h_{(0)j}^n = h_{(0)}^n \text{ const. in space}$$
$$h_{(1)j}^{n\dagger} = h_{(1)j}^n = h_{(1)}^n \text{ const. in space}$$

Next, we continue with the $\mathcal{O}(1/\epsilon^2)$ terms in the momentum equation, both $\overrightarrow{w}_j$ and $\overleftarrow{w}_j$. Simply, one can see that $\overrightarrow{w}^{n\dagger}_{(0)j}$ and $\overleftarrow{w}^{n\dagger}_{(0)j}$ are constant in space. Thus $\pi^{n\dagger}_{(0)j}$ and $u^{n\dagger}_{(0)j}$ are constant, so the *div*-free condition is also fulfilled. The similar argument can be done for $\pi^{n\dagger}_{(1)j}$ and $u^{n\dagger}_{(1)j}$. Hence, it turns out that the Lagrange step is AC.

For the projection step, we show the asymptotic consistency for the case, $\tilde{u}^{n\dagger}_{j-1/2} < 0$ and $\tilde{u}^{n\dagger}_{j+1/2} < 0$. The other cases can be done in a very similar way. The update for density is

$$h_j^{n+1} = h_j^{n\dagger} - \frac{\Delta t}{2a\Delta x}\left(h^{n\dagger}_{j+1} - h^{n\dagger}_j\right)\left(a\left(u^{n\dagger}_{j+1} + u^{n\dagger}_j\right) - \left(\pi^{n\dagger}_{j+1} - \pi^{n\dagger}_j\right)\right).$$

So, if we gather $\mathcal{O}(1)$ and $\mathcal{O}(\epsilon)$ terms, one can simply see that $h^{n+1}_{(0)j} = h^{n\dagger}_{(0)j} = h^n_{0c}$. Similar argument holds for $h^{n+1}_{(1)j}$ and thus $p^{n+1}_{(0)j}$ and $p^{n+1}_{(1)j}$.

To show the div-free condition, one should consider $\mathcal{O}(1)$ terms of the momentum update:

$$h^{n+1}_{(0)j}u^{n+1}_{(0)j} = h^{n\dagger}_{(0)j}u^{n\dagger}_{(0)j} - \frac{\Delta t}{2a\Delta x}\left(h^{n\dagger}_{(0)j+1} - h^{n\dagger}_{(0)j}\right)\left(a\left(u^{n\dagger}_{(0)j+1} + u^{n\dagger}_{(0)j}\right) - \left(\pi^{n\dagger}_{(0)j+1} - \pi^{n\dagger}_{(0)j}\right)\right).$$

$u^{n+1}_{(0)j} = u^{n\dagger}_{(0)j} = u^n_{0c}$, and the zeroth order component of velocity filed is solenoidal. Hence, combining the asymptotic consistency results for the Lagrange and projection steps together, it is obvious that the limit properties is satisfied.

The consistency of the discretization in the limit can be shown similar to Section 3.1, so the scheme is AC.

4.1.2. *Proof of height positivity (iii)*

In this section, we show that the height is positive due to a time restriction which is not restrictive for small ϵ. Also we show that the computed height has an ϵ-uniform lower-bound. Very similarly to Section 3.1, one can again show (3.19). Having (3.19) provides the following theorem.

Theorem 4.4. *For some Δt satisfying*

$$\frac{\Delta t}{\Delta x} \le \frac{2a/\epsilon^2}{\left(\overrightarrow{M}^n - \overleftarrow{m}^n\right)^+ - \left(\overrightarrow{m}^n - \overleftarrow{M}^n\right)^-}, \tag{4.13}$$

the LP-IMEX scheme preserves the positivity of height provided that $h^0_j > 0$ for all $j \in \mathbb{S}$. Also (4.13) *does not vanish as the Mach number goes to zero, so it non-restrictive for $\epsilon \ll 1$.*

Proof. The proof is similar as in Section 3.1; finally one gets

$$\lim_{\epsilon\to 0}\left[\frac{2a/\epsilon^2}{\left(\overrightarrow{M}^n - \overleftarrow{m}^n\right)^+ - \left(\overrightarrow{m}^n - \overleftarrow{M}^n\right)^-}\right] \ge \frac{2a/\epsilon^2}{\mathcal{O}(\frac{1}{\epsilon^2}) + \mathcal{O}(\frac{1}{\epsilon}) + \mathcal{O}(1)} \ge C. \tag{4.14}$$

Hence, the condition (4.13), which provides positivity of the height, is uniform in ϵ. □

Now, one can prove the following lemma similarly to Section 3.1.

Lemma 4.5. *Under the condition* (4.13)*, the computed height $h^{n\dagger}_j$ is bounded away from zero in a finite time, where the lower-bound is given by*

$$h^{n+1}_{LB} := \frac{\min_j h^n_j}{1 + \frac{\Delta t \epsilon^2}{2a\Delta x}\left[\left(\overrightarrow{M}^n + \overleftarrow{M}^n\right) - \left(\overrightarrow{m}^n + \overleftarrow{m}^n\right)\right]} > 0.$$

4.1.3. *Proof of local energy inequality (iv)*

The proof here is similar as in Section 3.1, only one should use $s(w) := \frac{\epsilon^4 w^2}{4a^2}$, which gives

$$\eta(\overrightarrow{w}, \overleftarrow{w}) = \frac{1}{2}\left(u^2 + \frac{\pi^2}{a^2}\right) = E - \frac{\mathcal{E}}{\epsilon^2} + \frac{\pi^2}{2a^2}. \tag{4.15}$$

Following the same procedure, one can easily obtain

$$\mathcal{R}_j^{n\dagger} = \left(\tau_j^{n\dagger} - \tau_j^n\right)^2 \int_0^1 \left(-p_\tau(\tau_j^{n+1/2}) - (a/\epsilon)^2\right)(1-\zeta)\mathrm{d}\zeta \le 0, \tag{4.16}$$

which yields

$$a^2 \ge \ \epsilon^2 \max\left(\|h^{n\dagger}\|_{\ell_\infty}^3, \|h^n\|_{\ell_\infty}^3\right), \tag{4.17}$$

which satisfies the sub-characteristic condition as well as the energy inequality for the Lagrange step.

There is no change in the proof of entropy inequality for the projection step, thus the result holds under the ϵ-uniform time restrictions (4.13) and sub-characteristic condition (4.17).

Remark 4.6. It is clear from (4.16) that the energy dissipation is very large for small ϵ; in fact this is more than enough. Thus one expects the scheme to be diffusive. Note also that although the sub-characteristic condition suggests the parameter a to be of $\mathcal{O}(\epsilon^2)$, such a choice ruins the ϵ-uniformity of the time step restrictions. So, as has been already done in the asymptotic consistency analysis, $a = \mathcal{O}(1)$ should be chosen.

4.1.4. *Proof of ℓ_∞-stability (v)*

Similarly to Section 3.1, one can show that for the well-prepared initial data, the computed height, momentum and velocity are stable in ℓ_∞-norm uniformly in ϵ. We skip the proof here.

Remark 4.7. Note that since the Theorem 3.10 applies to this case as well with $\mu = 2$, the asymptotic consistency analysis we have presented is rigorous. Also, one can conclude AP property very similarly.

4.2. Numerical analysis of LP-IMEX scheme for non-flat bottom

For the non-flat bottom case, one can summarize the stability properties in the following theorem.

Theorem 4.8. *The LP-IMEX scheme* (4.3)–(4.5) *and* (3.5), *applied to the shallow water equations with well-prepared initial datum, satisfies the following properties.*

(i) It can be expressed in locally conservative form for the density.
(ii) The scheme is AC up to $\mathcal{O}(\Delta x)$, i.e., it preserves the div-free condition for the zero-Froude limit has and (approximately) correct asymptotic expansion for the computed solution in terms of Definition 4.1, up to $\mathcal{O}(\Delta x)$.
(iii) Under some non-ϵ-restrictive CFL constraint (4.19), *the scheme is positivity preserving, i.e., $h_j^n > 0$ provided that $h_j^0 > 0$ for all $j \in \mathbb{S}$.*
(iv) Moreover, the scheme preserves the Lake at Rest (LaR), so it is well-balanced.

Now, we go through the proof of each part briefly, since the proofs are very similar to the case of flat bottom, Section 4.1.

4.2.1. *Proof of asymptotic consistency (ii)*

As in Definition 4.1. the argument for the asymptotic consistency is very similar to the flat bottom case, though with two important differences. The shallow water equations with non-flat bottom have different limit as $\epsilon \to 0$: the limit density (or height) is no longer constant, but the surface elevation is constant. Also rather than *div*-free velocity field, the momentum field should be solenoidal. Here we start with a well-prepared initial datum in the sense of Definition 4.2.

For the Lagrange step, from (4.12) it can be obtained that

$$h^{n\dagger}_{(0)j} = h^{n}_{(0)j}, \qquad h^{n\dagger}_{(1)j} = h^{n}_{(1)j} = h^{n}_{(1)} = \text{ const. in space.}$$

For the velocity, with the help of $\mathcal{O}(1/\epsilon^4)$ terms of (4.4) and (4.5) we obtain that $u^{n\dagger}_{(0)}$ should be constant and

$$\left(\pi^{n\dagger}_{(0)j+1} - \pi^{n\dagger}_{(0)j}\right) + \Delta z^{n}_{(0)j+1/2} b_{x,j+1/2} = 0. \tag{4.18}$$

Remark 4.9. If one uses $\alpha = \frac{a}{\epsilon}$, the asymptotic consistency of the water height $h^{n\dagger}$ cannot be concluded since (4.12) should be replaced by (3.13). This is the reason we have used the diffusive relaxation system with $\alpha = \frac{a}{\epsilon^2}$.

We now show the asymptotic consistency of the projection step for the case, $\tilde{u}^{n\dagger}_{j-1/2} < 0$ and $\tilde{u}^{n\dagger}_{j+1/2} < 0$. The other cases can be done in a very similar way. The $\mathcal{O}(1)$ height update gives

$$\begin{aligned} h^{n+1}_{(0)j} &= h^{n\dagger}_{(0)j} - \frac{\Delta t}{2a\Delta x}\left(h^{n\dagger}_{(0)j+1} - h^{n\dagger}_{(0)j}\right)\left(a\left(u^{n\dagger}_{(0)j+1} + u^{n\dagger}_{(0)j}\right) - \left(\pi^{n\dagger}_{(0)j+1} - \pi^{n\dagger}_{(0)j}\right) - \Delta z^{n}_{(0)j+1/2} b_{x,j+1/2}\right) \\ &= h^{n}_{(0)j} + \frac{\Delta t}{\Delta x} u^{n\dagger}_{(0)} \left(b_{j+1} - b_j\right) \\ &= h^{n}_{(0)j} + \mathcal{O}(\Delta x) \end{aligned}$$

for smooth bottom functions, due to (4.18) and since $u^{n\dagger}_{(0)}$ is a constant as mentioned above in the asymptotic consistency of Lagrange step. So, it yields that the scheme is not exactly AC for $h_{(0)}$, as long as the limit velocity is non-zero. But for small bed slope assumption, the scheme is AC up to $\mathcal{O}(\Delta x)$, for $h_{(0)j}$. So, it is AC up to $\mathcal{O}(\Delta x)$. Regarding $h_{(1)j}$, due to asymptotic analysis (4.7), one can see that the situation is similar, i.e., $h_{(1)j}$ is constant in time and space up to $\mathcal{O}(\Delta x)$-deviations.

Similarly, one can obtain that for the momentum the situation is the same, since

$$\begin{aligned} (hu)^{n+1}_{(0)j} &= (hu)^{n\dagger}_{(0)j} - \frac{\Delta t}{2a\Delta x}\left((hu)^{n\dagger}_{(0)j+1} + (hu)^{n\dagger}_{(0)j}\right)\left(a\left(u^{n\dagger}_{(0)j+1} + u^{n\dagger}_{(0)j}\right) - \left(\pi^{n\dagger}_{(0)j+1} - \pi^{n\dagger}_{(0)j}\right) - \Delta z^{n}_{(0)j+1/2} b_{x,j+1/2}\right) \\ &= h^{n}_{(0)j} u^{n\dagger}_{(0)j} - \frac{\Delta t}{\Delta x} u^{n\dagger}_{(0)j}\left(h^{n}_{(0)j+1} - h^{n}_{(0)j}\right) \\ &= h^{n}_{(0)j} u^{n\dagger}_{(0)j} + \mathcal{O}(\Delta x). \end{aligned}$$

Note that $h^{n}_{(0)j} u^{n\dagger}_{(0)j}$ variations are as small as $\mathcal{O}(\Delta x)$, so the momentum is consistent with the limit system up to the order $\mathcal{O}(\Delta x)$.

Regarding the consistency of the discretization in the limit for the Lagrange step, one can, similar to Section 3.1, show that

$$\begin{aligned} \frac{(hu)^{n\dagger}_{(0)j} - (hu)^{n}_{(0)j}}{\Delta t} - \frac{1}{2\Delta x}\left(\pi^{n\dagger}_{(2)j+1} - \pi^{n\dagger}_{(2)j-1}\right) + \frac{a}{2\Delta x}\left(u^{n\dagger}_{(2)j+1} - 2u^{n\dagger}_{(2)j+1} + u^{n\dagger}_{(2)j+1}\right) \\ = -\frac{1}{2\Delta x}\left(\Delta z^{n}_{(2)j-1/2}\frac{b_j - b_{j-1}}{\Delta x} + \Delta z^{n}_{(2)j+1/2}\frac{b_{j+1} - b_j}{\Delta x}\right), \end{aligned}$$

which is a consistent discretization of $\partial_t u + \partial_z \pi = -\frac{b_x}{\epsilon^2}$ in the limit. We refer the reader to consult with [8, Sect. 3.2.2] for this equation of the continuous Lagrange step; the only difference with [8, eq. (11)] is about the source term which has been gravity and friction in [8]. The proof for the projection step is clear using the definition of

$\tilde{u}^{n\dagger}_{j+1/2}$ and the relation (4.18). Hence the scheme is AC up to the $\mathcal{O}(\Delta x)$-deviations from the zero-Froude limit manifold.

4.2.2. *Proof of height positivity (iii)*

Compared to the time step restriction for the flat bottom case, there is an additional contribution due to the source terms, which has the same singularity as in the convection part. Thus, it is not difficult to show that due to (4.18) and since $u^{n\dagger}_{(0)}$ is constant in space, the condition (3.22) — with the interface velocity as defined in (4.6) — can be fulfilled uniformly in ϵ. So under the (stronger) following CFL condition (4.19), the scheme is positivity preserving, and it is non-restrictive for small ϵ.

$$\frac{\Delta t}{\Delta x} \leq \frac{2a/\epsilon^2}{\left(\overrightarrow{M}^{n\dagger} - \overleftarrow{m}^{n\dagger}\right)^+ - \left(\overrightarrow{m}^{n\dagger} - \overleftarrow{M}^{n\dagger}\right)^- + \frac{1}{\epsilon^2}\left(\Delta z^n_{j+1/2}\left|b_{x,j+1/2}\right| + \Delta z^n_{j-1/2}\left|b_{x,j-1/2}\right|\right)}. \tag{4.19}$$

Note that for this case, $\overrightarrow{M}$, $\overrightarrow{m}$, $\overleftarrow{M}$ and $\overleftarrow{m}$ are calculated at the time step $n\dagger$, due to the lack of a relation like (3.19).

4.2.3. *Proof of well-balancing (C-property) (vi)*

Lake at Rest or LaR is a very important equilibrium state which every shallow water scheme has to preserve. It simply describes a steady water state with flat surface and zero velocity. The failure in satisfying LaR at the discrete level, leads to spurious oscillations. To show that the scheme is well-balanced, i.e., it preserves LaR, at first we show that the scheme *may* have such a solution if one starts with LaR initial datum. Then, we argue that since the solution of the scheme is unique, then this should be the only case which can happen.

Looking at the projection step, one can see that to have a discrete solution at LaR equilibrium, the solution on the manifold

$$\mathcal{U}^{\Delta}_{LaR} := \left\{ \mathbb{U}_j = \begin{bmatrix} h_j \\ (hu)_j \end{bmatrix} \mid h_j = \eta - b_j, u_j = 0, \forall j \in \mathbb{S} \right\},$$

with zero velocity and flat water surface, it is sufficient if one has $\underline{\mathbb{U}}^{n\dagger} \in \mathcal{U}^{\Delta}_{LaR}$ and $\tilde{u}^{n\dagger}_{j+1/2} = 0$ for all $j \in \mathbb{S}$. Then, we can see that such a state is compatible with the scheme, i.e., the scheme may have such a solution. It is clear for the projection step, but let us clarify it for the Lagrange step. For the Lagrange step, we have three equations. The density equation is compatible with zero interface velocity and steady density. From the discussion on the asymptotic consistency, we have (4.12) which clearly shows that for steady density, the relaxation pressure would be also steady, so the solution remains at the equilibrium. It only remains to show the compatibility condition for the velocity after the Lagrange update. One can find that $u^{n\dagger} = u^n_j \equiv 0$ from

$$u^{n\dagger}_j = u^n_j - \frac{\Delta t}{2\epsilon^2 \Delta z^n_j}\left(\pi^{n\dagger}_{j+1} - \pi^{n\dagger}_{j-1} + h^{n\dagger}_{j-1/2}(b_j - b_{j-1}) + h^{n\dagger}_{j+1/2}(b_{j+1} - b_j)\right)., \tag{4.20}$$

due to what we got above about the $\pi^{n\dagger}$ and the flat surface.

Up to now, we have shown the existence of such a solution. By Theorem 3.10, one can simply show that the solution is also unique (which should be the well-balanced solution), thus the well-balancing is concluded.

4.3. **Rigorous analysis of asymptotic consistency**

For the case of non-flat bottom, Theorem 3.10 also applies. But, since

$$\underline{\overrightarrow{y}} = \underline{\overrightarrow{y}}\left(\underline{U}^n, b\right) = \underline{\Delta z^n} \odot \underline{\overrightarrow{w}}^n - \frac{a\Delta t}{2\epsilon^4}\underline{\Delta z^n_-} \odot \underline{\Delta b_-},$$

where $\underline{\Delta z^n}$ and $\underline{\Delta b_-}$ are the vectors of $\Delta z^n_{j-1/2}$ and $(b_j - b_{j-1})$ respectively, one can see that $\|\underline{\overrightarrow{y}}\| = \mathcal{O}(1/\epsilon^4)$. So using the boundedness of $\| \overrightarrow{J}_2 \|$ is futile to rigorously show asymptotic consistency. However, one can use the structure of $\| \overrightarrow{J}_\mu^{-1} \|$, which proposes the following lemma.

Lemma 4.10. *Denote* $\widetilde{\overrightarrow{J}}_\mu := \frac{\epsilon^\mu}{a\Delta t} \overrightarrow{J}_\mu$. *Then,*

(i) Denote the adjugate matrix of $\widetilde{\overrightarrow{J}}_\mu$ *by* $\operatorname{adj}\left(\widetilde{\overrightarrow{J}}_\mu\right)$, *and the all-ones matrix of size* $\mathcal{N}$ *by* $\mathbf{1}_\mathcal{N}$. *Then*

$$\operatorname{adj}\left(\widetilde{\overrightarrow{J}}_\mu\right) = (1 + \mathcal{O}(\epsilon^\mu))\, \mathbf{1}_\mathcal{N},$$

(ii) $\det\left(\widetilde{\overrightarrow{J}}_\mu\right) = \mathcal{O}(\epsilon^\mu)$.

Proof. It is known that the inverse of a circulant matrix is also circulant [24]. So, it is enough if we show that the entries of the first column of $\operatorname{adj}\left(\widetilde{\overrightarrow{J}}_\mu\right)$ are $1 + \mathcal{O}(\epsilon^2)$, which correspond to the first row of the cofactor matrix. We denote $\chi_j := \frac{\epsilon^\mu \Delta z^n_j}{a\Delta t}$ and for simplicity of the notation, we assume that $\chi_j = \chi$ is constant; the proof is similar for the non-constant case. For the cofactor matrix, one can see that the entry of the first row and j-th column is

$$\operatorname{cof}\left(\widetilde{\overrightarrow{J}}_\mu\right)_{1j} = (-1)^{j+1} \det \begin{bmatrix} K_1 & O_{(j-1)\times(\mathcal{N}-j)} \\ O_{(\mathcal{N}-j)\times(j-1)} & K_2 \end{bmatrix},$$

where $O_{q\times r}$ is a zero matrix of size $q \times r$, and K_1 and K_2 are defined as

$$K_1 = \begin{bmatrix} -1 & 1+\chi & 0 & \cdots & 0 \\ 0 & -1 & 1+\chi & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & -1 \end{bmatrix}_{(j-1)\times(j-1)}, \qquad K_2 = \begin{bmatrix} 1+\chi & 0 & \cdots & 0 \\ -1 & 1+\chi & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & -1 & 1+\chi \end{bmatrix}_{(\mathcal{N}-j)\times(\mathcal{N}-j)}.$$

Then, it is clear that

$$\operatorname{cof}\left(\widetilde{\overrightarrow{J}}_\mu\right)_{1j} = (-1)^{2j}(1+\chi)^{\mathcal{N}-j} = (1+\chi)^{\mathcal{N}-j}.$$

Hence, all the entries of $\operatorname{cof}\left(\widetilde{\overrightarrow{J}}_\mu\right)$ and so $\operatorname{adj}\left(\widetilde{\overrightarrow{J}}_\mu\right)$ are $1 + \mathcal{O}(\epsilon^\mu)$. As we mentioned, the proof for the scheme (4.3)-(4.5) with non-constant χ_j is similar.

For the point (ii), at first we show that the circulant matrix of central upwind discretization δ_+ defined as

$$\delta_+ := \mathbf{Circ}(1, \underbrace{0, \ldots, 0}_{\mathcal{N}-2 \text{ times}}, -1)$$

has rank $\mathcal{N} - 1$. To show that, we recall the explicit relation for the eigenvalues of a circulant matrix [24, eq. (3.7)], which for δ_+ gives $1 - \omega_j^{\mathcal{N}-1}$ as the eigenvalues for $j = 0, 1, \ldots, \mathcal{N} - 1$, where $\omega_j := e^{\frac{2\pi i j}{\mathcal{N}}}$ is the $\mathcal{N}$th-root of unity. So, for the eigenvalues to be zero, $e^{\frac{2\pi i j \mathcal{N}}{\mathcal{N}}} = 1$, which is only possible for $j = 0$, so there is exactly one zero eigenvalue, and the rank of δ_+ is $\mathcal{N} - 1$. From this result, and by assuming χ_j to be constant, among the eigenvalues of $\widetilde{\overrightarrow{J}}_\mu$ denoted by λ_j for $j = 1, 2, \ldots, \mathcal{N}$, there is exactly one which has been shifted to $\mathcal{O}(\epsilon^\mu)$, while all other eigenvalues are $\mathcal{O}(1)$; thus

$$\det\left(\overset{\simeq}{J}_\mu\right) = \prod_{j=1}^{\mathcal{N}} \lambda_j = \mathcal{O}(\epsilon^\mu). \tag{4.21}$$

For the case of non-constant χ_j, one can use the perturbation theory of eigenvalues like [43, Appendix K] and [2]. For example, using [43, Appendix K], one can simply find that the eigenvalues of δ_+ would be perturbed by $(\mathcal{N}+2)\epsilon^{\mu/\mathcal{N}}$ which vanishes as $\epsilon \to 0$. So, the perturbation is small in terms of ϵ, and the same result as 4.21 holds. ☐

Lemma 4.10 implies that the implicit operator is almost constant up to some deviations of order $\mathcal{O}(\epsilon^\mu)$. Using this and periodic boundary conditions, one can show the following theorem.

Theorem 4.11. *The norm of the updated Lagrange solution $\underline{\overrightarrow{w}}^{n\dagger}$, is at most as large as $\|\underline{\overrightarrow{w}}^n\| = \mathcal{O}(1/\epsilon^2)$, for periodic or compactly supported boundary conditions.*

Proof. Set $\mu = 2$. Then by Lemma 4.10, the leading part of $\overset{\simeq}{J}_2^{-1}$ is constant. For the statement of the theorem to be true, we should show that such a structure makes $\overset{\simeq}{J}_2^{-1}\underline{\overrightarrow{y}}$ of $\mathcal{O}(1/\epsilon^2)$. In other words, it filters the $\mathcal{O}(1/\epsilon^4)$ part of the $\underline{\overrightarrow{y}}$; we denote it as $\underline{\overrightarrow{y}}^* := -\frac{a\Delta t}{2\epsilon^4}\underline{\underline{\Delta z^n}} \odot \underline{\underline{\Delta b_-}}$. So, one can find that

$$\begin{aligned} y_j^* &= -\frac{a\Delta t}{2\epsilon^4}\left(\rho_j^n + \rho_{j-1}^n\right)(b_j - b_{j-1}) \\ &= -\frac{a\Delta t}{2\epsilon^4}\left(2\eta_{(0)j}^n - b_j - b_{j-1} + \mathcal{O}(\epsilon^2)\right)(b_j - b_{j-1}) \\ &= -\frac{a\Delta t}{2\epsilon^4}\left(2\eta_{(0)c}^n - b_j - b_{j-1}\right)(b_j - b_{j-1}) + \mathcal{O}(1/\epsilon^2), \end{aligned}$$

due to the well-preparedness of initial datum.

Now, it is enough to show that $\underline{\overrightarrow{y}}^*$ belongs to the kernel of the leading order of $\overset{\simeq}{J}_\mu^{-1}$, i.e., $\|\overset{\simeq}{J}_\mu^{-1}\underline{\overrightarrow{y}}^*\| = \mathcal{O}(1/\epsilon^2)$. One can show it simply by making a spatial summation and use the boundary condition, i.e.,

$$\sum_j y_j^* = -\frac{a\Delta t}{2\epsilon^4}\sum_j \left[2\eta_{(0)c}^n (b_j - b_{j-1}) - \left(b_j^2 - b_{j-1}^2\right)\right] = 0.$$

Hence, the $\mathcal{O}(1/\epsilon^4)$ terms vanish and one is left with the contributions of order $\mathcal{O}(1/\epsilon^2)$. ☐

Note that the similar results of Lemma 4.10 and Theorem 4.11, hold for $\overleftarrow{J}\mu$ and $\overleftarrow{w}$ respectively. This implies the following corollary.

Corollary 4.12. *The asymptotic consistency analysis for the scheme LP-IMEX scheme,* (4.3)-(4.5) *and* (3.5) *is rigorous, i.e, the asymptotic expansion is valid.*

Proof. Due to Theorem 4.11, the implicit step preserves the order of $\|\underline{\overrightarrow{w}}\|$. So, the order of u^{n+1} and π^{n+1} are $\mathcal{O}(1)$, and the asymptotic consistency analysis is rigorous. This concludes the rigorous asymptotic consistency of the implicit step, and also the whole scheme since the explicit step has already been shown to be AC. ☐

Appendix A. Entropy stability in the zero-Mach limit

In this section, we discuss the consequences of entropy stability for $\epsilon \to 0$ to show the stability of the solution and, due to the compactness, its strong convergence to some limit. Note that this is not the classical convergence result for $\Delta x \to 0$. Here the main objective is to discuss the stability region which entropy stability provides. It also gives some a priori information about the asymptotic consistency, although it cannot recover it completely since it does not use any detail of spatial or time integration, and even the splitting. This analysis assumes the positivity of density and energy inequality, so it is not limited to LP-IMEX scheme.

The procedure is as follows: Firstly, we recall that positivity and energy inequality gives boundedness of the density and, but not directly in the limit $\epsilon \to 0$. We then show this boundedness $\epsilon \to 0$, thus the existence of a converging subsequence due to the compactness. Then, we show that the limit density is the incompressible limit solution.

Denote entropy function $\mathcal{J} := \rho E$ and assume a fixed grid of size $\mathcal{N}$. We make a spatial summation on (3.8) to get

$$\sum_j \mathcal{J}(\underline{\mathbb{U}}_j^{n+1}) \leq \sum_j \mathcal{J}(\underline{\mathbb{U}}_j^{n}) \Longrightarrow \sum_j \mathcal{J}(\underline{\mathbb{U}}_j^{n+1}) \leq \sum_j \mathcal{J}(\underline{\mathbb{U}}_j^{0}) \leq C_\epsilon < \infty. \tag{A.1}$$

If in addition we assume positivity, then since $\mathcal{J}(\mathbb{U}) = \frac{1}{2}\frac{(\rho u)^2}{\rho} + \frac{\kappa/\epsilon^2}{\gamma-1}\rho^\gamma$ is always positive, thus

$$0 < \mathcal{J}(\underline{\mathbb{U}}_j^{n+1}) \leq C_\epsilon \qquad \forall j \in \mathbb{S}.$$

One immediate result, for fixed ϵ, is the ℓ_∞-boundedness, i.e., $\underline{\rho} \in \ell_\infty(\mathbb{S})$ and $\underline{u} \in \ell_\infty(\mathbb{S})$. So the energy inequality accompanied with positivity, provides a stability region Ξ_ϵ^0 which depends on the initial condition as well as ϵ. But how does the stability region Ξ_ϵ^0 change if rather than fixed ϵ, we consider $\epsilon \to 0$? If one keeps the grid fixed and considers $\epsilon \to 0$, the boundedness of the density is rather clear, either due to positivity accompanied with the conservation of scheme, or due to the boundedness of the entropy. But it is not straightforward to conclude the boundedness of the velocity. This is the first question we want to answer in this section.

Also note that the boundedness of the density sequence w.r.t. ϵ provides strong convergence. In other words, solutions with positive density, owing to the compactness of the space, have a converging sequence of vectors $\{\rho^{\epsilon_k,n}\}$ for any time step n that converges strongly to some limit $\underline{\rho}^{\epsilon_\infty,n}$, i.e.,

$$\lim_{k\to\infty} \left\| \underline{\rho}^{\epsilon_k,n} - \underline{\rho}^{\epsilon_\infty,n} \right\|_{\ell_p} = 0.$$

Here ϵ_k is a sequence approaching 0, to the incompressible limit. But it is not clear whether or not the limit is in the space of incompressible solutions. To determine whether the limit is the correct limit, is the second question we discuss in this section.

In what follows, with the help of energy inequality we show that the computed density by the scheme actually converges to its incompressible limit. We then show that the same assumptions are not enough to prove asymptotic consistency of the velocity, i.e., the *div*-free condition; nonetheless the boundedness of the velocity sequence, so its convergence to some limit, can be obtained. We discuss the results in the following lemma.

Lemma A.1. *Consider $\mathcal{N}$-vector sequence $\{\underline{\rho}^{\epsilon_k,n}\}$ in k, accompanied with well-prepared initial data, as the discrete density solution of the barotropic Euler equations* (2.1)-(2.2). *Assume that the scheme satisfies density positivity and energy inequality. Then as $\epsilon \to 0$ the sequence is bounded and approaches the incompressible limit $\{\underline{\rho}^{0,n}\}$ with the rate of $\mathcal{O}(\epsilon)$. Moreover the sequence of velocity $\{\underline{u}^{\epsilon_k,n}\}$ is bounded in ℓ_∞.*

Remark A.2. Both from formal asymptotic expansion and rigorous analysis [35], one expects to see the convergence of density to its incompressible limit with $\mathcal{O}(\epsilon^2)$ rate. However the convergence rate of Lemma A.1 is not optimal. We see that exactly due to this issue, the asymptotic consistency of the velocity cannot be obtained by these assumptions.

Proof. Consider the well-prepared initial datum as $\rho_i^{\epsilon,0} := \rho^{0,0} + \delta_i^{\epsilon,0}$ with $\delta_i^{\epsilon,0} = \mathcal{O}(\epsilon^2)$. Then write the density at the time step n as $\rho_i^{\epsilon,n} := \rho^{0,0} + \delta_i^{\epsilon,n}$. Then due to conservation of scheme $\|\rho^{\epsilon,n}\|_{\ell_1} = \|\rho^{\epsilon,0}\|_{\ell_1}$, one can simply get

$$\sum_i \delta_i^{\epsilon,n} = \sum_i \delta_i^{\epsilon,0} = \mathcal{N}\mathcal{O}(\epsilon^2). \tag{A.2}$$

So, it seems that in the limit the density, in general, oscillates around a constant state. But this does not give us convergence since the perturbations have no sign. By energy inequality one can see that

$$\sum_i \left(\rho^{0,0} + \delta_i^{\epsilon,n}\right)^2 \leq \sum_i \left(\rho^{0,0} + \delta_i^{\epsilon,0}\right)^2 + C_0\epsilon^2, \qquad C_0 := \sum_i \left(\rho^{0,0} + \delta_i^{\epsilon,0}\right)\left(u^{0,0} + \mu_i^{\epsilon,0}\right)^2, \tag{A.3}$$

where $\mu_i^{\epsilon,0} = \mathcal{O}(\epsilon)$ to fulfill the well-preparedness of initial datum. Then combining (A.2) and (A.3) yields

$$\|\underline{\delta}^{\epsilon,n}\|_{\ell_2}^2 = \left\|\underline{\delta}^{0,n}\right\|_{\ell_2}^2 + C_0\epsilon^2 = \mathcal{O}(\epsilon^2),$$

which shows that each component converges to the incompressible limit with— at least —$\mathcal{O}(\epsilon)$ rate; thus, the sequence is bounded in terms of ϵ. However, this rate is smaller that what one would expect; this is all can be got by these assumptions. Furthermore, by straightforward calculations, one can show that

$$\left\|\underline{\rho}^{\epsilon,n}\right\|_{\ell_2}^2 - \left\|\underline{\rho}^{\epsilon,0}\right\|_{\ell_2}^2 = \|\underline{\delta}^{\epsilon,n}\|_{\ell_2}^2 + \mathcal{O}(\epsilon^4) = \mathcal{O}(\epsilon^2)$$

and then by the complete energy inequality, not (A.3), one can obtain

$$\left\|\underline{\left(\rho u^2\right)}^{\epsilon,n}\right\|_{\ell_1} - \left\|\underline{\left(\rho u^2\right)}^{\epsilon,0}\right\|_{\ell_1} \leq \mathcal{O}(1),$$

thus $\left\|\underline{\left(\rho u^2\right)}^{\epsilon,n}\right\|_{\ell_1}$ is bounded, and since the density converges to the incompressible limit uniformly which is away from zero, the velocity is bounded as well. □

Remark A.3. It is not difficult to show, with the additional assumption of $\|\underline{\delta}^{\epsilon,n}\|_{\ell_2} = \mathcal{O}(\epsilon^2)$, that $\left\|\underline{\mu}^{\epsilon,n}\right\|_{\ell_2} = \mathcal{O}(\epsilon)$. Thus the asymptotic consistency would be obtained completely.

The similar analysis can be done for the case of shallow water equations. Here we state the main result and the sketch of the proof.

Lemma A.4. *Consider $\mathcal{N}$-vector sequence $\{\underline{h}^{\epsilon_k,n}\}$ in k, accompanied with well-prepared initial data, as the discrete height solution of the shallow water equations* (4.1)-(4.2)*. Assume that the sequence satisfies height positivity and energy inequality. Then as $\epsilon \to 0$ the sequence is bounded and approaches the zero-Froude limit $\left\{\underline{h}^{0,n}\right\}$ with the rate of $\mathcal{O}(\epsilon)$. Moreover the sequence of velocity $\{\underline{u}^{\epsilon_k,n}\}$ is bounded in ℓ_∞.*

Proof. Consider the well-prepared initial datum as $h_i^{\epsilon,0} := \eta - b_i + \delta_i^{\epsilon,0}$ with $\delta_i^{\epsilon,0} = \mathcal{O}(\epsilon)$. Then write the height at the time step n as $h_j^{\epsilon,n} := \eta - b_i + \delta_i^{\epsilon,n}$. Using mass conservation one can find $\sum_i \delta_i^{\epsilon,n} = \sum_i \delta_i^{\epsilon,o}$ and from energy inequality

$$\sum_i \left(\left(\eta - b_i + \delta_i^{\epsilon,n}\right)^2 + 2b_i\left(\eta - b_i + \delta_i^{\epsilon,n}\right)\right) \leq \sum_i \left(\left(\eta - b_i + \delta_i^{\epsilon,0}\right)^2 + 2b_i\left(\eta - b_i + \delta_i^{\epsilon,0}\right)\right) + C_0\epsilon^2,$$

with $C_0 := \sum_i \left(h^{0,0} + \delta_i^{\epsilon,0}\right)\left(u^{0,0} + \mu_i^{\epsilon,0}\right)^2$, one concludes that $\|\underline{\delta}^{\epsilon,n}\|_{\ell_2} = \mathcal{O}(\epsilon)$, thus by arguments similar to Lemma A.1 the velocity is also bounded. □

2. Conclusion and future works

In this paper, we have investigated the stability results of the LP-IMEX scheme for one-dimensional barotropic Euler equations, which has been presented in [14], regarding the uniformity in terms of Mach number with well-prepared initial data. We have shown that the scheme is asymptotically consistent (not only formally but also rigorously), also have obtained a Mach-uniform time step restriction which provides entropy inequality, density positivity, as well as stability of the solution in ℓ_∞-norm. Also, we have extended the analysis to the shallow water equations with non-flat bottom as an important example of balance laws.

The natural next step would be to extend this analysis to the full Euler equations or multiple space dimensions, which are formidable tasks, particularly the latter as has been discussed to some extent in [9, 20]. Also as been done in [37] by the compensated compactness approach, it is of interest to prove the convergence of the scheme to the unique entropy solution.

Acknowledgments

The author would like to gratefully thank Sebastian Noelle from RWTH Aachen for many helpful discussions and suggestions in the preparation of this paper. He also would like to acknowledge Christophe Chalons and his group in Université Versailles Saint-Quentin-en-Yvelines for giving the motivation for this work and the subsequent discussions.

References

1. Saul Abarbanel, Pravir Duth, and David Gottlieb, *Splitting methods for low Mach number Euler and Navier–Stokes equations*, Computers & fluids **17** (1989), no. 1, 1–12.
2. Rajendra Bhatia, *Perturbation bounds for matrix eigenvalues*, vol. 53, SIAM, 1987.
3. Georgij Bispen, *IMEX finite volume methods for the shallow water equations*, Ph.D. thesis, Johannes Gutenberg-Universität, 2015.
4. Georgij Bispen, Koottungal Revi Arun, Maria Lukáčová-Medvi'ová, and Sebastian Noelle, *IMEX large time step finite volume methods for low Froude number shallow water flows*, Communication in Computational Physics **16** (2014), 307–347.
5. François Bouchut, *Nonlinear stability of finite volume methods for hyperbolic conservation laws: And well-balanced schemes for sources*, Springer Science & Business Media, 2004.
6. Francois Bouchut and Michael Westdickenberg, *Gravity driven shallow water models for arbitrary topography.*
7. Christophe Chalons and Jean-François Coulombel, *Relaxation approximation of the Euler equations*, Journal of Mathematical Analysis and Applications **348** (2008), no. 2, 872–893.
8. Christophe Chalons, Mathieu Girardin, and Samuel Kokh, *Large time step and asymptotic preserving numerical schemes for the gas dynamics equations with source terms*, SIAM Journal on Scientific Computing **35** (2013), no. 6, A2874–A2902.
9. ———, *An all-regime Lagrange-rojection like scheme for the gas dynamics equations on unstructured meshes*, submitted, hal-01007622, June 2014.
10. Christophe Chalons, P. Kestener, Samuel Koch, and Maxime Stauffert, *A large time-step and well-balanced Lagrange–projection type scheme for the shallow water equations*, Preprint, 2016.
11. Christophe Chalons, Marc Massot, and Aymeric Vié, *On the Eulerian Large Eddy Simulation of disperse phase flows: An asymptotic preserving scheme for small Stokes number flows*, Multiscale Modeling & Simulation **13** (2015), no. 1, 291–315.
12. Gui-Qiang Chen, Charles D. Levermore, and Tai-Ping Liu, *Hyperbolic conservation laws with stiff relaxation terms and entropy*, Communications on Pure and Applied Mathematics **47** (1994), no. 6, 787–830.
13. Frédéric Coquel, Edwige Godlewski, and Nicolas Seguin, *Regularization and relaxation tools for interface coupling*, Proceedings of XX Congresso de Ecuaciones Differenciales Y Aplicaicones, CEDYA, 2007.
14. Frédéric Coquel, Quang Nguyen, Marie Postel, and Quang Tran, *Entropy-satisfying relaxation method with large time-steps for Euler IBVPs*, Mathematics of Computation **79** (2010), no. 271, 1493–1533.
15. Floraine Cordier, Pierre Degond, and Anela Kumbaro, *An Asymptotic–Preserving all–speed scheme for the Euler and Navier–Stokes equations*, Journal of Computational Physics **231** (2012), no. 17, 5685–5704.
16. Gianni Dal Maso, Philippe G. Lefloch, and François Murat, *Definition and weak stability of nonconservative products*, Journal de mathématiques pures et appliquées **74** (1995), no. 6, 483–548.
17. Pierre Degond, Jian-Guo Liu, and Marie-Hélène Vignal, *Analysis of an asymptotic preserving scheme for the Euler–Poisson system in the quasineutral limit*, SIAM Journal on Numerical Analysis **46** (2008), no. 3, 1298–1322.
18. Pierre Degond and Min Tang, *All speed scheme for the low Mach number limit of the isentropic Euler equation*, Communications in Computational Physics **10** (2011), no. 1, 1–31.

19. Stéphane Dellacherie, *Analysis of Godunov type schemes applied to the compressible Euler system at low Mach number*, Journal of Computational Physics **229** (2010), no. 4, 978–1016.
20. Stéphane Dellacherie, Jonathan Jung, Pascal Omnes, and Pierre-Arnaud Raviart, *Construction of modified Godunov type schemes accurate at any Mach number for the compressible Euler system*, submitted, hal-00776629, November 2015.
21. Stéphane Dellacherie, Pascal Omnes, and Felix Rieper, *The influence of cell geometry on the Godunov scheme applied to the linear wave equation*, Journal of Computational Physics **229** (2010), no. 14, 5315–5338.
22. Jean Donea, Antonio Huerta, J.-Ph. Ponthot, and A. Rodríguez-Ferran, *Arbitrary Lagrangian-Eulerian methods*, John Wiley & Sons, Ltd, 2004.
23. Edwige Godlewski and Pierre-Arnaud Raviart, *Numerical approximation of hyperbolic systems of conservation laws*, vol. 118, Springer, 1996.
24. Robert M. Gray, *Toeplitz and circulant matrices: A review*, now publishers inc, 2006.
25. Hervé Guillard and Angelo Murrone, *On the behavior of upwind schemes in the low Mach number limit: II. Godunov type schemes*, Computers & fluids **33** (2004), no. 4, 655–675.
26. Hervé Guillard and Cécile Viozat, *On the behaviour of upwind schemes in the low Mach number limit*, Computers & fluids **28** (1999), no. 1, 63–86.
27. Jeffrey Haack, Shi Jin, and Jian-Guo Liu, *An all-speed asymptotic-preserving method for the isentropic Euler and Navier–Stokes equations*, Communications in Computational Physics **12** (2012), no. 4, 955–980.
28. Helge Holden, Kenneth H. Karlsen, Knut-Andreas Lie, and Nils Henrik Risebro, *Splitting methods for partial differential equations with rough solutions*, European Math. Soc. Publishing House, 2010.
29. Roger A. Horn and Charles R. Johnson, *Matrix analysis*, Cambridge University Press, New York, NY, USA, 1986.
30. Shi Jin, *Runge–Kutta methods for hyperbolic conservation laws with stiff relaxation terms*, Journal of Computational Physics **122** (1995), no. 1, 51–67.
31. ———, *Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations*, SIAM Journal on Scientific Computing **21** (1999), no. 2, 441–454.
32. ———, *Asymptotic preserving (AP) schemes for multiscale kinetic and hyperbolic equations: A review*, Lecture Notes for Summer School on Methods and Models of Kinetic Theory(M&MKT), Porto Ercole (Grosseto, Italy) (2010), 177–216.
33. Shi Jin and Zhouping Xin, *The relaxation schemes for systems of conservation laws in arbitrary space dimensions*, Communications on Pure and Applied Mathematics **48** (1995), no. 3, 235–276.
34. Sergiu Klainerman and Andrew Majda, *Singular limits of quasilinear hyperbolic systems with large parameters and the incompressible limit of compressible fluids*, Communications on Pure and Applied Mathematics **34** (1981), no. 4, 481–524.
35. ———, *Compressible and incompressible fluids*, Communications on Pure and Applied Mathematics **35** (1982), no. 5, 629–651.
36. Edward W. Larsen, J. E. Morel, and Warren F. Miller, *Asymptotic solutions of numerical transport problems in optically thick, diffusive regimes*, Journal of Computational Physics **69** (1987), no. 2, 283–324.
37. Corrado Lattanzio and Denis Serre, *Convergence of a relaxation scheme for hyperbolic systems of conservation laws*, Numerische Mathematik **88** (2001), no. 1, 121–134.
38. Philippe G. LeFloch and Michael Westdickenberg, *Finite energy solutions to the isentropic Euler equations with geometric effects*, Journal de mathématiques pures et appliquées **88** (2007), no. 5, 389–429.
39. Tai-Ping Liu, *Hyperbolic conservation laws with relaxation*, Communications in Mathematical Physics **108** (1987), no. 1, 153–175.
40. Nader Masmoudi, *Examples of singular limits in hydrodynamics*, Handbook of Differential Equations: Evolutionary Equations **3** (2007), 195–275.
41. Andreas Meister, *Asymptotic single and multiple scale expansions in the low Mach number limit*, SIAM Journal on Applied Mathematics **60** (1999), no. 1, 256–271.
42. Sebastian Noelle, Georgij Bispen, Koottungal Revi Arun, Maria Lukáčová-Medvi'ová, and C.-D. Munz, *A weakly asymptotic preserving low Mach number scheme for the Euler equations of gas dynamics*, SIAM Journal on Scientific Computing **36** (2014), no. 6, B989–B1024.
43. Alexander M. Ostrowski, *Solution of equations and systems of equations*, vol. 9, Academic Press New York, 1966.
44. Felix Rieper, *On the dissipation mechanism of upwind–schemes in the low Mach number regime: A comparison between Roe and HLL*, Journal of Computational Physics **229** (2010), no. 2, 221–232.
45. ———, *A low–Mach number fix for Roe's approximate Riemann solver*, Journal of Computational Physics **230** (2011), no. 13, 5263–5287.
46. Felix Rieper and Georg Bader, *The influence of cell geometry on the accuracy of upwind schemes in the low Mach number regime*, Journal of Computational Physics **228** (2009), no. 8, 2918–2933.
47. Jochen Schütz and Sebastian Noelle, *Flux splitting for stiff equations: A notion on stability*, Journal of Scientific Computing (2014), 1–19 (English).
48. Jim M. Varah, *A lower bound for the smallest singular value of a matrix*, Linear Algebra and its Applications **11** (1975), no. 1, 3–5.
49. Hermann Weyl, *Shock waves in arbitrary fluids*, Communications on Pure and Applied Mathematics **2** (1949), no. 2-3, 103–122.