A high-order method for weakly compressible flows

Klaus Kaiser[*†] and Jochen Schütz[†]

Institut für Geometrie und Praktische Mathematik
Templergraben 55, 52062 Aachen, Germany

* IGPM, RWTH Aachen University, Templergraben 55, D-52062 Aachen
† Faculty of Sciences, Hasselt University, Agoralaan Gebouw D, BE-3590 Diepenbeek

updated version of February 1, 2017

A high-order method for weakly compressible flows

Klaus Kaiser*,†, Jochen Schütz†

March 6, 2017

In this work, we introduce an IMEX discontinuous Galerkin solver for the weakly compressible isentropic Euler equations. The splitting needed for the IMEX temporal integration is based on the recently introduced *reference solution* splitting [1, 2], which makes use of the *incompressible* solution. We show that the overal method is *asymptotic preserving*.

*IGPM, RWTH Aachen University, Templergraben 55, 52062 Aachen
†Faculty of Sciences, Hasselt University, Agoralaan Gebouw D, BE-3590 Diepenbeek

1. Introduction

In this work, we consider the (weakly-)compressible isentropic Euler equations [3, 4] in dimensionless form,

$$
\rho_t + \nabla \cdot (\rho \boldsymbol{u}) = 0
$$
$$
(\rho \boldsymbol{u})_t + \nabla \cdot (\rho \boldsymbol{u} \otimes \boldsymbol{u}) + \frac{1}{\varepsilon^2} \nabla p = 0.
$$
(1)

The wave speeds in normal direction $\boldsymbol{n}$ of this (assumed two-dimensional) problem are

$$
\lambda_1 = \boldsymbol{u} \cdot \boldsymbol{n} \qquad \text{and} \qquad \lambda_{2,3} = \boldsymbol{u} \cdot \boldsymbol{n} \pm \frac{c}{\varepsilon},
$$
(2)

which means that there is a convective and two acoustic waves. In what follows, we assume that the reference Mach number ε is small, i.e., $\varepsilon \ll 1$, and all the other quantities are $\mathcal{O}(1)$, which physically means that the solution is a small disturbance of the incompressible solution. Indeed, it can be shown that under suitable requirements on initial and boundary data ("well-preparedness"), there is convergence of density and momentum $(\rho, \rho \boldsymbol{u})$ towards its incompressible counterpart as $\varepsilon \to 0$, see [5, 6, 7] and the references therein. Furthermore, it is obvious that this problem constitutes a *singularly perturbed equation* in ε, as the equations change type in the limit.

Due to the change of type as $\varepsilon \to 0$, the equations get extremely stiff and therefore it is highly non-trivial to design efficient and stable algorithms. Explicit-in-time solving techniques have the drawback that they lead to a CFL condition in which the time step size Δt must be proportional to $\varepsilon \Delta x$, where Δx is a measure for the spatial grid size. If it is not the goal to accurately resolve all the features, but only to resolve the convective part of the flow, this condition is extremely restrictive, and a so called *convective CFL condition*

$$
\Delta t \lesssim \frac{\Delta x}{\|\boldsymbol{u}\|}
$$
(3)

is preferable. Fully implicit-in-time methods, on the other hand, which are stable under such a CFL condition, tend to add too much spurious diffusion [8].

In the past few years, so called IMEX (implicit-explicit) splitting schemes got more and more popular for solving compressible flow problems, especially for low Mach numbers, see e.g. [9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19] and the references therein. Optimally, such a scheme should be designed in a way that slow waves are handled with an explicit (thus efficient) and fast waves are handled with an implicit (thus unconditionally stable) method. Of course such a strict splitting of waves is only possible in the linear one-dimensional case [20], and therefore, a suitable splitting for the nonlinear multidimensional case has to be defined very carefully.

Over the past few years, many famous splittings for the Euler equations at low Mach number have been designed, beginning by the ground-breaking work of Klein [15]. For a non-exhaustive list, we refer to [10, 12, 14] and the references therein. However, many of those splittings have their shortcomings. It has been reported [21] that Klein's splitting seems to be unstable in some instances. (Which does not include Klein's original algorithm as it is based on a semi discrete decoupling of the pressure.) Furthermore, all of the mentioned splittings need a physical intuition and are not directly extendable to other singularly perturbed differential equations.

To partly overcome these shortcomings, we have over the past few years developed a new type of splitting that is based on the $\varepsilon = 0$ ("incompressible") solution of the problem. The splitting, termed RS-IMEX (see Sec. 3), is generic in the sense that it can in principle be applied to any type of singularly perturbed equation, including singularly perturbed ODEs [1] and the isentropic Euler equations [2]. Related ideas have already been published earlier, for the shallow water equations in [10, 22] and for kinetic equations in [23], a stability analysis of the splitting has been done in [21] and [24].

In [1], we have applied the splitting idea to singularly perturbed ordinary differential equations with high-order IMEX discretizations, namely IMEX linear multistep methods [25, 26, 27] and IMEX Runge-Kutta methods [28, 29, 30, 31, 32, 33, 34]. In [2], we have applied the splitting idea to a low-order finite volume scheme for the isentropic Euler equations. In both publications, we have seen that the newly developed splitting can be highly advantageous. This present work is a 'natural' extension of those previous works: We combine a high-order-in-time IMEX Runge-Kutta scheme with a high-order-in-space discontinuous Galerkin (DG) method (see [35, 36, 37, 38, 39] for classical DG and [40, 41, 42] for IMEX DG) using the newly developed splitting. The difficulty herein lies in the subtle interplay of the stiffness induced by the singular character of the equation and the stiffness induced by the high-order approximation of both spatial and temporal variables. We show how to choose the numerical viscosities in such a way that the resulting method is *asymptotically consistent*, see e.g. [43], which means that its $\varepsilon \to 0$ limit is a consistent discretization of the corresponding incompressible equations. Numerical results show the convergence of the method. It turns out that the overal scheme is indeed stable under a convective CFL condition (3), order degradation is not observed.

This paper is organized as follows: The governing equations are discussed in Sec. 2, the splitting and corresponding IMEX time integration are presented in Sec. 3. The fully discrete method is introduced in Sec. 4, with its asymptotic consistency property being discussed in Sec. 5. Numerical results are shown in Sec. 6. As usual, the paper ultimately gives some conclusion and outlook in the last Sec. 7. To make the paper more self-consistent, the Butcher tableaux for the used IMEX Runge-Kutta schemes are shown in the appendix in Sec. A.

2. Governing equations

Let $\Omega \subset \mathbb{R}^2$ be a two-dimensional domain, and consider the isentropic Euler equations as in (1), with $\rho \in \mathbb{R}$ density and $\boldsymbol{u} = (u, v)^T \in \mathbb{R}^2$ velocity in $x-$ and $y-$direction, respectively. p denotes pressure given for polytropic fluids as $p(\rho) := \kappa \rho^\gamma$ with a $\kappa > 0$ and a $\gamma \geq 1$. Note that the (scaled) characteristic Mach

number ε is given by

$$\varepsilon := \frac{u^*}{\sqrt{p(\rho^*)/\rho^*}},$$

where u^* and ρ^* are the corresponding characteristic values for velocity and density, respectively, used to nondimensionalize the equation. The isentropic Euler equations can directly be rewritten as a conservation law in divergence form

$$\boldsymbol{w}_t + \nabla \cdot \boldsymbol{f}(\boldsymbol{w}) = 0, \qquad \forall x \in \Omega, \ t \in (0, T) \tag{4}$$

$$\boldsymbol{w}(x, t = 0) = \boldsymbol{w}^0(x), \quad \forall x \in \Omega \tag{5}$$

with

$$\boldsymbol{w} := \begin{pmatrix} \rho \\ \rho\boldsymbol{u} \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{f}(\boldsymbol{w}) := \begin{pmatrix} \rho\boldsymbol{u} \\ \rho\boldsymbol{u} \otimes \boldsymbol{u} + \frac{1}{\varepsilon^2} p \cdot \text{Id} \end{pmatrix},$$

where Id denotes the two dimensional identity matrix. $\boldsymbol{w}^0$ are given initial data. Computing the eigenvalues of $\partial_{\boldsymbol{w}} \boldsymbol{f}(\boldsymbol{w}) \cdot \boldsymbol{n}$ gives the characteristic wave speeds

$$\lambda_1 = \boldsymbol{u} \cdot \boldsymbol{n} \qquad \text{and} \qquad \lambda_{2,3} = \boldsymbol{u} \cdot \boldsymbol{n} \pm \frac{c}{\varepsilon}, \tag{2}$$

where $c = \sqrt{\frac{\gamma p}{\rho}}$ denotes the speed of sound of the system. Obviously, these eigenvalues are on different scales w.r.t. ε. Scales can be best understood by considering an *asymptotic expansion* of every quantity, namely

$$\boldsymbol{w} = \boldsymbol{w}_{(0)} + \varepsilon\boldsymbol{w}_{(1)} + \varepsilon^2\boldsymbol{w}_{(2)} + \mathcal{O}(\varepsilon^3). \tag{6}$$

Inserting this expansion into the isentropic Euler equations (1), collecting terms with equal power of ε and taking the limit $\varepsilon \to 0$ leads to the incompressible Euler equations [5]

$$\rho_{(0)} \equiv \text{const} > 0, \quad \nabla \cdot \boldsymbol{u}_{(0)} = 0,$$

$$(\boldsymbol{u}_{(0)})_t + \nabla \cdot (\boldsymbol{u}_{(0)} \otimes \boldsymbol{u}_{(0)}) + \frac{\nabla p_{(2)}}{\rho_{(0)}} = 0. \tag{7}$$

The existence of a limit necessitates the use of specially designed initial data, see e.g. [5, 6, 7] and the references therein, which we introduce in the sequel for the isentropic Euler equations:

Definition 1 (Well prepared initial conditions)**.** *We call initial data $\boldsymbol{w}^0 = (\rho^0, \rho^0\boldsymbol{u}^0)^T$ for the compressible equation* well prepared *if they can be represented by an asymptotic expansion as in (6) and fulfill*

$$\rho^0 = \text{const} + \mathcal{O}(\varepsilon^2), \quad \nabla \cdot \boldsymbol{u}^0 = \mathcal{O}(\varepsilon).$$

Well prepared initial data, together with sufficient smoothness, guarantee the convergence of the solution as $\varepsilon \to 0$ [5].

3. RS-IMEX time integration

The core idea of IMEX schemes is to separate *stiff* and *non-stiff* parts, and then to treat the former ones implicitly, and the latter ones explicitly. For the ease of presentation, we start by considering the simplest setting of all IMEX frameworks, the IMEX-Euler semi discretization to define the splitting. Then, we extend the proceeding to IMEX Runge-Kutta methods.

The temporal domain is given by $[0, T]$ with $T \in \mathbb{R}^+$. To define our methods, we have to split this domain into $N + 1$ time instances $t^n := n\Delta t$,

$$0 = t^0 < \ldots < t^n < \ldots < t^N = T.$$

Uniform time slabs are not a necessity, but are used for notational convenience.

3.1. RS-IMEX splitting

We assume for the moment that a splitting of the convective flux into

$$\boldsymbol{f}(\boldsymbol{w}) = \widetilde{\boldsymbol{f}}(\boldsymbol{w}) + \widehat{\boldsymbol{f}}(\boldsymbol{w}) \tag{8}$$

is already given. Then, applied to (4), the time-discrete IMEX-Euler scheme is defined by

$$\boldsymbol{w}^{n+1} - \boldsymbol{w}^n + \Delta t \nabla \cdot \left(\widetilde{\boldsymbol{f}}(\boldsymbol{w}^{n+1}) + \widehat{\boldsymbol{f}}(\boldsymbol{w}^n) \right) = 0. \tag{9}$$

Note that $\widetilde{\boldsymbol{f}}$ is the part that is treated implicitly, while $\widehat{\boldsymbol{f}}$ is the part that is treated explicitly. In the following the upper index n of $\boldsymbol{w}^n$ corresponds to the numerical solution - with respect to time - at time instance t^n.

As already pointed out in the introduction, the choice of a splitting of the convective flux function $\boldsymbol{f}$ is a core ingredient to obtain a stable and efficient numerical method in the low-Mach range. This work relies on the recently introduced RS-IMEX splitting [1, 2, 21], where RS stands for *reference solution* and denotes the limit-solution $\boldsymbol{w}_{(0)} = (\rho_{(0)}, \rho_{(0)} \boldsymbol{u}_{(0)})^T$, see (7). The splitting relies on a linearization of the flux function $\boldsymbol{f}$ around $\boldsymbol{w}_{(0)}$ being used in the stiff part. For a more detailed derivation of the RS-IMEX splitting we refer to [1, 2], but also to [23, 10, 22] for earlier applications of a similar idea. More formally the RS-IMEX splitting is given in the following definition.

Definition 2 (RS-IMEX). *The* RS-IMEX *splitting is defined by*

$$\widetilde{\boldsymbol{f}}(\boldsymbol{w}) = \boldsymbol{f}(\boldsymbol{w}_{(0)}) + \partial_{\boldsymbol{w}} \boldsymbol{f}(\boldsymbol{w}_{(0)}) \cdot (\boldsymbol{w} - \boldsymbol{w}_{(0)}),$$
$$\widehat{\boldsymbol{f}}(\boldsymbol{w}) = \boldsymbol{f}(\boldsymbol{w}) - \widetilde{\boldsymbol{f}}(\boldsymbol{w}),$$

where $\boldsymbol{w}_{(0)}$ denotes the asymptotic solution solution $\boldsymbol{w}_{(0)} = (\rho_{(0)}, \rho_{(0)} \boldsymbol{u}_{(0)})^T$ from (7), $\partial_{\boldsymbol{w}} \boldsymbol{f}$ denotes the Jacobian of $\boldsymbol{f}$.

Due to its definition, $\widetilde{f}$ is linear in $\boldsymbol{w}$ and, as this part is treated implicitly, the resulting system can be solved efficiently by a linear solution technique. Note that, although the idea stems from a linearization, there is no second-order linearization error of the flux, because remaining terms are collected in $\widehat{\boldsymbol{f}}$. Applying the definition of $\boldsymbol{f}$ given in (4) to the RS-IMEX splitting, one can directly compute the flux functions $\widehat{\boldsymbol{f}}$ and $\widetilde{\boldsymbol{f}}$ for the isentropic Euler equations:

Definition 3 (RS-IMEX splitting for the isentropic Euler equations). *The RS-IMEX splitting for the istentropic Euler equations is given by*

$$\widetilde{\boldsymbol{f}}(\boldsymbol{w}) = \begin{pmatrix} \rho \boldsymbol{u} \\ -\rho \boldsymbol{u}_{(0)} \otimes \boldsymbol{u}_{(0)} + \rho \boldsymbol{u} \otimes \boldsymbol{u}_{(0)} + \rho \boldsymbol{u}_{(0)} \otimes \boldsymbol{u} + \frac{1}{\varepsilon^2} \left(p(\rho_{(0)}) + p'(\rho_{(0)})(\rho - \rho_{(0)}) \right) \cdot \mathrm{Id} \end{pmatrix},$$
$$\widehat{\boldsymbol{f}}(\boldsymbol{w}) = \begin{pmatrix} 0 \\ \rho(\boldsymbol{u} - \boldsymbol{u}_{(0)}) \otimes (\boldsymbol{u} - \boldsymbol{u}_{(0)}) + \frac{1}{\varepsilon^2} \left(p(\rho) - p(\rho_{(0)}) - p'(\rho_{(0)})(\rho - \rho_{(0)}) \right) \cdot \mathrm{Id} \end{pmatrix}.$$

Remark 1. *Note that the RS-IMEX splitting idea as given in Def. 2 can directly be extended to a wide range of different singularly perturbed equations.*

Remark 2. *In contrast to $\boldsymbol{f}$, both $\widehat{\boldsymbol{f}}$ and $\widetilde{\boldsymbol{f}}$ depend - through the use of the reference quantity $\boldsymbol{w}_{(0)}$ - explicitly on t. We do not add t as an additional variable to the fluxes to keep the notation short. It will become important in the definition of the IMEX-Runge-Kutta scheme, because technically, we do not treat an autonomous differential equation any more.*

That $\widehat{\boldsymbol{f}}$ is indeed 'non-stiff' is indicated by the following lemma:

Lemma 1. *1. The eigenvalues of $\partial_{\boldsymbol{w}}\widehat{\boldsymbol{f}}(\boldsymbol{w})\cdot\boldsymbol{n}$ of are given by*

$$\widehat{\boldsymbol{\lambda}} = \begin{pmatrix} 0 \\ (\boldsymbol{u}-\boldsymbol{u}_{(0)})\cdot\boldsymbol{n} \\ 2(\boldsymbol{u}-\boldsymbol{u}_{(0)})\cdot\boldsymbol{n} \end{pmatrix}.$$

2. The in magnitude largest eigenvalues of the Jacobian of the implicit part are in $\mathcal{O}\left(\frac{1}{\varepsilon}\right)$.

We can conclude two different things from La. 1. First, the stiffness of the equation is completely hidden in the implicit part. Second, if we take the limit $\varepsilon \to 0$, the influence of the explicit part vanishes.

3.2. IMEX Runge-Kutta method

We have discussed the RS-IMEX splitting in the context of a straightforward IMEX-Euler discretization, see (9). The extension to higher-order methods is evident, methods of choice are, e.g., high-order IMEX Runge-Kutta methods [28, 29, 30, 31, 32, 33, 34] or high-order IMEX linear multistep methods [25, 26, 27]. In this work, we consider IMEX Runge-Kutta methods, where we restrict ourselves to a (relatively large) subclass which we identified as important in our previous work [1]:

- We only consider IMEX Runge-Kutta methods which are *globally stiffly accurate* (GSA), see e.g. [29]. In short this is fulfilled if the update step is equal to the last internal stage of the Runge-Kutta method. This corresponds to the *first same as last* property for an explicit and the *stiffly accurate* property for an implicit Runge-Kutta method. The IMEX Runge-Kutta methods are fully defined by the two Butcher tableaux $\widetilde{A}$ and $\widehat{A}$ and the corresponding temporal coefficients $\widetilde{c}$ and $\widehat{c}$.

- We only consider IMEX Runge-Kutta methods where the implicit matrix $\widetilde{A}$ is a lower triangular one, such that in every internal stage only one implicit variable occurs. This is mostly due to efficiency reasons.

- We only consider IMEX Runge-Kutta methods of type A or type CK. This is given if the implicit matrix $\widetilde{A}$ is invertible (type A) or the first entry of the implicit matrix equals 0 and the remaining submatrix is invertible (type CK). See Def. 4 for more details. For a more detailed classification of IMEX Runge-Kutta methods we refer to [44].

In the following, we first introduce such a Runge-Kutta scheme for the semi-discrete-in-time discretization of the Euler equations (4).

Definition 4 (GSA IMEX Runge-Kutta scheme for (4))**.** *For every $t^{n+1} = t^n + \Delta t$ do the following:*

1. For $i = 1,\dots,s$ solve

$$\boldsymbol{w}^{n,i} - \boldsymbol{w}^n + \Delta t \left(\sum_{j=1}^{i} \widetilde{A}_{i,j} \nabla \cdot \widetilde{\boldsymbol{f}}(\boldsymbol{w}^{n,j}) + \sum_{j=1}^{i-1} \widehat{A}_{i,j} \nabla \cdot \widehat{\boldsymbol{f}}(\boldsymbol{w}^{n,j}) \right) = 0, \qquad (10)$$

where $\boldsymbol{w}^{n,i}$ denotes the solution of the ith internal stage. Note that $\widetilde{f}$ is evaluated at time $\widetilde{t}^{n,j}$, and $\widehat{f}$ at $\widehat{t}^{n,j}$, with

$$\widetilde{t}^{n,j} := t^n + \widetilde{c}_j \Delta t, \qquad \widehat{t}^{n,j} := t^n + \widehat{c}_j \Delta t,$$

see also Rem. 2.

2. Set $\boldsymbol{w}^{n+1} := \boldsymbol{w}^{n,s}$.

The coefficients of the IMEX RK method are given by two Butcher tableaux, the one with overhats referring to the explicit, the other to the implicit part. Because of our restrictions on the Runge-Kutta method, the implicit coefficient matrix has to fulfill $\widetilde{A}_{ii} \neq 0$ for $i = 2 \ldots s$. For a type A method, there even holds $\widetilde{A}_{11} \neq 0$ in addition.

Based on our work in [1], we use the IMEX Runge-Kutta methods presented in Tbl. 2, 3, 4 and 5; also given in [28, 33, 34]. A classification of these methods can be seen in Tbl. 1.

4. IMEX DG method

High-order temporal integration has to be coupled to a high-order spatial discretization. The method of choice of the latter in this work is a combination of a high-order IMEX Runge-Kutta method with a high-order discontinuous Galerkin (DG) discretization [35, 36, 37, 38, 39], yielding an IMEX DG method [40, 41, 42].

4.1. Preliminary definitions

We assume that the periodic domain $\Omega \subset \mathbb{R}^2$ is divided into ne $\in \mathbb{N}$ non-overlapping cells Ω_k as

$$\bigcup_{k=1}^{ne} \overline{\Omega_k} = \Omega \qquad \text{and} \qquad \Omega_k \cap \Omega_i = \emptyset \quad \forall k \neq i.$$

The boundary of the cell Ω_k is denoted by $\partial \Omega_k$ and $\boldsymbol{n}_k$ denotes the corresponding outward normal vector. On this triangulation $\{\Omega_k\}$ we define a broken polynomial space by

$$V_q := \{v \in L^2(\Omega) \; : \; v_{|\Omega_k} \in \mathbb{P}_q(\Omega_k) \; \forall \; k = 1, \ldots, ne\},$$

where $\mathbb{P}_q(\Omega_k)$ denotes the space of all polynomial functions with maximum degree q on cell Ω_k. For system-valued functions (there are three components in the Euler equations) we define the corresponding space

$$V_q^3 := V_q \times V_q \times V_q.$$

Of course an adaptive choice of q is possible. For a value $x \in \partial \Omega_k$, we define the interior $(-)$ and exterior $(+)$ value, respectively, of a function $\sigma \in V_q$ by

$$\sigma^{\mp}(x) := \lim_{0 < \delta \to 0} \sigma(x \mp \delta \boldsymbol{n}_k). \tag{11}$$

If a boundary is considered independently of a specific cell, we can in a similar way define a value of $\sigma^{\mp}$ based on an arbitrary, but fixed direction of edge normal vectors.

4.2. IMEX Runge-Kutta Discontinuous Galerkin method

Following the common steps [39], we can define the DG residual of both $\nabla \cdot \widehat{f}(w)$ and $\nabla \cdot \widetilde{f}(w)$ by the quantities

$$\widehat{\mathbf{R}}(\boldsymbol{w}_{\Delta x}; \boldsymbol{\varphi}) \; := \; -\int_{\Omega} \widehat{\boldsymbol{f}}(\boldsymbol{w}_{\Delta x}) \cdot \nabla \boldsymbol{\varphi} \mathrm{d}x + \sum_{k=1}^{ne} \int_{\partial \Omega_k} \widehat{\boldsymbol{h}}(\boldsymbol{w}_{\Delta x}^-, \boldsymbol{w}_{\Delta x}^+) \boldsymbol{\varphi} \cdot \boldsymbol{n_k} \mathrm{d}s, \qquad \text{and}$$

$$\widetilde{\mathbf{R}}(\boldsymbol{w}_{\Delta x}; \boldsymbol{\varphi}) \; := \; -\int_{\Omega} \widetilde{\boldsymbol{f}}(\boldsymbol{w}_{\Delta x}) \cdot \nabla \boldsymbol{\varphi} \mathrm{d}x + \sum_{k=1}^{ne} \int_{\partial \Omega_k} \widetilde{\boldsymbol{h}}(\boldsymbol{w}_{\Delta x}^-, \boldsymbol{w}_{\Delta x}^+) \boldsymbol{\varphi} \cdot \boldsymbol{n_k} \mathrm{d}s,$$

respectively. Note that integration over Ω is to be understood in the cell-wise sense. $\widetilde{\boldsymbol{h}}$ and $\widehat{\boldsymbol{h}}$ are stiff and non-stiff numerical flux function, respectively, given by

$$\widetilde{\boldsymbol{h}}(\boldsymbol{w}^-, \boldsymbol{w}^+) := \frac{1}{2}\left(\widetilde{\boldsymbol{f}}(\boldsymbol{w}^-) + \widetilde{\boldsymbol{f}}(\boldsymbol{w}^+)\right) + \frac{1}{2}\operatorname{Diag}\left(\frac{1}{\varepsilon^2}, 1, 1\right)(\boldsymbol{w}^- - \boldsymbol{w}^+) \cdot \boldsymbol{n}, \tag{12}$$

$$\widehat{\boldsymbol{h}}(\boldsymbol{w}^-, \boldsymbol{w}^+) := \frac{1}{2}\left(\widehat{\boldsymbol{f}}(\boldsymbol{w}^-) + \widehat{\boldsymbol{f}}(\boldsymbol{w}^+)\right) + \varepsilon\left(\boldsymbol{w}^- - \boldsymbol{w}^+\right) \cdot \boldsymbol{n}. \tag{13}$$

Remark 3. *The numerical flux is of Rusanov-type.*

- *Let us note that a somewhat similar choice of the stiff stabilization, for the equations in primitive variables, has been made in [45], motivated by the fundamental work of Turkel [46], who introduced preconditioning of the time derivative to enhance steady-state computations for low-Mach flows.*

- *The choice of the non-stiff stabilization is motivated by La. 1, as the eigenvalues of $\partial_{\boldsymbol{w}}\widehat{\boldsymbol{f}}(\boldsymbol{w}) \cdot \boldsymbol{n}$ are in $\mathcal{O}(\varepsilon)$ if one assumes that $\boldsymbol{u} = \boldsymbol{u}_{(0)} + \mathcal{O}(\varepsilon)$.*

- *As observed in [10] and [2], the choice of the numerical flux function affects asymptotic consistency. The choice here guarantees the latter important property.*

The extension of Def. 4 to the fully discrete DG scheme can be done in a straightforward way by replacing fluxes $\boldsymbol{f}$ by discrete fluxes $\mathbf{R}$. To get the notation right, we shortly review this discretization here:

Definition 5 (High-order method for weakly compressible flows). *For every $t^{n+1} = t^n + \Delta t$ do the following:*

1. *For $i = 1, \ldots, s$ solve*

$$\int_{\Omega}\left(\boldsymbol{w}_{\Delta x}^{n,i} - \boldsymbol{w}_{\Delta x}^n\right)\boldsymbol{\varphi}\,\mathrm{d}x + \Delta t\left(\sum_{j=1}^{i}\widetilde{A}_{i,j}\widetilde{\mathbf{R}}(\boldsymbol{w}_{\Delta x}^{n,j}; \boldsymbol{\varphi}) + \sum_{j=1}^{i-1}\widehat{A}_{i,j}\widehat{\mathbf{R}}(\boldsymbol{w}_{\Delta x}^{n,j}; \boldsymbol{\varphi})\right) = 0 \quad \forall\boldsymbol{\varphi} \in V_q^3, \tag{14}$$

where $\boldsymbol{w}_{\Delta x}^{n,i}$ denotes the solution of the ith internal stage. Also $\widetilde{\mathbf{R}}$ and $\widehat{\mathbf{R}}$ depend on time t and are evaluated at

$$\widetilde{t}^{n,j} := t^n + \widetilde{c}_j\Delta t, \qquad \text{and} \qquad \widehat{t}^{n,j} := t^n + \widehat{c}_j\Delta t,$$

respectively.

2. *Set $\boldsymbol{w}_{\Delta x}^{n+1} := \boldsymbol{w}_{\Delta x}^{n,s}$.*

In Def. 5, we have summarized the final algorithm to be used in this work. With this, we are now ready to prove asymptotic consistency.

5. Asymptotic consistency

As mentioned in Sec. 1, our aim is to develop a method whose $\varepsilon \to 0$ limit is a consistent discretization of the limit equation (7), which means that it preserves the asymptotic behavior of the corresponding equation. We prove that our method is *asymptotically consistent*, for the ease of presentation in two steps:

1. First, we consider the semi discrete (discrete in time) setting (10).

2. Then, we consider the fully discrete setting (14).

Unfortunately, the methods we have introduced require a lot of notation. The following list gives an overview of the terms we use.

Remark 4 (Notation). *1. An upper index n, e.g., $\boldsymbol{u}^n$, indicates that the quantity is given at time level $t = t^n$.*

2. *An additional upper index i, e.g., $\boldsymbol{u}^{n,i}$, denotes the ith internal stage of an IMEX Runge-Kutta method.*

3. *A lower index Δx, e.g., $\boldsymbol{u}_{\Delta x}$, denotes a variable which belongs to a discontinuous Galerkin discretization.*

4. *An additional upper index $-$ or $+$, e.g., $\boldsymbol{u}_{\Delta x}^{\mp}$, denotes the interior or exterior value corresponding to an edge, see (11).*

5. *A lower index in brackets (i), e.g., $\boldsymbol{u}_{(i)}$, denotes a variable which belongs to the ith component of an asymptotic expansion, see (6)*

5.1. Semi discrete setting

We start by considering the RS-IMEX splitting for the isentropic Euler equation (1) coupled to an IMEX Runge-Kutta temporal discretization as in (10).

Theorem 1. *The RS-IMEX splitting, given in Def. 3, coupled to an IMEX Runge-Kutta temporal discretization as given in Def. 4 is asymptotically consistent if well prepared initial data at time $t = 0$ and periodic boundary conditions are used.*

Proof. We first show that, given $\boldsymbol{w}^n$ is well-prepared, also $\boldsymbol{w}^{n+1}$ is well-prepared. Because $\boldsymbol{w}^0$ is well-prepared, one can then inductively prove that all $\boldsymbol{w}^n$ are well-prepared.

We assume that all the (discrete) quantities can be represented with an asymptotic expansion as in (6), e.g.,

$$(\rho\boldsymbol{u})^{n,j} = (\rho\boldsymbol{u})_{(0)}^{n,j} + \varepsilon(\rho\boldsymbol{u})_{(1)}^{n,j} + \dots.$$

If $\widetilde{A}_{1,1} = 0$, which happens for type CK methods, then the first internal stage is equal to the previous time instance $\boldsymbol{w}_{\Delta x}^n$. It is therefore directly well prepared. Therefore, we consider the ith internal stage with $\widetilde{A}_{ii} \neq 0$.

Because the numerical density is constant up to $\mathcal{O}(\varepsilon^2)$, we know that its zeroth-order expansion is equal to the reference density $\rho_{(0)}$. Therefore, considering the $\mathcal{O}(\varepsilon^{-2})$ terms of the momentum equation, we obtain

$$0 = \nabla \frac{\widetilde{A}_{i,i}}{\varepsilon^2} \left(p(\rho_{(0)}) + p'(\rho_{(0)})(\rho_{(0)}^{n,i} - \rho_{(0)}) \right)$$

$$\Leftrightarrow 0 = \nabla \frac{\widetilde{A}_{i,i}}{\varepsilon^2} \left(p'(\rho_{(0)})\rho_{(0)}^{n,i} \right)$$

$$\Leftrightarrow 0 = \nabla \rho_{(0)}^{n,i}.$$

Thus the limit density is constant in space. Next we consider the $\mathcal{O}(1)$ terms of the first equation and integrate over the whole domain. Using the periodic boundary conditions we get

$$\int_\Omega \rho_{(0)}^{n,i} - \rho_{(0)}^n \mathrm{d}x = 0.$$

Since both values are constant in space, we can conclude that $\rho_{(0)}^{n,i}$ is constant in i, and therefore it is equal to $\rho_{(0)}$. Considering again the $\mathcal{O}(1)$ terms of the mass equation we now obtain

$$\sum_j \widetilde{A}_{ij} \nabla \cdot (\rho\boldsymbol{u})_{(0)}^{n,j} = 0.$$

8

Since $\boldsymbol{w}^n$ is well-prepared, one can then inductively show that all stage values $(\rho\boldsymbol{u})^{n,i}_{(0)}$ are solenoidal, and one can directly obtain $\nabla \cdot (\rho\boldsymbol{u})^{n,i}_{(0)} = 0$ if $\widetilde{A}_{ii} \neq 0$. This means that all stage values are well-prepared. Because the underlying IMEX RK method is globally stiffly accurate, the update step is equal to the last stage. This automatically proves that the values $\boldsymbol{w}^{n+1}$ are well-prepared.

The proof is finalized by the remark that the discrete limit momentum equation is a consistent discretization of the limit momentum equation. $\qquad\square$

A question which arises from the use of the RS-IMEX splitting is how to compute the limit solution. In an ideal case this solution is given, but generally we need a numerical method for its computation. It is useful to compute the limit solution in such a way that it corresponds to the solution of the limit method.

Theorem 2. *The limit of the semi discrete method* (10) *is a discretization that is fully implicit-in-time.*

Proof. We consider the ith internal stage and add a zero as

$$\rho_{(0)}\boldsymbol{u}^{n,j}_{(0)} \otimes \boldsymbol{u}^{n,j}_{(0)} - \rho_{(0)}\boldsymbol{u}^{n,j}_{(0)} \otimes \boldsymbol{u}^{n,j}_{(0)},$$

then the limit numerical method reads (see also Def. 3)

$$
\begin{pmatrix} 0 \\ \rho_{(0)}\boldsymbol{u}^{n,i}_{(0)} \end{pmatrix} = \begin{pmatrix} 0 \\ \rho_{(0)}\boldsymbol{u}^{n}_{(0)} \end{pmatrix}
$$
$$
- \Delta t \sum_j \widetilde{A}_{ij} \nabla \cdot \begin{pmatrix} \boldsymbol{u}^{n,j}_{(0)} \\ -\rho_{(0)}(\boldsymbol{u}^{n,j}_{(0)} - \boldsymbol{u}_{(0)}(\widetilde{t}^{n,j})) \otimes (\boldsymbol{u}^{n,j}_{(0)} - \boldsymbol{u}_{(0)}(\widetilde{t}^{n,j})) + \rho_{(0)}\boldsymbol{u}^{n,j}_{(0)} \otimes \boldsymbol{u}^{n,j}_{(0)} + p'(\rho_{(0)})\rho^{n,j}_{(2)} \cdot \mathrm{Id} \end{pmatrix}
$$
$$
- \Delta t \sum_j \widehat{A}_{ij} \nabla \cdot \begin{pmatrix} 0 \\ \rho_{(0)}(\boldsymbol{u}^{n,j}_{(0)} - \boldsymbol{u}_{(0)}(\widehat{t}^{n,j})) \otimes (\boldsymbol{u}^{n,j}_{(0)} - \boldsymbol{u}_{(0)}(\widehat{t}^{n,j})) + \left(p^{n,j}_{(2)} - p'(\rho_{(0)})\rho^{n,j}_{(2)} \right) \cdot \mathrm{Id} \end{pmatrix} \cdot
$$

We show that the limit method corresponds to a fully implicit method with the help of mathematical induction. Therefore we assume that the reference solution equals to the limit numerical solution for the $i-1$ previous stages (which should be given for the first instance due to the initial data). Additionally, from the asymptotic expansion one concludes

$$p^{n,i}_{(2)} = p'(\rho_{(0)})\rho^{n,i}_{(2)}.$$

Finally, this all together simplifies to

$$
\begin{pmatrix} 0 \\ \rho_{(0)}\boldsymbol{u}^{n,i}_{(0)} \end{pmatrix} = \begin{pmatrix} 0 \\ \rho_{(0)}\boldsymbol{u}^{n}_{(0)} \end{pmatrix}
$$
$$
+ \Delta t \sum_j \widetilde{A}_{ij} \nabla \cdot \begin{pmatrix} \boldsymbol{u}^{n,i}_{(0)} \\ -\rho_{(0)}(\boldsymbol{u}^{n,i}_{(0)} - \boldsymbol{u}_{(0)}(\widetilde{t}^{n,i})) \otimes (\boldsymbol{u}^{n,i}_{(0)} - \boldsymbol{u}_{(0)}(\widetilde{t}^{n,i})) + \rho_{(0)}\boldsymbol{u}^{n,i}_{(0)} \otimes \boldsymbol{u}^{n,i}_{(0)} + p^{n,i}_{(2)} \cdot \mathrm{Id} \end{pmatrix},
$$

which is a fully implicit discretization of the incompressible equation with additional terms in $(\boldsymbol{u}^{n,i}_{(0)} - \boldsymbol{u}_{(0)}(\widetilde{t}^{n,i}))$. If $\boldsymbol{u}_{(0)}(\widetilde{t}^{n,i})$ has been computed by a fully implicit method (which takes only the implicit part of the used IMEX Runge-Kutta method), the two solutions correspond to each other. This concludes the proof. $\qquad\square$

5.2. Fully discrete setting

Here, we consider the fully discrete setting, i.e., temporal discretization with an IMEX Runge-Kutta method and spatial discretization with a DG method, see (14). To clarify the choice of the numerical diffusion coefficients in (12), we start with the following lemma:

Lemma 2. *Let the function $\sigma_{\Delta x} \in V_q$ be such that*

$$\int_{\partial \Omega_k} (\sigma_{dx}^- - \sigma_{\Delta x}^+) \varphi^- \mathrm{d}s = 0, \qquad \forall \varphi \in V_q, \ \forall k = 1, \dots, \mathrm{ne} \,. \tag{15}$$

Then, $\sigma_{\Delta x}$ is continuous.

Proof. We can choose $\varphi = \sigma_{\Delta x}$ in (15) and obtain

$$\int_{\partial \Omega_k} (\sigma_{\Delta x}^- - \sigma_{\Delta x}^+) \sigma_{\Delta x}^- \mathrm{d}s = 0$$

on every cell Ω_k. Summing up over the whole domain and rearranging terms leads to

$$0 = \sum_k \int_{\partial \Omega_k} (\sigma_{\Delta x}^- - \sigma_{\Delta x}^+) \sigma_{\Delta x}^- \mathrm{d}s = \sum_e \int_e (\sigma_{\Delta x}^- - \sigma_{\Delta x}^+)^2 \mathrm{d}s.$$

This means that $\sigma_{\Delta x}^- = \sigma_{\Delta x}^+$ and therefore the quantity $\sigma_{\Delta x}$ is continuous over every cell boundary. $\square$

This lemma has a direct consequence for the numerical solution, namely, if we can show that the numerical stabilization of one quantity lives on a different scale (with respect to ε) than the rest of the corresponding equation, the $\varepsilon = 0$ limit of this quantity is continuous. We will apply this to the momentum equation and show that the discrete approximation to $\rho_{(0)}$ is continuous, and one can then easily prove that it is constant.

Theorem 3. *The RS-IMEX splitting, given in Def. 3, coupled to an IMEX Runge-Kutta temporal discretization as given in Def. 5 is asymptotically consistent if periodic boundary conditions and discretely well prepared initial data, see (16), at time $t = 0$ are used.*

Proof. The proof is similar as before. We show inductively, starting from $n = 0$, that given well-prepared values $\boldsymbol{w}_{\Delta x}^n$, the algorithm preserves the well-preparedness. More precisely, we show that the internal stage $\boldsymbol{w}_{\Delta x}^{n,i}$ of the RS-IMEX DG method fulfills $\rho_{\Delta x}^{n,i} = \rho_{(0)} + \mathcal{O}(\varepsilon^2)$, and $\nabla \cdot u_{\Delta x,(0)} = 0$ in a discrete sense, see (16) (i.e., it is well prepared in a discrete sense) if all the previous internal stages and the previous time instances are also discretely well prepared, and there holds that $\rho_{\Delta x}^n = \rho_{(0)} + \mathcal{O}(\varepsilon^2)$. Together with the well-preparedness at time $t = 0$ and the GSA property, this yields the well-preparedness of all discrete quantities.

Note that if $\widetilde{A}_{1,1} = 0$ then the first internal stage is equal to the previous time instance $\boldsymbol{w}_{\Delta x}^n$, thus it is directly well prepared. Therefore we now consider a given i such that $\widetilde{A}_{i,i} \neq 0$. We assume that every quantity can be expressed by an asymptotic expansion as in (6), e.g.,

$$\rho_{\Delta x}^n = \rho_{\Delta x,(0)}^n + \varepsilon \rho_{\Delta x,(1)}^n + \varepsilon^2 \rho_{\Delta x,(2)}^n + \mathcal{O}(\varepsilon^3).$$

Due to the numerical stabilization the only terms in $\mathcal{O}(\varepsilon^{-2})$ in the momentum equation are the pressure terms, thus

$$\begin{aligned} 0 = &\widetilde{A}_{i,i} \int_{\Omega_k} \left(p(\rho_{(0)}) + p'(\rho_{(0)})(\rho_{\Delta x,(0)}^{n,i} - \rho_{(0)}) \right) \nabla \boldsymbol{\varphi} \mathrm{d}x \\ &- \widetilde{A}_{i,i} \frac{1}{2} \int_{\partial \Omega_k} \left(p(\rho_{(0)}) + p'(\rho_{(0)})(\rho_{\Delta x,(0)}^{n,i,-} - \rho_{(0)}) + p(\rho_{(0)}) + p'(\rho_{(0)})(\rho_{\Delta x,(0)}^{n,i,+} - \rho_{(0)}) \right) \boldsymbol{\varphi} \boldsymbol{n}_k \mathrm{d}x \end{aligned}$$

for every test-function $\varphi \in V_q^2$. Note that we have directly used the fact that the initial values and all previous stages are well prepared. Therefore, there are no explicit contributions. Using integration by parts and changing signs leads to

$$
\begin{aligned}
0 =& \widetilde{A}_{i,i} \int_{\Omega_k} \nabla \left(p(\rho_{(0)}) + p'(\rho_{(0)})(\rho_{\Delta x,(0)}^{n,i} - \rho_{(0)}) \right) \varphi \mathrm{d}x \\
& - \widetilde{A}_{i,i} \frac{1}{2} \int_{\partial\Omega_k} \left(p(\rho_{(0)}) + p'(\rho_{(0)})(\rho_{\Delta x,(0)}^{n,i,-} - \rho_{(0)}) - p(\rho_{(0)}) - p'(\rho_{(0)})(\rho_{\Delta x,(0)}^{n,i,+} - \rho_{(0)}) \right) \varphi \boldsymbol{n}_k \mathrm{d}s \\
=& \widetilde{A}_{i,i} \int_{\Omega_k} \nabla p'(\rho_{(0)}) \rho_{\Delta x,(0)}^{n,i} \varphi \mathrm{d}x - \widetilde{A}_{i,i} \frac{1}{2} \int_{\partial\Omega_k} p'(\rho_{(0)})(\rho_{\Delta x,(0)}^{n,i,-} - \rho_{\Delta x,(0)}^{n,i,+}) \varphi \boldsymbol{n}_k \mathrm{d}s.
\end{aligned}
$$

Due to La. 2 and the choice of the implicit stabilization, which is in $\mathcal{O}(\varepsilon^{-2})$ for the first equation, we know that $\rho_{\Delta x,(0)}^{n,i}$ is continuous over the cell boundary of Ω_k. Therefore we obtain

$$
0 = \int_{\Omega_k} \nabla \rho_{\Delta x,(0)}^{n,i} \varphi \mathrm{d}x.
$$

This holds true on every cell Ω_k and for every test-function φ and therefore $\rho_{\Delta x,(0)}^{n,i}$ must be a cell-wise constant. Since it is also continuous it is constant over the whole domain. Similarly, one can also conclude that $\rho_{\Delta x,(1)}^{n,i}$ is constant over the whole domain. Next we consider the $\mathcal{O}(1)$ terms of the conservation of mass equation. Note that, because this part is purely implicit, the reference solution does not occur, so for all $\varphi \in V_q$ there holds

$$
\begin{aligned}
0 =& \int_{\Omega_k} \left(\rho_{\Delta x,(0)}^{n,i} - \rho_{\Delta x,(0)}^{n} \right) \varphi \mathrm{d}x - \Delta t \sum_j \widetilde{A}_{i,j} \int_{\Omega_k} \rho_{\Delta x,(0)}^{n,j} \boldsymbol{u}_{\Delta x,(0)}^{n,j} \cdot \nabla \varphi \mathrm{d}x \\
& + \Delta t \sum_j \widetilde{A}_{i,j} \frac{1}{2} \int_{\partial\Omega_k} \left(\rho_{\Delta x,(0)}^{n,j} \boldsymbol{u}_{\Delta x,(0)}^{n,j,-} + \rho_{\Delta x,(0)}^{n,j} \boldsymbol{u}_{\Delta x,(0)}^{n,j,+} \right) \boldsymbol{n}_k \varphi \mathrm{d}s \\
& + \Delta t \sum_j \widetilde{A}_{i,j} \frac{1}{2} \int_{\partial\Omega_k} \left(\rho_{\Delta x,(2)}^{n,j,-} - \rho_{\Delta x,(2)}^{n,j,+} \right) \varphi \mathrm{d}s.
\end{aligned}
$$

With the help of periodicity we can now choose $\varphi \equiv 1$ as the test function and summing over the whole domain. This leads to

$$
0 = \left(\rho_{\Delta x,(0)}^{n,i} - \rho_{\Delta x,(0)}^{n} \right) |\Omega|.
$$

Consequently, $\rho_{\Delta x,(0)}^{n,i}$ is also constant in time and is equal to $\rho_{(0)}$ because of the requirements on the previous stages and initial conditions. Considering again conservation of mass, this equation can now be written as

$$
\begin{aligned}
0 =& \sum_j \widetilde{A}_{i,j} \int_{\Omega_k} \rho_{(0)} \boldsymbol{u}_{\Delta x,(0)}^{n,j} \cdot \nabla \varphi \mathrm{d}x \\
& - \sum_j \widetilde{A}_{i,j} \frac{1}{2} \int_{\partial\Omega_k} \left(\rho_{(0)} \boldsymbol{u}_{\Delta x,(0)}^{n,j,-} + \rho_{(0)} \boldsymbol{u}_{\Delta x,(0)}^{n,j,+} \right) \boldsymbol{n}_k \varphi \mathrm{d}s \\
& - \sum_j \widetilde{A}_{i,j} \frac{1}{2} \int_{\partial\Omega_k} \left(\rho_{\Delta x,(2)}^{n,j,-} - \rho_{\Delta x,(2)}^{n,j,+} \right) \varphi \mathrm{d}s.
\end{aligned} \tag{16}
$$

This is a consistent discretization of $\nabla \cdot \boldsymbol{u} = 0$ with stabilization terms in $\rho_{\Delta x,(2)}^{n,i}$. This corresponds to a stabilization with the pressure, since $p_{\Delta x,(2)}^{n,i} = \gamma \kappa \rho_{(0)}{}^{\gamma-1} \rho_{\Delta x,(2)}^{n,i}$. Stabilizing the divergence equation with

the pressure is also used in literature for discontinuous Galerkin methods for incompressible equations, see e.g. [47].

As for the semi discrete case, it is straightforward to see that the limit momentum equation is a consistent discretization of the corresponding equation. Thus the method is asymptotically consistent. $\square$

This section is finalized with some remarks:

Remark 5. • *The choice of the numerical flux function is essential for the previous theorem. Taking implicit stabilization coefficients in $\mathcal{O}(1)$, periodic boundary conditions and polynomial degree $q = 0$ results in a method which is not guaranteed to be AC. In [2], this problem has been solved by using a different type of boundary condition, based on the work of [14]. In [48], this problem is solved by adding implicit diffusion to the mass equation, which is similar to the choice of the numerical flux functions presented in this work.*

• *The proof of the asymptotic consistency does not rely on the fact that the equations are two-dimensional. In fact, the three-dimensional case is also covered.*

6. Numerical results

In this section we consider an example with exact solution to investigate the numerical method in terms of stability and accuracy. The *high-order vortex* is given by a pressure function $p(\rho) = \frac{1}{2}\rho^2$ and periodic initial conditions

$$\rho^0(x,y) = 2 + 250,000\varepsilon^2 \begin{cases} \frac{1}{2}e^{\frac{2}{\Delta r}}\Delta r - \mathrm{Ei}\left(\frac{2}{\Delta r}\right) & r < \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$$

$$\boldsymbol{u}^0(x,y) = \begin{pmatrix} 1/2 \\ 0 \end{pmatrix} + 500 \begin{pmatrix} \frac{1}{2} - y \\ x - \frac{1}{2} \end{pmatrix} \cdot \begin{cases} e^{\frac{1}{\Delta r}} & r < \frac{1}{2} \\ 0 & \text{otherwise} \end{cases},$$

where $r := \sqrt{(x - \frac{1}{2})^2 + (y - \frac{1}{2})^2}$ and $\Delta r := r^2 - \frac{1}{4}$. The solution is a transport of the vortex in x-direction, i.e.

$$\rho(x,y,t) = \rho^0\left(x - \frac{1}{2}t, y\right), \quad \boldsymbol{u}(x,y,t) = \boldsymbol{u}^0\left(x - \frac{1}{2}t, y\right).$$

The high-order vortex can be seen as a high-order extension to a vortex defined by Bispen et al. [10]. Note that the vortex is defined with the help of the exponential integral function

$$\mathrm{Ei}(x) := \int_{-\infty}^{x} \frac{e^t}{t} \mathrm{d}t.$$

This exponential integral function is, amongst others, implemented in the boost package [49, 50], which is used in this implementation. The finite element code is based on the software Netgen [51]; linear systems are solved through PETSc [52, 53, 54].

Remark 6. *In the following, if not stated otherwise, we use an "exact" reference solution. This means that we project the exact reference solution onto¡ the given DG space and use this projection to compute the splitting.*
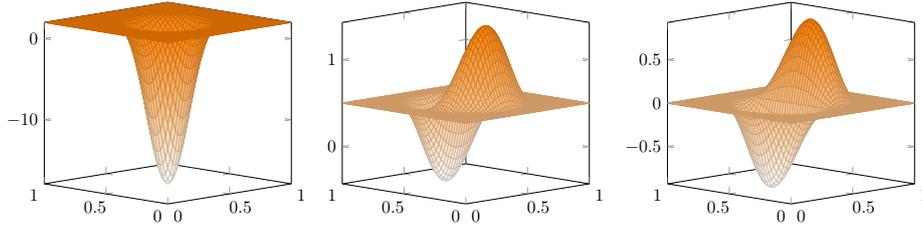
Figure 1: Initial values of the high order vortex for $\varepsilon = 1$. Left: initial density ρ^0. Middle and right: components of $\boldsymbol{u}^0$.

6.1. Choice of the CFL number

The stable use of IMEX schemes should be possible under a *convective* CFL number, see (3). In this section, we try to numerically determine a proper ratio of $\frac{\Delta t}{\Delta x}$ that produces a stable method. The investigation here will be purely numerically. For preliminary analytical work in this direction, we refer to the work of Zakerzadeh [24] and Zakerzadeh and Noelle [21].

We choose a fixed grid (ne = 64) and perform 500 steps with the numerical method for different polynomial degrees, advective CFL numbers (more precisely for $\Delta t/\Delta x = \mathrm{CFL}/\max \|\boldsymbol{u}^0\|_\infty$) and values of ε. We compute the L^2-error of the numerical approximation in every step and if this error raises over a threshold (1000) we can say that the combination of CFL number and ε is instable. Of course such a test can only be a rough indication of stability, and not replace a proof.
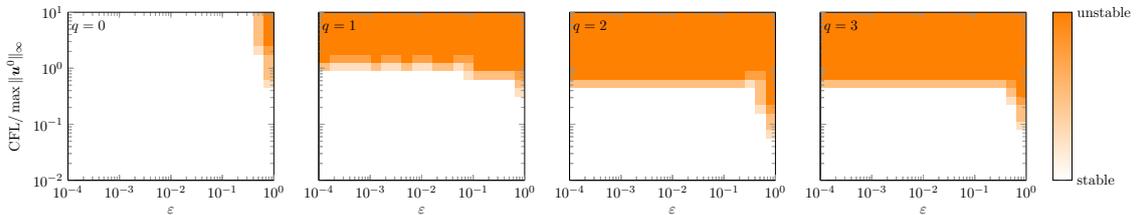

Figure 2: Numerical stability analysis of the RS-IMEX DG method for $q = 0$ with the IMEX-Euler method (left), $q = 1$ with the IMEX-DPA-242 method (middle-left), $q = 2$ with the IMEX-ARS-443 method (middle-right) and $q = 3$ with the IMEX-ARK-4A2 method (right). In all cases a fixed grid was chosen (ne = 64), and 500 time-steps were performed. If the L^2-error raises over a specific threshold we call the method unstable (orange) if it keeps below a specific threshold we call the method stable (white).

In Fig. 2 we summarized the results of this analysis. Note that in this example

$$\|\boldsymbol{u}^0\|_\infty \approx 1.43.$$

For the low order ($q = 0$) case, one can see that stability is very pronounced. This is a result of the relatively large numerical diffusion in the numerical flux. For the higher order case the influence of the numerical flux function is much less pronounced. There is a threshold in the CFL number below which the method is stable. Fortunately, this threshold is independent of ε; it gets smaller with q increasing. (This is of course for standard DG known quite well [55].) Furthermore, we can observe that the method seems to be less stable for larger ε, which is not surprising since with $\varepsilon \to 0$, the influence of the implicit part gets more pronounced.
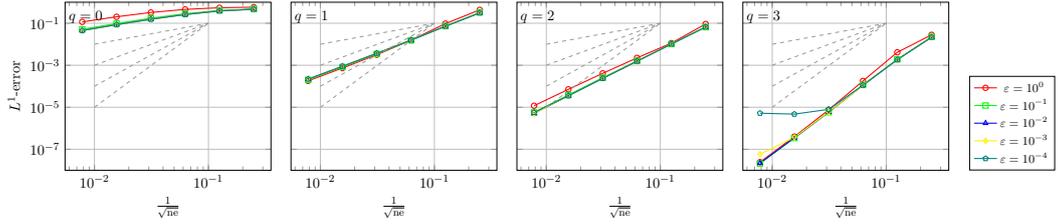
Figure 3: Convergence of the RS-IMEX DG method for the high-order vortex with an exact reference solution: Different values of ε and for $q = 0$ with the IMEX-Euler method (left), $q = 1$ with the IMEX-DPA-242 method (middle-left), $q = 2$ with the IMEX-ARS-443 method (middle-right) and $q = 3$ with the IMEX-ARK-4A2 method (right). As an error measure, we chose the L^1 error between the numerical solution and the exact solution. The dashed lines give the different optimal convergence order, from first order up to fourth order.

Overall these results give us an indication on how to choose the advective CFL number in the following numerical results. To be completely away from the unstable points we choose

$$\frac{\Delta t}{\Delta x} = \frac{\text{CFL}}{\max \|\boldsymbol{u}^0\|_\infty} = 0.05.$$

Since $\|\boldsymbol{u}^0\|_\infty \approx 1.43$ this corresponds to an advective CFL number of

$$\text{CFL} \approx 0.05 \cdot 1.43 \approx 0.0715.$$

6.2. Convergence study

In this section we compute the convergence order for the previously defined example. Grids have been generated with quadratic cells and results, presented in the following, are compared using the L^1-norm of the error at the time instance $T = 0.125$.

The computations are summarized in Fig. 3. In the following we discuss the results for the various polynomial degrees q.

q = 0 and q = 1   For both low order cases we obtain the desired convergence order. Just for the first order case ($q = 0$) the convergence order is not reached before some refinements are done. We believe that this is due to the large numerical diffusion we add in the conservation of mass equation. Choosing a higher polynomial degree reduces the influence of the numerical flux and therefore the second order method gives the desired results.

q = 2   The convergence order of this formally third order method is only ≈ 2.7. Since all other methods deliver the desired results, we believe that this effect is not due to the low Mach number, but insufficient grid resolution. To justify this assumption we also computed the convergence of a third order explicit DG method for $\varepsilon = 1$, see Fig. 4. Also this method starts with a convergence order of about 2.7; after several refinements the convergence order gets close to 3. Such a highly refined grid is unfortunately at this moment not feasible for our solver and an implicit method.

q = 3   This is the most interesting case. For large values of ε the correct convergence order is given but for $\varepsilon = 10^{-3}$ the order reduces in the last given refinement and for $\varepsilon = 10^{-4}$ the error gets even constant.
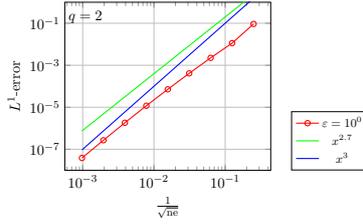
Figure 4: Convergence of a third order explicit DG method for the high order vortex.
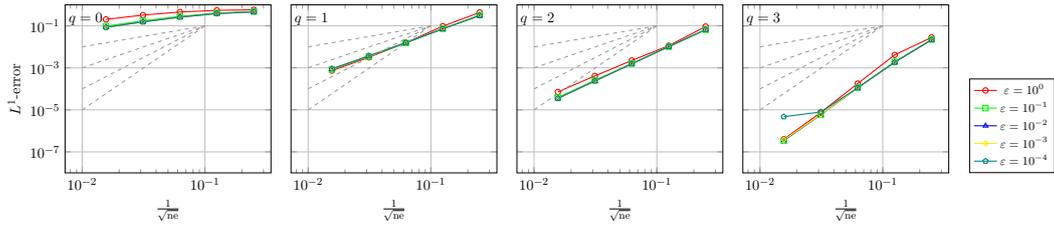

Figure 5: Convergence of the RS-IMEX DG method for the high-order vortex with an implicit method for computing the reference solution: Different values of ε and for $q = 0$ with the IMEX-Euler method (left), $q = 1$ with the IMEX-DPA-242 method (middle-left), $q = 2$ with the IMEX-ARS-443 method (middle-right) and $q = 3$ with the IMEX-ARK-4A2 method (right). As an error measure, we chose the L^1 error between the numerical solution and the exact solution. The dashed lines give the different optimal convergence order, from first order up to fourth order.

We do not believe that this effect is due to order reduction, as presented in [44] for IMEX RK methods, because this would happen for a time step Δt depending on ε and therefore the effect of order reduction would occur for $\varepsilon = 10^{-3}$ first and then for $\varepsilon = 10^{-4}$, not the other way around. Furthermore, it is also not a stability issue, as with decreasing ε, the method is stable for more values of Δt. We believe that with this example, we are hitting the machine accuracy: $\varepsilon = 10^{-4}$, so the term in front of the pressure gradient is $\frac{1}{\varepsilon^2} = 10^8$. Furthermore, the error level is about 10^{-6}. Multiplying already yields machine accuracy of around $2 \cdot 10^{-14}$. Similar issues and how to solve them are discussed in the works [56, 57].

6.3. Convergence study: discrete reference solution

Up to this point we only used a given reference solution. Of course in more complex examples an exact solution is not given. Therefore we also computed the same numerical example with an approximate reference solution. Due to Theorem 2 we use a fully implicit discontinuous Galerkin method to solve the incompressible equation.

In Fig. 5 the results for the same setting as in the previous section are summarized, showing that they are pretty similar to the ones from the previous section.

7. Conclusion and outlook

In the current paper we have coupled the RS-IMEX splitting with a high-order temporal and spatial discretization. The resulting method has been shown to be asymptotically consistent. Furthermore, numerical results give rise to the conjecture that the method is asymptotically stable and asymptotically accurate.

The next important steps in the development of the RS-IMEX splitting are inherent. First, a more detailed stability analysis is desirable to prove analytically that the method is stable under a convective CFL restriction. Second, the identification of more complex test-cases or equations is useful to test the method in a large range of settings. Furthermore, reducing the computational effort is extremely important, especially compared to other numerical methods given in literature. Therefore our aim is to figure out in which way the reference solution can be computed most efficiently, especially if a less accurate reference solution can also be employed. Another step is the use of more efficient numerical methods for the implicit part, e.g., the hybridized discontinuous Galerkin method for spatial discretization (see e.g. [58, 59, 60, 61, 62]).

Up to now, we have only considered IMEX Runge-Kutta methods. Unfortunately, those methods are difficult to construct when going to orders larger than four. A very interesting class of IMEX schemes are the IMEX general linear methods (GLM), see, e.g., [63, 64] and the references therein. They can be more easily constructed to higher order while preserving properties such as A-stability. An investigation of an IMEX GLM is therefore of high interest.

8. Acknowledgment

A. IMEX Runge-Kutta methods

For the sake of completeness, we list the employed Runge-Kutta methods in this appendix section. They are listed in the standard Butcher-tableau form $\left(\begin{array}{c|c||c|c} \widetilde{c} & \widetilde{A} & \widehat{c} & \widehat{A} \\ \hline & \widetilde{b} & & \widehat{b} \end{array} \right)$.

	IMEX-Euler	IMEX-DPA-242	IMEX-ARS-443	IMEX-ARK-4A2
Order	1	2	3	4
GSA	Yes	Yes	Yes	Yes
Type	CK	A	CK	CK
Butcher Tbl.	Tbl. 2	Tbl. 3	Tbl. 4	Tbl. 5

Table 1: Classification of the used IMEX Runge-Kutta methods concerning their order, structure and type.

$$\begin{array}{c|cc||c|cc} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 \\ \hline & 0 & 1 & & 1 & 0 \end{array}$$

Table 2: A first order IMEX RK method called IMEX-Euler [28]. Left: implicit, right: explicit.

References

[1] J. Schütz, K. Kaiser, A new stable splitting for singularly perturbed ODEs, Applied Numerical Mathematics 107 (2016) 18–33.

1/2	1/2	0	0	0		0	0	0	0	0
2/3	1/6	1/2	0	0		1/3	1/3	0	0	0
1/2	-1/2	1/2	1/2	0		1	1	0	0	0
1	3/2	-3/2	1/2	1/2		1	1/2	0	1/2	0
	3/2	-3/2	1/2	1/2			1/2	0	1/2	0

Table 3: A second order IMEX RK method called IMEX-DPA-242 [33]. Left: implicit, right: explicit.

0	0	0	0	0	0		0	0	0	0	0
1/2	0	1/2	0	0	0		1/2	1/2	0	0	0
2/3	0	1/6	1/2	0	0		2/3	11/18	1/18	0	0
1/2	0	-1/2	1/2	1/2	0		1/2	5/6	-5/6	1/2	0
1	0	3/2	-3/2	1/2	1/2		1	1/4	7/4	3/4	-7/4	0
	0	3/2	-3/2	1/2	1/2			1/4	7/4	3/4	-7/4	0

Table 4: A third order IMEX RK method called IMEX-ARS-443[28]. Left: implicit, right: explicit.

[2] K. Kaiser, J. Schütz, R. Schöbel, S. Noelle, A new stable splitting for the isentropic Euler equations, Journal of Scientific Computing (in press).
URL https://www.igpm.rwth-aachen.de/forschung/preprints/442

[3] J. D. Anderson, Fundamentals of Aerodynamics, 3rd Edition, McGraw-Hill New York, 2001.

[4] P. Wesseling, Principles of Computational Fluid Dynamics, Vol. 29 of Springer Series in Computational Mechanics, Springer Verlag, 2001.

[5] S. Klainerman, A. 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) 481–524.

[6] S. Schochet, Fast singular limits of hyperbolic PDEs, Journal of Differential Equations 114 (2) (1994) 476–512.

[7] W.-A. Yong, A note on the zero Mach number limit of compressible Euler equations, Proceedings of the American Mathematical Society 133 (10) (2005) 3079–3085.

[8] D. Kröner, Numerical Schemes for Conservation Laws, Wiley Teubner, 1997.

[9] G. Bispen, M. Lukáčová-Medvid'ová, L. Yelash, Asymptotic preserving IMEX finite volume schemes for low mach number euler equations with gravitation, Journal of Computational Physics 335 (2017) 222–248.

0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
1/3	-1/6	1/2	0	0	0	0	0		1/3	1/3	0	0	0	0	0	0
1/3	1/6	-1/3	1/2	0	0	0	0		1/3	1/6	1/6	0	0	0	0	0
1/2	3/8	-3/8	0	1/2	0	0	0		1/2	1/8	0	3/8	0	0	0	0
1/2	1/8	0	3/8	-1/2	1/2	0	0		1/2	1/8	0	3/8	0	0	0	0
1	-1/2	0	3	-3	1	1/2	0		1	1/2	0	-3/2	0	2	0	0
1	1/6	0	0	0	2/3	-1/2	2/3		1	1/6	0	0	0	2/3	1/6	0
	1/6	0	0	0	2/3	-1/2	2/3			1/6	0	0	0	2/3	1/6	0

Table 5: A fourth order IMEX RK method called IMEX-ARK-4A2[34]. Left: implicit, right: explicit.

[10] G. Bispen, K. Arun, M. Lukáčová-Medvid'ová, S. Noelle, IMEX large time step finite volume methods for low Froude number shallow water flows, Communications in Computational Physics 16 (2014) 307–347.

[11] F. Cordier, P. Degond, A. Kumbaro, An asymptotic-preserving all-speed scheme for the Euler and Navier-Stokes equations, Journal of Computational Physics 231 (2012) 5685–5704.

[12] P. Degond, M. Tang, All speed scheme for the low Mach number limit of the isentropic Euler equation, Communications in Computational Physics 10 (2011) 1–31.

[13] F. Giraldo, M. Restelli, M. Läuter, Semi-implicit formulations of the Navier-Stokes equations: Application to nonhydrostatic atmospheric modeling, SIAM Journal on Scientific Computing 32 (6) (2010) 3394–3425.

[14] J. Haack, S. Jin, J.-G. Liu, An all-speed asymptotic-preserving method for the isentropic Euler and Navier-Stokes equations, Communications in Computational Physics 12 (2012) 955–980.

[15] R. Klein, Semi-implicit extension of a Godunov-type scheme based on low Mach number asymptotics I: One-dimensional flow, Journal of Computational Physics 121 (1995) 213–237.

[16] A. Müller, J. Behrens, F. Giraldo, V. Wirth, Comparison between adaptive and uniform discontinuous Galerkin simulations in dry 2d bubble experiments, Journal of Computational Physics 235 (2013) 371–393.

[17] S. Noelle, G. Bispen, K. Arun, M. Lukáčová-Medvid'ová, 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) B989–B1024.

[18] M. Restelli, Semi-lagrangian and semi-implicit discontinuous Galerkin methods for atmospheric modeling applications, PhD thesis Politecnico di Milano.

[19] L. Yelash, A. Müller, M. Lukáčová-Medvid'ová, F. X. Giraldo, V. Wirth, Adaptive discontinuous evolution Galerkin method for dry atmospheric flow, Journal of Computational Physics 268 (2014) 106–133.

[20] J. Schütz, S. Noelle, Flux splitting for stiff equations: A notion on stability, Journal of Scientific Computing 64 (2) (2015) 522–540.

[21] H. Zakerzadeh, S. Noelle, A note on the stability of implicit-explicit flux splittings for stiff hyperbolic systems, IGPM Preprint Nr. 449.

[22] F. X. Giraldo, M. Restelli, High-order semi-implicit time-integrators for a triangular discontinuous Galerkin oceanic shallow water model, International Journal for Numerical Methods in Fluids 63 (9) (2010) 1077–1102.

[23] F. Filbet, S. Jin, A class of asymptotic-preserving schemes for kinetic equations and related problems with stiff sources, Journal of Computational Physics 229 (20) (2010) 7625–7648.

[24] H. Zakerzadeh, Asymptotic analysis of the RS-IMEX scheme for the shallow water equations in one space dimension, IGPM Preprint Nr. 455.

[25] U. M. Ascher, S. Ruuth, B. Wetton, Implicit-Explicit methods for time-dependent partial differential equations, SIAM Journal on Numerical Analysis 32 (1995) 797–823.

18

[26] W. Hundsdorfer, J. Jaffré, Implicit–explicit time stepping with spatial discontinuous finite elements, Applied Numerical Mathematics 45 (2) (2003) 231–254.

[27] W. Hundsdorfer, S.-J. Ruuth, IMEX extensions of linear multistep methods with general monotonicity and boundedness properties, Journal of Computational Physics 225 (2) (2007) 2016–2042.

[28] U. M. Ascher, S. Ruuth, R. Spiteri, Implicit-explicit Runge-Kutta methods for time-dependent partial differential equations, Applied Numerical Mathematics 25 (1997) 151–167.

[29] S. Boscarino, L. Pareschi, G. Russo, Implicit-explicit Runge–Kutta schemes for hyperbolic systems and kinetic equations in the diffusion limit, SIAM Journal on Scientific Computing 35 (1) (2013) A22–A51.

[30] C. A. Kennedy, M. H. Carpenter, Additive Runge-Kutta schemes for convection-diffusion-reaction equations, Applied Numerical Mathematics 44 (2003) 139–181.

[31] L. Pareschi, G. Russo, Implicit-explicit Runge-Kutta schemes for stiff systems of differential equations, Recent Trends in Numerical Analysis 3 (2000) 269–289.

[32] G. Russo, S. Boscarino, IMEX Runge-Kutta schemes for hyperbolic systems with diffusive relaxation, European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS 2012).

[33] G. Dimarco, L. Pareschi, Asymptotic preserving implicit-explicit Runge–Kutta methods for nonlinear kinetic equations, SIAM Journal on Numerical Analysis 51 (2) (2013) 1064–1087.

[34] H. Liu, J. Zou, Some new additive Runge–Kutta methods and their applications, Journal of Computational and Applied Mathematics 190 (1-2) (2006) 74–98.

[35] B. Cockburn, S. Hou, C.-W. Shu, The Runge–Kutta local projection discontinuous Galerkin finite element method for conservation laws IV: The multidimensional case, Mathematics of Computation 54 (1990) 545–581.

[36] B. Cockburn, S. Y. Lin, C.-W. Shu, TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws III: One dimensional systems, Journal of Computational Physics 84 (1989) 90–113.

[37] B. Cockburn, C.-W. Shu, TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws II: General framework, Mathematics of Computation 52 (1988) 411–435.

[38] B. Cockburn, C.-W. Shu, The Runge-Kutta local projection p^1-discontinuous Galerkin finite element method for scalar conservation laws, RAIRO Mathematical modelling and numerical analysis 25 (1991) 337–361.

[39] B. Cockburn, C.-W. Shu, The Runge-Kutta discontinuous Galerkin Method for conservation laws V: Multidimensional Systems, Mathematics of Computation 141 (1998) 199–224.

[40] A. Kanevsky, M. H. Carpenter, D. Gottlieb, J. S. Hesthaven, Application of implicit-explicit high order Runge-Kutta methods to discontinuous-Galerkin schemes, Journal of Computational Physics 225 (2) (2007) 1753–1781.

[41] P.-O. Persson, High-order LES simulations using implicit-explicit Runge-Kutta schemes, AIAA Paper 11-684 (2011).

[42] H. Wang, C.-W. Shu, Q. Zhang, Stability and error estimates of local discontinuous Galerkin methods with implicit-explicit time-marching for advection-diffusion problems, SIAM Journal on Numerical Analysis 53 (1) (2015) 206–227.

[43] S. Jin, Asymptotic preserving (AP) schemes for multiscale kinetic and hyperbolic equations: A review, Rivista di Matematica della Universita Parma 3 (2012) 177–216.

[44] S. Boscarino, Error analysis of IMEX Runge-Kutta methods derived from differential-algebraic systems, SIAM Journal on Numerical Analysis 45 (2007) 1600–1621.

[45] H. Guillard, C. Viozat, On the behavior of upwind schemes in the low Mach number limit, Computers and Fluids 28 (1) (1999) 63–86.

[46] E. Turkel, Preconditioned methods for solving the incompressible and low speed compressible equations, Journal of Computational Physics 72 (2) (1987) 277 – 298.

[47] D. D. Pietro, A. Ern, Mathematical aspects of discontinuous Galerkin Methods, Vol. 69, Springer Science & Business Media, 2011.

[48] G. Bispen, IMEX finite volume methods for the shallow water equations, Ph.D. thesis, Johannes Gutenberg-Universität (2015).

[49] B. Schling, The Boost C++ Libraries, XML Press, 2011.

[50] BOOST C++ Libraries, http://www.boost.org.

[51] J. Schöberl, Netgen - an advancing front 2d/3d-mesh generator based on abstract rules, Computing and Visualization in Science 1 (1997) 41–52.

[52] S. Balay, W. D. Gropp, L. C. McInnes, B. F. Smith, Efficient management of parallelism in object oriented numerical software libraries, in: E. Arge, A. M. Bruaset, H. P. Langtangen (Eds.), Modern Software Tools in Scientific Computing, Birkhäuser Press Boston, 1997, pp. 163–202.

[53] S. Balay, J. Brown, K. Buschelman, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, B. F. Smith, H. Zhang, PETSc Web page, http://www.mcs.anl.gov/petsc (2011).

[54] S. Balay, J. Brown, K. Buschelman, V. Eijkhout, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, B. F. Smith, H. Zhang, PETSc users manual, Tech. Rep. ANL-95/11 - Revision 3.1, Argonne National Laboratory (2010).

[55] B. Cockburn, C. W. Shu, Runge-Kutta discontinuous Galerkin methods for convection-dominated problems, Journal of Scientific Computing 16 (2001) 173–261.

[56] J. Sesterhenn, B. Müller, H. Thomann, On the cancellation problem in calculating compressible low mach number flows, Journal of Computational Physics 151 (2) (1999) 597 – 615. doi:http://dx.doi.org/10.1006/jcph.1999.6211.
URL http://www.sciencedirect.com/science/article/pii/S0021999199962113

[57] S.-H. Lee, Cancellation problem of preconditioning method at low mach numbers, Journal of Computational Physics 225 (2) (2007) 1199 – 1210. doi:http://dx.doi.org/10.1016/j.jcp.2007.04.001.
URL http://www.sciencedirect.com/science/article/pii/S0021999107001544

[58] N. Nguyen, J. Peraire, B. Cockburn, A hybridizable discontinuous Galerkin method for the incompressible navier-stokes equations, AIAA Paper 2010-362.

[59] J. Peraire, N. C. Nguyen, B. Cockburn, A hybridizable discontinuous Galerkin method for the compressible Euler and Navier-Stokes equations, AIAA Paper 10-363 (2010).

[60] A. Jaust, J. Schütz, M. Woopen, An HDG method for unsteady compressible flows, in: R. M. Kirby, M. Berzins, J. S. Hesthaven (Eds.), Spectral and High Order Methods for Partial Differential Equations ICOSAHOM 2014, Vol. 106 of Lecture Notes in Computational Science and Engineering, Springer International Publishing, 2015, pp. 267–274.

[61] A. Jaust, J. Schütz, M. Woopen, A hybridized discontinuous Galerkin method for unsteady flows with shock-capturing, AIAA Paper 2014-2781.

[62] J. Schütz, M. Woopen, G. May, A hybridized DG/mixed scheme for nonlinear advection-diffusion systems, including the compressible Navier-Stokes equations, AIAA Paper 2012-0729.

[63] P. E. Vos, C. Eskilsson, A. Bolis, S. Chun, R. M. Kirby, S. J. Sherwin, A generic framework for time-stepping partial differential equations (PDEs): general linear methods, object-oriented implementation and application to fluid problems, International Journal of Computational Fluid Dynamics 25 (3) (2011) 107–125.

[64] H. Zhang, A. Sandu, S. Blaise, Partitioned and implicit–explicit general linear methods for ordinary differential equations, Journal of Scientific Computing 61 (1) (2014) 119–144. doi:10.1007/s10915-014-9819-z.
URL http://dx.doi.org/10.1007/s10915-014-9819-z