A note on the stability of implicit-explicit flux-splittings for stiff systems of hyperbolic conservation laws

Hamed Zakerzadeh† and Sebastian Noelle†

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

† Institut für Geometrie und Praktische Mathematik, RWTH Aachen University, Templergraben 55, 52056 Aachen, Germany. email:(h.zakerzadeh, noelle)@igpm.rwth-aachen.de

* The first author was supported by the scholarship of RWTH Aachen university through Graduiertenförderung nach Richtlinien zur Förderung des wissenschaftlichen Nachwuchses (RFwN).
A NOTE ON THE STABILITY OF IMPLICIT-EXPLICIT FLUX-SPLITTINGS FOR STIFF SYSTEMS OF HYPERBOLIC CONSERVATION LAWS∗

HAMED ZAKERZADEH† AND SEBASTIAN NOELLE†

Abstract. We analyze the stability of implicit-explicit flux-splitting schemes for stiff systems of conservation laws. In particular, we study the modified equation of the corresponding linearized systems. We first prove that symmetric splittings are stable, uniformly in the singular parameter ε. Then we study non-symmetric splittings. We prove that for the barotropic Euler equations, the Degond–Tang splitting [Degond & Tang, Comm. Comp. Phys. 10 (2011), pp. 1–31] and the Haack–Jin–Liu splitting [Haack, Jin & Liu, Comm. Comp. Phys. 12 (2012), pp. 955 - 980], and for the shallow water equations the recent RS-IMEX splitting are strictly stable in the sense of Majda–Pego. For the full Euler equations, we find a small instability region for a flux splitting introduced by Klein [Klein, J. Comp. Phys. 121 (1995), pp. 213–237], if this splitting is combined with an IMEX scheme as in [Noelle, Bispen, Arun, Lukáčová, Munz, SIAM J. Sci. Comp. 36 (2014), pp. B989–B1024].

Key words. stiff hyperbolic systems, flux-splitting, IMEX scheme, asymptotic preserving (AP) property, modified equation, stability analysis

AMS subject classifications. 35L65, 65M08, 76M45

1. Introduction The efficient and stable approximation of solutions to stiff differential equations is a longstanding challenge in numerical analysis. For systems of ordinary differential equations (ODEs), stiffness may be defined as the simultaneous occurrence of eigenvalues of different orders of magnitude. In the context of conservation laws, the key example is the low Mach number flow in gas dynamics, where the speed of acoustic waves is much bigger than the advection speed. In the limit, the equations change type from hyperbolic to hyperbolic-elliptic, or from compressible to incompressible flow. The numerical challenge is to establish a scheme which is efficient (i.e., it has a time step independent of the Mach number ε), uniformly consistent and stable as ε → 0. Moreover, the limit scheme should be a consistent and stable approximation of the incompressible Euler equations. In [20,22], Jin introduced the term asymptotic preserving (AP) for such schemes.

In the context of ODEs, implicit-explicit (or IMEX) schemes are a method of choice. They split the system into a fast and a slow part, and treat them implicitly respectively explicitly (see the classic textbook [17]). For hyperbolic conservation laws, we refer to [7] for a concise review of classical approaches, such as preconditioning proposed by Chorin [3] and Turkel [43]. More recently, some pioneering papers such as Klein [25], Degond and Tang [9], and Haack, Jin and Liu [16] split the flux function f(u) into fast and slow fluxes, f(u) and ̂f(u), in such a way that the Jacobians ̂A(u) := ̂f′(u) and A(u) := ̂f′(u) are hyperbolic, and then they apply an IMEX time discretization to

∗The first author was supported by the scholarship of RWTH Aachen university through Graduiertenförderung nach Richtlinien zur Förderung des wissenschaftlichen Nachwuchses (RFwN).
†Institut für Geometrie und Praktische Mathematik, RWTH Aachen University, Templergraben 55, 52056 Aachen, Germany. email: {h.zakerzadeh, noelle}@igpm.rwth-aachen.de
On the stability of IMEX splittings

the split system (see also \[2,3,7,33\] and the references therein). Note that most of these works prove the scheme to be uniformly (asymptotically) consistent. However, there are only few results regarding the uniform (asymptotic) stability - among them \[8\] for the Euler–Poisson system and \[10,45\] for the barotropic Euler equations. Also for kinetic and transport equations there exist few papers providing rigorous analysis for the asymptotic stability, e.g., based on von Neumann analysis or energy methods (see \[\[13,19,23,27,29\]\]).

Recently, the second author and collaborators analyzed and applied the following splittings for singularly-perturbed conservation laws. In \[3\] they studied a linearly implicit method (see \[36\] and \[17\] Chap IV.7) proposed by Giraldo and Restelli \[12\] for the shallow water equations. They proved asymptotic consistency, and numerical experiments demonstrated \(\varepsilon\)-uniform stability. In \[33\], they used a variant of Klein’s auxiliary splitting \[25\] for the full Euler equations. Unexpectedly, the latter required an \(\varepsilon\)-dependent time step for stability. Motivated by this, Schütz and Noelle \[38\] began a stability study of the modified equation of the linearized system in Fourier variables. Computations showed the instability of some Fourier modes for Klein’s auxiliary splitting, when used within the context of flux-splitting IMEX schemes (note that this is not the context of Klein’s original algorithm, cf. Remark 5.3). In \[38\] Remark 9] the authors conjectured that the culprit is a resonance between the implicit and explicit parts of the algorithm, as expressed by the commutator \(\hat{A}\hat{A} - \hat{A}\hat{A}\), which in general is \(O(\varepsilon^{-1})\). They proved the stability of characteristic splittings in terms of the modified equations, for which the commutator vanishes. To go beyond characteristic splittings, the authors and collaborators developed a generalized version of linearly implicit methods, called Reference Solution IMEX (RS-IMEX) which de-singularizes the commutator; see \[24,34\] for the derivation, \[24,44\] for the AP analysis for barotropic flows, and \[37\] and \[24\] for the application of Van der Pol and two-dimensional barotropic Euler systems, respectively.

In the present paper, we analyze the stability of the modified equation for a general class of splittings. We first point out a stability result for symmetric splittings and relate this stability result to the linear stability in the sense of Majda–Pego \[30\]. Then, given a general background state, we study Fourier symbols for linearized modified equations for non-symmetric flux-splittings and apply this to several well-known IMEX schemes. Note that our analysis applies to a general background state and any frequency variable for which the modified equation is valid, while the previous work \[38\] evaluated the Fourier symbols numerically using fixed background states and frequencies.

The paper is organized as follows: in Section 2 we review the algorithm, modified equation and Fourier analysis for linear systems, all as in \[38\]. In Sections 3 and 4, we assemble a number of stability results for symmetric and general non-symmetric splittings, respectively. Using the results of Section 4, we prove in Section 5 that the modified equations resulting from the splittings in \[9,16\], as well as the RS-IMEX splitting \[24,34,44\], are stable in the sense of Majda–Pego. We also study Klein’s auxiliary splitting \[25\], and discover a small instability region for the example of two colliding pulses \[25,33\], for a moderate CFL number. This seems to give a hint at the numerical difficulties observed in \[33\].

Acknowledgments. The authors would like to thank Rupert Klein for stimulating discussions on low Mach number flows. H.Z. thanks Jochen Schütz for a helpful discussion on \[38\]. Also he gratefully acknowledges an illuminating discussion with Charlotte
Perrin about the construction of symmetric splittings. S.N. would like to thank Arun K.R., Georgij Bispen, Mária Lukáčová-Medviďová, Claus-Dieter Munz, and Jochen Schütz for the collaborations leading to [3,33,38], which provided first insights into the stability of flux splittings.

2. Flux-splittings, IMEX schemes and the modified equation

Let us first review the linear framework introduced in [38]. We consider the linear system of hyperbolic conservation laws

\[ u_t + Au_x = 0, \quad u(0, \cdot) = u_0(x), \tag{2.1} \]

where \( u : \Omega \times \mathbb{R}_+ \to \mathbb{R}^m \) are conservative variables and \( A \in \mathbb{R}^{m \times m} \) is a constant matrix (depending only on the parameter \( \varepsilon \)) which is real diagonalizable with eigenvalues of \( \lambda_1 \geq \ldots \geq \lambda_m \). We assume well-prepared initial data so that the time derivatives of the solution, \( \partial_t^k u \), are bounded uniformly in \( \varepsilon \) for \( k \in \mathbb{N} \) (see [14,26,31]).

**Definition 2.1 (Admissible splitting [38]).** The splitting \( A = \tilde{A} + \hat{A} \) is admissible provided that

(i) both \( \tilde{A} \) and \( \hat{A} \) induce a hyperbolic system, i.e., they have real eigenvalues and a complete set of eigenvectors;

(ii) the eigenvalues of \( \hat{A} \) are bounded independently of \( \varepsilon \), e.g., \( \mathcal{O}(1) \), and at least one of the eigenvalues of \( \tilde{A} \) is \( \mathcal{O}(\frac{1}{\varepsilon}) \).

As in [38], we choose a Rusanov-type scheme for both implicit and explicit parts on the domain \( \Omega \), and with the time step \( \Delta t \) and the spatial step \( \Delta x := \frac{\|\Omega\|}{N} \), where \( N \) is the number of computational cells. Such a scheme can be written as

\[ u_j^{n+1} = u_j^n - \frac{\Delta t}{\Delta x} \left( \tilde{F}_{j+1/2}^{n+1} - \tilde{F}_{j-1/2}^{n+1} + \hat{F}_{j+1/2}^n - \hat{F}_{j-1/2}^n \right), \]

where the numerical fluxes are defined as

\[ \tilde{F}_{j+1/2}^{n+1} := \frac{1}{2} \tilde{A} \left( u_{j+1}^{n+1} + u_j^{n+1} \right) - \frac{\tilde{\alpha}}{2} \left( u_{j+1}^{n+1} - u_j^{n+1} \right), \]
\[ \hat{F}_{j+1/2}^n := \frac{1}{2} \hat{A} \left( u_{j+1}^n + u_j^n \right) - \frac{\hat{\alpha}}{2} \left( u_{j+1}^n - u_j^n \right), \]

with numerical diffusion coefficients \( \tilde{\alpha} \) and \( \hat{\alpha} \) for stiff and non-stiff parts, respectively. Then, the modified equation (see [38, eq. (10)]) reads

\[ u_t + Au_x = \frac{\Delta t}{2} \left( \frac{\alpha \Delta x}{\Delta t} I_m - \tilde{A}^2 + \hat{A}^2 + [\tilde{A}, \hat{A}] \right) u_{xx}, \tag{2.2} \]

where \( \alpha := \tilde{\alpha} + \hat{\alpha} \) and \( [\tilde{A}, \hat{A}] := \tilde{A}\hat{A} - \hat{A}\tilde{A} \) is the commutator of the stiff and non-stiff Jacobians.

Applying the Fourier transform to (2.2) leads to

\[ \hat{u}_t + \left( -i\xi A - \frac{\xi^2 \Delta t}{2} \left( \frac{\alpha \Delta x}{\Delta t} I_m - \tilde{A}^2 + \hat{A}^2 + [\tilde{A}, \hat{A}] \right) \right) \hat{u} = 0, \tag{2.3} \]
which gives the following convenient stability result:

**Lemnma and Definition 2.2 (Corollary 1, [38]).** The modified equation (2.2) is \(L_2\)-stable if the frequency matrix

\[
\mathcal{P}(\xi) := -i A \xi - \xi^2 D, \quad D := \frac{\Delta t}{2} \left( \frac{\alpha \Delta x}{\Delta t} I_m - \tilde{A}^2 + \tilde{A}^2 + [\tilde{A}, \hat{A}] \right)
\]

only has eigenvalues with negative real parts. In this case we call \(\mathcal{P}(\xi)\) a stable matrix, and say that the IMEX splitting satisfies condition (A). Note that \(\xi\) denotes the frequency variable.

**Remark 2.3.** (i) From now on, whenever we talk about stability, it means stability in the sense of condition (A), unless explicitly stated otherwise.

(ii) The modified equation (2.4) is derived formally by truncating Taylor expansions in space and time. We conjecture that a rigorous justification will have to rely on a low-frequency assumption such as

\[\|\xi^{k+1} A^{k+1}\| = O(1) \quad \text{for} \ k = 2, 3, \ldots,\]

together with a suitable CFL condition.

(iii) Recalling a famous result of Gel’fand [11], if one considers a convection-diffusion system of equations like (2.4), the well-posedness requires the viscosity matrix (in this case \(D\)) to be parabolic, i.e., its eigenvalues should have positive real parts. So for the general case, parabolicity is a necessary condition for stability. Nonetheless in the context of this paper, since the frequencies are small due to a low-frequency assumption, the argument in [11] does not hold and parabolicity is not a necessary condition.

Unfortunately, without any additional structural assumption, obtaining a general stability condition for \(\mathcal{P}\) is very delicate. For example, in [38] the authors introduce a characteristic splitting, for which the Jacobians are simultaneously diagonalizable and hence the commutator \([\hat{A}, \tilde{A}]\) vanishes. This immediately provides stability of the modified equation; see also Remark 1.4(i) below for \(\ell_2\)-stability. In Section 3 we study the eigenvalues of \(\mathcal{P}\) assuming symmetry of the system and its splitting. This seems to be a promising framework for stability. In Section 4 we study the general, non-symmetric case and review the stability conditions of Majda and Pego [30]. In Section 5 we verify these conditions to obtain linearized stability or instability for a number of recent splittings.

3. Stability of \(\mathcal{P}\) for symmetric splittings In this section, we assume that \(A, \tilde{A}\) and \(\hat{A}\) are symmetric, and point out the stability of such a splitting in Corollary 3.4. Then, introducing the notion of strict stability in the sense of Majda–Pego [30] we generalize condition (A) to linearized systems. For symmetric splittings, this notion gives a more general stability result (see Theorem 3.7). Non-symmetric splittings are treated in Section 4.

Note that for any symmetric matrix \(A\), a symmetric splitting is always possible, e.g., if one chooses \(\hat{A} = \text{diag}(A|_{\xi=1})\). For any symmetric splitting, the commutator is a
skew-Hermitian matrix, therefore
\[
P(\xi) = - \left[ i A \xi + \xi^2 \frac{\Delta t}{2} [\tilde{A}, \hat{A}] + \xi^2 \frac{\Delta t}{2} \left( \frac{\alpha \Delta x}{\Delta t} I_m - \tilde{A}^2 + \tilde{A}^2 \right) \right],
\]
where \(\mathcal{A}\) and \(\mathcal{H}\) are skew-Hermitian and Hermitian respectively. One may conjecture that the eigenvalues of \(\mathcal{H}\) would be positive. The following lemma verifies this conjecture.

**Lemma 3.1.** The Hermitian matrix \(\mathcal{H}\) is positive-definite under a non-restrictive CFL condition, independently of \(\varepsilon\).

**Proof.** We start with the eigenvalue stability inequality (see \cite{40} eq. 1.64), which is a result of Courant–Fischer–Weyl min-max principle. It states that for two Hermitian matrices \(L, M \in \mathbb{H}_m\), the following holds
\[
|\lambda_k(L + M) - \lambda_k(L)| \leq \|M\|_{\text{op}}, \quad k = 1, 2, \ldots, m,
\]
where the operator norm \(\|\cdot\|_{\text{op}}\) is defined as
\[
\|M\|_{\text{op}} := \max \left( |\lambda_1(M)|, |\lambda_m(M)| \right), \quad \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m.
\]
So if one puts \(L = \tilde{A}^2\) and \(M = -\hat{A}^2\) in (3.2), it yields
\[
-\lambda < \lambda_k(\tilde{A}^2) - \|\tilde{A}\|_{\text{op}}^2 \leq \lambda_k(-\tilde{A}^2 + \tilde{A}^2) \leq \lambda_k(\tilde{A}^2) + \|\tilde{A}\|_{\text{op}}^2,
\]
with \(\lambda \geq 0\). Due to the order of magnitude of eigenvalues, \(\lambda\) can be chosen to be positive and \(\mathcal{O}(1)\), namely \(\lambda > \|\tilde{A}\|_{\text{op}}^2\) which implies the time step restriction \(\Delta t < \frac{\alpha \Delta x}{\|\tilde{A}\|_{\text{op}}^2}\). This CFL condition shifts the eigenvalues to the right (by \(\frac{\alpha \Delta x}{\Delta t} I_m\)), so that the eigenvalues of \(\mathcal{H}\) are positive. \(\square\)

**Remark 3.2.** Another way to prove Lemma 3.1 is to use the numerical range or equivalently the Rayleigh quotient, since we are working with symmetric matrices. Using sub-additivity property of the numerical range, it is enough to show that the numerical range of \(\tilde{A}^2\) is positive, and to put the numerical range of \(\frac{\alpha \Delta x}{\Delta t} I_m - \tilde{A}^2\) in the right half-plane under some CFL condition.

Given these properties of \(\mathcal{A}\) and \(\mathcal{H}\), there is a sum of a Hermitian and a skew-Hermitian matrix in (3.1), and one can use the Bendixon’s theorem in \cite{18} (see \cite{1} for the original work which is limited to real matrices), which shows that given a Hermitian matrix with stable eigenvalues in the left half-plane and a skew-Hermitian matrix, the sum will have stable eigenvalues, i.e., the eigenvalues have negative real parts. To recall, we restate the theorem from \cite{18}; see also \cite{4} for a nice review.

**Theorem 3.3** (Theorem II, \cite{18}). Consider the matrix \(M \in \mathbb{K}^{m \times m}\) with \(\mathbb{K} = \mathbb{C}\) or \(\mathbb{R}\), when \(\lambda_k^H := \lambda_k(\mathcal{H}(M)) = p_k \in \mathbb{R}\) for \(k = 1, \ldots, m\), where \(\mathcal{H}\) stands for the Hermitian part. Then the following holds
\[
\min_k p_k \leq \Re \left[ \lambda_k(M) \right] \leq \max_k p_k.
\]
Hence from Lemma 3.1 and Theorem 3.3 one can conclude immediately the following corollary.
On the stability of IMEX splittings

Corollary 3.4. Under a non-restrictive CFL condition, an admissible symmetric splittings is stable, i.e., it satisfies condition (A).

Remark 3.5. One could also use an energy estimate to show that for the hyperbolic-parabolic system with symmetric $A$, the positive-definiteness of the viscosity matrix $D$ is necessary and sufficient for $L_2$-stability.

We would like to conclude this section with the notion of strict stability in the sense of Majda–Pego [30] as defined below.

Definition 3.6 ([30]). For the non-linear system $u_t + f(u)_x = (D(u))_x$, the viscosity matrix $D$ is strictly stable at $u_0$ if and only if there exists a $\delta > 0$ such that the eigenvalues $\lambda_k(\xi)$ of the matrix $P(\xi) := -f'(u_0)i\xi - \xi^2 D(u_0)$ satisfy the following algebraic condition

$$\Re[\lambda_k(\xi)] \leq -\delta|\xi|^2, \quad \text{for all } \xi \in \mathbb{R}. \quad (3.5)$$

This definition is in fact the generalization of Lemma 2.2 for systems linearized around an arbitrary state $u_0$. Using this framework, one can also find the generalization of the stability of symmetric splittings at $u_0$, as in [30,32].

Theorem 3.7. Consider (2.4) and let $M(u_0)$ be a real symmetric positive-definite matrix, symmetrizing $A(u_0)$ from the left, i.e., $(MA)|_{u_0}$ is symmetric. Then if $(MD)|_{u_0}$ is positive-definite, the modified equation in Fourier space (2.4) is strictly stable at $u_0$, i.e., there exists a $\delta > 0$ such that $\Re[\lambda_k(P(\xi))] \leq -\delta|\xi|^2$.

It is clear that for symmetric splittings, the identity matrix can play the role of $M$ and Theorem 3.7 is reduced to the arguments we have presented above.

4. Stability of $P$ for non-symmetric splittings In this section, we study the stability of $P$ without symmetry assumption, so that the commutator contributes to the real parts of the eigenvalues of $P$ and hence to the stability. Note that Majda–Pego’s notion of strict stability refers to a given state $u_0$, around which the system is linearized. To keep the notation simpler, in the following we suppress the dependence of $A$, $\hat{A}$, and $\hat{A}$ on $u_0$. Thus, from now on, the state $u_0$ is arbitrary, but fixed.

Let us denote the spectrum of $P$ as $\sigma(P)$. Then, by the theorem of spectral inclusion [15 Theorem 1.2-1], this spectrum (and in particular its convex hull) is contained in the closure of the numerical range of $P$. In other words, $\text{Conv}(\sigma(P)) \subseteq \text{w}(P)$, where the numerical range $w(P)$ is defined as $w(P) := \{\langle v, P v \rangle, v \in \mathbb{C}^m, \|v\|_2 = 1\}$. In fact, the real part of the numerical range of $P$ is bounded by the spectrum of its Hermitian part, i.e.

$$\Re[w(P)] = -\text{Conv}(\sigma((iA)^H \xi + \xi^2 D^H)).$$

Due to the eigenvalue stability inequality, one can find the upper-bound of the numerical range as

$$\Re[w(P)] \leq -\left(\xi^2 \lambda_m(D^H) - \xi \| (iA)^H \|_{op} \right), \quad (4.1)$$

where $\lambda_m(D^H)$ denotes the smallest eigenvalue. So, for the stability of $P$, it is sufficient to set the numerical range to be in the left half-plane. Thus, the following proposition can be concluded immediately from (4.1).
Proposition 4.1. For symmetric systems of size \( m \), the positive-definiteness of \( D \) implies strict stability with \( \delta = \lambda_m(D^H) \), the smallest eigenvalues of \( D^H \).

However for a non-symmetric \( A \), as shown in [38] by an example, the positive-definiteness of \( D \) does not necessarily imply condition (\( A \)). For the other direction, one can also cook up several examples with positive-definite \( D \) which are unstable. Nonetheless, the stability can be provided by a modified version of Proposition 4.1. As we will see, this is the same condition as [30, Thm. 2.1]. However, let us restate it regarding the context of this paper.

Theorem 4.2. For a strictly hyperbolic system with \( A \in \mathbb{R}^{m \times m} \), and with eigenvector matrix \( V \), the positivity of \( V^{-1}DV \) is sufficient for the stability in terms of condition (\( A \)).

Proof. By construction, \( A \) is hyperbolic and can be diagonalized as \( A = V \Lambda V^{-1} \), where \( V \) is the matrix of eigenvectors. Substituting this into the definition of \( P \) yields

\[
P(\xi) = -i\xi A - \xi^2 D = V \left( -i\xi \Lambda - \xi^2 \tilde{D} \right) V^{-1},
\]

where \( \tilde{D} := V^{-1}DV \). Since similarity transformations do not change the spectrum, we instead study the eigenvalues of \( \tilde{P}(\xi) \) defined as

\[
\tilde{P}(\xi) := -i\xi \Lambda - \xi^2 \tilde{D}.
\]

One can decompose \( \tilde{D} \) as the sum of Hermitian and skew-Hermitian matrices, i.e., \( \tilde{D}^H + \tilde{D}^\# \). From positivity \( -\xi^2 \tilde{D}^H \) is stable and by Theorem 3.3, the addition of skew-Hermitian matrices, \( -\xi^2 \tilde{D}^\# \) and \( -i\xi \Lambda \) cannot destabilize a stable Hermitian matrix. So \( P \) is stable. \( \Box \)

The following lemma allows us to study \( \tilde{D}' := V^{-1}(\tilde{A} - \tilde{\Lambda})(\tilde{A} + \tilde{\Lambda})V \) instead of \( \tilde{D} \):

Lemma 4.1. Consider the linear system (2.1). Let \( V \) to be the eigenvector matrix of \( A \), and suppose that the splitting \( A = \tilde{A} + \hat{A} \) is admissible in the sense of Definition 2.1. If there exists a lower-bound \( \lambda_H(\tilde{D}') \) for the eigenvalues of the Hermitian part of \( \tilde{D}' := V^{-1}(\tilde{A} - \tilde{\Lambda})(\tilde{A} + \tilde{\Lambda})V \), such that \( \lambda_H(\tilde{D}') = O(1) \), then the splitting is strictly stable in the sense of Majda–Pego.

Remark 4.3. Compared to the full frequency matrix \( P(\xi) \) used in Lemma 2.2, Lemma 4.1 is a convenient simplification since \( \tilde{D}'^H \) does not depend on \( \xi \).

One can go one step further and use the structure of the viscosity matrix, which is known for (2.2) unlike most of the literature in the hyperbolic-parabolic systems, cf. [6,11,30]. By the assumption, \( A \) and \( \tilde{A} \) are real diagonalizable, i.e.

\[
A = V \Lambda V^{-1}, \quad \tilde{A} = U \tilde{\Lambda} U^{-1},
\]

where \( V \) and \( U \) are matrices of eigenvectors and \( U = VJ \), where \( J \) stands for the change of basis matrix. Substituting these into the definition of \( P \) gives \( \tilde{P} \) as

\[
\tilde{P}(\xi) = -i\xi \Lambda - \xi^2 \Delta t \left[ \frac{\alpha \Delta x}{\Delta t} I_m + \Lambda^2 - 2J^{-1}\tilde{\Lambda}J\Lambda \right] = \tilde{D}.
\]
This form of $\tilde{D}$ reveals more of its structure. As one can see, for the positivity of $\tilde{D}$ the role of $J$ is crucial. For example for an admissible characteristic splitting, $J = I_m$ and the positivity (and stability) is clear since the components of $\tilde{D}$ are diagonal. So, it is plausible to claim that the splittings whose eigenspaces are close to each other are more likely to be stable. This indeed matches with results of [28, Sect. D.7], that if the eigenspaces of the split matrices coincide, the power-boundedness of each step is enough for the stability of the whole scheme. Although this form seems to be useful, the analysis of eigenspaces is a very delicate issue and we leave it open for now.

We wish to conclude this section with some remarks.

**Remark 4.4.** (i) The characteristic splitting decouples the system into $m$ scalar equations

$$\partial_t w_k + \lambda_k \partial_x w_k = 0, \quad k = 1 \ldots m.$$  

Then, using von Neumann stability analysis, it is not difficult to show that the explicit step (from time step $n$ to some intermediate step $n^\dagger$) and the implicit step (from the intermediate step $n^\dagger$ to the new time step $n+1$) are $\ell_2$-stable, respectively under an appropriate (and $\varepsilon$-uniform) CFL condition and unconditionally (see [41, Sect 3.3.4]).

(ii) In the light of [16, Lemma 3.1], the stability of each step is clearly enough for the stability of the whole scheme; however, it is far from being necessary in most cases, and almost often not practical to be fulfilled. For instance notice that the example in [38, Sect. 7] does not have stable steps. One could confirm numerically that for both stable and unstable settings (with $\varepsilon_1 = 10^{-1}$ and $\varepsilon_2 = 10^{-2}$ respectively) the implicit operator $\tilde{S}$ is power-bounded while the explicit operator $\hat{S}$ is not. Nonetheless, their multiplication $\tilde{S}\hat{S}$ makes one case stable and the other unstable. For further details about the stability of the difference equations, the reader can consult with [28, Appendix D] and [42, Chap. 4].

(iii) It seems plausible to conjecture that if the commutator is $O(1)$, then the viscosity matrix is parabolic under a suitable and non-restrictive choice of CFL condition, using the continuity of eigenvalues [35, Appendix K]. On the other hand, it is not even clear if the parabolicity is a relevant condition to be used for low frequency modes, as we mentioned earlier in Remark 2.3.

5. Applications  In this section, we show that Lemma 4.1 provides the linearized stability at any given state $u_0$ of several splittings used in practice, namely the splitting of Haack–Jin–Liu [16] (abbreviated as HJL hereinafter), Degond–Tang [9] (DT hereinafter) and the RS-IMEX splitting. We also discuss the numerical instability which has been reported in [33] for Klein’s auxiliary splitting of the Euler equations [25]. Recall that our analysis is based on the modified equation (2.2) and hence on the implicit-explicit Euler time integration accompanied with Rusanov-type numerical fluxes.
5.1. Haack–Jin–Liu splitting For the barotropic Euler equations, the HJL splitting decomposes the Jacobian of the flux function as

\[ A = \begin{bmatrix} 0 & 1 \\ -u^2 + \frac{p'(\varrho)}{\varepsilon^2} & 2u \end{bmatrix}, \]

\[ \hat{A} = \begin{bmatrix} 0 & \beta \\ -u^2 + \frac{p'(\varrho) - a(t)}{\varepsilon^2} & 2u \end{bmatrix}, \]

\[ \bar{A} = \begin{bmatrix} 0 & 1 - \beta \\ \frac{a(t)}{\varepsilon^2} & 0 \end{bmatrix}, \]

where \( \varrho, u, \) and \( p(\varrho) = \kappa \varrho^\gamma \) are the density, velocity, and pressure. \( \beta \in [0,1] \) is a parameter to be chosen (note that it is called \( \alpha \) in [16]) and \( a(t) := \min_x p'(\rho(x,t)) \). With these settings, the splitting is admissible in the sense of Definition 2.1. For further details see [16].

Assume that the system has been linearized around an arbitrary state \((\varrho, u) = (\varrho_0, u_0)\). Then, in light of Lemma 4.1, we have to study the positivity of \( \bar{D}' \). With the aid of Maple\textsuperscript{TM}, one can get

\[
\lim_{\varepsilon \to 0} \left( \varepsilon^2 \lambda_{H(\bar{D}')}^{1,2} \right) = \lim_{\varepsilon \to 0} \left[ \varepsilon^2 (\beta - 2) u^2 + \left( a - \beta p' \pm ((\beta - 1)p' + a) \right) \right] = \lim_{\varepsilon \to 0} \left[ \varepsilon^2 (\beta - 2) u^2 + \left( a \pm (p' + a) + \beta (-p' \pm p') \right) \right].
\]

Owing to the formal analysis for \( \varepsilon \ll 1 \), the asymptotic expansion gives \( p' - a = O(\varepsilon^2) \), so

\[
\lim_{\varepsilon \to 0} \left( \varepsilon^2 \lambda_{H(\bar{D}')}^{1,2} \right) = \begin{cases} a \\ (1 - 2\beta)a \end{cases}.
\]

Thus, if we set \( \beta \leq 1/2 \) and since \( a > 0 \), both eigenvalues are nonnegative, and \( \lambda_{H(\bar{D}')} = O(1) \). So, due to Lemma 4.1, the scheme is strictly stable in the sense of Majda–Pego under a non-restrictive CFL condition. Note that for the numerical experiments presented in [16], \( \beta \) is chosen as \( \beta \leq 1/2 \) and often of \( O(\varepsilon^2) \).

5.2. Degond–Tang splitting In [9] and for the barotropic Euler equations with the pressure function \( p(\varrho) = \kappa \varrho^\gamma \) (like the case of HJL splitting), the following splitting has been proposed:

\[ \hat{A} = \begin{bmatrix} 0 & 0 \\ -u^2 + \theta p'(\varrho) & 2u \end{bmatrix}, \quad \hat{\lambda} = 0, 2u, \]

\[ \bar{A} = \begin{bmatrix} 0 & \frac{1}{(1 - \theta \varepsilon^2) p'(\varrho)} \\ \frac{\varrho p'(\varrho)}{\varepsilon^2} & 0 \end{bmatrix}, \quad \bar{\lambda} = \pm \sqrt{\frac{(1 - \theta \varepsilon^2) p'(\varrho)}{\varepsilon}}, \]

where \( \theta \) is an ad-hoc parameter to be chosen between 0 and \( 1/\varepsilon^2 \). Note that it is discussed in [7,9,39] that taking \( \theta = O(1) \) leads to the AP property; so, we assume \( \theta \) to be \( O(1) \). Then, one can clearly confirm that this splitting is admissible in the sense of Definition 2.1.
As for the HJL splitting, we study the positivity of \( \tilde{D}' \). With the aid of Maple™, one can get
\[
\lim_{\varepsilon \to 0} \left( \varepsilon^2 \lambda_1^{1,2}_{\mathcal{H}(\tilde{D}')} \right) = \lim_{\varepsilon \to 0} \left[ -\varepsilon^2 \left( \beta + 2u^2 \right) p' + p' \pm O(\varepsilon^2) \right] = p'.
\]
(5.3)

Thus, both eigenvalues are positive in the limit, and due to Lemma 4.1, the scheme is strictly stable in the sense of Majda–Pego under a non-restrictive CFL condition.

5.3. RS-IMEX splitting

We apply the RS-IMEX splitting to the shallow water equations with flat bottom and the lake at rest reference solution as in [34, 44] (see also [3]), which gives the splitting
\[
\hat{A} = \begin{bmatrix} 0 & -m^2 \varepsilon^2/(z\varepsilon^2 - b) & 0 \\ z\varepsilon^2 & (z\varepsilon^2 - b)^2 & 2m \\ -m^2 \varepsilon^2 - b & \end{bmatrix}, \quad \hat{\lambda} = 0, \quad \frac{2m}{z\varepsilon^2 - b},
\]
\[
\tilde{A} = \begin{bmatrix} 0 & \frac{1}{z\varepsilon^2} \\ -b & 0 \\ \end{bmatrix}, \quad \tilde{\lambda} = \pm \sqrt{-b}/\varepsilon^2,
\]
where \( z \) denotes a scaled perturbation of the height from a constant reference, \( h + b = \varepsilon^2 z \), \( m \) is the momentum and \( b \) is the bottom function which is negative and constant. So, it can be concluded that this splitting is admissible in the sense of Definition 2.1. We refer the reader to [24, 34, 44] for details of this splitting.

As for the HJL splitting, one can obtain that
\[
\lim_{\varepsilon \to 0} \left( \varepsilon^2 \lambda_1^{1,2}_{\mathcal{H}(\tilde{D}')} \right) = \lim_{\varepsilon \to 0} \frac{-b^5 \pm \varepsilon |b| b^2 z - m^2}{(z\varepsilon^2 - b)^4} = -b > 0,
\]
(5.4)
since \( b < 0 \). Hence, using Lemma 4.1 and similarly as the case of HJL splitting, the splitting is strictly stable in the sense of Majda–Pego under a non-restrictive CFL condition. Note that the leading orders \( \varepsilon^2 \lambda_1^{1,2}_{\mathcal{H}(\tilde{D}')} \) have been the same for the HJL (with \( \beta = O(\varepsilon^2) \)), DT and RS-IMEX splittings.

Remark 5.1. It would be interesting to extend the stability result to equations with variable bottom. Nonetheless it is not clear how to linearize the Jacobian matrices \( \hat{A} \) and \( \tilde{A} \) (by freezing \( b \)), and also simultaneously the source term (by freezing \( b_x \)). Thus, it is more difficult to understand the linearization error, and hence the validity of the stability analysis.

Example 5.2. In addition to the previous analysis of the modified equation in the low Mach/Froude number limit (\( \varepsilon \ll 1 \)), we now study \( \lambda_1^{1,2}_{\mathcal{H}(\tilde{D}')} \) for all \( \varepsilon \in (0, 1] \). For this we consider the pressure law \( p(\varrho) = \varrho^2/2 \) and choose \( (\varrho_0, u_0) = (1, 1) \) for HJL and DT splittings, and equivalently \( (z_0, b, u_0) = (0, -1, 1) \) for the RS-IMEX splitting. We also set the ad-hoc parameters of HJL and DT splittings as \( \beta = \varepsilon^2 \) and \( \theta = 1 \). Figure 5.1 shows that \( \lambda_1^{1,2}_{\mathcal{H}(\tilde{D}')} \) are bounded from below. Indeed, \( \lambda_1^{1,2}_{\mathcal{H}(\tilde{D}')} \) is always positive, while \( \lambda_2^{1,2}_{\mathcal{H}(\tilde{D}')} \) is positive to the left of the kink around \( \varepsilon \in (0.4, 0.6) \), and negative to the right, but uniformly bounded. Thus, owing to Lemma 4.1, all these splittings are asymptotically stable. Note that the plots of RS-IMEX, HJL and DT splittings are hardly distinguishable for small \( \varepsilon \).
5.4. Klein’s auxiliary splitting

In his influential paper [25], Klein introduced two flux-splittings for the full Euler equations. The main splitting introduces two sub-systems, called system (I) and (II), given by [25, eqs. (3.1)-(3.2)]. In the second splitting, system (I) is replaced by the so-called auxiliary system (I*), which is given by [25, eq. (3.8)].

In this section, we analyze the stability of a flux-splitting IMEX scheme, which uses Klein’s auxiliary splitting as a building block (cf. [33]).

Here, the background state for the linearization is \((\varrho, \varrho u, \varrho E) = (\varrho_0, \varrho_0 u_0, \varrho_0 E_0)\), where the total energy \(\varrho E\) satisfies the dimensionless equation of state \(\varrho E = \frac{p}{\gamma - 1} + \frac{\varepsilon^2}{2} \varrho |u|^2\). Following the derivation in [33], the auxiliary splitting is given by

\[
\begin{align*}
\hat{A} &= \begin{bmatrix}
0 & 1 & 0 \\
\frac{(\gamma - 1)\varepsilon^2}{2} - 1 & u^2 (2 - (\gamma - 1)\varepsilon^2) & u \\
-\hat{A}_{31} & \hat{A}_{32} & \hat{A}_{33}
\end{bmatrix}, \\
\hat{A}_{31} &:= -u \left[ (1 + \varepsilon^2(\gamma - 1)) (\frac{\varepsilon^2}{2} u^3 - (\gamma - 1)\varepsilon^2 u^2 + (1 - \varepsilon^2) \frac{p_{\text{inf}}}{\varrho}) \right] + \frac{(\gamma - 1)\varepsilon^4}{2} u^3, \\
\hat{A}_{32} &:= E + \varepsilon^2(\gamma - 1) \left[ (\frac{\varepsilon^2}{2} u^2 - (\gamma - 1)\varepsilon^2 u^2 + (1 - \varepsilon^2) \frac{p_{\text{inf}}}{\varrho}) \right] - (\gamma - 1)\varepsilon^4 u^2, \\
\hat{A}_{33} &:= (1 + \varepsilon^2(\gamma - 1)) u,
\end{align*}
\]

\[
\tilde{A} = \begin{bmatrix}
0 & 0 & 0 \\
\frac{1}{2}(\gamma - 1)u^2 & -\frac{1}{2}(\gamma - 1)u & 0 \\
-\frac{p_{\text{inf}}}{\varepsilon^2} + \frac{\gamma - 1}{2} \varrho \varrho^2 + (\gamma - 1)\varepsilon^2 u^3 & -\varepsilon^2 (\gamma - 1)u^2 & 0
\end{bmatrix},
\]

with \(1 < \gamma < \frac{5}{3}\). The choice of the parameter \(p_{\text{inf}} := \min_x p(x,t)\) guarantees the hyperbol-
On the stability of IMEX splittings

icity of split systems, whose eigenvalues read

\[ \hat{\lambda} = u, u \pm c^*, \quad c^* := \sqrt{\frac{p + (\gamma - 1)\Pi}{\varrho}}, \quad \Pi := (1 - \varepsilon^2)p_{\text{inf}} + \varepsilon^2p, \]

\[ \hat{\lambda} = 0, \pm \frac{(1 - \varepsilon^2)}{\varepsilon} \sqrt{\frac{(\gamma - 1)(p - p_{\text{inf}})}{\varrho}}. \]

So, the splitting is admissible in the sense of Definition 2.1 (see [33]).

Our attempts to compute the eigenvalues of \( \tilde{D}' \mathcal{H} \) for this full Euler case with Maple™ failed. Thus we study the full frequency matrix \( P \) for the example of two colliding pulses [25], where we had previously found an \( \varepsilon \)-dependent timestep [33]. The domain is \([-L,L]\) with \( L := 2/\varepsilon \), for \( \gamma = 1.4 \). The initial data are

\[ \varrho(x,0) = \varrho^{(0)} + \frac{\varepsilon}{2}\varrho^{(1)} \left( 1 - \cos \left( \frac{2\pi x}{L} \right) \right), \quad \varrho^{(0)} = 0.955, \quad \varrho^{(1)} = 2, \quad (5.5) \]

\[ p(x,0) = p^{(0)} + \frac{\varepsilon}{2}p^{(1)} \left( 1 - \cos \left( \frac{2\pi x}{L} \right) \right), \quad p^{(0)} = 1, \quad p^{(1)} = 2\gamma, \quad (5.6) \]

\[ u(x,0) = \frac{1}{2}u^{(0)} \text{sign}(x) \left( 1 - \cos \left( \frac{2\pi x}{L} \right) \right), \quad u^{(0)} = 2\sqrt{\gamma}. \quad (5.7) \]

In order to apply the Majda-Pego stability framework, we linearize around

\[ \varrho_0 = \varrho^{(0)} + \frac{\varepsilon}{2}\varrho^{(1)} = 0.955 + \varepsilon, \]

\[ p_0 = p^{(0)} + \frac{\varepsilon}{2}p^{(1)} = 1 + \varepsilon\gamma \]

\[ u_0 = \sqrt{\gamma} \]

and \( p_{\text{inf}} = 1 \). Note that we have replaced \( 1 - \cos \left( \frac{2\pi x}{L} \right) \) by its mean value 1. The numerical diffusion and the grid parameters are chosen as in [33],

\[ \alpha = \sqrt{\frac{\varrho_0}{\varrho_0} + \max_x (u(x,0))}, \quad \Delta x = 0.05, \quad \Delta t = \frac{\text{CFL}}{\max_x (u(x,0))} \Delta x. \]

We compute the real parts of the eigenvalues of the frequency matrix \( P \) of the modified equation numerically. Figure 5.2 displays \( \Re(\lambda_1) \) for different CFL numbers and frequency variable \( \xi = \varepsilon \pi \). The three subfigures are zooms in \( \varepsilon \). The figures reveal a small instability region near \( \varepsilon \in (0.02,0.06) \) and for CFL = 0.45. This seems to correspond closely to some of the numerical experiments in [33], where the CFL number needed to be reduced when changing \( \varepsilon \) from 0.1 to 0.05.

Note, however, that the lack of uniform stability in [33] is much stronger than the one seen in Figure 5.2, since in [33] the CFL-number needed to decrease linearly with the Mach number, while in Figure 5.2 \( \Re(\lambda_1) \leq 0 \) uniformly in \( \varepsilon \), for fixed CFL = 0.02. This discrepancy may possibly be due to a fundamental difference between the Fourier analysis in the present paper and the real computation in [33]: based on Lemma 2.2, Figure 5.2 studies a single Fourier mode. On the other hand, due to the \( \text{sign}(x) \) function
in (5.7), the initial data for the velocity contain a superposition of all Fourier modes, which may trigger instabilities not explained by the present analysis.

**Remark 5.3.** It is important to point out some differences between the algorithms in [33] and [25]. Klein develops his approach using the more complex setting of multiple space variables and multiple pressures. Algorithmically, he “combines explicit predictor steps for long wave linear acoustics or global compression with a single implicit scalar Poisson-type corrector scheme” [25, p.3]. Thus, our stability analysis has no direct implication for the scheme proposed in [25]. Rather, it should be seen as a comment to [33].

![Graphs](image1.png)

**Fig. 5.2.** $\Re(\lambda_p)$ for Klein’s auxiliary splitting w.r.t. $\varepsilon$, in different regions of $\varepsilon$ and for $\xi = \varepsilon \pi$.

### 6. Concluding remarks

We have studied the stability of several flux-splitting IMEX schemes for stiff systems of hyperbolic conservation laws by modified equation analysis. First we reviewed and slightly expanded the stability framework of Majda and Pego. We showed that for symmetric splittings the viscosity matrix is positive, which gives strict stability. Furthermore, we discussed a general class of splittings, and showed that positivity of the
viscosity matrix, after being transformed by the matrix of eigenvectors of $A$, is sufficient for strict stability (this matches the results of [30]). This criterion has been used to show the stability of the Haack–Jin–Liu splitting [16] and the Degond–Tang splitting [9] for the isentropic Euler equations, and the recent RS-IMEX splitting for the shallow water equations (with flat bottom). For the full Euler equation, we discovered a small region of instability for Klein’s so-called auxiliary splitting [25] for the two colliding pulses example. This seems to confirm computational results in [33].

REFERENCES

[22] ———. 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), pp. 177–216.


