TEST EQUATIONS AND LINEAR STABILITY OF
IMPLICIT-EXPLICIT GENERAL LINEAR METHODS

ANDREW J. STEYER

Abstract. Eigenvalue perturbation theory is applied to justify using complex-valued linear scalar test equations to characterize the stability of implicit-explicit general linear methods (IMEX GLMs) solving autonomous linear ordinary differential equations (ODEs) when the implicitly treated term is sufficiently stiff relative to the explicitly treated term. The stiff and non-stiff matrices are not assumed to be simultaneously diagonalizable or triangularizable and neither matrix is assumed to be symmetric or negative definite. The stability of IMEX GLMs solving complex-valued scalar linear ODEs displaying parabolic and hyperbolic stiffness is analyzed and related to the higher dimensional theory. The utility of the theoretical results is highlighted with a stability analysis of a family of IMEX Runge-Kutta methods solving IVPs of a linear 2D shallow-water model and a linear 1D advection-diffusion model.

Key words. general linear methods, implicit-explicit methods, IMEX methods, Runge-Kutta methods, linear multistep methods, stability, stability analysis

AMS subject classifications. 65L04, 65L05, 65L06, 65L20, 65M12, 65M20, 65F15

1. Introduction. Stability analysis is a critical step in the derivation of efficient and accurate methods for the time-integration of initial value problems (IVPs) of ordinary and partial differential equations (ODEs and PDEs respectively). For stiff IVPs the time-step of explicit methods is often limited more by stability rather than by accuracy, while the use of implicit methods with less restrictive stability properties requires solving algebraic equations. Implicit-explicit (IMEX) methods are a compromise between pure explicit and pure implicit methods that can avoid some of the stability restrictions of pure explicit methods and at the same time simplify or reduce the dimension of the algebraic equations arising from implicitly treated terms. The stability analysis of IMEX methods is challenging since the stability regions of the implicit and explicit part of the method can couple in complex and counter-intuitive ways and a method can have different stability properties for different IMEX splittings. The focus of this paper is on further developing the linear stability theory for IMEX general linear methods (GLMs) and in particular providing justification for the use of complex-valued scalar linear test equations to characterize their stability.

Our contribution is to develop a theory for the stability of IMEX GLMs solving IVPs of autonomous linear ODEs of the form \( \dot{x} = Ax \) where the only hypothesis placed on the IMEX splitting of \( A \) into nonstiff and stiff components \( A = N + S \) is that \( S \) is stiff relative to \( N \). In particular we do not assume that the splitting is such that \( N \) and \( S \) are, as in [2] and [1], simultaneously diagonalizable or triangularizable (SD or ST respectively) or that either of \( S \) or \( N \) is, as in [12] or [36], symmetric or negative definite. Our motivation for weakening these standard assumptions made in
the literature is the spatially discrete linear shallow water model (Equation (32) in Section 5.2) which can be viewed as a simplification of the horizontal explicit vertically implicit (HEVI) IMEX splitting used in nonhydrostatic atmosphere models (see e.g. [39]). The IMEX splitting used in Equation (32) is not ST and the coefficient matrices are not symmetric (or even skew-symmetric unless \( g = h_0 \)) or negative definite.

Our main result (Theorem 9) uses eigenvalue perturbation theory to show that the stability of an IMEX GLM solving an IVP of \( \dot{x} = Ax \) with an IMEX splitting \( A = N + S \) where \( S \) is sufficiently stiff relative to \( N \) can be approximately characterized by complex-valued scalar linear test equations of the form \( \dot{z} = \lambda z + \mu z \) where the coefficients \( \lambda \) and \( \mu \) are eigenvalues of \( N \) and \( S \) respectively. This is followed by Proposition 14 which states a necessary condition and a sufficient condition for stability in terms of the step-size and a term measuring how well test equations characterize stability. Theorem 9, Proposition 14, and the other results of Section 3 do not assume that \( N \) and \( S \) are ST or SD or that either of \( N \) or \( S \) is symmetric or negative definite. Note that (see e.g. Example 4 in Section 2 below) the stability of an IMEX GLM is not always correctly characterized by analysis of the stability regions of its explicit and implicit method in isolation. This motivates the analysis in Section 4, approximately justified by the results of Section 3, of several joint IMEX stability regions arising from complex-valued scalar linear test equations with either hyperbolic or parabolic stiffness. In Section 5 we use the results of Sections 3-4 in the stability analysis of spatially discrete linear 1D advection-diffusion and linear 2D shallow water models. We derive a family of IMEX Runge-Kutta (RK) methods and show that in the advection-diffusion model, which has an ST IMEX splitting, the IMEX stability regions accurately predict the maximum step-size of several members of this IMEX RK family. For the linear shallow water model, whose IMEX splitting is not ST, the maximum stable time-step of two methods of this IMEX RK family is the opposite of what is predicted from analysis of their stability regions and can be explained using the perturbation stability theory developed in Section 3.

The analysis and application of IMEX methods, including IMEX RK methods, IMEX linear multistep methods (LMMs), and IMEX GLMs, for the approximate solution of IVPs has a long and rich history (see e.g. the 1968 paper [42] where the Strang splitting is derived). We refer readers to the book by Jackiewicz [25] for an extensive treatment of the theory and analysis of GLMs. Order conditions for partitioned RK methods were derived by Hairer in [21] (see also [28] and for a more recent presentation see [22]). The accuracy of IMEX LMMs was analyzed in Section 2 of [2]. Order conditions for diagonally implicit IMEX GLMs are given in Theorem 2.1 of [5] and order conditions for IMEX GLMs with high stage order are given in Theorem 2 of [4]. In [3] the stability and contractivity of a family of semi-implicit RK methods was investigated using the logarithmic norm. The stability of IMEX RK methods solving linear IVPs whose IMEX splitting is ST was analyzed in Section 3 of [1]. IMEX RK methods with large IMEX stability regions are constructed in [27]. The stability of LMMs solving complex-valued test equations in two variables and linear IVPs whose IMEX splitting is ST was analyzed in [43], [2], [18], and somewhat more recently in [29]. In [36], the unconditional stability of IMEX LMMs solving IVPs of linear ODEs whose IMEX splitting has a matrix that is symmetric and negative definite was analyzed. Analysis of the stability of IMEX RK methods solving partial differential equations (PDEs) with relaxation is found in [33], [37], and [34]. The stability of IMEX GLMs solving complex-valued linear scalar ODEs was investigated in [26] and [38].

The stability of numerical methods for IVPs of ODEs and PDEs and the closely
related topic of stiffness are classic subjects in numerical analysis. The earliest work on these subjects dates at least to the 1952 paper of Curtiss and Hirschfelder [13] and to the PhD thesis and subsequent seminal works of Dahlquist [14, 15, 16]. Since then many other stability theories for the solution of IVPs, including monotonically contracting nonlinear and nonautonomous IVPs, have emerged such as B-stability [9] or algebraic stability and AN-stability [7]. Equivalences amongst these various stability theories are investigated in [10]. We refer readers to [11] and the references therein for an overview of the history of the theory of stiffness. Recently, time-dependent theories of stiffness and stability have been developed [11, 40, 41] for the stability analysis of nonlinear and nonautonomous IVPs that are not monotonically contracting.

The remainder of this paper is organized as follows. In Section 2 we cover some preliminary notation and definitions. Theorem 9 and Proposition 14 are proved in Section 3 where the higher dimensional theory is developed. Complex-valued scalar linear test equations are analyzed in detail in Section 4. In Section 5 we analyze the stability of spatially discrete linear advection-diffusion and shallow water models to illustrate our theoretical results and explain the stability properties of a family of IMEX RK methods solving IVPs of these models. The paper is concluded in Section 6 with some final remarks and acknowledgments.

2. Implicit-explicit general linear methods. For \( w \in \mathbb{N} \) we let \( \| \cdot \| \) denote a norm on \( \mathbb{R}^w \) and use the same symbol for the induced matrix norm on \( \mathbb{R}^{w \times w} \). Consider a nonlinear and time-dependent (non-autonomous) ODE of the following form:

\[
\dot{x} = f(x, t) \equiv n(x, t) + s(x, t),
\]

where \( f, n, s : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d \) for some positive integer \( d \). Assume for the remainder of this paper that the solution of each IVP of (1) with initial condition \( x(t_0) = x_0 \) exists and is unique for every \( x_0 \in \mathbb{R}^d \) and \( t_0 \in \mathbb{R} \). We refer to the additive splitting \( f(x, t) = n(x, t) + s(x, t) \) as an IMEX splitting. An \( r \)-stage, \( k \)-step additive GLM for the numerical approximation of the solution of IVPs of (1) using the additive splitting \( f = n + s \), initial time \( t_0 \in \mathbb{R} \), initial values \( X_{0,1}, \ldots, X_{0,k} \in \mathbb{R}^d \), and step-size \( h > 0 \) is defined by the following equations where \( m \in \mathbb{N}, i = 1, \ldots, r, \) and \( w = 1, \ldots, k \):

\[
\begin{align*}
G_{m,i} &= \sum_{j=1}^{k} U_{i,j} X_{m,j} + h \sum_{j=1}^{r} C_{i,j} n(G_{m,j}, t_{m,j}) + h \sum_{j=1}^{r} \hat{C}_{i,j} s(G_{m,j}, \hat{t}_{m,j}) \\
X_{m+1,w} &= \sum_{j=1}^{k} V_{w,j} X_{m,j} + h \sum_{j=1}^{r} C_{i,j} n(G_{m,j}, t_{m,j}) + h \sum_{j=1}^{r} \hat{C}_{i,j} s(G_{m,j}, \hat{t}_{m,j})
\end{align*}
\]

where \( t_m := t_0 + hm, t_{m,j} := t_m + c_j h \) for \( m \in \mathbb{N} \) and \( j = 1, \ldots, r \) and the subsets of \( \mathbb{R} \) given by \( \{U_{i,j}\}_{i=1,j=1}^{r,k}, \{V_{i,j}\}_{i=1,j=1}^{r,k}, \{C_{i,j}\}_{i=1,j=1}^{r,r}, \{\hat{C}_{i,j}\}_{i=1,j=1}^{r,r}, \{D_{i,j}\}_{i=1,j=1}^{k,r}, \{\hat{D}_{i,j}\}_{i=1,j=1}^{k,r}, \{c_i\}_{i=1}^{r}, \) and \( \{\hat{c}_i\}_{i=1}^{r} \) are method defining coefficients. The external stages \( \{X_{m,i}\}_{i=1}^{k} \) approximate linear combinations of solutions of IVPs of (1).

We denote the bilinear Kronecker product for matrices \( A = (A_{i,j})_{i,j=1}^{r,k} \in \mathbb{R}^{q \times q} \) and \( B \in \mathbb{R}^{w \times w} \) by

\[
A \otimes B := \begin{bmatrix} A_{1,1} B & \cdots & A_{1,q} B \\ \vdots & \ddots & \vdots \\ A_{q,1} B & \cdots & A_{q,q} B \end{bmatrix} \in \mathbb{R}^{qw \times qw}.
\]

Define the following matrices in terms of the method defining coefficients: \( U = (U_{i,j}) \in \mathbb{R}^{r \times k}, V = (V_{i,j}) \in \mathbb{R}^{k \times k}, C = (C_{i,j}), \hat{C} = (\hat{C}_{i,j}) \in \mathbb{R}^{r \times r}, D = (D_{i,j}), \hat{D} = (\hat{D}_{i,j}) \in \mathbb{R}^{k \times k} \).
\((\hat{D}_{i,j}) \in \mathbb{R}^{k \times r}, c = (c_i) \in \mathbb{R}^{r \times 1}, \hat{c} = (\hat{c}_i) \in \mathbb{R}^{r \times 1}\) and let \(X_m := (X_{m,1}^T, \ldots, X_{m,k}^T)^T \in \mathbb{R}^{d_k}\) and \(G_m := (G_{m,1}^T, \ldots, G_{m,r}^T)^T \in \mathbb{R}^{d_r}\). This gives the following compact formulation of (2):

\[
G_m = (U \otimes I_d)X_m + h(C \otimes I_d)\pi_m + h(\hat{C} \otimes I_d)\pi_m \quad X_{m+1} = (V \otimes I_d)X_m + h(D \otimes I_d)\pi_m + h(\hat{D} \otimes I_d)s_m, \quad m \in \mathbb{N}
\]

where we define \(\pi_m := (n(G_{m,1}, t_{m,1})^T, \ldots, n(G_{m,r}, t_{m,r})^T)^T \in \mathbb{R}^{d_r}\) and \(s_m := (s(G_{m,1}, t_{m,1})^T, \ldots, s(G_{m,r}, t_{m,r})^T)^T \in \mathbb{R}^{d_r}\). We represent IMEX GLMs following the style of [6]:

\[
\begin{array}{c|c|c|c|c|c|c}
\hline
& c & \hat{c} & C & \hat{C} & D & U \\
\hline
\end{array}
\]

Henceforth we shall always assume that the GLM \(\begin{array}{c|c|c|c|c|c|c}
\hline
& c & \hat{c} & A & \hat{A} & 1_r & 1 \\
\hline
\end{array}\), referred to as the explicit method of (3) is explicit \((C\) is strictly lower triangular\) and that the GLM \(\begin{array}{c|c|c|c|c|c|c}
\hline
& c & \hat{c} & \hat{C} & D & \hat{U} & V \\
\hline
\end{array}\), referred to as the implicit method, is implicit \((\hat{C}\) is not strictly lower triangular\) so that the additive GLM defined by (3) is an IMEX GLM. Unless otherwise noted we always assume that for an IMEX splitting \(f(x, t) = n(x, t) + s(x, t)\) that the term \(n(x, t)\) on the left is the explicitly treated term and the term \(s(x, t)\) on the right is the implicitly treated term and reserve the letters \(N, n\) for non-stiff terms and \(S, s\) for stiff terms.

We now discuss two important classes of IMEX GLMs. The first class are the \(r\)-stage IMEX RK methods \(\begin{array}{c|c|c|c|c|c|c}
\hline
& c & \hat{c} & A & \hat{A} & 1_r & 1 \\
\hline
\end{array}\) (we define \(1_w := (1, \ldots, 1)^T \in \mathbb{R}^w\) for any positive integer \(w\)) where \(\begin{array}{c|c|c|c|c|c|c}
\hline
& c & A & b^t & \hat{A} & b^t & 1 \\
\hline
\end{array}\) is an \(r\)-stage explicit RK method and \(\begin{array}{c|c|c|c|c|c|c}
\hline
& c & \hat{c} & \hat{A} & b^t & \hat{b}^t & 1 \\
\hline
\end{array}\) is an \(r\)-stage implicit RK method. The second class of IMEX GLMs are IMEX LMMs of the following form:

\[
x_{m+k} + \sum_{j=0}^{k-1} \alpha_i x_{m+i} = \sum_{j=0}^{k} \beta_i n(x_{m+i}, t_{m+i}) + \sum_{j=0}^{k} \hat{\beta}_i s(x_{m+i}, t_{m+i}), \quad m \in \mathbb{N}
\]

where \(\{\beta_i\}_{i=0}^k, \{\hat{\beta}_i\}_{i=0}^k, \{\alpha_i\}_{i=0}^{k-1} \subset \mathbb{R}\) and \(t_{m+i} = mh + ih\). We express the \(k\)-step method (5) as an IMEX GLM following the approach used in [8] to represent LMMs as irreducible GLMs:

\[
(6) \quad c = \hat{c} = [k], \quad C = [\beta_k], \quad \hat{C} = [\hat{\beta}_k], \quad U = [0, \ldots, 0, 1] \in \mathbb{R}^{1 \times k} \quad D = [\beta_0 - \alpha_0 \beta_k, \ldots, -\alpha_k \hat{\beta}_k] \quad \hat{D} = [\beta_0 - \alpha_0 \hat{\beta}_k, \ldots, -\alpha_k \hat{\beta}_k] \quad V = \begin{bmatrix} 0 & -\alpha_0 \\ 1 & -\alpha_1 \\ \vdots & \vdots \\ 0 & -\alpha_k \end{bmatrix} \in \mathbb{R}^{k \times k}
\]
The following complex-valued scalar linear test equation in two variables has been used (see e.g. Section 4.4 of [4]) to characterize the stability of IMEX GLMs (3):

\[
\dot{z} = \lambda z + \mu z, \quad \lambda, \mu \in \mathbb{C}.
\]

Test equations of the form (7) have often been used in the context of IMEX RK methods [1] and IMEX LMMs [2, 18] to define and characterize the stability regions of these methods. The test equation (7) is meant to serve as caricature of the higher dimensional linear case when \( n(x,t) = Nx \), \( s(x,t) = Sx \) for matrices \( N, S \in \mathbb{R}^{d \times d} \) where \( \lambda \) and \( \mu \) are eigenvalues of \( N \) and \( S \) respectively. This caricature is justified in the case where \( N \) and \( S \) are ST (see e.g. the end of Section 1 on page 2 of [18] or Proposition 5 below).

Applying the method (3) to solve (7) with step-size \( h > 0 \) and initial value \( Z_0 \in \mathbb{C}^k \) results in the following iteration:

\[
Z_{m+1} = R(h\lambda, h\mu)Z_m, \quad m \in \mathbb{N}
\]

where the stability function \( R : \mathbb{C}^2 \rightarrow \mathbb{C}^{k \times k} \) of (3) is defined by the following formula:

\[
R(w, z) = V + [Dw + \hat{D}z][I_r - Cw - \hat{C}z]^{-1}U.
\]

Stability regions for IMEX GLMs are defined analogously to the standard GLM case by taking into account the fact that the stability function is now a function of two complex variables as opposed to one. For any matrix \( A \) we let \( \text{eig}(A) \) denote the set of eigenvalues of \( A \) and let \( |\text{eig}(A)| := \max\{|\lambda| : \lambda \in \text{eig}(A)\} \).

**Definition 1.** A matrix \( A \in \mathbb{R}^{w \times w} \) is said to be power bounded if given any induced matrix norm \( \|\cdot\| \) on \( \mathbb{R}^w \) there exists \( M \geq 0 \) so that \( \|A^j\| \leq M \) for all \( j \in \mathbb{N} \).

**Definition 2.** The stability region of an IMEX GLM (3) is the following set:

\[
\{(w, z) \in \mathbb{C}^2 : R(w, z) \text{ is power bounded}\}.
\]

The region of absolute stability or absolute stability region of an IMEX GLM is the following set:

\[
\{(w, z) \in \mathbb{C}^2 : |\text{eig}(R(w, z))| < 1\}.
\]

The region of absolute stability is contained in the stability region since a sufficient condition for power boundedness is that the eigenvalues of a matrix are all strictly less than one in modulus. On the other hand, if \( R(w, z) \) has an eigenvalue greater than 1 in modulus, then \( (w, z) \) cannot be in the stability region.

The basic stability theory for IMEX GLMs is a compromise between the purely explicit and the purely implicit cases. The restriction \( R(w, z)|_{w=0} \) is the stability function of the implicit method of (3) and the restriction \( R(w, z)|_{z=0} \) is the stability function of the explicit method of (3). By taking \( w \) with \( \text{Re}(w) < 0 \) outside of the stability region of the explicit method so that \( |R(w, 0)| > 1 \) it follows no IMEX GLM can be A-stable in the sense that \( (w, z) \) is in the stability region for all \( (w, z) \in \mathbb{C}^2 \) such that \( \text{Re}(w), \text{Re}(z) \leq 0 \).

One approach to investigate the stability region of an IMEX GLM is to relate it to the stability regions of its implicit and explicit methods. Denote the closed left half complex plane as \( \mathbb{C}^- := \{z \in \mathbb{C} : \text{Re}(z) \leq 0\} \). The following definition, based on concepts introduced in [18], take into account the fact that the stability region of (3) depends on jointly on both the implicit and explicit method of (3).
Definition 3. A method (3) is implicitly A-stable if \((w, z)\) is in the linear stability region for all \(z \in \mathbb{C}^-\) whenever \(w \in \mathbb{C}\) is in the stability region of the explicit method of (3). The implicit stability region of (3) is the set of all \(z \in \mathbb{C}\) such that \((w, z)\) is in the stability region of (3) for all \(w\) in the stability region of the explicit method of (3).

Both the IMEX Euler method and the Crank-Nicolson leap-frog method (Equation 2.5 of [18]) are implicitly A-stable. However, there are serious limitations to using implicit A-stability and the implicit stability region to understand the stability region of (3). As illustrated in the following example many methods are not implicitly A-stable and it can be the case that \((w, z)\) is outside the linear stability region of a method (3) even if \(w\) is in the linear stability region of the explicit method and \(z\) is in the linear stability region of the implicit method.

Example 4. The four-stage, fourth order, L-stable DIMSIM4 method (see page A1438 of [38]) is not implicitly A-stable since \(|R(0.1 + 0i, 0)| \leq 0.850\) and \(|R(0.1 + 0i, -0.4 + 0i)| \geq 1.06\). The three-stage, third-order, L-stable, DIRK (3,4,3) method (see Section 2.7 of [1]) is also not implicitly A-stable since \(|R(0, 1.1i)| \leq 0.98\), and \(|R(2i, 1.1i)| \leq 0.59\).

Example 4 illustrates a point that must be stressed: we cannot in general draw conclusions about the stability region of an IMEX GLM by separately considering the stability regions of its implicit or explicit methods.

### 3. Justification of scalar test equations.

In this section we focus on the justification of using scalar linear test equations of the form (7) to characterize the stability of (3) applied to solve a linear autonomous ODE of the following form:

\[
\dot{x} = Ax, \quad A \in \mathbb{R}^{d \times d}, \quad x \in \mathbb{R}^d
\]

with an IMEX splitting \(A = N + S\) where \(N\) and \(S\) are the nonstiff and stiff coefficient matrices respectively:

\[
\dot{x} = Nx + Sx, \quad N, S \in \mathbb{R}^{d \times d}, \quad x \in \mathbb{R}^d.
\]

It is often useful (see e.g. Section 1 of [33]) to consider an equation of the form (10) depending on parameter \(\delta > 0\) that controls the stiffness of the stiffly treated term:

\[
\dot{x} = Nx + \delta^{-1}Sx, \quad \delta > 0
\]

The term \(\delta^{-1}S\) becomes very stiff in the “stiff limit” \(\delta \to 0\). Applying the IMEX GLM (3) to solve the linear ODE (10) yields the following linear difference equation:

\[
X_{m+1} = R(hN, hS)X_m, \quad m \geq 0
\]

where the stability matrix \(R(hN, hS)\) is defined as follows (recall that \(\otimes\) denotes the Kronecker product):

\[
R(hN, hS) := V \otimes I_d + [D \otimes hN + \hat{D} \otimes hS][I_d - C \otimes hN - \hat{C} \otimes hS]^{-1}[U \otimes I_d].
\]

Although the symbol \(R\) is used to represent the stability matrix (13) and the stability function (8) of scalar linear test equations it will always be clear from context what \(R\) is representing.

The following proposition shows that if \(N\) and \(S\) are ST, then the eigenvalues of \(R(hN, hS)\) are the given the union of eigenvalues of stability functions of the method (3) applied with step-size \(h > 0\) to solve \(d\) scalar test equations of the form (7) where \(\lambda \in \text{eig}(N)\) and \(\mu \in \text{eig}(S)\).
Proposition 5. Assume that $N$ and $S$ are ST with $P \in \mathbb{C}^{d \times d}$ such that $U_N = P^{-1}N^P$ and $U_S = P^{-1}S^P$ both upper triangular. Then the set of $d$ eigenvalues of the stability matrix $R(hS, hN)$ is exactly equal to the union $\bigcup_{i=1}^d \text{eig}(R(h(U_N)_{i,i}, h(U_S)_{i,i}))$ where $(U_N)_{i,i}$ and $(U_S)_{i,i}$ denote the $i$th diagonal entries of $U_N$ and $U_S$ respectively for $i = 1, \ldots, d$.

Proof. The method (3) applied to solve (10) takes the following form where $i = 1, \ldots, r$, $m \in \mathbb{N}$, and $w = 1, \ldots, k$:

\begin{equation}
\begin{cases}
G_{m,i} = \sum_{j=1}^{k} U_{i,j} X_{m,j} + \frac{1}{h} \sum_{j=1}^{r} C_{i,j} NG_{m,j} + \frac{1}{h} \sum_{j=1}^{r} \hat{C}_{i,j} SG_{m,j}, \\
X_{m+1,w} = \sum_{j=1}^{k} V_{w,j} X_{m,j} + \frac{1}{h} \sum_{j=1}^{r} D_{w,j} NG_{m,j} + \frac{1}{h} \sum_{j=1}^{r} D_{w,j} SG_{m,j}.
\end{cases}
\end{equation}

Under the change of variables $G_{m,i} = PK_{m,i}$, $X_{m,w} = PY_{m,w}$ the system (14) is transformed to the following system:

\begin{equation}
\begin{cases}
K_{m,i} = \sum_{j=1}^{k} U_{i,j} Y_{m,j} + \frac{1}{h} \sum_{j=1}^{r} C_{i,j} U_N K_{m,j} + \frac{1}{h} \sum_{j=1}^{r} \hat{C}_{i,j} U_S K_{m,j}, \\
Y_{m+1,w} = \sum_{j=1}^{k} V_{w,j} Y_{m,j} + \frac{1}{h} \sum_{j=1}^{r} D_{w,j} U_N K_{m,j} + \frac{1}{h} \sum_{j=1}^{r} D_{w,j} U_S K_{m,j}.
\end{cases}
\end{equation}

If $Y_m := (Y_{m,1}^T, \ldots, Y_{m,k}^T)^T$, then $Y_{m+1} = (I_k \otimes P^{-1})R(hU_N, hU_S)(I_k \otimes P)Y_m$. We express the variables $K_{m,i}$ and $Y_{m,w}$ using component-wise notation where $K_{m,i} = (K_{m,i,1}, \ldots, K_{m,i,k})^T$ and $Y_{m,w} = (Y_{m,1,w}, \ldots, Y_{m,k,w})^T$ and for $q = 1, \ldots, d$ we let $K_{m,i}^q := (K_{m,i,1}^q, \ldots, K_{m,i,k}^q)^T$ and $Y_{m,w}^q := (Y_{m,1,w}^q, \ldots, Y_{m,k,w}^q)^T$. For $q = 1, \ldots, d$, the fact that $U_N$ and $U_S$ are upper triangular together with Equation (15) imply that each $Y_{m+1,w}^q$ depends only on $\{Y_{m,w}^q\}_{w=1,\ldots,k}$. Define a permutation of $Y_m$ by the following:

$\tilde{Y}_m := (Y_{m,1}^1, Y_{m,1}^2, Y_{m,1}^3, \ldots, Y_{m,1}^d, Y_{m,2}^1, Y_{m,2}^2, Y_{m,2}^3, \ldots, Y_{m,2}^d, \ldots, Y_{m,k}^1, Y_{m,k}^2, Y_{m,k}^3, \ldots, Y_{m,k}^d)^T = (\tilde{Y}_{m,1}^T, \ldots, \tilde{Y}_{m,d}^T)^T$

where $\tilde{Y}_{m,q} := (Y_{m,1}^q, Y_{m,2}^q, \ldots, Y_{m,k}^q)^T \in \mathbb{R}^k$ for $q = 1, \ldots, d$. Since the components of $\tilde{Y}_m$ are a permutation of the components of $Y_m$, there exists $Q \in \mathbb{R}^{dk \times dk}$ so that

$\tilde{Y}_{m+1} = Q^{-1}(I_k \otimes P^{-1})R(hU_N, hU_S)(I_k \otimes P)Q\tilde{Y}_m \equiv \tilde{R}(hU_N, hU_S)\tilde{Y}_m$.

For $q = 1, \ldots, d$, the fact that each $Y_{m+1,w}^q$ depends only on $\{Y_{m,w}^q\}_{w=1,\ldots,k}$ implies that $\tilde{Y}_{m+1,q}$ depends only on $\{\tilde{Y}_{m,d}^q\}_{d=q}^k$. It follows that $\tilde{R}(hU_N, hU_S)$ is a block upper triangular matrix of the following form:

\begin{equation}
\tilde{R}(hU_N, hU_S) = \begin{bmatrix}
\tilde{R}_{1,1} & \cdots & \tilde{R}_{1,d} \\
\vdots & \ddots & \vdots \\
\tilde{R}_{d,1} & \cdots & \tilde{R}_{d,d}
\end{bmatrix}, \quad \tilde{R}_{i,j} \in \mathbb{R}^{k \times k}, \quad i \leq j.
\end{equation}

Equation (15) and the fact that $U_N$ and $U_S$ are upper triangular imply that for $i = 1, \ldots, d$ each diagonal block $\tilde{R}_{i,i}$ is exactly the coefficient function of the method (3) applied to solve $\dot{z} = (U_N)_{i,i} z + (U_S)_{i,i} z$ with step-size $h > 0$. It follows that the set of eigenvalues $R(hN, hS)$ is the set of eigenvalues of $\tilde{R}(hU_N, hU_S)$ which is given by the union $\bigcup_{i=1}^d \text{eig}(R(h(U_N)_{i,i}, h(U_S)_{i,i}))$. $\square$

If the method (3) is a one-step method ($k = 1$), and $S$ and $N$ are ST then the matrix $\tilde{R}$ in Equation (16) is upper triangular and the (Lyapunov) stability of the zero solution of the linear system $X_{m+1} = R(hN, hS)X_m$ is determined (we ignore transient growth due to non-normality as discussed in [23]) by the moduli of the stability functions of $d$ complex-valued linear scalar test equations of the form (7)
where \( \lambda \in \text{eig}(N) \) and \( \mu \in \text{eig}(S) \). However this is not necessarily the case if \( k > 1 \) and \( N \) and \( S \) are ST but not SD. If \( S \) and \( N \) are SD, then \( R(hN, hS) \) is block diagonal (\( R_{i,j} = 0 \) if \( i \neq j \)) and therefore is power bounded if and only if each stability function \( R_{i,i} \) is power bounded for \( i = 1, \ldots, d \).

If \( N \) and \( S \) are normal and commuting, then \( N \) and \( S \) are SD and the conclusion of Proposition 5 holds, which for IMEX LMMs was noted in [18]. McCoy’s Theorem (see [32] or Theorem 1.3.4 of [35]) implies that \( N \) and \( S \) are ST if and only if \( p(N, S)(NS - SN) \) is nilpotent for every noncommutative polynomial \( p \) (a noncommutative polynomial \( p \) is one for which \( p(x, y) \neq p(y, x) \) for some pair of matrices \( x, y \)). Even for splittings \( A = N + S \) where \( N \) and \( S \) arise from the spatial discretizations of two separate variables of a PDE we cannot guarantee that \( N \) and \( S \) will be ST (see Equation (31) below). The following example shows that even if the moduli of the eigenvalues of \( R(h\lambda, h\mu) \) are strictly less than 1 for every pair \( (h\lambda, h\mu) \) where \( \lambda \in \text{eig}(S) \) and \( \mu \in \text{eig}(N) \) this does not imply that the eigenvalues of \( R(hN, hS) \) are bounded by 1 in modulus.

**Example 6.** Consider the IMEX Euler method \( x_{m+1} = x_m + hN(x_m, t_m) + hS(x_{m+1}, t_{m+1}) \) applied to solve the following two-dimensional linear ODE:

\[
\dot{x} = Ax = \begin{bmatrix} -4 & \alpha \\ \gamma & -\beta - 1 \end{bmatrix} x = \begin{bmatrix} -1 & \alpha \\ 0 & -1 \end{bmatrix} x + \begin{bmatrix} -3 & 0 \\ \gamma & -\beta \end{bmatrix} x \equiv Nx + Sx
\]

where \( \alpha, \beta > 0 \). Note that eigenvalues of \( A \) have real parts less than zero implying exponential stability of the origin for this ODE and also note that \( S \) and \( N \) are not uniformly triangularizable by McCoy’s Theorem since \((NS - SN)^2 \) is not nilpotent. The eigenvalue of \( N \) is \(-1 \) and the eigenvalues of \( S \) are \(-3 \) and \(-\beta \). This yields two potential test equations:

\[
\dot{z} = -1z - 3z, \quad \dot{z} = -1z - \beta z.
\]

Applying the IMEX Euler method with step-size \( h > 0 \) to solve the test equations with some initial condition \( z(0) = z_0 \in \mathbb{C} \) results in the following linear difference equations:

\[
z_{m+1} = \left( \frac{1 - h}{1 + 3h} \right) z_m, \quad z_{m+1} = \left( \frac{1 - h}{1 + h\beta} \right) z_m, \quad m \in \mathbb{N}.
\]

If we apply the IMEX Euler method with step-size \( h > 0 \) and some initial condition \( 0 \neq x_0 \in \mathbb{R}^2 \) to solve \( \dot{x} = Nx + Sx \) we obtain the following two-dimensional linear difference equation:

\[
x_{m+1} = \begin{bmatrix} \frac{1 - h}{1 + 3h} & \frac{\alpha h}{1 + 3h} \\ \frac{\gamma h^2(1-h)}{1 + 3h} & \frac{\alpha h^2}{1 + 3h} \end{bmatrix} x_m \equiv A(h; \alpha, \beta, \gamma) x_m, \quad m \in \mathbb{N}.
\]

If \( h = 2 \), then the coefficients of the test equations satisfy that \( |1 - h|/|1 + 3h|, |1 - h|/|1 + h\beta| \leq 1 \) for all \( \beta > 0 \) and hence they are stable for \( h = 2.0 \) and all \( \beta > 0 \). However, if \( \beta = \gamma = \alpha = 1 \), then \( A(2; 1, 1, 1) \) has an eigenvalue with modulus greater than 1.

For the remainder of this section fix a maximal step-size \( h_0 > 0 \) and maximal \( \delta_0 > 0 \) and assume that the following matrices arising from the internal state equations of (3) are always invertible for all \( h \in (0, h_0) \) and \( \delta \in (0, \delta_0) \):

\[
I_{dr} - C \otimes hN, \quad I_{dr} - \hat{C} \otimes hS, \quad I_{dr} - C \otimes hN - \hat{C} \otimes hS, \quad I_{dr} - \hat{C} \otimes h\delta^{-1}S, \quad I_{dr} - C \otimes hN - \hat{C} \otimes h\delta^{-1}S
\]
The following theorem gives two ways of characterizing the stability of \( R(hN, h\delta^{-1}S) \) in the stiff limit \( \delta \to 0 \): in terms of the product of the explicit and implicit stability matrices \( R(hN,0)R(0,h\delta^{-1}) \) and in terms of the implicit stability matrix \( R(0,h\delta^{-1}S) \).

**Theorem 7.** The following two conclusions hold:

1. Express \( R(hN,hS) = R(hN,0)R(0,hS) + \delta R \). Then \( \delta R = \delta R(hN,hS) \) is given by

\[
\delta R = [V \otimes I_d] - [V \otimes I_d]^2 + (D \otimes hN)[I_{dr} - C \otimes hN]^{-1}((U - UV) \otimes I_d) + ((\hat{D} - V\hat{D}) \otimes hS)[I_{dr} - \hat{C} \otimes hS]^{-1}(U \otimes I_d) + \mathcal{O}(h^2)
\]

If \( V^2 = V \) and \( \hat{C} \otimes S \neq 0 \), then \( \lim_{h_0, \delta \to 0} \delta R(hN,\delta^{-1}hS) = 0 \) and \( \delta R = \mathcal{O}(h^2) \) for all \( h \in (0,h_0] \).

2. Express \( R(hN,hS) = R(0,hS) + \delta R \). Then \( \delta R = \delta R(hN,hS) \) is given by

\[
\delta R(hN,hS) = [(D \otimes hN)[I_{dr} - C \otimes hN - \hat{C} \otimes hS]^{-1} + (\hat{D} \otimes hS)[I_{dr} - \hat{C} \otimes hS]^{-1}(D \otimes hN)[I_{dr} - C \otimes hN - \hat{C} \otimes hS]^{-1}](U \otimes I_d).
\]

If \( \hat{C} \otimes S \neq 0 \), then \( \lim_{h_0, \delta \to 0} \delta R(hN,\delta^{-1}hS) = 0 \) for all \( h \in (0,h_0] \).

**Proof.** We prove the first conclusion since the proof of the second is very similar. Using the definitions of \( R(hN,0) \), \( R(0,hS) \), and \( R(hN,hS) \) together with the properties of the Kronecker product discussed in Section 2 we find that:

\[
\delta R = R(hN,hS) - R(hN,0)R(0,hS)
\]

\[
= (V \otimes I_d) + (D \otimes hN + \hat{D} \otimes hS)[I_{dr} - C \otimes hN - \hat{C} \otimes hS]^{-1}(U \otimes I_d)
\]

\[
- (V \otimes I_d)^2 - (V \otimes I_d)(D \otimes hN)[I_{dr} \hat{C} \otimes hS]^{-1}(U \otimes I_d)
\]

\[
- (D \otimes hN)[I_{dr} - C \otimes hN]^{-1}(U \otimes I_d)(V \otimes I_d)
\]

\[
- (D \otimes hN)[I_{dr} - C \otimes hN]^{-1}(U \otimes I_d)(D \otimes hN)[I_{dr} \hat{C} \otimes hS]^{-1}(U \otimes I_d)
\]

The assumption that \( V = V^2 \) which implies that \( (V \otimes I_d) = (V^2 \otimes I_d) = (V \otimes I_d)^2 \) then implies that

\[
(18) \quad \delta R = h \cdot ((\hat{D} - V\hat{D}) \otimes S)[I_{dr} - C \otimes hN - \hat{C} \otimes hS]^{-1} - [I_{dr} - \hat{C} \otimes hS]^{-1}(U \otimes I_d)
\]

\[
+ h \cdot (D \otimes N)[I_{dr} - C \otimes hN - \hat{C} \otimes hS]^{-1} - [I_{dr} - C \otimes hN]^{-1}((U - UV) \otimes I_d)
\]

Therefore if \( \hat{C} \otimes S \neq 0 \), then it follows that \( \delta R(hN,h\delta^{-1}S) \to 0 \) as \( \delta \to \infty \) for all \( h \in (0,h_0] \). The fact that \( \delta R = \mathcal{O}(h^2) \) follows from the following relations and Equation (18):

\[
[I_{dr} - C \otimes hN - \hat{C} \otimes hN]^{-1} - [I_{dr} - \hat{C} \otimes hN]^{-1} = \mathcal{O}(h),
\]

\[
[I_{dr} - C \otimes hN - \hat{C} \otimes hN]^{-1} - [I_{dr} - \hat{C} \otimes hS]^{-1} = \mathcal{O}(h).
\]

The approach used in Theorem 7 can be refined using eigenvalue perturbation theory. Let \( P_S, P_N \in \mathbb{R}^{1 \times d} \) be invertible matrices such that \( U_S = P_S^{-1}SP_S \) and \( U_N = P_N^{-1}NP_N \) are both upper triangular. We then interpret \( P_S^{-1}AP_S \) as a perturbed upper triangular matrix:

\[
P_S^{-1}AP_S = U_N + U_S + \delta N, \quad \delta N := P_S^{-1}NP_S - U_N.
\]
Under the change of variables \( X_m = (I_k \otimes P_S)Z_m \) and \( G_m = (I_r \otimes P_S)K_m \) the method (3) applied to solve (9) with \( U_N + \delta N \) as the nonstiff term and \( U_S \) as the stiff term is transformed to the following system:

\[
\begin{align*}
K_m &= (U \otimes I_d)Z_m + (C \otimes (hU_N + h\delta N))K_m + (\hat{C} \otimes hU_S)K_m \\
Z_{m+1} &= (V \otimes I_d)Z_m + (D \otimes (hU_N + h\delta N))K_m + (\hat{D} \otimes hU_S)K_m.
\end{align*}
\]

(19)

Note that the stability properties of the zero solution of the systems (3) and (19) are equivalent since stability is preserved by a time-independent linear change of variables and also note that (19) is equivalent to applying the method (3) to solve \( \dot{z} = (U_N + \delta N)z + U_S z \) with step-size \( h > 0 \). The following lemma is a consequence of Equation (13).

**Lemma 8.** Let \( h \in (0, h_0] \) and \( \delta \in (0, \delta_0) \). With \( U_N, U_S, \) and \( \delta N \) defined as above we have \( R(hU_N + h\delta N, hU_S) = R(hU_N, hU_S) + \delta R \) where \( \delta R = \delta R(hU_N, hU_S, h\delta N) \) satisfies the following equation:

\[
\delta R = [(D \otimes h\delta N)[M + \delta M]^{-1} - (D \otimes hN + \hat{D} \otimes hS)[M + \delta M]^{-1}(\delta M)M^{-1}](U \otimes I_d)
\]

where \( M = I_{dr} - C \otimes hU_N - \hat{C} \otimes hU_S \) and \( \delta M = -C \otimes h\delta N \). If \( (\hat{C} \otimes S) \neq 0 \), then \( \delta R \to 0 \) as \( \delta_0 > \delta \to 0 \) for all \( h \in (0, h_0] \).

We now state and prove the main eigenvalue perturbation result.

**Theorem 9.** Let \( \{\varphi_i(h)\}_{i=1}^{dk} \) denote the eigenvalues of \( R(hN, h\delta^{-1}S) \) which are identical to the eigenvalues of \( R(hU_N + h\delta N, h\delta^{-1}U_S) \) and assume that \( \hat{C} \otimes S \neq 0 \). Then given \( \varepsilon > 0 \) there exists \( \delta_1 \in (0, \delta_0) \) so that if \( \delta \in (0, \delta_1) \), then for \( i = 1, \ldots, dk \) there exists an eigenvalue \( \psi_i(h) \) of \( R(hU_N, h\delta^{-1}U_S) \) so that \( |\varphi_i(h) - \psi_i(h)| < \varepsilon \) for all \( h \in [0, h_0] \).

Proof. For each \( h \in [0, h_0] \) and \( \delta \in (0, \delta_0) \) we form a Schur decomposition \( K(h, \delta)^*R(hU_N, h\delta^{-1}U_S)K(h, \delta) = D(h, \delta) + T(h, \delta) \) where \( D(h, \delta) \) is diagonal, \( T(h, \delta) \) is strictly upper triangular, and \( K(h, \delta) \) is unitary. There exists \( M > 0 \) so that \( \|T(h, \delta)\|_2 \leq M \) for all \( h \in (0, h_0] \) and \( \delta \in (0, \delta_0) \) where \( \| \cdot \|_2 \) denotes the Euclidean 2-norm. The Bauer-Fike Theorem for non-diagonalizable matrices (Theorem 7.2.3 of [19]) implies that for each \( i = 1, \ldots, dk \), there exists an eigenvalue \( \psi_i(h) \) of \( R(hU_N, h\delta^{-1}U_S) \) so that \( |\varphi_i(h) - \psi_i(h)| \leq \max\{\theta, \theta^{1/q}\} \) where \( \theta \) is defined as \( \theta := \|\delta R(hU_N, h\delta^{-1}U_S, h\delta N)\|_2 \sum_{k=1}^{\infty} \|T(h, \delta)\|_2^k \) and \( q \in \{1, \ldots, dk\} \), which may depend on \( h \) and \( \delta \), is the smallest positive integer such that the strictly upper triangular and hence nilpotent \( T(h, \delta) \) satisfies \( T(h, \delta)^q = 0 \). Lemma 8 together with the fact that \( \|T(h, \delta)\|_2 \leq M \) for all \( h \in (0, h_0] \) and \( \delta \in (0, \delta_0) \) then implies that given \( \varepsilon > 0 \), there exists \( \delta_1 \in (0, \delta_0) \) so that \( \max\{\theta, \theta^{1/q}\} \leq \max_{i=1, \ldots, dk} \theta^{1/l} < \varepsilon \) for \( l = 1, \ldots, dk \) whenever \( \delta \in (0, \delta_1) \) and \( h \in (0, h_0] \).

We now make two remarks related to Theorem 9 and the theory developed in this section.

**Remark 10.** How small \( \delta > 0 \) must be taken in Theorem 9 to satisfy a given eigenvalue perturbation bound \( \varepsilon > 0 \) depends on the size of \( h_0 \). This has implications for practical problems where the step-size will be chosen as large as possible subject to e.g. constraints on accuracy and \( \delta > 0 \) will typically be fixed, but small.

**Remark 11.** Rather than consider Equation (11) in the stiff limit we could consider the equation \( \dot{x} = \varepsilon N x + Sx \) and take the non-stiff limit \( \varepsilon \to 0 \). One can then prove a theorem analogous to Theorem 9 where the eigenvalues of \( R(\varepsilon hN, hS) \) are approximately those of \( R(\varepsilon hU_N, hU_S) \) for all \( \varepsilon > 0 \) sufficiently small. This fact is used in the analysis of the linear shallow water model in Section 5.2.
Tables of values of the measure $E$ for Equation (21) solved by various IMEX GLMs and values of $\delta > 0$ for $\beta = 1$ (left) and $\beta = 100$ (right) using $h = 1$. See Section 4 for references on where to find definitions of the methods (DIMSIM3b denotes IMEX-DIMSIM3b).

| $\delta$ | $\beta = 1$ | $\beta = 100$ |
|----------|-------------|--------------|
|          | ARS343 | DIMSIM3b | IMEX-BDF2 | ARS343 | DIMSIM3b | IMEX-BDF2 |
| 1        | 4.48E-2 | 1.41E-1 | 1.82E-1 | 1.58E0 | 5.62E0 | 2.85E0 |
| 1/10     | 3.29E-2 | 4.68E-2 | 8.91E-2 | 1.16E0 | 8.83E-1 | 1.12E0 |
| 1/100    | 4.93E-3 | 2.14E-2 | 2.35E-2 | 5.55E-2 | 1.55E-1 | 2.24E-1 |
| 1/1000   | 5.11E-4 | 9.35E-3 | 6.71E-3 | 9.12E-3 | 6.21E-2 | 5.63E-2 |

The following example shows that if $S$ has a nonzero eigenvalue that is non-stiff relative to the eigenvalues of $N$, then the conclusion of Theorem 9 is may not be true.

**Example 12.** Consider the system $x_{n+1} = A(h; \beta, \gamma, \alpha)x_n$ defined as in Equation (17). Note that the eigenvalues of $S$ are $-3$ and $-\beta$ so that when $\beta$ becomes large, $S$ has a single stiff eigenvalue $-\beta$ and a relatively nonstiff eigenvalue $-3$. If $h = 2$ and $\alpha = 1$ and $\beta = \gamma$, then

$$
\lim_{\beta \to \infty} A(2, \beta, \beta, 1) = \begin{bmatrix} -1/7 & 2/7 \\ -1/7 & -2 \end{bmatrix} := A_\infty.
$$

The eigenvalues of $A_\infty$ are approximately $-1.9778$ and $-0.1651$ and hence $A(2, \beta, \beta, 1)$ has an eigenvalue of modulus greater than $1$ for all sufficiently large $\beta$. Notice when $h = 2$ the coefficients of the scalar test equations limit to $(1 - h)/(1 + 3h) \to -1/7$ and $(1 - h)/(1 + 3h) \to 0$ as $\beta \to \infty$ and therefore do not approximate the eigenvalues of $A(2, \beta, \beta, 1)$ for $\beta$ sufficiently large.

We now define a quantity $E = E(N, S; h) = \psi(h)$ to measure how well test equations characterize the stability of an IMEX GLM solving the linear ODE (9) with $\{h \lambda_n\}_{n=1}^{d}$ denote the eigenvalues of $N$ and $\{\delta \mu_i\}_{i=1}^{d}$ denote the eigenvalues of $S$. We let $\{\delta \lambda_n\}_{n=1}^{d} \equiv \bigcup_{n=1}^{d} \text{eig}(A(h \lambda_n, h \mu_i))$ and $\{\delta \mu_i\}_{i=1}^{d}$ to measure the stability of an IMEX GLM solving an IVP of an ODE of the form $\dot{x} = N x + \delta^{-1} S x$ based on how stiff $\delta^{-1} S$ is relative to $N$.

**Example 13.** Consider the following equation:

$$
\dot{x} = Q_N \begin{bmatrix} -1 & \beta \\ 0 & -2 \end{bmatrix} Q_N x + Q_S \begin{bmatrix} -1 & -1 \\ 0 & -2 \end{bmatrix} Q_S x = N x + c S x, \quad \delta > 0, \beta \in \mathbb{R}
$$

where

$$
Q_N = \begin{bmatrix} \cos(\sqrt{2}/2) & -\sin(\sqrt{2}/2) \\ \sin(\sqrt{2}/2) & \cos(\sqrt{2}/2) \end{bmatrix}, \quad Q_S = \begin{bmatrix} \cos(\sqrt{3}/2) & -\sin(\sqrt{3}/2) \\ \sin(\sqrt{3}/2) & \cos(\sqrt{3}/2) \end{bmatrix}.
$$

Note that $N$ and $S$ are not uniformly triangulariable for any $\delta > 0$ and $\beta \in \mathbb{R}$ by McCoy’s Theorem. In Table 1 we plot values of $E(N, \delta^{-1} S; h)$ for $h = 1$ to measure how well test equations describe the stability of several IMEX GLMs applied to solve IVPs of $\dot{x} = N x + S x$. 

![Table 1](attachment:image.jpg)
The following Proposition, whose proof follows from the definitions of $E$, $U_S$, and $U_N$, shows how to use $E$ to characterize the stability of an IMEX GLM solving a linear ODE of the form (10).

**Proposition 14.** Let $E = E(h)$ be defined as above and define the following two subsets:

\[ C_+ := \{(z, w) \in \mathbb{C}^2 : |R(z, w)| \leq 1 + E\}, \quad C_- := \{(z, w) \in \mathbb{C}^2 : |R(z, w)| \leq 1 - E\}. \]

If $h > 0$ is such that $(h(U_N)_{j,j}, h(U_S)_{j,j}) \in C_-$ for $j = 1, \ldots, d$, then the eigenvalues of $R(hN, hS)$ are all bounded by 1 in modulus. If $h > 0$ is such that $(h(U_N)_{j,j}, h(U_S)_{j,j}) \notin C_+$ for at least one $j \in \{1, \ldots, d\}$, then $R(hN, hS)$ has at least one eigenvalue of modulus strictly greater than 1.

Proposition 14 is used in the analysis of a linear shallow water model in Section (5.2).

We close this section with the following remark. The conclusions of Theorem 9 and Proposition 14 are bounds on the eigenvalues of the coefficient matrix of an autonomous linear difference equation. While eigenvalues of this coefficient matrix are a useful tool for stability analysis, in the case where the coefficient matrix has eigenvalues of modulus equal to 1, further analysis may be necessary to determine stability or instability of the zero solution.

**4. Stability of scalar test equations.** In this section we analyze the stability of IMEX GLMs solving complex-valued scalar linear test equations whose use (by Theorem 9) in characterizing the stability of IMEX GLMs solving IVPs of linear ODEs of the form (10) is justified when $S$ is sufficiently stiff relative to $N$. In addition to Equation (7) we consider test equations for hyperbolic-parabolic problems:

\[ \dot{z} =inz + \mu z, \quad n \in \mathbb{R}, \mu \in \mathbb{C}^- \]

and also for hyperbolic-hyperbolic problems:

\[ \dot{z} =inz + isz, \quad n, s \in \mathbb{R}. \]

Test equations of the form (22) are considered in the study of spatial discretizations of stiff hyperbolic-parabolic PDEs [2, 18] where the eigenvalues of $N$ are pure imaginary and the eigenvalues of $S$ have non-positive real parts. Test equations of the form (23) were used in the analysis of IMEX RK methods [30, 31] and IMEX LMMs [17] solving spatially discretized hyperbolic PDEs such as those arising in models of geophysical fluid flow. The stability function for an IMEX GLM applied to solve (22) is:

\[ L(n, z) := V + [Din + \hat{D}z][I_r - Cin - \hat{C}z]^{-1}U, \quad n \in \mathbb{R}, z \in \mathbb{C}^- \]

and the stability function for an IMEX GLM method (3) applied to solve (23) is:

\[ M(n, s) := V + [Din + \hat{D}is][I_r - Cin - \hat{C}is]^{-1}U = L(n, is), \quad n, s \in \mathbb{R}. \]

We use the functions $L(n, z)$ and $M(n, s)$ to define stability regions of (3) based on the test equations (22) and (23). The HP-stability region is the following set:

\[ S_{HP} := \{(n, z) : n \in \mathbb{R}, z \in \mathbb{C} \text{ with } \text{Re}(z) \leq 0, \text{ and } L(n, z) \text{ is power bounded}\}, \]

For $s_0 \geq 0$ the HP($s_0$)-stability region is the following set:

\[ S_{HP}(s_0) = \{(n, z) : n \in \mathbb{R}, z \in \mathbb{C} \text{ with } \text{Re}(z) \leq -s_0, \text{ and } L(n, z) \text{ is power bounded}\}. \]
The H-stability region is the following set
\[ S_H := \{ (n, s) \in \mathbb{R} : M(n, s) \text{ is power bounded} \}. \]

Analogous to Definition 2 we can also define absolute H-, HP-, and \( HP(s_0) \)-stability regions based on when the coefficient matrices \( M(n, s) \) and \( L(n, z) \) have eigenvalues bounded by 1 in modulus.

The H-stability region is easy to plot and analyze relative to the HP-stability region since it is two dimensional rather than three dimensional. The following proposition shows how to use vertical lines in the H- and HP\((s_1)\)-stability region to map out subsets of the HP- and HP\((s_0)\)-stability regions when \( 0 \leq s_1 \leq s_0 \).

**Proposition 15.** For any \( n \in \mathbb{R} \) and \( s \geq 0 \) define the following set:
\[ l(n, s) := \{ (n, z) : n \in \mathbb{R}, \text{Re}(z) \leq -s \}. \]

Assume that \( L(n, z) \) is holomorphic for all \( z \in \mathbb{C}^- \). The following three conclusions hold:

1. \( l(n, 0) \subseteq S_{HP} \) if and only if \( M(n, s) \) is power bounded for all \( s \in \mathbb{R} \).
2. \( L(n, s_0 + i s) \) is power bounded for all \( s \in \mathbb{R} \) if and only if \( l(n, s_0) \in HP(s_0) \).
3. \( l(n, s_1) \in S_{HP}(s_1) \) implies that \( l(n, s_0) \in S_{HP}(s_1) \) whenever \( s_1 \leq s_0 \).

**Proof.** The maximum principal implies that the equality \( \sup_{z \in \mathbb{C}^-} \| L(n, z)^m \| = \sup_{s \in \mathbb{R}} \| M(n, s)^m \| \) holds for all \( m \in \mathbb{N} \) which then implies that \( l(n, 0) \subseteq S_{HP} \) if and only if \( (n, s) \in S_H \) for all \( s \in \mathbb{R} \) which proves the first conclusion. The second conclusion is similarly a consequence of the maximum principal. The third conclusion follows from the fact that the fact that \( S_0 \subseteq S_1 \) where \( S_q := \{ z : \text{Re}(z) \leq s_q \} \) for \( q = 0, 1 \).

Figure 1 highlights the utility of Proposition 15. It shows that e.g. \( l(n, s_0) \in S_{HP}(s_0) \) for the ARS443 method whenever \( n \in (0, 1.8) \) and \( s_0 \leq -0.1 \) and similar results for the IMEX-BDF3 and IMEX-DIMSIM3b methods (see [1] for definitions of the ARS methods, Tables 3-4 on page 137 of [4] for the IMEX-DIMSIM3b method, Figure 1 on page A1438 of [38] for the IMEX-DIMSIM4 method, Examples 2.1, 2.2, and 2.3 of [18] for respectively the CN-LF, IMEX-BDF2, and IMEX-Adams methods, and Equation 3.11 of [24] for the IMEX-BDF3 method.). In Section 5.1 we use the \( HP(s_0) \)-stability region to explain the difference in maximum stable step-size of several methods from a parametrized family of IMEX RK methods used for the time-integration of a spatially discrete advection-diffusion model.

Let \( I_{exp} \) denote the intersection of the stability region of the explicit method of (3) with the imaginary axis and let \( R_{exp} = -i \cdot I_{exp} := \{ -i \lambda : \lambda \in I_{exp} \} \). We can now repeat the definitions of implicit I- and A-stability for the test equations (22) and (23).

**Definition 16.** We say that (3) is implicitly A-stable for (22) (resp. I-stable for (23)) if \( L(n, z) \) is power bounded for all \( z \in \mathbb{C}^- \) whenever \( n \in R_{exp} \) (resp. \( M(n, s) \) is power bounded for all \( s \in \mathbb{R} \) whenever \( n \in R_{exp} \)).

The following proposition is a consequence of Proposition 15.

**Proposition 17.** If (3) is implicitly A-stable for (23), then (3) is implicitly A-stable for (22).

As shown in Section 2, despite the apparent attractiveness of implicit A- and I-stability, these concepts can exclude IMEX GLMs with otherwise good stability properties and performance (see Figures 2, 3, and 4 below and the IMEX-KG232a-c methods defined in Section 5.1).
Fig. 1. Plot of $H^p(s_0)$-stability region for ARS443 with $s_0 = 0.1$ (top left), IMEX-BDF3 with $s_0 = 0.5$ (top right), IMEX-DIMSIM3b with $s_0 = 0.1$ (bottom).

Unlike implicit A- and I-stability, stiff accuracy is a property that translates from the implicit to the IMEX setting. The following is a corollary of the second conclusion of Theorem 7.

**Corollary 4.1.** If (3) is an IMEX GLM whose implicit method is stiffly accurate, then $R(w, -\infty) = 0$ for every $w \in \mathbb{C}$.

The following theorem, which extends the results of the first conclusion of Theorem 7 to scalar test equations, shows that if the stability function of an IMEX RK method accurately resolves the exponential $e^{h\mu + h\lambda}$ when applied to solve the scalar linear test equation (7), then the stability function approximately decouples into a product of the stability functions of the explicit and implicit method.

**Theorem 18.** Consider an IMEX RK method with local truncation error of order $p \geq 1$ and $\lambda, \mu \in \mathbb{C}$. If $h > 0$ is such that the stability function $R(\cdot, \cdot)$ is holomorphic at $(z, w) = (h\lambda, h\mu)$, then $R(h\lambda, h\mu) = R(h\lambda, 0)R(0, h\mu) + O(h^{p+1})$.

**Proof.** Let $\mathcal{M}$ be an IMEX RK method with local truncation error of order $p \geq 1$ and let $\lambda, \mu \in \mathbb{C}$. If $h > 0$ is such that the stability function $R(w, z)$ of $\mathcal{M}$ is analytic
Fig. 2. H-stability (n vs. s) regions of two IMEX RK methods: ARS343 (left) and ARS443 (right). Shaded regions denote regions of stability while unshaded regions denote regions of instability.

at \((h\lambda, h\mu)\), then

\[
R(h\lambda, h\mu) = e^{h\lambda + h\mu} + \mathcal{O}(h^{p+1}).
\]  

(24)

For an IMEX RK method, if \(R(w, z)\) is holomorphic at \((w_0, z_0)\), then it is holomorphic at \((0, z_0)\) and \((w_0, 0)\). It therefore follows that the stability function \(R(w, z)\) is holomorphic at \((0, h\mu)\) and \((h\lambda, 0)\). This implies that the following relations hold:

\[
R(h\lambda, 0) = e^{h\lambda} + \mathcal{O}(h^{p+1}), \quad R(0, h\mu) = e^{h\mu} + \mathcal{O}(h^{p+1}).
\]  

(25)

Combining Equations (24) and (25) implies that

\[
R(h\lambda, h\mu) = R(h\lambda, 0)R(0, h\mu) + \mathcal{O}(h^{p+1}).
\]

Note that Theorem 18 is not a corollary of Theorem 7 since for dimension \(d > 1\) the Baker-Campbell-Hausdorff formula implies that for matrices \(N, S \in \mathbb{R}^{d \times d}\) we only have that \(e^{hN + hS} = e^{hN}e^{hS} + \mathcal{O}(h^2)\) unless \(N\) and \(S\) commute. We close this section by remarking that a result analogous to Theorem 18 holds for the underlying one-step method of an IMEX GLM with \(k \geq 2\), although the statement and proof of such a theorem is beyond the scope of this paper.

5. Applications. In this section we analyze two examples to highlight the theory developed in Sections 3 and 4. In Section 5.1 we present a linear advection-diffusion model with a stiff diffusion term. An SD IMEX splitting of the model is defined by letting the advection term be the explicitly treated term and the diffusion term be the implicitly treated term. A family of second order IMEX RK methods, the IMEX-KG232 family, is applied to solve IVPs of this model and their performance and stability properties can be explained entirely terms of the theory developed in Section 4 and their stability regions.

In Section 5.2 we present a linear shallow water model with a HEVI IMEX splitting that is not ST as a contrast to the SD splitting used in the advection-diffusion model. The performance and stability of methods of the IMEX-KG232 family solving
IVPs of the linear shallow water model can not be entirely explained in terms of their stability regions which in fact can give misleading predictions for the maximum stable step-size. The theory developed in Section 3, Theorem 9 and Proposition 14, are used to explain the performance and stability of the IMEX-KG232 family applied to solve IVPs of the linear shallow water model.

The maximum stable step-size $h_{\text{max}}$ used in Sections 5.1 and 5.2 is approximated as follows. We first compute the maximal value $h^*_{\text{max}}$ for which the eigenvalues of the stability matrix $R$ have all have modulus bounded by $1 + 10^{-8}$. The quantity $h_{\text{max}}$ is then computed from $h^*_{\text{max}}$ by decreasing $h^*_{\text{max}}$ until the set $\{R^j\}_{j=0}^{K}$ does not become unbounded as $K \rightarrow \infty$.

5.1. Linear advection-diffusion model. Consider the following scalar advection-diffusion PDE initial boundary value problem posed on the unit interval $[0, 1]$ with periodic boundary conditions:

\begin{equation}
\begin{cases}
    u_t = u_x + \alpha u_{xx}, & x \in (0, 1), \quad t > 0, \quad \alpha > 0 \\
    u(0, t) = u(1, t), & u(x, 0) = u_0(x).
\end{cases}
\end{equation}
Fig. 4. \(H\)-stability regions (\(n\ vs. \(s\)) of two IMEX GLMs: the IMEX-DIMSIM3B from page A1438 of [4] (left), and IMEX DIMSIM4 from page 137 of [38](right). Shaded regions denote regions of stability while unshaded regions denote regions of instability.

The diffusion term \(\alpha u_{xx}\) becomes stiff relative to the advective term \(u_x\) for \(\alpha >> 1\). For a given positive integer \(J\) we take a uniform mesh \(0 = x_0 < x_1 < \ldots < x_{J-1} = 1\), let \(u_j\) be the approximate value of \(u(x_j, t)\) for \(j = 0, \ldots, J-1\), and discretize (26) using centered finite differences \(u_x(x_j, t) \approx \frac{1}{\Delta x} (u_{j+1} - u_{j-1})\) and \(u_{xx}(x_j, t) \approx \frac{1}{(\Delta x)^2} (u_{j+1} + u_{j-1} - 2u_j)\) for \(j = 0, \ldots, J-1\) and \(\Delta x := x_1 - x_0\). This results in the following ODE system:

\[
\dot{u}_j = N_J u_j + S_J u_j, \quad u_J = (u_0, \ldots, u_J)^T
\]

where \(N_J, S_J \in \mathbb{R}^{J \times J}\) are defined as follows:

\[
N_J = \frac{1}{2\Delta x} \begin{bmatrix}
0 & 1 & 0 & \cdots & 0 & -1 \\
-1 & 0 & 1 & \cdots & 0 & \vdots \\
0 & \ddots & \ddots & \vdots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & -1 & 0 \\
1 & 0 & \cdots & 0 & -1
\end{bmatrix}, \quad S_J = \frac{\alpha}{(\Delta x)^2} \begin{bmatrix}
-2 & 1 & 0 & \cdots & 0 & 1 \\
1 & -2 & 1 & \cdots & 0 & \vdots \\
0 & \ddots & \ddots & \vdots & \vdots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & 1 & -2 & 1 \\
1 & 0 & \cdots & 0 & 1 & -2
\end{bmatrix}.
\]

Notice that \(N_J\) and \(S_J\) are normal and commuting and they are therefore SD. Specifically (see e.g. Section 3 of [2]), applying a discrete Fourier transform diagonalizes the system (27) to one of the form

\[
\dot{v}_j = \lambda_j v_j + \mu_j v_j, \quad j = 0, \ldots, J-1
\]

where

\[
\lambda_j = \frac{i}{\Delta x} \sin(2\pi j \Delta x), \quad \mu_j = \frac{2\alpha}{(\Delta x)^2} (\cos(2\pi j \Delta x) - 1), \quad j = 0, \ldots, J-1.
\]

Since the eigenvalues of \(N_J\) are pure imaginary and the eigenvalues of \(S_J\) are real and nonpositive it follows that the stability of an IMEX GLM applied to solve IVPs of (27) is determined by its HP-stability region. Furthermore, since \(\mu_j = 0\) if and only if \(\lambda_j = 0\) for \(j = 0, \ldots, J-1\), it follows that if there is some positive minimal step-size
The stability polynomial of the explicit method is the second order Kinnmark and Gray accurate method. We refer to this family of methods as IMEX-KG232 family since the arbitrary. Setting $\Delta t = 10^J$ of the HP the denote the explicit method of (30) as EX-KG232.

Now consider the following family of IMEX RK methods:

$$
\begin{array}{l}
\begin{array}{c|c|c}
0 & 1/2 & 1/2 \\
1/2 & 1/2 & 1 \\
1 & 1 & 1 \\
\end{array}
\end{array}
\begin{array}{l}
\begin{array}{c|c|c}
0 & \hat{a}_1 & \hat{d}_1 \\
\hat{a}_2 & \hat{d}_2 & 1 \\
& 1 & 1 \\
\end{array}
\end{array}
$$

The coefficients of the implicit method of (30) are chosen as follows. Let $\hat{d}_2 > 0$ be arbitrary. Setting $\hat{a}_2 = 1/2 - \hat{d}_2$ and $\hat{d}_1 = (\frac{1}{2} - \hat{d}_2)/(1 - \hat{d}_2)$ results in a second order accurate method. We refer to this family of methods as IMEX-KG232 family since the stability polynomial of the explicit method is the second order Kinnmark and Gray polynomial (see Table 1 of [20]) with optimal stability on the imaginary axis, they are second order, and there are three explicit stages and two implicit stages. We consider three L-stable methods from this family: IMEX-KG232a with $\hat{d}_1 = \frac{1}{2}(2 - \sqrt{2}) = \hat{d}_2$, IMEX-KG232b with $\hat{d}_1 = \frac{1}{2}(2 + \sqrt{2}) = \hat{d}_2$, and IMEX-KG232c with $\hat{d}_2 = 2, \hat{d}_1 = 1.5$. The denote the explicit method of (30) as EX-KG232.

We display (Table 2) the maximum stable step-size, $h_{\text{max}}$, for the IMEX-KG232a-c and EX-KG232 methods using $\alpha = 1000$ and $J = 10, 100, 1000$. In all tested cases $h_{\text{max}}$ for IMEX-KG232b is about 1.21 times larger than $h_{\text{max}}$ for IMEX-KG232c and about 5.83 times larger than $h_{\text{max}}$ for IMEX-232a. This can be explained from the $HP(s_0)$-stability regions of the IMEX-KG232 methods (Figure 5). In each case of $J = 10, 100, 1000$ we have $\mu_{\text{min}} > 3.8E4$ (defined in Equation (29)) and we can choose a minimal step-size of $\Delta t_0 = 8E - 4$ since IMEX-KG232a-c are each stable for $j = 10, 100, 1000$ and step-size $h = 8E - 4$. Therefore we use $s_0 = 30$ for a comparison of the $HP(s_0)$-stability regions. Let $l(n) := \{(i, n, z) : \text{Re}(z) \leq s_0\}$. Figure 5 shows that for $n \in \mathbb{R}$ the stability region of IMEX-KG232a contains vertical lines $l(n)$ whenever $|n| < 0.8$ and therefore by the second and third conclusions of Proposition 15 it contains the set $\{(i, n, z) : \text{Re}(z) \leq s_0, |n| < 0.8\}$. Similarly the stability region of IMEX-KG232b contains the set $\{(i, n, z) : \text{Re}(z) \leq s_0, |n| < 4.8\}$ and the stability region of IMEX-KG232c contains the set $\{(i, n, z) : \text{Re}(z) \leq s_0, |n| < 4.0\}$. Notice that $4.8/4 = 6$ and $4.8/4 = 1.2$ so that the ratio of the $h_{\text{max}}$ attained by IMEX-KG232b relative to IMEX-KG232a and IMEX-KG232c can be explained by taking the ratios of the maximum value of $n$ such that $l(n)$ is contained the stability region of the method.

**5.2. Linear shallow water model.** In this section we present a linearized shallow water model with a HEVI IMEX splitting. Scalar test equations of the form
Fig. 5. H-stability (left column) and $HP(s_0)$-stability (right column) regions of the IMEX-KG232a (top row), IMEX-KG232b (middle row), and IMEX-KG232c (bottom row) methods with $s_0 = 30$. Shaded regions denote regions of stability while unshaded regions denote regions of instability.

\[(23)\] where $\lambda$ is an eigenvalue of the nonstiff matrix and $\mu$ is an eigenvalue of the stiff matrix are approximately (and not exactly!) eigenvalues of the stability matrix of an IMEX GLM solving an IVP of this model when the vertical resolution is much finer than the horizontal resolution. The stability regions of IMEX-KG232a-c family do not make as accurate of a prediction of the maximum stable step-size as in the SD example in Section 5.1. However, the stability behavior of these methods can still be characterized using the perturbation theory developed in Section 3.
Consider the non-rotating shallow water equations linearized around an equilibrium \((u_0, v_0, h_0)^T = (0, 0, h_0)^T\) posed in a channel of horizontal length \(K\) and height \(L\) with periodic horizontal boundary conditions and with \(u, v\) and \(p\) vanishing at the vertical boundary:

\[
\begin{align*}
\begin{aligned}
u_t + gp_x &= 0, & \quad (x, y) \in [0, K] \times [O, L], & \quad K, L > 0 \\
v_t + gp_y &= 0, \\
p_x + h_0(u_x + v_x) &= 0 \\
u(0, y, t) &= u(L, y, t), \\
v(0, y, t) &= v(L, y, t), \\
\quad (x, y) \in [0, K] \times [O, L], \quad K, L > 0 \\
p(0, y, t) &= p(L, y, t) \\
u(x, 0, t) &= 0 = u(x, K, t), \\
v(x, 0, t) &= 0 = v(x, K, t), \\
\quad (x, y) \in [0, K] \times [O, L], \quad K, L > 0 \\
p(x, 0, t) &= p(x, K, t).
\end{aligned}
\end{align*}
\]

The constant \(g > 0\) is the gravitational constant and \(h_0 > 0\) is some initial height. Choose uniform meshes \(0 = x_1 < x_2 < \ldots < x_I = L\) and \(0 = y_1 < y_2 < \ldots < y_J = K\) where \(\Delta x \equiv x_{i+1} - x_i\) and \(\Delta y \equiv y_{j+1} - y_j\) are constant for \(i = 1, \ldots, I\) and \(j = 1, \ldots, J\) and let \(u_{i,j} \approx u(x_i, y_j)\), \(v_{i,j} \approx v(x_i, y_j)\), and \(p_{i,j} \approx p(x_i, y_j)\). Equation (31) is discretized in space using centered finite differences e.g.

\[
\begin{align*}
\begin{aligned}
u_t &\equiv u(0, y, t), \\
v_t &\equiv v(0, y, t), \\
p(0, y, t) &\equiv p(0, y, t) \\
u(x, 0, t) &\equiv u(x, K, t), \\
v(x, 0, t) &\equiv v(x, K, t), \\
p(x, 0, t) &\equiv p(x, K, t).
\end{aligned}
\end{align*}
\]

This leads to the following 3\(IJ\)-dimensional linear ODE:

\[
\begin{align*}
\begin{aligned}
u_t + gp_x &= 0, & \quad (x, y) \in [0, K] \times [O, L], & \quad K, L > 0 \\
v_t + gp_y &= 0, \\
p_x + h_0(u_x + v_x) &= 0 \\
u(0, y, t) &= u(L, y, t), \\
v(0, y, t) &= v(L, y, t), \\
\quad (x, y) \in [0, K] \times [O, L], \quad K, L > 0 \\
p(0, y, t) &= p(L, y, t) \\
u(x, 0, t) &= 0 = u(x, K, t), \\
v(x, 0, t) &= 0 = v(x, K, t), \\
\quad (x, y) \in [0, K] \times [O, L], \quad K, L > 0 \\
p(x, 0, t) &= p(x, K, t).
\end{aligned}
\end{align*}
\]

where \(A_{IJ} := \tilde{D}_J \otimes I_I, B_{IJ} := I_J \otimes D_I, \) and \(\tilde{D}_J \in \mathbb{R}^{J \times J}, D_I \in \mathbb{R}^{I \times I}\) are defined by the following equations:

\[
\tilde{D} := \frac{K}{2\Delta x} \begin{bmatrix}
0 & -1 & 0 & \ldots & 0 & 1 \\
1 & 0 & -1 & 0 & \ldots & \ldots \\
0 & 1 & 0 & -1 & \ldots & \ldots \\
\vdots & \vdots & \vdots & \ddots & \ddots & \ddots \\
0 & 1 & 0 & -1 & \ldots & \ldots \\
-1 & 0 & \ldots & 0 & 1 & 0
\end{bmatrix}
\]

\[
D := \frac{L}{2\Delta y} \begin{bmatrix}
0 & -1 & 0 \\
1 & 0 & -1 \\
\vdots & \vdots & \vdots \\
1 & 0 & -1 \\
0 & 1 & 0
\end{bmatrix}
\]

We use the following IMEX splitting:

\[
M_{IJ} = \begin{bmatrix}
0 & 0 & \frac{gA_{IJ}}{h_0} \\
0 & 0 & \frac{gB_{IJ}}{h_0} \\
\frac{h_0A_{IJ}}{h_0} & 0 & 0
\end{bmatrix} + \begin{bmatrix}
0 & 0 & \frac{gA_{IJ}}{h_0} \\
0 & 0 & \frac{gB_{IJ}}{h_0} \\
\frac{h_0A_{IJ}}{h_0} & 0 & 0
\end{bmatrix} \equiv N_{IJ} + S_{IJ}.
\]

To mimic the geometry of more complex atmosphere and ocean models, where the thinness of the fluid height relative to its horizontal surface area causes the vertical resolution to be much higher than the spatial resolution, we choose \(K \geq L\). McCoy’s Theorem implies that \(N_{IJ}\) and \(S_{IJ}\) do not satisfy the hypotheses of Proposition 5. However, Remark 11 and Theorem 9 together with the fact that \(N_{IJ}\)
and $S_{I,J}$ each have pure imaginary eigenvalues, imply that for fixed height $L$ the hyperbolic-hyperbolic test equations (23) asymptotically approximate the eigenvalues of $R(hN_{I,J}, hS_{I,J})$ as $K \to \infty$ (See Figure 6).

In Table 3 we show the maximum stable time-step $h_{\text{max}}$ achieved by the methods IMEX-KG232a-c applied to solve IVPs of (32) and give the approximate value of $\mathcal{E}(h)$ (defined in Section 3) at $h = h_{\text{max}}$. As expected based on Table 2 and the analysis in Section 5.1, $h_{\text{max}}$ for the IMEX-KG232a method is much smaller than $h_{\text{max}}$ of the IMEX-KG232b-c methods. However, notice that $h_{\text{max}}$ is larger for IMEX-KG232c than for IMEX-KG232b which is the opposite of what one would expect based on their stability regions (see Table 2) and the analysis in Section 5.1. Thus, in the non-ST case, the stability region is not an exact predictor of what method will be able to take the largest stable time-steps, even if the differential equation is linear.

We now explain this counter-intuitive stability phenomenon using the theory developed in Section 3. Take $S = S_{I,J}$ and $N = N_{I,J}$ and let $P_S, P_N \in \mathbb{C}^{3IJ}$ be invertible matrices so that $U_S := P_S^{-1}SP_S$ and $U_N := P_N^{-1}NP_N$ are upper triangular and note that the eigenvalues of $N$ and $S$ are pure imaginary. Define the following sets:

$$E_{S,N} := \{(-i(U_N)_{j,j}, -i(U_S)_{j,j}) : j = 1, \ldots, 3IJ, i = \sqrt{-1}\}$$

$$h_{\text{max}}E_{N,S} := \{(h_{\text{max}}n, h_{\text{max}}s) : (n, s) \in E_{N,S}\}.$$  

The set $E_{S,N}$ represents all the pairs of coefficients of the hyperbolic-hyperbolic scalar test equations (23) of the ODE (32). Define the following three sets (note that $M(n, s)$ is a complex scalar for an IMEX RK method):

$$C_- := \{(n, s) : |M(n, s)| \leq 1 - \mathcal{E}(h_{\text{max}})\}, \quad C_+ := \{(n, s) : |M(n, s)| \leq 1 + \mathcal{E}(h_{\text{max}})\},$$

$$C_0 := \{(n, s) : |M(n, s)| \leq 1\} = S_{HP}.$$  

Proposition 14 implies that if $h_{\text{max}}E_{S,N}$ is contained in $C_-$, then the eigenvalues of $R(h_{\text{max}}N, h_{\text{max}}S)$ are all bounded by 1 in modulus and on the other hand, that if $h_{\text{max}}E_{N,S}$ is not contained in $C_+$, then $R(h_{\text{max}}N, h_{\text{max}}S)$ must have an eigenvalue of modulus greater than 1. Theorem 9 and Remark 11 imply if we restrict $h$ so that $h \in [0, h_0]$ for some $h_0 > 0$, then as $K \to 0$ we have $\mathcal{E}(h) \to 0$ for all $h \in [0, h_0]$. However, since we are seeking out the maximal stable time-step $h_{\text{max}}$, for a given fixed $K > 0$ we cannot expect that $\mathcal{E}(h_{\text{max}}) \approx 0$, although empirically (see Table 3) the values of $\mathcal{E}(h_{\text{max}})$ are typically not very large at $h = h_{\text{max}}$. Table 3 shows that $\mathcal{E}(h_{\text{max}}) < 2$ in all experiments and $\mathcal{E}(h_{\text{max}}) < 0.5$ for all experiments with IMEX-KG232b-c.

Figure 7 displays contour plots of the H-stability function $M(n, s)$ of the IMEX-KG232a-c methods on the contour lines where $|M(n, s)| = 1, 1 - \mathcal{E}(h_{\text{max}}), 1 + \mathcal{E}(h_{\text{max}})$ (denoted as $\partial C_0$, $\partial C_-$, and $\partial C_+$ respectively) together with a point plot of the set $h_{\text{max}}E_{S,N}$. Notice that in each case the set $h_{\text{max}}E_{S,N}$ is contained in $C_+$ as predicted by Proposition 14. For IMEX-KG232b-c, the set $C_-$ slightly underestimates the size of $h_{\text{max}}$ and $h_{\text{max}}E_{N,S}$ is contained in H-stability region for both methods. For IMEX-KG232a, the set $C_-$ substantially underestimates $h_{\text{max}}$ and the method performs much better than expected. Tables 3 and Figure 6 show that $\mathcal{E}(h_{\text{max}})$ is larger for IMEX-KG232a than for IMEX-KG232b-c which results in a low estimate of $h_{\text{max}}$ based on its stability region. Notice that in each case the maximal step-size for the explicit method EX-KG232 is significantly smaller than any of the tested IMEX-KG232 methods including IMEX-KG232a.
Fig. 6. Plots of $\mathcal{E}(h)$ vs. $K$ for various IMEX RK, IMEX LMM, and IMEX GLMs, using $L = 1$, $h = 1$, $g = 0.981$, $h_0 = 0.1$ and $I = J = 10$.

Table 3

| Method         | $h_{\text{max}}$ | $\mathcal{E}(h_{\text{max}})$ | $h_{\text{max}}$ | $\mathcal{E}(h_{\text{max}})$ | $h_{\text{max}}$ | $\mathcal{E}(h_{\text{max}})$ |
|----------------|-------------------|-------------------------------|-------------------|-------------------------------|-------------------|-------------------------------|
| EX-KG232       | 5.25E-1           | N/A                           | 7.3E-1            | N/A                           | 7.3E-1            | N/A                           |
| IMEX-KG232a    | 7.61E-1           | 1.93E0                        | 4.06E1            | 7.41E-1                       | 4.04E3            | 8.07E-1                       |
| IMEX-KG232b    | 7.62E-1           | 3.99E-1                       | 9.75E1            | 4.59E-1                       | 9.75E3            | 4.58E-1                       |
| IMEX-KG232c    | 7.62E-1           | 3.68E-1                       | 1.06E2            | 4.37E-1                       | 1.06E4            | 4.39E-1                       |

6. Afterward. The stability of a $k$-step and $r$-stage IMEX GLM solving an autonomous linear ODE of dimension $d$ can often be determined by a family of $dk$ eigenvalues parametrized by the time-step. We have used perturbation theory to approximate these eigenvalues in terms of the stability functions of the solution of complex-valued linear scalar test equations whose coefficients are eigenvalues of the coefficient matrices of the IMEX splitting. The perturbation bound can be made arbitrarily small by making the stiff term in the IMEX splitting sufficiently stiff rel-
Fig. 7. Plot of $h_{\text{max}}E_{N,S}$ (represented as dots) together with the contour lines $\partial C_-$ (labeled A), $\partial C_0$ (labeled B) and $\partial C_+$ (labeled C). Values of $h_{\text{max}}$ and $E(h_{\text{max}})$ are found in Table 3.

ative to the nonstiff term. The theory we have developed is able to explain why the maximal time-step achieved by the IMEX-KG232a-c methods solving a linear shallow water model whose IMEX splitting is not ST is not what would be predicted from the stability regions of these methods.

There are several avenues for further investigation of the stability of IMEX GLMs solving IVPs whose IMEX splitting is not ST. The first avenue is getting better estimates of $E(h)$ and characterizing its dependence on method coefficients. This would allow the development of methods whose stability regions better predict the maximum stable time-step for linear IMEX splittings that are not ST. Another avenue is on developing a nonlinear and time-dependent stability theory that is not restricted only to autonomous linear IVPs. The third apparent area for additional exploration is the stability of more general partitioned methods.

REFERENCES

[1] Ascher, U., Ruuth, S., and Spiteri, R., Implicit-explicit Runge-Kutta methods for time-
dependent partial differential equations, Appl. Numer. Math., 25 (1997), pp. 151–167, https://doi.org/10.1016/S0168-9274(97)00056-1.

2. Ascher, U., Ruuth, S., and Wetton, B., Implicit-explicit methods for time-dependent partial differential equations, SIAM J. Numer. Anal., 32 (1995), pp. 797–823, https://doi.org/10.1137/0732037.

3. Bader, G., Hairer, E., and Lubich, C., On the stability of semi-implicit methods for ordinary differential equations, BIT, 22 (1982), pp. 211–222, https://doi.org/10.1007/BF01944478.

4. Blaise, S., Sandu, A., and Zhang, H., Partitioned and implicit-explicit general linear methods for ordinary differential equations, J. Sci. Comput., 61 (2014), pp. 119–144, https://doi.org/10.1007/s10915-014-9819-z.

5. Brá, M., Cardone, A., Jackiewicz, Z., and Pierzchalla, P., Error propagation for implicit-explicit general linear methods, Appl. Numer. Math., 131 (2018), pp. 207–231, https://doi.org/10.1016/j.apnum.2018.05.004.

6. Burrage, K. and Butcher, J., Non-linear stability of a general class of differential equation methods, BIT, 20 (1980), pp. 185–203, https://doi.org/10.1007/BF01933191.

7. Burrage, K. and Butcher, J.C., Stability criteria for implicit Runge-Kutta methods, SIAM J. Numer. Anal., 16 (1979), pp. 46–57, https://doi.org/10.1137/0716004.

8. Butcher, J. and Hill, A., Linear multistep methods as irreducible general linear methods, BIT, 46 (2006), pp. 5–19, https://doi.org/10.1007/s10543-006-0046-3.

9. Butcher, J.C., A stability property of implicit Runge-Kutta methods, BIT, 15 (1975), pp. 358–361, https://doi.org/10.1007/BF01931672.

10. Butcher, J.C., The equivalence of algebraic stability and AN-stability, BIT, 27 (1987), pp. 510–533, https://doi.org/10.1007/BF01937275.

11. Calvo, M., Jay, L., and Söderlind, G., Stiffness 1952-2012: Sixty years in search of a definition, BIT, 55 (2015), pp. 531–558, https://doi.org/10.1007/s10543-014-0503-3.

12. Crouzeix, M., Une méthode multipas implicite-explicite pour l’approximation des équations d’évolution paraboliques, Numer. Math., 35 (1980), pp. 257–276, https://doi.org/10.1007/BF01396412.

13. Curtiss, C. and Hirschfelder, J., Integration of stiff equations, Proc. Natl. Acad. Sci. USA, 38 (1952), pp. 235–243.

14. Dahlquist, G., Convergence and stability in the numerical integration of ordinary differential equations, Math. Scand., 4 (1956), pp. 33–53, https://doi.org/10.7146/math.scand.a-10454.

15. Dahlquist, G., Stability and error bounds in the numerical integration of ordinary differential equations, Trans. Royal Inst. Technol., Stockholm, Sweden, Nr. 130 (1959), p. 87, https://doi.org/10.1002/zamm.19610410611.

16. Dahlquist, G., A special stability problem for linear multistep methods, BIT, 3 (1963), pp. 27–43, https://doi.org/10.1007/BF01963532.

17. Durran, D. and Blossey, P., Implicit-explicit multistep methods for fast-wave-slow-wave problems, Mon. Wea. Rev., 140 (2012), pp. 1307–1325, https://doi.org/10.1175/MWR-D-11-00088.1.

18. Frank, J., Hundsdorfer, W., and Verwer, J., On the stability of implicit-explicit linear multistep methods, Appl. Numer. Math., 25 (1997), pp. 193–205, https://doi.org/10.1016/S0168-9274(97)00055-7.

19. Golub, G. and Van Loan, C., Matrix computations. Fourth edition., Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, MD, 2013.

20. Gray, W. and Kinnmark, I., One step integration methods with maximum stability regions, Math. Comput. Simulation, XXVI (1984), pp. 84–92, https://doi.org/10.1016/0378-4719(84)90039-9.

21. Hairer, E., Order conditions for numerical methods for partitioned ordinary differential equations, Numer. Math., 36 (1981), pp. 431–445, https://doi.org/10.1007/BF01395056.

22. Herity, M., Pareschi, L., and Steffensen, S., Implicit-explicit Runge-Kutta schemes for numerical discretization of optimal control problems, SIAM J. Numer. Anal., 51 (2013), pp. 1875–1899, https://doi.org/10.1137/12086504.

23. Higham, D. and Trefethen, L., Stiffness of ODEs, BIT, 33 (1993), pp. 285–303, https://doi.org/10.1007/BF01989751.

24. Hundsdorfer, W. and Ruuth, S., IMEX extensions of linear multistep methods with general monotonicity and boundedness properties, J. Comp. Physics, 225 (2007), pp. 2016–2042, https://doi.org/10.1016/j.jcp.2007.03.003.

25. Jackiewicz, Z., General linear methods for ordinary differential equations, vol. 146, John Wiley & Sons Inc., Hoboken, N.J., USA, 2009.

26. Jackiewicz, Z., Cardone, A., Sandu, A., and Zhang, H., Construction of highly stable
implicit-explicit general linear method, Dynamical Systems and Differential Equations, AIMS Proceedings 2015, Proceedings of the 10th AIMS International Conference (Madrid, Spain), (2015), pp. 185–194, https://doi.org/10.3934/proc.2015.0185.

27 JACkiewicz, Z. AND Izzo, G., Highly stable implicit-explicit Runge-Kutta methods, Appl. Numer. Math., 113 (2017), pp. 71–92, https://doi.org/10.1016/j.apnum.2016.10.018.

28 Jackiewicz, Z. AND Vermiglio, R., Order conditions for partitioned runge-kutta methods, Applications of Mathematics, 45 (2000), pp. 301–316, https://doi.org/10.1023/A:1022323529349.

29 Kato, T., Stability of implicit-explicit linear multistep methods for ordinary and delay differential equations, Front. Math. China, 113 (2009), pp. 113–129, https://doi.org/10.1007/s11464-009-0005-9.

30 Lock, S.-J., Weller, H., and Wood, N., Runge-Kutta IMEX schemes for the horizontally explicit/vertically implicit (HEVI) solution of wave equations, J. Comput. Phys., 252 (2013), pp. 365–381, https://doi.org/10.1016/j.jcp.2013.06.025.

31 Lock, S.-J., Weller, H., and Wood, N., Numerical analyses of Runge-Kutta implicit-explicit schemes for horizontally explicit, vertically implicit solutions of atmospheric models, Q.J.R. Meteorol. Soc., 140 (2014), pp. 1654–1669, https://doi.org/10.1002/qj.2246.

32 Mc Cov, N., On the characteristic roots of matrix polynomials, Bull. Amer. Math. Soc., 42 (1936), pp. 592–600.

33 Pareschi, L AND Russo, L., Implicit-explicit Runge-Kutta schemes and applications to hyperbolic systems with relaxation, J. Sci. Comput., 25 (2005), pp. 129–155, https://doi.org/10.1007/s10915-004-4636-4.

34 Pareschi, L, Russo, L., and Boscarino, S., A unified IMEX Runge-Kutta approach for hyperbolic systems with multscale relaxation, SIAM J. Numer. Anal., 55 (2017), pp. 2085–2109, https://doi.org/10.1137/M1111449.

35 Rajavi, H. and Rosenthal, P., Simultaneous Triangularization, Springer-Verlag New York, 2000, https://doi.org/10.1007/978-1-4612-1200-3.

36 Rosales, R., Seibold, B., Shirokoff, D., and Zhou, D., Unconditional stability for multistep ImEx schemes: Theory, SIAM J. Numer. Anal., 55 (2017), pp. 2336–2360, https://doi.org/10.1137/16M1094324.

37 Russo, L. and Boscarino, S., Flux-explicit IMEX Runge-Kutta schemes for hyperbolic to parabolic relaxation problems, SIAM J. Numer. Anal., 51 (2013), pp. 163–190, https://doi.org/10.1137/110850803.

38 Sandu, A., Zhang, H., and Blaise, S., High order implicit-explicit general linear methods with optimized stability regions, SIAM J. Sci. Comput., 38 (2016), pp. A1430–A1453, https://doi.org/10.1137/15M1018897.

39 Satoh, M., Conservative scheme for the compressible nonhydrostatic models with the horizontally explicit and vertically implicit time integration scheme, Mon. Wea. Rev., 19 (2002), pp. 1227–1245, https://doi.org/10.1175/1520-0493(2002)130⟨1227:CSFTCN⟩2.0.CO;2.

40 Steyer, A. and Van Vleck, E.S., A Lyapunov and Sacker-Sell spectral stability theory for one-step methods, BIT, (2018), https://doi.org/10.1007/s10543-018-0704-2.

41 Steyer, A. and Van Vleck, E.S., Underlying one-step methods and nonautonomous stability of general linear methods, Discrete Contin. Dyn. Syst. Ser. B, (2018), https://doi.org/10.3934/dcdsb.20181108.

42 Strang, G., On the construction and comparison of difference schemes, SIAM J. Numer. Anal., 5 (1968), pp. 506–517, https://doi.org/10.1137/0705041.

43 Varah, J., Stability restrictions on second-order, three level finite difference schemes for parabolic equations, SIAM J. Numer. Anal., 32 (1980), pp. 300–309, https://doi.org/10.1137/0717025.