

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/230865813>

Implicit–Explicit Runge–Kutta schemes for stiff systems of differential equations

Chapter · January 2001

CITATIONS

76

READS

1,178

2 authors:



Lorenzo Pareschi

University of Ferrara

172 PUBLICATIONS 3,731 CITATIONS

[SEE PROFILE](#)



Giovanni Russo

University of Catania

252 PUBLICATIONS 4,294 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Lattice Boltzmann method for hydrodynamics [View project](#)



Kinetic and Fokker-Planck equations modeling socio-economic phenomena in multi-agent systems [View project](#)

IMPLICIT-EXPLICIT RUNGE-KUTTA SCHEMES FOR STIFF SYSTEMS OF DIFFERENTIAL EQUATIONS

LORENZO PARESCHI* AND GIOVANNI RUSSO†

Abstract. We present new implicit-explicit (IMEX) Runge Kutta methods suitable for time dependent partial differential systems which contain stiff and non stiff terms (i.e. convection-diffusion problems, hyperbolic systems with relaxation). Here we restrict to diagonally implicit schemes and emphasize the relation with splitting schemes and asymptotic preserving schemes. Accuracy and stability properties of these schemes are studied both analytically and numerically.

Key words. Runge-Kutta methods, splitting schemes, stiff systems.

AMS subject classifications. 65C20, 82D25

1. Introduction. Many stiff systems of differential equations can be written in the form

$$(1) \quad y' = f(y) + \frac{1}{\varepsilon}g(y),$$

where $y = y(t) \in R^N$, $f, g : R^N \rightarrow R^N$ and $\varepsilon > 0$ is the stiffness parameter.

System (1) may represent a system of N Ode's or a discretization, obtained by the method of lines, of a system of Pde's, such as, for example, convection-diffusion equations, discrete velocity models in kinetic equations, hyperbolic systems with relaxation.

In this work we consider the latter case. These systems have a special structure, that allows us to rewrite them into two coupled subsystems as

$$(2) \quad \begin{aligned} u' &= f_1(u, v), \\ v' &= f_2(u, v) + \frac{1}{\varepsilon}g_2(u, v), \end{aligned}$$

where $y = (u, v)$, $u = u(t) \in R^m$, $v = v(t) \in R^{N-m}$, $m < N$, $f = (f_1, f_2)$, and $g = (0, g_2)$.

In the stiff limit $\varepsilon \rightarrow 0$, system (2) reduces to a differential-algebraic system. We shall assume that the algebraic equation $g_2(u, v) = 0$, can be uniquely solved in v , giving

$$(3) \quad v = e(u),$$

and therefore system (2) reduces to

$$(4) \quad u' = f_1(u, e(u)) \equiv F(u).$$

A very general and commonly used approach to the solution of this problem is based on splitting methods. A simple splitting of system (2) consists in solving separately the non-stiff problem

$$(5) \quad \begin{aligned} u' &= f_1(u, v), \\ v' &= f_2(u, v), \end{aligned}$$

*Department of Mathematics, University of Ferrara, via Machiavelli 35, I-44100 Ferrara, ITALY. pareschi@dm.unife.it

†Department of Mathematics, University of L'Aquila L'Aquila, Italy russo@univaq.it

applying an explicit scheme and, using an implicit scheme, the stiff problem

$$(6) \quad \begin{aligned} u' &= 0, \\ v' &= \frac{1}{\varepsilon} g_2(u, v). \end{aligned}$$

This splitting is restricted to first order accuracy in time, nevertheless its simple structure presents several advantages:

- Some properties of the solution are easily maintained (positivity, TVD property, other physically relevant properties).
- Consistency with the stiff limit (4) as $\varepsilon \rightarrow 0$ can be easily checked (asymptotic preservation).
- In many cases the implicit scheme for g can be explicitly solved thanks to the fact that u is unchanged during this step. This is of paramount importance in many applications since it avoids the use of iterative solvers.

Higher order splitting can be constructed using suitable combinations of the two previous steps [23, 8, 9]. Unfortunately all these higher order extensions present a severe loss of accuracy when the g term is stiff [15]. Second order Runge-Kutta splitting which maintain the accuracy in the stiff limit have been constructed recently [15, 5]. There is also a third-order scheme that uses splitting and extrapolation but it is rather inefficient [6].

The aim of this paper is to have a unified approach of splitting schemes, and to provide a framework for the derivation of more general, accurate and efficient schemes. In particular, we show that these schemes are strictly related with the recently developed IMEX Runge-Kutta schemes [1]. An IMEX scheme consists of applying an implicit discretization for g and an explicit one for f . Since systems of the form (1) arise from a Pde in more than one space variable, the simplicity and efficiency of solving the algebraic equations corresponding to the implicit part of the discretization at each step is of fundamental importance. To this aim, it is natural to consider diagonally implicit Runge-Kutta (DIRK) schemes for g .

We show that most of the splitting schemes can be written in the formalism of IMEX Runge-Kutta schemes, where the solver for g is a DIRK scheme. Similarly it is easy to write an IMEX Runge-Kutta scheme in splitting form. In particular, we show that the asymptotic preserving requirement [18], i.e. the consistency of the scheme with (4), is satisfied every time we use an L -stable scheme for g and that the implicit step can be solved, in many cases, every time we use a DIRK scheme for g . Thanks to the IMEX Runge-Kutta formalism we derive new second and third order schemes which maintain the order in the stiff limit.

The rest of the paper is organized as follows. In section 2 we introduce the class of IMEX Runge-Kutta schemes under study, and emphasize the analogy with splitting schemes. Section 3 is devoted to second order schemes which maintain the order of accuracy in the limit. Accuracy and stability analysis are also presented. In section 4 we present some third order extension of the schemes. Finally in section 5 we investigate the performance of the schemes for a simple test equation.

2. IMEX Runge-Kutta schemes. An Implicit-Explicit Runge-Kutta scheme for system (1) is of the form

$$(1) \quad Y_i = y_0 + h \sum_{j=1}^{i-1} \tilde{a}_{ij} f(t_0 + \tilde{c}_j h, Y_j) + h \sum_{j=1}^{\nu} a_{ij} \frac{1}{\varepsilon} g(t_0 + c_j h, Y_j),$$

$$(2) \quad y_1 = y_0 + h \sum_{i=1}^{\nu} \tilde{w}_i f(t_0 + \tilde{c}_i h, Y_i) + h \sum_{i=1}^{\nu} w_i \frac{1}{\varepsilon} g(t_0 + c_i h, Y_i).$$

The matrices $\tilde{A} = (\tilde{a}_{ij})$, $\tilde{a}_{ij} = 0$ for $j \geq i$ and $A = (a_{ij})$ are $\nu \times \nu$ matrices such that the resulting scheme is explicit in f , and implicit in g . An IMEX Runge-Kutta scheme is characterized by these two matrices and the coefficient vectors $\tilde{c} = (\tilde{c}_1, \dots, \tilde{c}_\nu)^T$, $\tilde{w} = (\tilde{w}_1, \dots, \tilde{w}_\nu)^T$, $c = (c_1, \dots, c_\nu)^T$, $w = (w_1, \dots, w_\nu)^T$. They can be represented by a double *tableau* in the usual Butcher notation,

$$\begin{array}{c|c} \tilde{c} & \tilde{A} \\ \hline & \tilde{w}^T \end{array}, \quad \begin{array}{c|c} c & A \\ \hline & w^T \end{array}.$$

A sufficient condition to guarantee that f is always evaluated explicitly is that the scheme for g is diagonally implicit i.e. $a_{ij} = 0$, for $j > i$. More general conditions will be considered in [4].

2.1. DIRK schemes and splitting methods. As already observed in the introduction, here we will restrict to diagonally implicit Runge-Kutta (DIRK) schemes. One step of such a scheme can be performed in splitting form as follows:

For $i = 1, \dots, \nu$,

- Evaluate

$$(3) \quad Y_i^* = y_0 + h \sum_{j=1}^{i-2} \tilde{a}_{i,j} f_j + h \tilde{a}_{i,i-1} f(t_0 + \tilde{c}_{i-1} h, Y_{i-1}),$$

where

$$f_j = f(t_0 + \tilde{c}_j h, Y_j).$$

- Solve for Y_i

$$(4) \quad Y_i = Y_i^* + h \sum_{j=1}^{i-1} a_{i,j} \frac{1}{\varepsilon} g_j + h a_{i,i} \frac{1}{\varepsilon} g(t_0 + \tilde{c}_i h, Y_i),$$

where

$$g_j = g(t_0 + \tilde{c}_j h, Y_j).$$

Finally evaluate

$$(5) \quad y_1 = y_0 + h \sum_{i=1}^{\nu} \tilde{w}_i f_i + h \sum_{i=1}^{\nu} w_i \frac{1}{\varepsilon} g_i.$$

Note that, in many cases the implicit step (4) based on a DIRK solver for g can be solved explicitly for a system of the form (2). This is of paramount importance for the simplicity and efficiency of the resulting schemes.

Similarly, every Runge-Kutta splitting scheme can be written as an IMEX Runge-Kutta scheme based on a DIRK solver for g . For example we report in Tables 1-3 the tableau corresponding to the standard splitting scheme SP(1,1,1), Jin [15] splitting Jin(2,2,2) defined only in the stiff regime $h \gg \varepsilon$ and the L-stable Caflisch-Jin-Russo [5] splitting CJR(2,2,2). We will also report the IMEX Midpoint(1,2,2) scheme [1] in Table 4 which has the nice feature of being second order accurate in the nonstiff regime using only one evaluation of the stiff source term.

As usual, the triplet (s, σ, p) characterizes the number s of stages of the implicit scheme, the number σ of stages of the explicit scheme and the order p of the scheme.

0	0	1	1
1	1	1	1

TABLE 1

Tableau for the explicit (left) implicit (right) L -stable $SP(1,1,1)$ scheme

0	0	0	-1	-1	0
1	1	0	2	1	1
1/2	1/2		1/2	1/2	

TABLE 2

Tableau for the explicit (left) implicit (right) $Jin(2,2,2)$ scheme

2.2. Order conditions. These schemes can be viewed as a particular class of partitioned Runge-Kutta schemes [14], and therefore their order conditions can be derived from the general theory of order conditions for partitioned schemes [11], however for low order schemes it is probably preferable to treat these systems differently, since partitioned schemes are usually applied to systems where the stiffness is associated to some components of the unknown vector, while in our case the stiffness is associated to the additive term of the right hand side.

The general technique to derive order conditions for Runge-Kutta schemes is based on the Taylor expansion of the exact and numerical solution. In particular, conditions for schemes of order p are obtained by imposing that the solution of a system

$$y' = f(t, y), \quad y(t_0) = y_0$$

at time $t = t_0 + h$ agrees with the numerical solution obtained by one step of a Runge-Kutta scheme up to order h^{p+1} . For a detailed account of the order conditions for Runge-Kutta schemes see, for example, [3] or [12]. Order conditions for partitioned Runge-Kutta schemes have been derived by Hairer [11].

Here we construct the order conditions for IMEX Runge-Kutta schemes up to order 3. The general treatment of the order condition for IMEX Runge-Kutta will be discussed in [4].

We apply scheme (1, 2) to system (1), with $\varepsilon = 1$. We assume that the coefficients \tilde{c}_i , c_i , \tilde{a}_{ij} , a_{ij} , and the weights \tilde{w}_i , w_i satisfy the conditions

$$\tilde{c}_i = \sum_j \tilde{a}_{i,j}, \quad c_i = \sum_j a_{i,j}, \quad \sum_i \tilde{w}_i = 1, \quad \sum_i w_i = 1.$$

In this way it is enough to analyze the order conditions for autonomous systems,

$$(6) \quad y' = f(y) + g(y), \quad y(0) = y_0.$$

The Taylor expansion of the exact solution of system (6) can be written in terms of the elementary differentials as

$$y(h) = y_0 + (f + g)h + \frac{1}{2}(f'f + f'g + g'f + g'g)h^2$$

0	0	0	0	0	0	0	0	0	0
$\tilde{\alpha}$	$\tilde{\alpha}$	0	0	0	0	0	0	0	0
$\tilde{\alpha}$	$\tilde{\alpha}$	0	0	0	0	γ	γ	0	β
$\frac{\eta(\tilde{\alpha} + \tilde{\beta})}{\eta(\tilde{\alpha} + \beta)}$	$\eta\tilde{\alpha}$	$\eta\tilde{\beta}$	0	0	$\frac{\eta(\beta + \gamma)}{\eta(\beta + \gamma)}$	$\gamma\eta$	0	$\beta\eta$	μ
	$\eta\tilde{\alpha}$	$\eta\beta$	0	0		$\gamma\eta$	0	$\beta\eta$	μ

$$\beta = \frac{2\mu - 1}{2(\mu - 1)}, \gamma = -\frac{2\mu^2 - 2\mu + 1}{2\mu(\mu - 1)}, \tilde{\alpha} = \frac{1}{2\mu}, \tilde{\beta} = -\frac{1}{2(\mu - 1)}, \eta = -2\mu(\mu - 1)$$

TABLE 3
Tableau for the explicit (left) implicit (right) L-stable CJR(3,2,2) scheme

0	0	0	0	0
1/2	1/2	0	0	1/2
	0	1	0	1

TABLE 4
Tableau for the explicit (left) implicit (right) Midpoint(1,2,2) scheme

$$(7) \quad + \frac{1}{6}(f''ff + 2f''fg + f''gg + g''ff + 2g''fg + g''gg + f'f'f + f'g'f + g'f'f + g'g'f + f'f'g + f'g'g + g'f'g + g'g'g)h^3 + O(h^4).$$

The Taylor expansion of the Runge-Kutta solution after one step is given by

$$y_1 = y_0 + y_1^1 h + y_1^2 h^2 + y_1^3 h^3 + O(h^4),$$

where the coefficients of the elementary differentials appearing in the expansion depend on the matrices \tilde{A} , A and on the vectors \tilde{c} , c , \tilde{w} , w . By equating the coefficients of the elementary differentials one obtains the following order conditions:

First order:.

$$(8) \quad \sum_i \tilde{w}_i = 1, \quad \sum_i w_i = 1,$$

Second order:.

$$(9) \quad \sum_i \tilde{w}_i \tilde{c}_i = 1/2, \quad \sum_i w_i c_i = 1/2,$$

$$(10) \quad \sum_i \tilde{w}_i c_i = 1/2, \quad \sum_i w_i \tilde{c}_i = 1/2,$$

Third order:.

$$(11) \quad \sum_{ij} \tilde{w}_i \tilde{a}_{ij} \tilde{c}_j = 1/6, \quad \sum_i \tilde{w}_i \tilde{c}_i \tilde{c}_i = 1/3, \\ \sum_{ij} w_i a_{ij} c_j = 1/6, \quad \sum_i w_i c_i c_i = 1/3,$$

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \gamma & \gamma & 0 & 0 \\ 1 & \delta & 1-\delta & 0 \end{array}, \quad \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \gamma & 0 & \gamma & 0 \\ 1 & 0 & 1-\gamma & \gamma \end{array}, \quad \gamma = 1 - \frac{\sqrt{2}}{2}, \delta = 1 - \frac{1}{2\gamma}$$

TABLE 1

Tableau for the explicit (left) implicit (right) L-stable ARS(2,2,2) scheme

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \gamma & \gamma & 0 & 0 \\ 1 & \delta & 1-\delta & 0 \end{array}, \quad \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \gamma & 0 & \gamma & 0 \\ 1 & 0 & 1-\gamma & \gamma \end{array}, \quad \gamma = 1 - \frac{\sqrt{2}}{2}, \delta = -2\frac{\sqrt{2}}{3}$$

TABLE 2

Tableau for the explicit (left) implicit (right) L-stable ARS(2,3,2) scheme

$$\begin{aligned} \sum_{ij} \tilde{w}_i \tilde{a}_{ij} c_j &= 1/6, & \sum_{ij} \tilde{w}_i a_{ij} \tilde{c}_j &= 1/6, & \sum_{ij} \tilde{w}_i a_{ij} c_j &= 1/6, \\ \sum_{ij} w_i \tilde{a}_{ij} c_j &= 1/6, & \sum_{ij} w_i a_{ij} \tilde{c}_j &= 1/6, & \sum_{ij} w_i \tilde{a}_{ij} \tilde{c}_j &= 1/6, \\ \sum_i \tilde{w}_i c_i c_i &= 1/3, & \sum_i \tilde{w}_i \tilde{c}_i c_i &= 1/3, \\ \sum_i w_i \tilde{c}_i \tilde{c}_i &= 1/3, & \sum_i w_i \tilde{c}_i c_i &= 1/3. \end{aligned} \tag{12}$$

Conditions (8), (9), (11) are the standard order conditions for the two *tableau*, each of them taken separately. Conditions (10) and (12) are new conditions that arise because of the coupling of the two schemes.

The order conditions will simplify a lot if $\tilde{c} = c$. For this reason only such schemes are considered in [1] and [4]. In particular, we observe that, if the two tableau differ only for the value of the matrices A , i.e. if $\tilde{c}_i = c_i$ and $\tilde{w}_i = w_i$, then the standard order conditions for the two schemes are enough to ensure that the combined scheme is third order. Note, however, that this is true only for schemes up to third order.

3. L-stable second order DIRK schemes. In this section we deal with the accuracy of IMEX schemes. Since the schemes are used to solve a system that depends on the stiffness parameter ε , L-stability of the implicit Runge-Kutta integrator is a highly desirable property. In particular we expect that the accuracy of the resulting IMEX schemes depends on this stiffness parameter. Actually, even the regularity properties of the solution of the system (2) may depend on ε . In many cases of physical relevance, the solution of the initial value problem for system (2) is smooth, uniformly in ε , provided the initial condition is a *local equilibrium*, i.e. satisfies

$$(1) \quad g_2(u_0, v_0) = 0.$$

In such cases one might require that the numerical solution after one time step remains “close” to the exact solution uniformly in the stiffness parameter ε .

Here we shall limit to analyze the accuracy properties of the schemes for $\varepsilon = 1$ and for small values of ε . Uniform accuracy of the schemes will be investigated numerically.

0	0	0	0	0	0	0	0	0	0
1/2	1/2	0	0	0	1/2	0	1/2	0	0
1/3	1/3	0	0	0	1/3	0	0	1/3	0
1	0	1	0	0	1	0	0	3/4	1/4
	0	1	0	0	1	0	0	3/4	1/4

TABLE 3
Tableau for the explicit (left) implicit (right) L-stable LRR(3,2,2) scheme

0	0	0	1 - C	1 - C	0
1	1	0	C	C - δ	δ
	1/2	1/2		1/2	1/2

 $\delta = 1 - \frac{1}{2C}$

TABLE 4
Tableau for the explicit (left) implicit (right) L-stable PR(2,2,2) scheme

3.1. Accuracy analysis. In this section we perform the accuracy analysis of the different L-stable schemes presented in Table 3 and Tables 1-4. Schemes 1-2 were presented in [1], scheme 3 was derived in [20], whereas scheme 4 is new.

All these schemes have second order accuracy in the nonstiff limit, i.e. for $\varepsilon = 1$. Here we are interested in the stiff accuracy, i.e. in the accuracy of the numerical solution in the limit as $\varepsilon \rightarrow 0$. In addition, we consider the asymptotic behavior of the solution for small values of ε .

Let us consider the system (2), in the particular case where $g_2(u, v) = s(u) - v$. In this case, the asymptotic behavior of the solution for small values of the parameter ε can be obtained by the so called Chapman-Enskog expansion. According to this expansion, the function u satisfies the equation

$$u' = f_1 + \varepsilon \frac{\partial f_1}{\partial v} (f_2 - s'(u)f_1) + O(\varepsilon^2),$$

where f_1 , f_2 , and $\partial f_1 / \partial v$ are evaluated at $(u, s(u))$. The function v is given by

$$v = s(u) + \varepsilon (f_2 - s'(u)f_1) + O(\varepsilon^2).$$

From this expansion it is possible to derive a Taylor expansion of the solution after one time step, in terms of ε and h . For simplicity we shall apply the different IMEX to system (2), with f_1 , f_2 , and g_2 linear, i.e.

$$(2) \quad \begin{aligned} u' &= a_1 u + b_1 v, \\ v' &= a_2 u + b_2 v + \frac{1}{\varepsilon} (cu - v). \end{aligned}$$

We shall assume that the initial state is in “equilibrium”, i.e. it satisfies the condition

$$v_0 = s(u_0) = cu_0.$$

After one time step, the exact solution of the system satisfies

$$(3) \quad \begin{aligned} u &= u_0(1 + \tilde{a}h + \frac{1}{2}\tilde{a}^2h^2 + \varepsilon b_1\tilde{b}h) + O(h^3, \varepsilon^2h), \\ v &= u_0(c(1 + \tilde{a}h + \frac{1}{2}\tilde{a}^2h^2) + \varepsilon\tilde{b}(1 + (\tilde{a} + b_1c)h)) + O(h^3, \varepsilon^2h), \end{aligned}$$

where $\tilde{a} = a_1 + b_1 c$, and $\tilde{b} = a_2 + (b_2 - a_1)c - b_1 c^2$.

Now we apply the different schemes for one time step. The results are summarized in the sequel.

CJR(3,2,2):. The Taylor expansion of the first component is

$$u_1 = u_0 \left(1 + \tilde{a}h + \frac{1}{2}\tilde{a}^2 h^2 + \frac{1-\mu}{1-2\mu}\varepsilon b_1 \tilde{b}h \right) + O(h^3, \varepsilon^2 h).$$

Therefore the scheme is second order accurate in the limit $\varepsilon = 0$, but it does not provide the correct limit for small values of ε , unless one takes $\mu = 1$. The value $\mu = 1$, however, is excluded, because for this value of the parameter the scheme is no longer L -stable (this can be seen by the expression of the stability function, and its limit as $z \rightarrow \infty$).

This can be expressed by the formula

$$u_1(\varepsilon, h) - u(\varepsilon, h) = O(h^3, \varepsilon h).$$

Similarly, for the second component one has

$$v_1(\varepsilon, h) - v(\varepsilon, h) = O(h^3, \varepsilon h).$$

ARS(2,2,2):. The Taylor expansion of the first component is

$$u_1 = u_0 \left(1 + \tilde{a}h + \frac{1}{2}\tilde{a}^2 h^2 + \frac{1}{2-\sqrt{2}}\varepsilon b_1 \tilde{b}h \right) + O(h^3, \varepsilon^2 h).$$

Therefore the scheme is second order accurate in the limit $\varepsilon = 0$, but it is not accurate for small values of ε . The result of the analysis can be summarized by the formulas

$$\begin{aligned} u_1(\varepsilon, h) - u(\varepsilon, h) &= O(h^3, \varepsilon h), \\ v_1(\varepsilon, h) - v(\varepsilon, h) &= O(h^3, \varepsilon h). \end{aligned}$$

ARS(2,3,2):. In this case the Taylor expansion of the numerical solution for u agrees with the expansion of the exact solution up to terms $O(h^2, \varepsilon h)$. However, the scheme suffers a degradation of accuracy in the second variable. The result may be summarized by the formulas

$$\begin{aligned} u_1(\varepsilon, h) - u(\varepsilon, h) &= O(h^3, \varepsilon^2 h), \\ v_1(\varepsilon, h) - v(\varepsilon, h) &= O(h^2, \varepsilon h). \end{aligned}$$

Note, however, that the loss of accuracy in the variable v does not affect the accuracy in the variable u . This is easily understood in the limit $\varepsilon \rightarrow 0$. In this case the scheme becomes a second order scheme for the limit equation in u . In the exact solution the second variable is just a function of the first. The numerical scheme does not provide the correct projection, but computes the second variable by solving an equation for it. Although the local truncation error in v is $O(h^2)$, the scheme is essentially second order even in v in the limit $\varepsilon \rightarrow 0$, because the variable u is second order accurate, and so is v . This analysis is in agreement with the numerical results shown in Section 5.

PR(2,2,2):. The Taylor expansion of the numerical solution for u is given by

$$u_1 = u_0(1 + \tilde{a}h + \frac{1}{2}\tilde{a}^2h^2 + \frac{C}{2C-1}\varepsilon b_1\tilde{b}h) + O(h^3, \varepsilon h).$$

From this expression it follows that the scheme, similarly to CJR(3,2,2), is second order in u , in the limit $\varepsilon = 0$, but it is not accurate for small values of ε

$$u_1(\varepsilon, h) - u(\varepsilon, h) = O(h^3, \varepsilon h).$$

The expansion of v_1 gives

$$v_1(\varepsilon, h) - v(\varepsilon, h) = O(h^2, \varepsilon h),$$

therefore the local truncation error is second order in h . Note that, because u is accurate, we expect the global truncation error in v to be of the same order of the local one.

LRR(3,2,2):. The scheme provides second order accuracy, both in u and v , in the limit $\varepsilon = 0$, and the first order term in ε is first order in u . The results are summarized by the expressions

$$\begin{aligned} u_1 - u(\varepsilon, h) &= O(h^3, \varepsilon h^2), \\ v_1 - v(\varepsilon, h) &= O(h^3, \varepsilon h). \end{aligned}$$

We remark that the above analysis is performed by assuming that the stiffness parameter is either $O(1)$ or quite small. Probably a more satisfactory approach to the same problem, in the linear case, is obtained by using the concept of B -consistency and B -convergence [7]. In this case, in fact, uniform bounds in ε for the local truncation error could be derived.

3.2. L-stability and asymptotic preservation. In this section we will show that the so called asymptotic preservation property for hyperbolic systems with relaxation, i.e. consistency of the scheme in the stiff limit $\varepsilon \rightarrow 0$, is guaranteed if the implicit step is solved by a L-stable scheme.

To this aim let us consider the general structure of an IMEX Runge-Kutta scheme written in splitting form (3)-(4) and (5) with $Y = (U, V)$, $f = (f_1, f_2)$ and $g = (0, g_2)$. We will also assume that the algebraic equation $g_2(u, v) = 0$ can be uniquely solved in v giving (3).

The L-stability property together with our assumptions implies that the implicit step (4) as $\varepsilon \rightarrow 0$ yields

$$(4) \quad g_2(t_0 + \tilde{c}_j h, U_j, V_j) = 0 \quad \Longleftrightarrow \quad V_j = e(U_j).$$

Hence, denoting by $F(u) = f_1(U, e(U))$, the explicit step (3) gives the explicit scheme for the limiting equilibrium equation (4)

$$(5) \quad U_i^* = u_0 + h \sum_{j=1}^{i-2} \tilde{a}_{ij} F_j + h \tilde{a}_{i,i-1} F(t_0 + \tilde{c}_{i-1} h, U_{i-1}),$$

where

$$F_j = F(t_0 + \tilde{c}_j h, U_j).$$

Finally we need to evaluate

$$(6) \quad u_1 = u_0 + h \sum_{i=1}^{\nu} \tilde{w}_i F_i.$$

It is remarkable that, under our assumptions, the tableau of the limiting scheme for (4) is always the tableau of the explicit part of the IMEX Runge-Kutta scheme. Note that in the limit the value of the v component can be recovered directly from u using (3).

3.3. Stability analysis. In this section we perform the stability analysis of the schemes. In particular, we use a generalization of the concept of A -stability to systems of equations which are the sum of a stiff and a non-stiff part.

Let us consider a system of the form (1) with $f(y) = \lambda_1 y$, $g(y) = \lambda_2 y$, and $y_0 = 1$. When applying scheme (1)-(2) to the above system, one obtains

$$(7) \quad Y_i = 1 + h \sum_{j=1}^{\nu} \tilde{a}_{ij} \lambda_1 Y_j + h \sum_{j=1}^{\nu} a_{ij} \lambda_2 Y_j,$$

$$(8) \quad y_1 = 1 + h \sum_{i=1}^{\nu} \tilde{w}_i \lambda_1 Y_i + h \sum_{i=1}^{\nu} w_i \lambda_2 Y_i.$$

Using a vector notation with $Y = (Y_1, \dots, Y_{\nu})^T$, Eq. (7) becomes,

$$Y = e + \lambda_1 h \tilde{A} Y + \lambda_2 h A Y$$

where $e = (1, \dots, 1)^T$.

Solving for Y and substituting in Eq. (8), one has

$$(9) \quad y_1 = \mathcal{R}(z_1, z_2) \equiv 1 + (z_1 \tilde{w}^T + z_2 w^T)(I - z_1 \tilde{A} - z_2 A)^{-1} e,$$

where $z_1 = \lambda_1 h$, $z_2 = \lambda_2 h$. The function $\mathcal{R}(z_1, z_2)$ is the *function of absolute stability*.

As an example, we report the function of absolute stability of some of the schemes we discussed before.

The function of absolute stability of scheme PR(2,2,2) is given by

$$\mathcal{R}(z_1, z_2) = 1 - \frac{(z_1 + z_2)(2C - 3z_2C + z_2 + z_1C + 2z_2C^2)}{(1 - z_2 + z_2C)(-2C + 2z_2C - z_2)}.$$

For $C = 1$ and $z_1 = 0$ this function becomes

$$\mathcal{R}(0, z_2) = -\frac{2 + z_2}{-2 + z_2},$$

which is not the stability function of an L -stable scheme, since $R(0, \infty) = 1$.

The function of absolute stability of scheme ARS(2,2,2) is given by

$$\mathcal{R}(z_1, z_2) = 1 - \frac{z_2^2 \gamma^2 - z_1 \gamma z_2 + z_1 \gamma^2 z_2 + z_1^2 \gamma \delta - z_1 - z_2 + z_1 \delta z_2 \gamma - z_1^2 \gamma}{(-1 + z_2 \gamma)^2},$$

and the one of scheme ARS(2,3,2) is given by

$$\mathcal{R}(z_1, z_2) = 1 - \frac{(z_1 + z_2)(-2z_1 \gamma + z_1 \gamma^2 - z_1^2 \gamma^2 + z_1^2 \gamma^2 \delta + z_1 \gamma^2 \delta z_2 - 1 + \gamma^2 z_2)}{(-1 + z_2 \gamma)^2}.$$

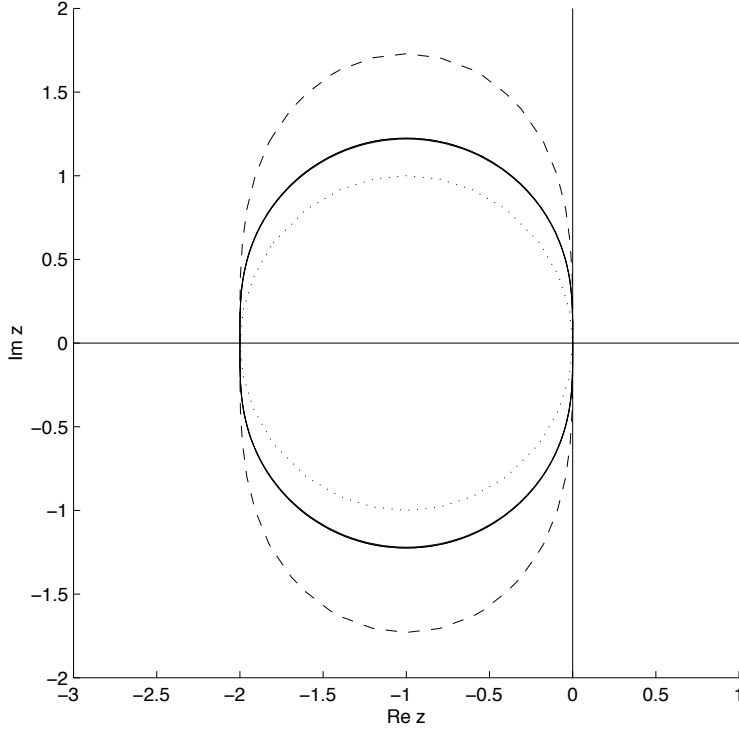


FIG. 1. Stability region S_1 for scheme CJR(3,2,2) with $\mu = 1/3$, ARS(2,3,2), PR(2,2,2) with $C = 1/\sqrt{2}$ (continuous line). The dotted and the dashed lines are the stability region of explicit Euler and explicit second order Runge-Kutta respectively.

The *region of absolute stability* S_A associated to scheme (1, 2) is defined as

$$S_A = \{(z_1, z_2) \in \mathbf{C}^2 : |\mathcal{R}(z_1, z_2)| \leq 1\}.$$

It is evident that the region does not contain the set $\mathbf{C}^- \times \mathbf{C}^-$. Our goal is to show that there exist two regions of the complex plane, $S_1 \subset \mathbf{C}$, $S_2 \subset \mathbf{C}$, with the following properties:

$$(10) \quad S_A \supset S_1 \times S_2,$$

$$(11) \quad S_2 \supset \mathbf{C}^- \equiv \{z \in \mathbf{C} : \Re(z) \leq 0\},$$

and to compute them. It is clear that such two sets, if they exist, are not unique. We shall compute explicitly the largest set S_1 for which $S_2 \supset \mathbf{C}^-$. Such region is defined by

$$S_1 = \{z_1 \in \mathbf{C} : \sup_{z_2 \in \mathbf{C}^-} |\mathcal{R}(z_1, z_2)| \leq 1\}.$$

In order to proceed we make use of the following

LEMMA 3.1. *For any fixed $z_1 \in \mathbf{C}$, the function $|\mathcal{R}(z_1, z_2)|$ assumes its maximum value in \mathbf{C}^- for some z_2 belonging to the imaginary axis.*

Proof.

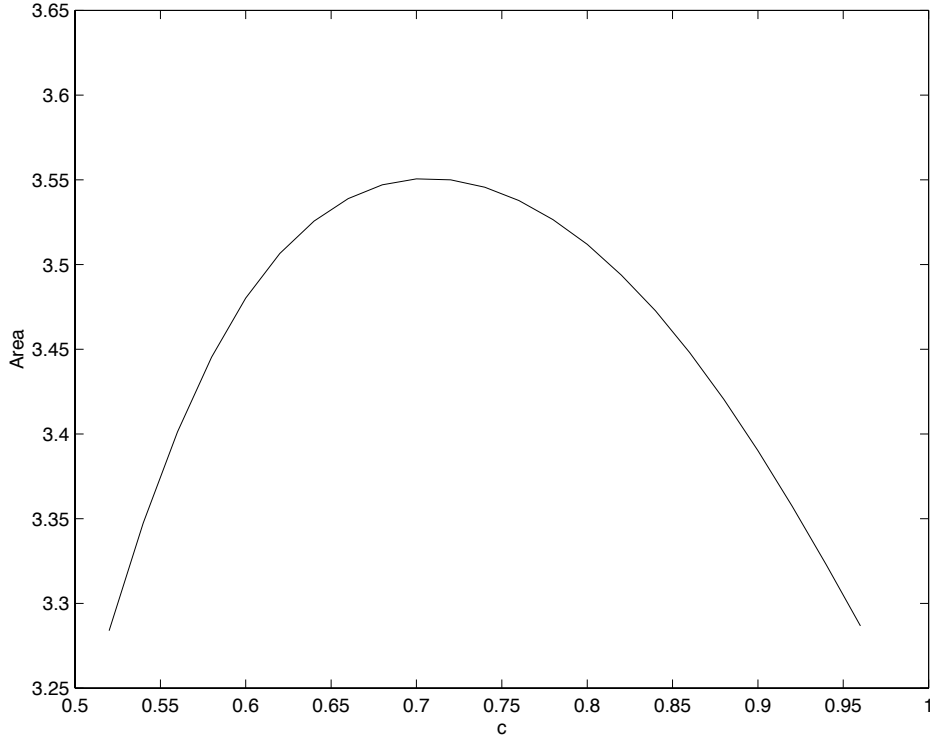


FIG. 2. Area of the stability region S_1 for scheme PR, as a function of the parameter c .

For any fixed value z_1 , the function $\mathcal{R}(z_1, z_2)$ is analytic in z_2 , holomorphic in \mathbf{C}^- . Therefore the maximum value of the function in \mathbf{C}^- is obtained for some value of $z_2 \in \partial\mathbf{C}^-$. Furthermore, we observe that

$$\lim_{z_2 \rightarrow \infty} \mathcal{R}(z_1, z_2) = 0,$$

and hence the maximum of $|\mathcal{R}(z_1, z_2)|$ occurs on the imaginary axis.

From the above lemma, we can write

$$S_1 = \{z_1 \in \mathbf{C} : \max_{y \in \mathbf{R}} |\mathcal{R}(z_1, iy)| \leq 1\}.$$

The boundary of this region (boundary locus) will be obtained as

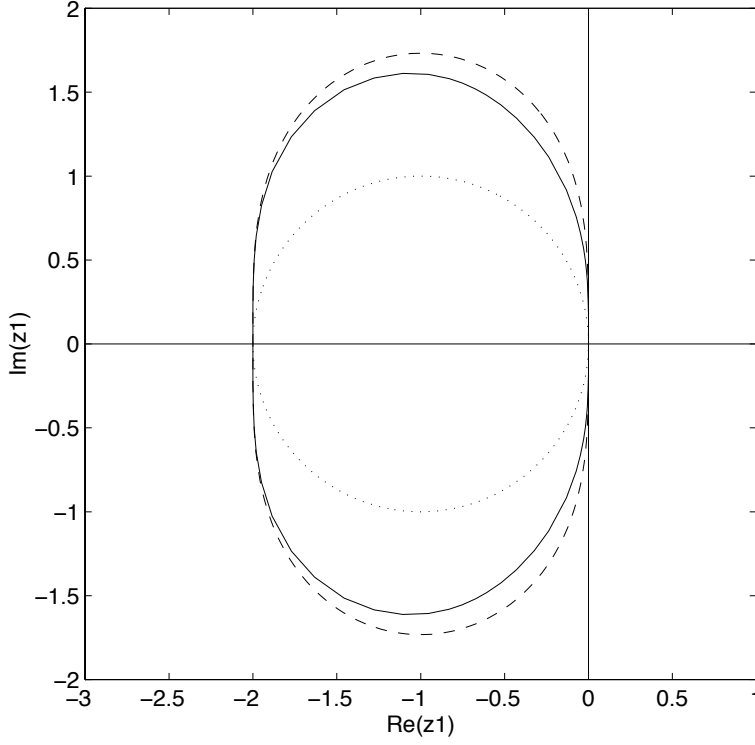
$$\partial S_1 = \{z_1 \in \mathbf{C} : \max_{y \in \mathbf{R}} |\mathcal{R}(z_1, iy)| = 1\}.$$

In order to compute the region S_1 we act as follows. Let us define the function

$$F(\theta, \rho, y) = |\mathcal{R}(-1 + \rho e^{i\theta}, iy)|^2, \quad \theta \in [-\pi, \pi), \quad y \in \mathbf{R}.$$

This function can be written in the form

$$F(\theta, \rho, y) = \frac{N(\theta, \rho, y)}{D(\theta, \rho, y)},$$

FIG. 3. Stability region S_1 for scheme LRR(3,2,2).

where N and D are polynomials in y and ρ , whose coefficients are trigonometric functions in θ . Now, let $F_s(\theta, \rho, y) = N(\theta, \rho, y) - D(\theta, \rho, y)$ and

$$\mathcal{H}(\theta, \rho) = \max_{y \in \mathbf{R}} F_s(\theta, \rho, y).$$

The boundary of S_1 can therefore be expressed in terms of $\mathcal{H}(\theta, \rho)$,

$$\partial S_1 = \{(\theta, \rho) : \mathcal{H}(\theta, \rho) = 0\}.$$

The boundary locus can be efficiently computed by a technique similar to the one used to compute the stability region of Runge-Kutta schemes. First observe that $\mathcal{H}(\theta, 0) < 0$, therefore $-1 \in S_1$. Now let us denote by \hat{y} the value of y for which $\mathcal{H}(\theta, \rho) = F_s(\theta, \rho, y)$, i.e.

$$\hat{y}(\theta, \rho) = \arg \max_{y \in \mathbf{R}} F_s(\theta, \rho, y),$$

and let $\hat{\rho}$ denote the value of ρ for which $F_s(\theta, \rho, \hat{y}(\theta, \rho)) = 0$. Then the boundary locus satisfies the equations

$$(12) \quad F_s(\theta, \rho, y) = 0,$$

$$(13) \quad G(\theta, \rho, y) \equiv \frac{\partial F_s}{\partial y}(\theta, \rho, y) = 0.$$

The second equation is a consequence of the fact that \hat{y} is an extremal, and the function $F_s(\theta, \rho, y)$ is smooth in y .

By differentiating equations (12) and (13) one obtains a set of differential equations that defines the boundary locus. By assuming that the quantities θ , ρ , and y are function of a parameter τ , one can derive the equations for the trajectories

$$(14) \quad \frac{d\theta}{d\tau} = -\alpha(\theta, \rho, y)F_1(\theta, \rho, y),$$

$$(15) \quad \frac{d\rho}{d\tau} = \alpha(\theta, \rho, y)F_2(\theta, \rho, y),$$

$$(16) \quad \frac{dy}{d\tau} = \alpha(\theta, \rho, y)F_3(\theta, \rho, y),$$

where the function α is arbitrary, and has the effect of changing the parametrization, and

$$(17) \quad F_1(\theta, \rho, y) = \frac{\partial F_s}{\partial \rho} \frac{\partial G}{\partial y},$$

$$(18) \quad F_2(\theta, \rho, y) = \frac{\partial F_s}{\partial \theta} \frac{\partial G}{\partial y},$$

$$(19) \quad F_3(\theta, \rho, y) = \frac{\partial G}{\partial \theta} \frac{\partial F_s}{\partial \rho} - \frac{\partial F_s}{\partial \theta} \frac{\partial G}{\partial \rho}.$$

The initial condition for the system is

$$\rho(0) = 1, \quad \theta(0) = 0, \quad y(0) = 0.$$

We decided to use the arc-length as parameter. In this case

$$\alpha = \frac{1}{\sqrt{F_1^2 + F_2^2 + F_3^2}}.$$

The system is solved forward in τ to obtain the upper branch of the boundary locus, up to a point so that $\theta = \pi$. Then the boundary locus is extended symmetrically below the real axis.

By using the above mentioned technique, we computed the stability region for the schemes shown in the previous section.

First we show the stability region S_1 for schemes CJR(3,2,2) with $\mu = 1/3$, ARS(2,3,2), PR(2,2,2) with $C = 1/\sqrt{2}$. The result is shown in Fig. 1. Note that the stability regions of the three schemes are almost identical and are smaller than the one of standard explicit RK2, but they include the stability region of explicit Euler.

The size of the stability region of the family of schemes PR(2,2,2) depends slightly on the parameter C , while its shape is almost unchanged. In Fig. 2 we plot the area of the stability region S_1 as a function of the parameter C . We observe that the maximum of the area is obtained near the value $C = 1/\sqrt{2}$.

We report also the stability region S_1 for scheme LRR(3,2,2) in Fig. 3. Note that this scheme has a slightly wider stability region.

4. Third order schemes. Third order IMEX Runge-Kutta schemes can be obtained directly using the order conditions we have seen in Section 2. We will report here, as examples, the ARS(2,3,3) and the L-stable ARS(3,4,3) and ARS(4,4,3) IMEX Runge-Kutta schemes (see [1]). In all these schemes the implicit scheme is a DIRK scheme and hence they can be written in splitting form accordingly to (3)-(4) and (5). In particular scheme ARS(3,4,3) is defined using the root of a third order polynomial. We report here its numerical equivalent.

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \gamma & \gamma & 0 & 0 \\ 1-\gamma & \gamma-1 & 2-2\gamma & 0 \end{array}, \quad \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \gamma & 0 & \gamma & 0 \\ 1-\gamma & 0 & 1-2\gamma & \gamma \end{array}, \quad \gamma = \frac{3+\sqrt{3}}{6}$$

$$\begin{array}{c|ccc} & 0 & 1/2 & 1/2 \end{array}, \quad \begin{array}{c|ccc} & 0 & 1/2 & 1/2 \end{array}$$

TABLE 5
Tableau for the explicit (left) implicit (right) third order ARS(2,3,3) scheme

$$\begin{array}{c|cccc} 0 & 0 & 0 & 0 & 0 \\ \gamma & \gamma & 0 & 0 & 0 \\ (1+\gamma)/2 & (1+\gamma)/2 - \eta & \eta & 0 & 0 \\ 1 & 1-2\mu & \mu & \mu & 0 \end{array}, \quad \begin{array}{c|cccc} 0 & 0 & 0 & 0 & 0 \\ \gamma & 0 & \gamma & 0 & 0 \\ (1+\gamma)/2 & 0 & (1-\gamma)/2 & \gamma & 0 \\ 1 & 0 & 1-\delta-\gamma & \delta & \gamma \end{array}$$

$$\begin{array}{c|cccc} & 0 & 1-\delta-\gamma & \delta & \gamma \end{array}, \quad \begin{array}{c|cccc} & 0 & 1-\delta-\gamma & \delta & \gamma \end{array}$$

$$\gamma = 0.4358665215, \delta = -0.644373171, \eta = 0.3966543747, \mu = 0.5529291479.$$

TABLE 6
Tableau for the explicit (left) implicit (right) third order L-stable ARS(3,4,3) scheme

5. Test equation. A simple prototype of stiff system of the form (2) is obtained taking

$$f_1(u, v) = \alpha u, \quad f_2(u, v) = \beta v, \quad g_2(u, v) = e(u) - v,$$

which corresponds to the system of Ode's

$$(1) \quad u' = \alpha v,$$

$$(2) \quad v' = \beta u + \frac{1}{\varepsilon}(e(u) - v).$$

The choice $\alpha = -1$, $\beta = 1$ guarantees that the eigenvalues of the explicit part are $\pm i$.

The convergence plots have been obtained first computing the relative error E_h by comparing the numerical solution with step h to the numerical solution with step $h/2$ with different values of ε ranging from 10^{-5} to 1. The value $h = 0.05$ has been used, and the system has been integrated for $t \in [0, 5]$. Then, the convergence rates have been computed as usual through the error curves E_h and $E_{h/2}$. We present the results by plotting the L_2 norm of the relative error versus ε using a log-scale on the x -axis.

To examine the behavior of the different IMEX Runge-Kutta schemes we have considered $e(u) = \sin(u)$ and two different set of initial data:

Equilibrium data.

$$u(0) = \pi/2, \quad v(0) = \sin(u(0)) = 1.$$

In Fig. 1 we report the convergence profiles of the different L-stable second order schemes together with scheme Midpoint(2,2,2) in the case of equilibrium initial data. All the L-stable schemes gives similar results with second order accuracy both for

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

TABLE 7

Tableau for the explicit (left) implicit (right) third order L-stable ARS(4,4,3) scheme

large and small values of ε , and with an intermediate region where $h = O(\varepsilon)$ in which we have a deterioration of the accuracy, in particular for the v component of the solution. Note that, in the stiff limit the accuracy on the v component in the PR(2,2,2) deteriorates to first order. This is not a strong drawback of the scheme since an accurate value of v can be always computed from u in the limit $\varepsilon \rightarrow 0$. As expected the Midpoint(2,2,2) scheme is unstable for very small values of ε and the v component has a slightly oscillating behavior in the stiff limit. It is remarkable, however, that for this scheme the accuracy on u is uniformly second order with respect to ε .

The corresponding results for the third order schemes are given in the left column of Fig. 3. Again the behavior of the schemes is rather similar although it can be observed that the accuracy of ARS(4,4,3) deteriorates in a wider region.

Non-equilibrium data.

$$u(0) = \pi/2, \quad v(0) = 1/2.$$

In this case $g_2(u(0), v(0)) \neq 0$ and hence the problem presents an initial layer on the v component when $\varepsilon \rightarrow 0$.

In Fig. 2 we present the same convergence profiles as before of the different second order IMEX Runge-Kutta schemes in the case of non equilibrium initial data. The accuracy of the schemes deteriorates dramatically in the intermediate region $h = O(\varepsilon)$ and again this is more evident in the v component. In the stiff limit, in particular, both CJR(3,2,2) and ARS(2,2,2) schemes deteriorate to first order. Note that scheme PR(2,2,2) seems to outperform the other schemes since the results are similar to the case of equilibrium initial data. For this initial data, the Midpoint(2,2,2) scheme has a strongly oscillating behavior on the v component for ε small and the scheme is clearly unstable in the stiff regime.

Finally in the right column of Fig. 3 we give the results of the third order schemes for non-equilibrium initial data. Scheme ARS(3,4,3) clearly outperforms the other two schemes. Note that scheme ARS(2,3,3) is not L-stable and hence, although no instability is observed, v does not have the correct asymptotic limit. Furthermore scheme ARS(4,4,3) reduces to first order accuracy in the stiff limit.

REFERENCES

- [1] U. ASHER, S. RUUTH, AND R. J. SPITERI, *Implicit-explicit Runge-Kutta methods for time dependent Partial Differential Equations*, Appl. Numer. Math. 25, (1997), pp. 151-167.
- [2] U. ASHER, S. RUUTH, AND B. WETTON, *Implicit-explicit methods for time dependent PDE's*, SIAM J. Numer. Anal., 32, (1995) pp. 797-823.
- [3] J. BUTCHER, *The numerical analysis of Ordinary differential equations. Runge-Kutta and general linear methods..* John Wiley & Sons, Chichester and New York (1987).
- [4] J. BUTCHER, N. GUGLIELMI, AND G. RUSSO, in preparation.
- [5] R. E. CAFLISCH, S. JIN AND G. RUSSO, *Uniformly accurate schemes for hyperbolic systems with relaxation*, SIAM J. Numer. Anal., 34, (1997) pp. 246-281.
- [6] M. CECCHI MORANDI, M. REDIVO-ZAGLIA, G. RUSSO, *Extrapolation methods for hyperbolic systems with relaxation*, Journal of Computational and Applied Mathematics 66 (1996) pp. 359-375.
- [7] K. DEKKER, AND J.G. VERWER, *Stability of Runge-Kutta Methods for Stiff Nonlinear Differential Equations*, North-Holland, Amsterdam (1984).
- [8] BOUN OUMAR DIA AND MICHELLE SCHATZMAN, *Estimation sur la formule de Strang*, C. R. Acad. Sci. Paris, t.320, Série I (1995) pp. 775-779.
- [9] BOUN OUMAR DIA AND MICHELLE SCHATZMAN, *Commutateur de certains semi-groupes holomorphes et applications aux directions alternées*, Mathematical Modelling and Numerical Analysis 30 (1996) pp. 343-383.
- [10] S.GOTTLIEB, C.-W. SHU, *Total Variation Diminishing Runge-Kutta schemes*, Math. Comp. 67 (1998) pp. 73-85.
- [11] E. HAIRER, *Order conditions for numerical methods for Partitioned ordinary differential equations*, Numerische Mathematik 36 (1981) pp. 431-445.
- [12] E. HAIRER, S.P.NØRSET AND G. WANNER, *Solving ordinary differential equations, Vol.1 Non-stiff problems*, Springer-Verlag, New York (1987).
- [13] E. HAIRER AND G. WANNER, *Solving ordinary differential equations, Vol.2 Stiff and differential-algebraic problems*, Springer-Verlag, New York (1987).
- [14] E. HOFER *A partially implicit method for large stiff systems of Ode's with only few equations introducing small time-constants*, SIAM J. Numer. Anal. 13, (1976) pp. 645-663.
- [15] S. JIN, *Runge-Kutta methods for hyperbolic systems with stiff relaxation terms* J. Comput. Phys., 122 (1995) pp. 51-67.
- [16] S. JIN AND C. D. LEVERMORE, *Numerical Schemes for hyperbolic conservation laws with stiff relaxation terms*, J. Comp. Physics, 126 (1996) pp. 449-467.
- [17] S. JIN, L. PARESCHI AND G. TOSCANI *Diffusive relaxation schemes for multiscale discrete-velocity kinetic equations*, SIAM J. Num. Anal., 35 (1999) pp. 2405-2439.
- [18] S.JIN, A.KLAR AND L.PARESCHI *Asymptotic preserving schemes for multiscale kinetic equations: a unified approach*, work in preparation.
- [19] R.J. LEVEQUE, *Numerical Methods for Conservation Laws*, Lectures in Mathematics, Birkhauser Verlag, Basel (1992).
- [20] S.F.LIOTTA, V.ROMANO, G.RUSSO, *Central schemes for balance laws of relaxation type*, submitted to SIAM J. Numer. Anal.
- [21] G.NALDI AND L.PARESCHI, *Numerical schemes for hyperbolic systems with stiff diffusive relaxation*, SIAM J. Numer. Anal. to appear.
- [22] L.PARESCHI AND G.RUSSO, *Implicit-Explicit Runge-Kutta schemes for hyperbolic systems with stiff relaxation*, work in progress
- [23] G.STRANG, *On the construction and comparison of difference schemes*, SIAM J. Numer. Anal. 5, (1968) pp. 506.
- [24] TADMOR E., *Approximate Solutions of Nonlinear Conservation Laws*, Cockburn, Johnson, Shu and Tadmor Eds., Lecture Notes in Mathematics, N. 1697 (1998).

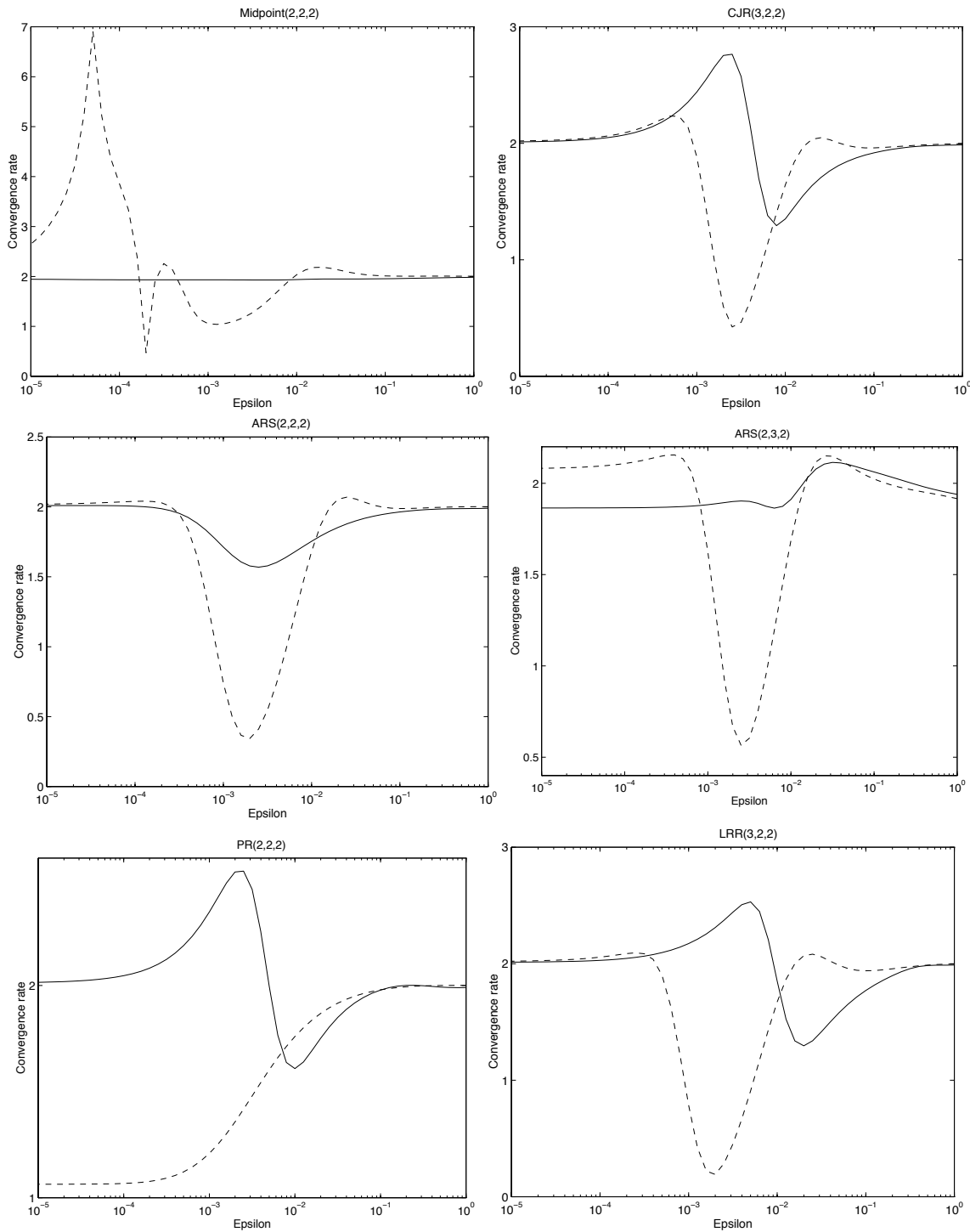


FIG. 1. Convergence rates of second order IMEX schemes for equilibrium initial data (dashed line: v component, continuous line: u component).

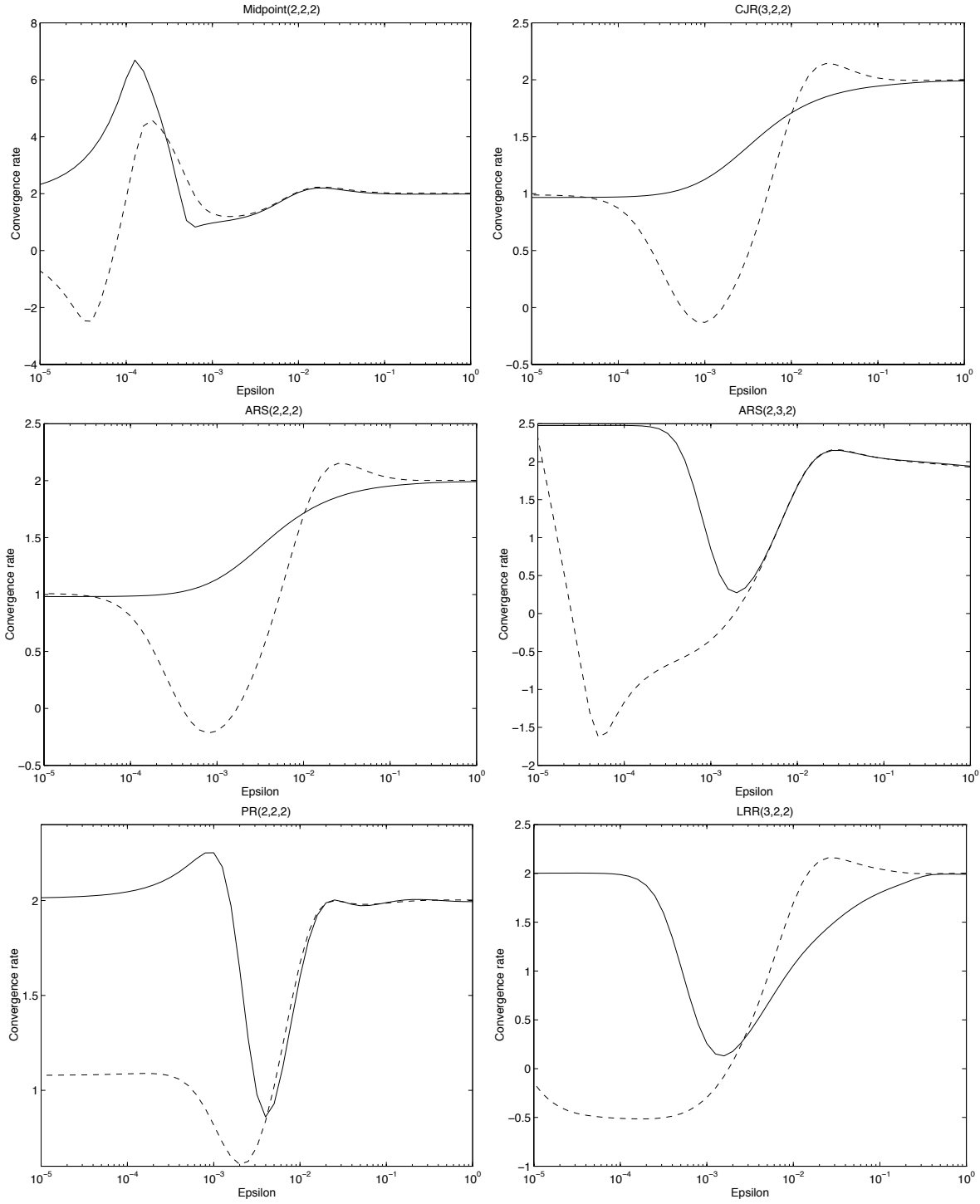


FIG. 2. Convergence rates of second order IMEX schemes for non equilibrium initial data (dashed line: v component, continuous line: u component).

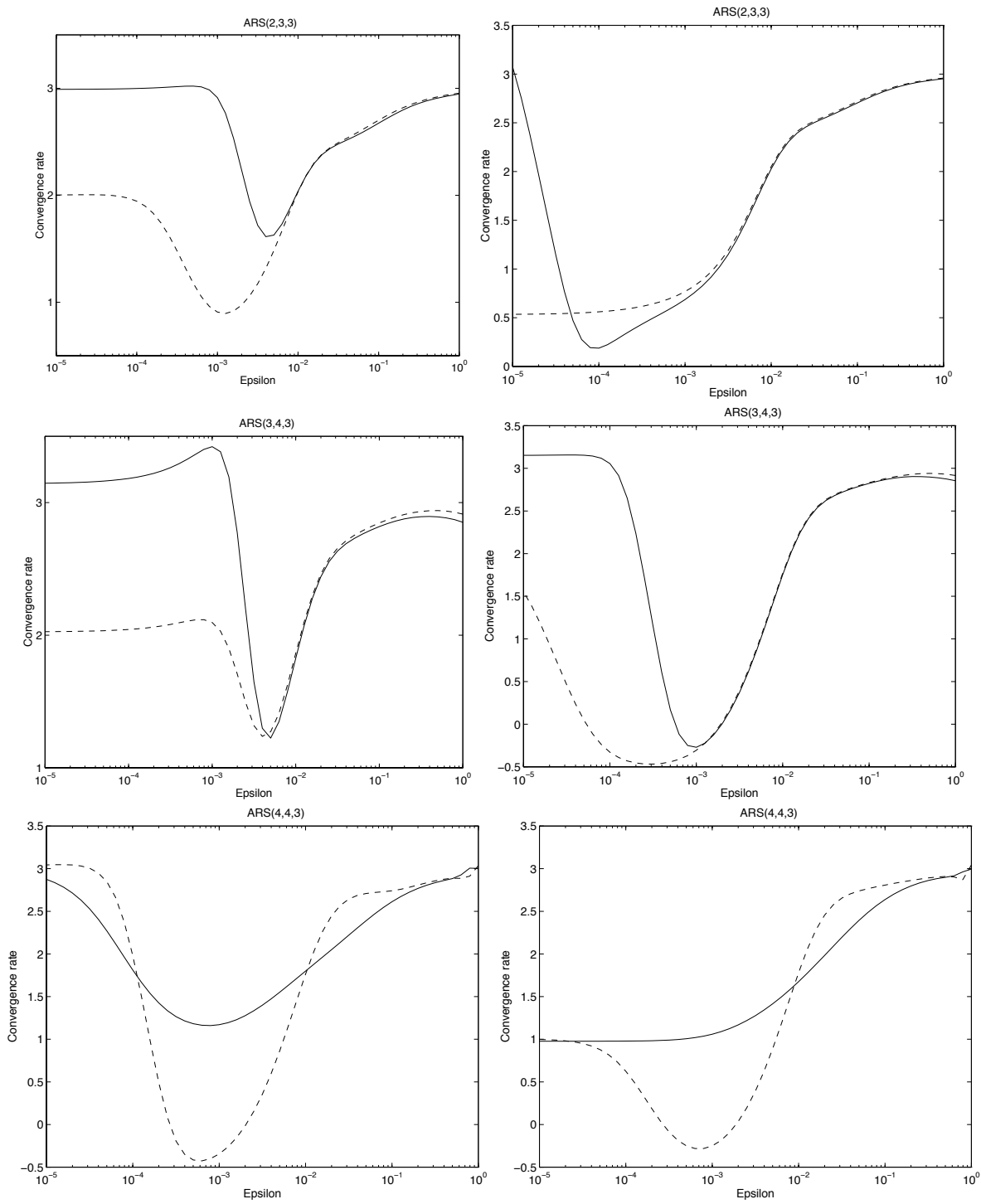


FIG. 3. Convergence rates of third order IMEX schemes for equilibrium initial data (left column) and non-equilibrium initial data (right column). Dashed line: v component, continuous line: u component.