JOURNAL OF
COMPUTATIONAL AND
APPLIED MATHEMATICS

ELSEVIER Journal of Computational and Applied Mathematics 140 (2002) 99–117

www.elsevier.com/locate/cam

Additive Runge–Kutta methods for the resolution of linear parabolic problems

B. Bujanda[a], J.C. Jorge[b, *]

[a] *Departamento de Matemáticas y Computación, Universidad de La Rioja, Edificio J. L. Vives, Calle Luis de Ulloa s/n, 26004 Logroño, Spain*

[b] *Departamento de Matemática e Informática, Universidad Pública de Navarra, Edificio Los Acebos, Campus Arrosadía s/n, 31006 Pamplona, Spain*

Received 14 August 2000

Abstract

We study the consistency for general additive Runge–Kutta methods in the integration of linear nonhomogeneous problems, obtaining necessary and sufficient conditions of order p, for arbitrary values of p. We use this result joined to some A-stability conditions for developing a third order additive Runge–Kutta method of type fractional steps and we show its efficiency in the numerical integration of a two-dimensional evolutionary convection–diffusion problem. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: RK additive methods; Order conditions

1. Introduction

Let us consider an initial value problem (IVP):

$$\begin{aligned} &y'(t) = f(t, y(t)),\\ &y(t_0) = y_0 \end{aligned} \tag{1}$$

and an arbitrary decomposition of f in n, usually simpler, addends of type

$$f(t, y(t)) = \sum_{i=1}^{n} f_i(t, y(t)). \tag{2}$$

* Corresponding author.

E-mail addresses: bbujanda@dmc.unirioja.es (B. Bujanda), jcjorge@unavarra.es (J.C. Jorge).

PII: S0377-0427(01)00475-7

Definition 1.1. An additive Runge–Kutta method of s stages and n levels is a one-step method such that, using partition (2), approaches the solution of the IVP (1) by means of the scheme

$$y_0$$

$$y_{m+1} = y_m + \tau \sum_{i=1}^{s} \sum_{k=1}^{n} b_i^k f_k(t_m + \tau c_i, Y^{m,i}) \quad \text{where}$$

$$Y^{m,i} = y_m + \tau \sum_{j=1}^{s} \sum_{k=1}^{n} a_{ij}^k f_k(t_m + \tau c_j, Y^{m,j}) \quad \text{for } i = 1, \ldots, s. \tag{3}$$

Here, y_m will be approximations to the solution $y(t_m)$, where $t_m = m\tau$. Typically, the vectors $Y^{m,i}$ are called stages of the method.

For simplicity, we shall consider only constant time step τ. The convergence results that we give in Section 3 are easily extended to variable time steps τ_m.

The scheme (3) suggests that there are n standard Runge–Kutta methods involved in an additive Runge–Kutta method, in such way that each one of them defines the contribution to the numerical solution of a part f_i of the derivative function f.

Different applications of additive Runge–Kutta methods can be found in the literature; in all of them, the main advantage is focused in reducing the computational cost of the numerical integration of certain IVPs, if a suitable partition of function f is considered. For example, in [6,7,13] some nonlinear stiff problems of type

$$y'(t) = J(t)y(t) + g(t, y),$$

$$y(t_0) = y_0,$$

where $g(t, y)$ is a perturbation of the linear term $J(t)y(t)$, are efficiently integrated by using some additive Runge–Kutta methods composed by an A-stable DIRK method for computing the contribution of the linear term and an explicit Runge–Kutta method for the perturbation $g(t, y)$, resulting a linearly implicit numerical method. Also, in [16] it is shown that certain class of semi-explicit additive Runge–Kutta methods provides robustness and low computational cost in the numerical integration of Hamiltonian systems with separable Hamiltonians, of type

$$\frac{\mathrm{d}p}{\mathrm{d}t} = f(q),$$

$$\frac{\mathrm{d}q}{\mathrm{d}t} = g(p),$$

where $p \equiv (p_1, \ldots, p_n)$, $q \equiv (q_1, \ldots, q_n)$, $f \equiv (-\partial H/\partial q_1, \ldots, -\partial H/\partial q_n)$ and $g \equiv (\partial H/\partial p_1, \ldots, \partial H/\partial p_n)$, being $H \equiv H(p_1, \ldots, p_n;\ q_1, \ldots, q_n) : \mathbb{R}^{2n} \to \mathbb{R}$ the Hamiltonian function. The scheme proposed in [16] is

$$\begin{aligned} p^{m+1} &= p^m + \tau \sum_{j=1}^{s} b_j^1 f_j, \\ q^{m+1} &= q^m + \tau \sum_{j=1}^{s} b_j^2 g_j, \end{aligned} \quad \text{with} \quad \begin{aligned} Y^{m,i} &= p^m + \tau \sum_{j=1}^{i} a_{ij}^1 f_j, \\ Z^{m,i} &= q^m + \tau \sum_{j=1}^{i-1} a_{ij}^2 g_j, \end{aligned} \quad \text{for } i = 1, \ldots, s,$$

where $f_j = f(Z^{m,j})$ and $g_j = g(Y^{m,j})$. Note that, for this kind of Hamiltonians, even though $a^1_{ii} \neq 0$, the last scheme is explicit, since $Y^{m,j}$ is not an argument of $f_j = f(Z^{m,j})$ and, for suitable choices of coefficients (a^k_{ij}, b^k_j), it preserves the symplectic structure of the phase space (such type of methods are called canonical).

In some classical papers (see [15,18]), and recently in others (see [3,5,11,12,17]), a special kind of additive Runge–Kutta methods, called fractional step Runge–Kutta, have been used efficiently in the time integration of multidimensional parabolic problems.

Definition 1.2. A fractional step Runge–Kutta method (abbreviated as FSRK), is an additive Runge–Kutta method verifying

$$a^k_{ii} \geqslant 0, \quad \forall i \in \{1,\ldots,s\},\ k \in \{1,\ldots,n\}, \qquad a^k_{ij} = 0, \quad \forall j > i,$$

$$|b^k_j| + \sum_{i=1}^{s} |a^k_{ij}| \neq 0 \Rightarrow |b^l_j| + \sum_{i=1}^{s} |a^l_{ij}| = 0, \quad \forall l \neq k,\ l,k \in \{1,\ldots,n\},\ i,j \in \{1,\ldots,s\},$$

$$a^l_{ii} a^k_{ii} = 0 \quad \text{if } k \neq l,\ i \in \{1,\ldots,s\},\ k,l \in \{1,\ldots,n\}.$$

We will assume that a choice of coefficients a^k_{ij}, b^k_i and c_i with $i,j = 1,\ldots,s$ and $k = 1,\ldots,n$ determines a unique additive Runge–Kutta method, although different partitions on the derivative function will give different numerical approaches y_m. Following the compact notations introduced by Butcher for the standard Runge–Kutta methods (see [4,10]), we will refer an additive Runge–Kutta method by means of their coefficients structured in a table as follows:

$$\begin{array}{c|c|c|c|c} Ce & \mathscr{A}^1 & \mathscr{A}^2 & \ldots & \mathscr{A}^n \\ \hline & (b^1)^{\mathrm{T}} & (b^2)^{\mathrm{T}} & \ldots & (b^n)^{\mathrm{T}}, \end{array}$$

where $\mathscr{A}^k = (a^k_{ij})$, $b^k = (b^k_i)$, $C = \operatorname{diag}(c_1,\ldots,c_s)$ and $e = (1,\ldots,1)^{\mathrm{T}} \in \mathbb{R}^s$ with $i,j = \{1,\ldots,s\}$ and $k = \{1,\ldots,n\}$.

We will focus our work in obtaining the order p conditions (for arbitrary values of p) that a general additive Runge–Kutta method must verify when it is applied to solve numerically a linear IVP of type

$$\begin{aligned} \frac{\mathrm{d}y(t)}{\mathrm{d}t} &= L(t)y(t) + g(t), \\ y(t_0) &= y_0, \end{aligned} \tag{4}$$

where $L : [t_0,T] \to \mathbb{R}^{d\times d}$ and $g : [t_0,T] \to \mathbb{R}^d$.

The use of Butcher's framework (see [4]) to study the consistency of (3) requires to transform (4) into an autonomous problem of type

$$\begin{aligned} \frac{\mathrm{d}\bar{y}(t)}{\mathrm{d}t} &= F(\bar{y}(t)), \\ \bar{y}(t_0) &= \bar{y}_0, \end{aligned} \tag{5}$$

where typically

$$\bar{y}(t) = \begin{pmatrix} t \\ y(t) \end{pmatrix} \in \mathbb{R}^{d+1}, \qquad F(\bar{y}(t)) = \begin{pmatrix} 1 \\ f(\bar{y}(t)) \end{pmatrix} \in \mathbb{R}^{d+1}, \qquad \bar{y}_0 = \begin{pmatrix} t_0 \\ y_0 \end{pmatrix} \in \mathbb{R}^{d+1}$$

and, before studying the consistency, some coefficient restrictions are usually imposed to ensure that the application of scheme (3) to resolve the IVP (4) provides the same numerical solution as that of the scheme applied to (5). In the case of considering only linear problems, such restrictions can be avoided and we do it in this paper.

In order to construct efficient methods of type FSRK for such problems we will combine the order p conditions that we obtain in this paper with some stability conditions according to the nature of the IVP to solve. For example, if we are interested in approaching the solution of a parabolic problem, we can carry out firstly a spatial semidiscretization of it, obtaining a one-parameter family of stiff IVPs of type

$$\begin{aligned} \frac{\mathrm{d}y_h(t)}{\mathrm{d}t} &= L_h(t)y_h(t) + g_h(t), \\ y_h(t_0) &= y_{0h}, \end{aligned} \tag{6}$$

where h is typically the mesh size, and where the stiffness grows to infinity when h tends to zero. Obviously, if we want an efficient time integration for the family of IVPs (6) it is convenient to have unconditional convergence results, or in other words, convergence for the numerical solutions without limitations between the mesh size h and the time step τ. If the numerical integration of these problems is performed with classical one-step explicit methods, a stable behaviour implies strong restrictions between τ and Lipschitz's constant of the problem, that, in this case, grows to infinity when h tends to zero. Therefore, if we want to obtain unconditional convergence, when we discretize (6), we need to consider implicit methods verifying certain properties of absolute stability. Similar restrictions between τ and h should be considered if we used explicit additive Runge–Kutta methods but such limitations can be avoided by using A-stable implicit additive Runge–Kutta methods.

This article is structured in four sections. In Section 2, we develop the study of the consistency, obtaining necessary and sufficient conditions for attaining order p.

In Section 3, we will use some stability results proved in [2,14], together with the consistency results developed here, for proving the convergence of the method.

Finally, in the last section we show the process for obtaining a third-order L-stable FSRK method and its efficiency in the numerical integration of a semidiscrete in space convection–diffusion problem.

2. Order conditions

In order to integrate the IVP (4) by means of a general additive Runge–Kutta method, we consider firstly a general decomposition for the linear operator $L(t)$ in n addends $\sum_{i=1}^{n} L_i(t)$ ($L_i : [t_0, T] \to \mathbb{R}^{d\times d}$) and also a decomposition for the source term $g(t)$ of the form $\sum_{i=1}^{n} g_i(t)$ ($g_i : [t_0, T] \to \mathbb{R}^d$), in such a way that we use the scheme (3) with the following decomposition for $f(t, y(t)) \equiv L(t)y(t) + g(t)$:

$$f(t, y(t)) = \sum_{i=1}^{n} f_i(t, y(t)) \quad \text{with } f_i(t, y(t)) = L_i(t)y(t) + g_i(t). \tag{7}$$

The use of FSRK methods for the integration of problems of type (6) with partition (7) is computationally interesting, because the initial problem can be reduced to a set of simpler problems, one in every stage $Y^{m,i}$ of scheme (3), in the form $(I - \tau k\, L_j(t_{m,i}))Y^{m,i} = F_i$, where F_i is computed explicitly from the result of the last stages, and where the linear operator L_j can be simpler than L in a certain way; for example, if the operators $L_j(t_{m,i})$ for $j = 1, \ldots, n$, are obtained by using a spatial discretization, via finite differences, of operators of type $d_j(\partial^2/\partial x_j^2) - v_j(\partial/\partial x_j) - k_j$, then $(I - \tau k L_j(t_{m,i}))$ are tridiagonal.[1] However, if we use classical implicit methods we should solve a family of block tridiagonal systems, enlarging strongly the computational cost, or, on the contrary, if we use classical explicit methods we should limit strongly the size of the time step when we use fine meshes.

In order to rewrite scheme (3) in a compact form we will use the following tensorial notations:

$$\hat{L}_i^m(\tau) = \begin{pmatrix} L_i(t_m + \tau c_1) & 0 & \ldots & 0 \\ 0 & L_i(t_m + \tau c_2) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & L_i(t_m + \tau c_s) \end{pmatrix} \in (\mathbb{R}^{d\times d})^{s\times s},$$

$$\hat{Y}^m = \begin{pmatrix} Y^{m,1} \\ Y^{m,2} \\ \vdots \\ Y^{m,s} \end{pmatrix} \in (\mathbb{R}^d)^s, \qquad \hat{G}_i^m(\tau) = \begin{pmatrix} g_i(t_m + \tau c_1) \\ g_i(t_m + \tau c_2) \\ \vdots \\ g_i(t_m + \tau c_s) \end{pmatrix} \in (\mathbb{R}^d)^s, \tag{8}$$

given $v \equiv (v_i) \in \mathbb{R}^s$, $M \equiv (m_{ij}) \in \mathbb{R}^{s\times s}$ and

$$I_{\mathbb{R}^d} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix} \in \mathbb{R}^{d\times d},$$

we denote

$$\bar{v} \equiv \begin{pmatrix} v_1 I_{\mathbb{R}^d} \\ \vdots \\ v_s I_{\mathbb{R}^d} \end{pmatrix} \in (\mathbb{R}^{d\times d})^s$$

and

$$\bar{M} \equiv \begin{pmatrix} m_{11} I_{\mathbb{R}^d} & \ldots & m_{1s} I_{\mathbb{R}^d} \\ \vdots & \ddots & \vdots \\ m_{s1} I_{\mathbb{R}^d} & \ldots & m_{ss} I_{\mathbb{R}^d} \end{pmatrix} \in (\mathbb{R}^{d\times d})^{s\times s}$$

[1] This is the main advantage of the classical alternating direction implicit methods.

and using these notations and partition (7), scheme (3) can be written compactly as follows:

$$
\begin{aligned}
y_{m+1} &= y_m + \tau \sum_{i=1}^{n} (\overline{b^i})^{\mathrm{T}} (\hat{L}_i^m(\tau) \hat{Y}^m + \hat{G}_i^m(\tau)), \\
\hat{Y}^m &= \bar{e} y_m + \tau \sum_{i=1}^{n} \overline{\mathscr{A}^i} (\hat{L}_i^m(\tau) \hat{Y}^m + \hat{G}_i^m(\tau)).
\end{aligned} \tag{9}
$$

To study the consistency of this scheme we introduce, as usual, the local error

$$e^{m+1} = y(t_{m+1}) - \hat{y}^{\,m+1},$$

where $\hat{y}^{\,m+1}$ is the approximation to $y(t_{m+1})$, obtained with one step of scheme (9) starting from $y(t_m)$ instead of y_m. We say that an additive Runge–Kutta method is consistent of order p if for data $L_i(t) \in \mathscr{C}^p([t_0, T]; \mathbb{R}^{d \times d})$, $g_i(t) \in \mathscr{C}^p([t_0, T]; \mathbb{R}^d)$, for $i = 1, \ldots, n$, it is verified that

$$\|e^{m+1}\| \leqslant C\tau^{p+1}, \quad \forall m \geqslant 0, \ \ \tau \to 0, \tag{10}$$

where C is a constant independent of τ.

In order to obtain the order conditions, we will develop in powers of τ and in terms of $L_i(t)$, $g_i(t)$ and their derivatives, the solution $y(t_{m+1})$ and its approach $\hat{y}^{\,m+1}$.

If the operator $(\bar{I} - \tau \sum_{i=1}^{n} \overline{\mathscr{A}^i}\, \hat{L}_i^m(\tau))$ is invertible,[2] then we can deduce that

$$
\begin{aligned}
\hat{y}_{\,m+1} = y(t_m) &+ \tau \sum_{i=1}^{n} (\overline{b^i})^{\mathrm{T}} \hat{L}_i^m(\tau) \left(\bar{I} - \tau \sum_{j=1}^{n} \overline{\mathscr{A}^j} \hat{L}_j^m(\tau) \right)^{-1} \left(\bar{e}\, y(t_m) + \tau \sum_{k=1}^{n} \overline{\mathscr{A}^k}\, \hat{G}_k^m(\tau) \right) \\
&+ \tau \sum_{i=1}^{n} (\overline{b^i})^{\mathrm{T}} \hat{G}_i^m(\tau).
\end{aligned} \tag{11}
$$

As we have assumed that $L_i(t) \in \mathscr{C}^p([t_0, T]; \mathbb{R}^{d \times d})$ and $g_i(t) \in \mathscr{C}^p([t_0, T]; \mathbb{R}^d)$, then the following Taylor's expansions for $\hat{L}_i^m(\tau)$ and $\hat{G}_i^m(\tau)$ can be used

$$
\begin{aligned}
\hat{L}_i^m(\tau) &= \sum_{j=0}^{p-1} \frac{\tau^j}{j!} (\bar{\mathrm{C}})^j \hat{L}_i^{m(j)}(0) + \mathcal{O}(\tau^p), \\
\hat{G}_i^m(\tau) &= \sum_{j=0}^{p-1} \frac{\tau^j}{j!} (\bar{\mathrm{C}})^j \hat{G}_i^{m(j)}(0) + \mathcal{O}(\tau^p),
\end{aligned} \tag{12}
$$

[2] Note that the operator $(\bar{I} - \tau \sum_{i=1}^{n} \overline{\mathscr{A}^i}\, \hat{L}_i^m(\tau))$ is invertible for τ small. Nevertheless, as we plan to integrate some stiff problems, it is convenient to ensure that the operator $(\bar{I} - \tau \sum_{i=1}^{n} \overline{\mathscr{A}^i}\, \hat{L}_i^m(\tau))$ is invertible for arbitrary τ. This is proved, for time-dependent maximal coercive operators $L_i(t)$, in [2].

where

$$\hat{L}_i^{m\,(j)}(0)=\frac{\mathrm{d}^j\hat{L}_i^m(0)}{\mathrm{d}t^j}\quad\text{and}\quad \hat{G}_i^{m\,(j)}(0)=\frac{\mathrm{d}^j\hat{G}_i^m(0)}{\mathrm{d}t^j}. \tag{13}$$

Besides, we can use the following development for $(\bar{I}-\tau\sum_{i=1}^{n}\overline{\mathscr{A}^i}\,\hat{L}_i^m(\tau))^{-1}$:

$$\left(\bar{I}-\tau\sum_{i=1}^{n}\overline{\mathscr{A}^i}\hat{L}_i^m(\tau)\right)^{-1}$$

$$=\sum_{i=0}^{p}\tau^i\left(\sum_{j=1}^{n}\overline{\mathscr{A}^j}\,\hat{L}_j^m(\tau)\right)^i+\tau^{p+1}\left(\sum_{i=1}^{n}\overline{\mathscr{A}^i}\,\hat{L}_i^m(\tau)\right)^{p+1}\left(\bar{I}-\tau\sum_{j=1}^{n}\overline{\mathscr{A}^j}\,\hat{L}_j^m(\tau)\right)^{-1}$$

$$=\sum_{i=0}^{p}\tau^i\left(\sum_{j=1}^{n}\overline{\mathscr{A}^j}\hat{L}_j^m(\tau)\right)^i+\mathscr{O}(\tau^{p+1}). \tag{14}$$

Using expansions (12) and (14), the next expression for $\hat{y}_{m+1}$ is deduced from (11):

$$\hat{y}_{m+1}=y(t_m)+\tau\sum_{i=1}^{n}(\overline{b^i})^{\mathrm{T}}\left(\sum_{j=0}^{p-1}\frac{\tau^j}{j!}(\bar{\mathrm{C}})^j\hat{L}_i^{m(j)}(0)\right)$$

$$\sum_{k=0}^{p-1}\tau^k\left(\sum_{l=1}^{n}\overline{\mathscr{A}^l}\sum_{r=0}^{E[p/k]}\frac{\tau^r}{r!}(\bar{\mathrm{C}})^r\hat{L}_l^{m(r)}(0)\right)^k\bar{e}\,y(t_m)$$

$$+\tau^2\sum_{i=1}^{n}(\overline{b^i})^{\mathrm{T}}\left(\sum_{j=0}^{p-2}\frac{\tau^j}{j!}(\bar{\mathrm{C}})^j\hat{L}_i^{m(j)}(0)\right)\sum_{k=0}^{p-2}\tau^k\left(\sum_{l=1}^{n}\overline{\mathscr{A}^l}\sum_{r=0}^{E[p/k]}\frac{\tau^r}{r!}(\bar{\mathrm{C}})^r\hat{L}_l^{m(r)}(0)\right)^k$$

$$\left(\sum_{s=1}^{n}\overline{\mathscr{A}^s}\left(\sum_{u=0}^{p-2}\frac{\tau^u}{u!}(\bar{\mathrm{C}})^u\,\hat{G}_s^{m\,(u)}(0)\right)\right)$$

$$+\tau\sum_{i=1}^{n}(\overline{b^i})^{\mathrm{T}}\left(\sum_{j=0}^{p-1}\frac{\tau^j}{j!}(\bar{\mathrm{C}})^j\,\hat{G}_i^{m(j)}(0)\right)+\mathscr{O}(\tau^{p+1}), \tag{15}$$

where $E[x]$ denotes the integer part of x (except in the case $k=0$ which, for convenience, we will consider $E[p/0]=p$).

In [8], an expansion for the exact solution of (4) in powers of τ and in function of $L(t)$, $g(t)$ and their derivatives are given. Such development is summarized in the following.

Lemma 2.1. *If $L(t)\in\mathscr{C}^p([t_0,T];\mathbb{R}^{d\times d})$ and $g(t)\in\mathscr{C}^p([t_0,T];\mathbb{R}^d)$, then the exact solution $y(t_{m+1})$ of problem* (4) *can be expanded as follows*:

$$y(t_{m+1})=y(t_m)+\sum_{i=1}^{p}\frac{\tau^i}{i!}y^{(i)}(t_m)+\mathscr{O}(\tau^{p+1}), \tag{16}$$

where

$$y^{(i)}(t) = \sum_{\substack{r=1,\dots,i \\ (\rho_1,\dots,\rho_r)\in\{0,\dots,i-1\}^r \\ \rho_1+\cdots+\rho_r=i-r}} c^{\rho_1,\dots,\rho_r}(F^{(\rho_1,\dots,\rho_r)}y(t) + G^{(\rho_1,\dots,\rho_r)}g(t)), \tag{17}$$

with

$$F^{(\rho_1,\dots,\rho_r)}y(t) = L^{(\rho_1)}(t)\dots L^{(\rho_r)}(t)y(t),$$
$$G^{(\rho_1,\dots,\rho_r)}g(t) = L^{(\rho_1)}(t)\dots L^{(\rho_{r-1})}(t)g^{(\rho_r)}(t)$$

and with

$$c^{\rho_1,\dots,\rho_r} = i!\prod_{j=1}^{r}\frac{1}{(r-j+1)+\sum_{k=j}^{r}\rho_k}\prod_{j=1}^{r}\frac{1}{\rho_j!}. \tag{18}$$

Now, if we take into account that

$$L^{(k)}(t) = \sum_{i=1}^{n} L_i^{(k)}(t), \quad g^{(k)}(t) = \sum_{i=1}^{n} g_i^{(k)}(t) \qquad \text{for } k = 0,\dots,p-1,$$

we can substitute the developments (17) by the following ones:

$$y^{(i)}(t) = \sum_{\substack{r=1,\dots,i \\ (i_1,\dots,i_r)\in\{1,\dots,n\}^r \\ (\rho_1,\dots,\rho_r)\in\{0,\dots,i-1\}^r \\ \rho_1+\cdots+\rho_r=i-r}} c^{\rho_1,\dots,\rho_r}(F_{i_1,\dots,i_r}^{(\rho_1,\dots,\rho_r)}y(t) + G_{i_1,\dots,i_r}^{(\rho_1,\dots,\rho_r)}g(t)),$$

$$F_{i_1,\dots,i_r}^{(\rho_1,\dots,\rho_r)}y(t) = L_{i_1}^{(\rho_1)}(t)L_{i_2}^{(\rho_2)}(t)\dots L_{i_r}^{(\rho_r)}(t)y(t) \quad \text{and}$$
$$G_{i_1,\dots,i_r}^{(\rho_1,\dots,\rho_r)}g(t) = L_{i_1}^{(\rho_1)}(t)L_{i_2}^{(\rho_2)}(t)\dots L_{i_{r-1}}^{(\rho_{r-1})}(t)g_{i_r}^{(\rho_r)}(t); \tag{19}$$

we call fractioned elementary differentials (FED) to the terms $F_{i_1,\dots,i_r}^{(\rho_1,\dots,\rho_r)}y(t)$ and $G_{i_1,\dots,i_r}^{(\rho_1,\dots,\rho_r)}g(t)$.

To find the order p conditions we impose that, in the developments (16)–(19) for $y(t_{m+1})$ and (15) for $\hat{y}_{m+1}$, the addends which contain the powers τ^i, with $0 \leqslant i \leqslant p$, are equal. This implies that we will make equal the coefficients that in expressions (15) and (16)–(19) multiply to the same FED $F_{i_1,\dots,i_r}^{(\rho_1,\dots,\rho_r)}y(t)$ or $G_{i_1,\dots,i_r}^{(\rho_1,\dots,\rho_r)}g(t)$, for every $(\rho_1,\dots,\rho_r)\in\{0,\dots,i-1\}^r$ such that $\rho_1+\cdots+\rho_r = i-r$, obtaining the following.

Theorem 2.1. *The necessary and sufficient order conditions for an additive Runge–Kutta method, applied to* (4), *to attain order p are*

$$(b^{i_1})^{\mathrm{T}}(\mathrm{C})^{\rho_1}\mathscr{A}^{i_2}(\mathrm{C})^{\rho_2}\dots\mathscr{A}^{i_r}(\mathrm{C})^{\rho_r}e = \prod_{j=1}^{r}\frac{1}{(r-j+1)+\sum_{k=j}^{r}\rho_k},$$

$$\textit{for all } r=1,\dots,p \quad \textit{for all } (\rho_1,\dots,\rho_r)\in\{0,\dots,p-1\}^r,$$

$$\textit{such that } 1 \leqslant r+\sum_{k=1}^{r}\rho_k \leqslant p \textit{ and for all } (i_1,i_2,\dots,i_r)\in\{1,2,\dots,n\}^r. \tag{20}$$

Proof. For proving that conditions (20) are sufficient to attain order p, we use the next commutativity relations among the operators $\hat{L}_i^{m(j)}(0)$, $\overline{\mathscr{A}^i}$ and $\bar{v}$, introduced in (8) and (13):

$$\overline{\mathscr{A}^i}\hat{L}_j^{m(k)}(0)=\hat{L}_j^{m(k)}(0)\overline{\mathscr{A}^i}, \quad (\bar{\mathrm{C}})^j\hat{L}_i^{m(k)}(0)=\hat{L}_i^{m(k)}(0)(\bar{\mathrm{C}})^j \quad \text{and} \quad \hat{L}_i^{m(k)}(0)\bar{v}=\bar{v}L_i^{(k)}(t_m).$$

These relations imply that

$$\begin{aligned}(\bar{b}^{i_1})^{\mathrm{T}}\hat{L}_{i_1}^{m(\rho_1)}(0)(\bar{\mathrm{C}})^{\rho_1}\overline{\mathscr{A}^{i_2}}\hat{L}_{i_2}^{m(\rho_2)}(0)\dots\overline{\mathscr{A}^{i_r}}\hat{L}_{i_r}^{m(\rho_r)}(0)(\bar{\mathrm{C}})^{\rho_r}\bar{v}\\
=(\bar{b}^{i_1})^{\mathrm{T}}(\bar{\mathrm{C}})^{\rho_1}\overline{\mathscr{A}^{i_2}}(\bar{\mathrm{C}})^{\rho_2}\dots\overline{\mathscr{A}^{i_r}}(\bar{\mathrm{C}})^{\rho_r}\bar{v}L_{i_1}^{(\rho_1)}(t_m)\dots L_{i_r}^{(\rho_r)}(t_m)\\
=(b^{i_1})^{\mathrm{T}}(\mathrm{C})^{\rho_1}\mathscr{A}^{i_2}(\mathrm{C})^{\rho_2}\dots\mathscr{A}^{i_r}(\mathrm{C})^{\rho_r}vL_{i_1}^{(\rho_1)}(t_m)\dots L_{i_r}^{(\rho_r)}(t_m),\end{aligned} \tag{21}$$

for all $(i_1,\dots,i_r)\in\{1,\dots,n\}^r$, for all $(\rho_1,\dots,\rho_r)\in\{0,\dots,p-1\}^r$ and for all $v\in\mathbb{R}^s$. Taking into account (21) together with the relation $\hat{G}_i^{m(k)}(0)=\bar{e}g_i^{(k)}(t_m)$, we expand (15), obtaining the following development:

$$\begin{aligned}\hat{y}_{m+1}=y(t_m)+\sum_{\substack{j=1,\dots,p\\ r=1,\dots,j\\ (i_1,\dots,i_r)\in\{1,\dots,n\}^r\\ (\rho_1,\dots,\rho_r)\in\{0,\dots,j-1\}^r\\ \rho_1+\dots+\rho_r=j-r}}\tau^j C_{i_1,\dots,i_r}^{(\rho_1,\dots,\rho_r)}F_{i_1,\dots,i_r}^{(\rho_1,\dots,\rho_r)}y(t_m)\\
+\sum_{\substack{j=1,\dots,p\\ r=1,\dots,j\\ (i_1,\dots,i_r)\in\{1,\dots,n\}^r\\ (\rho_1,\dots,\rho_r)\in\{0,\dots,j-1\}^r\\ \rho_1+\dots+\rho_r=j-r}}\tau^j C_{i_1,\dots,i_r}^{(\rho_1,\dots,\rho_r)}G_{i_1,\dots,i_r}^{(\rho_1,\dots,\rho_r)}g(t_m)\\
+\mathscr{O}(\tau^{p+1}),\end{aligned} \tag{22}$$

with

$$C_{i_1,\dots,i_r}^{(\rho_1,\dots,\rho_r)}=\frac{(b^{i_1})^{\mathrm{T}}(\mathrm{C})^{\rho_1}\mathscr{A}^{i_2}(\mathrm{C})^{\rho_2}\dots\mathscr{A}^{i_r}(\mathrm{C})^{\rho_r}e}{\rho_1!\rho_2!\dots\rho_r!}.$$

By imposing that the coefficient $C_{i_1,\dots,i_r}^{(\rho_1,\dots,\rho_r)}$ associated with FED $F_{i_1,\dots,i_r}^{(\rho_1,\dots,\rho_r)}y(t_m)$ in (22), is equal to the coefficient associated with the same FED in expressions (16)–(19) of $y(t_{m+1})$, which is $(1/i!)c^{\rho_1,\dots,\rho_r}$, we obtain the following order condition:

$$\begin{aligned}(b^{i_1})^{\mathrm{T}}(\mathrm{C})^{\rho_1}\mathscr{A}^{i_2}(\mathrm{C})^{\rho_2}\dots\mathscr{A}^{i_r}(\mathrm{C})^{\rho_r}e=\frac{\rho_1!\rho_2!\dots\rho_r!}{i!}c^{\rho_1,\dots,\rho_r}\\
=\prod_{j=1}^{r}\frac{1}{(r-j+1)+\sum_{k=j}^{r}\rho_k}.\end{aligned} \tag{23}$$

Secondly, making equal the coefficients that multiply in $y(t_{m+1})$ and $\hat{y}_{m+1}$ to the FED $G_{i_1,\dots,i_r}^{(\rho_1,\dots,\rho_r)}g(t_m)$, we deduce again the order condition (23).

Let us prove now that conditions (20) are also necessary. To get this, we choose the following IVP:

$$\begin{aligned}
&y_1'(t)=t^{\rho_1}y_2(t), && y_1(0)=0,\\
&y_2'(t)=t^{\rho_2}y_3(t), && y_2(0)=0,\\
&\vdots && \vdots\\
&y_{r-1}'(t)=t^{\rho_{r-1}}y_r(t), && y_{r-1}(0)=0,\\
&y_r'(t)=t^{\rho_r}, && y_r(0)=0,
\end{aligned}$$

which is obviously of type (4) and we also choose the following decomposition for $L(t)$ and $g(t)$:

$$L_i(t)=\begin{pmatrix} 0 & \delta_{i_1 i}t^{\rho_1} & 0 & \dots & 0\\ \vdots & \ddots & \ddots & \ddots & \vdots\\ \vdots & & \ddots & \ddots & 0\\ \vdots & & & \ddots & \delta_{i_{r-1}i}t^{\rho_{r-1}}\\ 0 & \dots & \dots & \dots & 0\end{pmatrix},\qquad g_i(t)=\begin{pmatrix}0\\0\\ \vdots\\0\\ \delta_{i_r i}t^{\rho_r}\end{pmatrix},$$

where

$$\delta_{ij}=\begin{cases}1 & \text{if } i=j,\\ 0 & \text{if } i\neq j.\end{cases}$$

For this problem it is easy to check that all the FED of type $F_{\bar{i}_1,\dots,\bar{i}_{\bar{r}}}^{(\bar{\rho}_1,\dots,\bar{\rho}_{\bar{r}})}y(t)$ and $G_{\bar{i}_1,\dots,\bar{i}_{\bar{r}}}^{(\bar{\rho}_1,\dots,\bar{\rho}_{\bar{r}})}g(t)$ are null, for $t=0$, excepting the FED $G_{i_1,\dots,i_r}^{(\rho_1,\dots,\rho_r)}g(0)$ which takes in its first component the value $\rho_1!\dots\rho_r!$. Note that the necessary order condition to attain order p in the first component of y_1 is

$$(b^{i_1})^{\mathrm{T}}(\mathrm{C})^{\rho_1}\mathscr{A}^{i_2}(\mathrm{C})^{\rho_2}\dots\mathscr{A}^{i_r}(\mathrm{C})^{\rho_r}e=\prod_{j=1}^{r}\frac{1}{(r-j+1)+\sum_{k=j}^{r}\rho_k}. \qquad \square$$

It is easy to observe an important growth in the number of order conditions that an additive Runge–Kutta method must verify, with respect to the corresponding number of conditions for standard Runge–Kutta methods, as long as the order p is increased. This is a great drawback for designing high order methods. Nevertheless, in some practical cases, the number of order conditions can be reduced. This fact occurs when some FED are functionally identical or null and, in these cases, their corresponding order conditions can be added or superseded, respectively. For example:

(i) (a) If $L_j^{(\sigma)}(t)L_k^{(\upsilon)}(t)=L_k^{(\upsilon)}(t)L_j^{(\sigma)}(t)$ for some $j,k\in\{1,\dots,n\}$ and for some $\sigma,\upsilon\in\{0,\dots,p-1\}$ then we can add the order conditions (20) associated with the (identical) FED

$$F_{i_1,\dots,i_m,j,k,i_{m+3},\dots,i_r}^{(\rho_1,\dots,\rho_m,\sigma,\upsilon,\rho_{m+3},\dots,\rho_r)}y(t),\qquad F_{i_1,\dots,i_m,k,j,i_{m+3},\dots,i_r}^{(\rho_1,\dots,\rho_m,\upsilon,\sigma,\rho_{m+3},\dots,\rho_r)}y(t)$$

and

$$G^{(\rho_1,\dots,\rho_m,\sigma,\upsilon,\rho_{m+3},\dots,\rho_r)}_{i_1,\dots,i_m,j,k,i_{m+3},\dots,i_r} g(t), \qquad G^{(\rho_1,\dots,\rho_m,\upsilon,\sigma,\rho_{m+3},\dots,\rho_r)}_{i_1,\dots,i_m,k,j,i_{m+3},\dots,i_r} g(t).$$

(b) If besides $g_j^{(\sigma)}(t)=0$, $g_k^{(\upsilon)}(t)=0$, then we can also add the order conditions associated with the FED

$$F^{(\rho_1,\dots,\rho_{r-2},\sigma,\upsilon)}_{i_1,\dots,i_{r-2},j,k}\, y(t)$$

and

$$F^{(\rho_1,\dots,\rho_{r-2},\upsilon,\sigma)}_{i_1,\dots,i_{r-2},k,j}\, y(t).$$

(ii) (a) If $L_{\bar{i}}^{(\bar{\rho})}(t)\equiv 0$, for a certain $\bar{i}$ and for some $\bar{\rho}\geqslant 0$ then we can supersede the order conditions (20) such that $i_j=\bar{i}$, and its associated index ρ_j is bigger than or equal to $\bar{\rho}$ for any $j\neq r$.

(b) Besides if $g_{\bar{i}}^{(\bar{\rho})}(t)\equiv 0$, then we can exclude also the order conditions (20) such that $i_r=\bar{i}$ and $\rho_r\geqslant\bar{\rho}$.

3. Stability and convergence

To study the convergence of scheme (3) we define the global error as follows:

$$E^\tau = \sup_{m\leqslant T/\tau} \|y(t_m)-y_m\|;$$

and we say that an additive Runge–Kutta method is convergent of order p if $E^\tau\leqslant C\tau^p$ for data $L_i(t)\in\mathscr{C}^p([t_0,T];\mathbb{R}^{d\times d})$ and $g_i(t)\in\mathscr{C}^p([0,T];\mathbb{R}^d)$.

The convergence of scheme (3) is deduced by combining the property of consistency (10) with some stability results. In Section 1, we pointed out that the Additive Runge–Kutta methods that we use must verify certain properties of absolute stability if we want to obtain unconditional convergence.

In order to introduce easily the concept of A-stability for an additive Runge–Kutta method, we apply scheme (3) to the test scalar IVP

$$y'(t)=\sum_{i=1}^n \lambda_i y(t) \quad \text{with } \mathrm{Re}(\lambda_i)\leqslant 0,\ i=1,\dots,n,$$

$$y(t_0)=y_0 \tag{24}$$

and we obtain the recurrence

$$y_{m+1}=\left(1+\sum_{i=1}^n \tau\lambda_i (b^i)^{\mathrm{T}}\left(I-\sum_{j=1}^n \tau\lambda_j\mathscr{A}^j\right)^{-1} e\right) y_m; \tag{25}$$

substituting $\tau\lambda_i$ by z_i, we deduce the amplification function associated with scheme (3):

$$R(z_1,\dots,z_n)=1+\sum_{i=1}^n z_i\,(b^i)^{\mathrm{T}}\left(I-\sum_{j=1}^n z_j\,\mathscr{A}^j\right)^{-1} e.$$

We say that an additive Runge–Kutta method is A-stable if and only if the recurrence (25) preserves the contractive behaviour of (24), i.e., $|y_{m+1}| \leqslant |y_m|$ for all m. This condition is obviously equivalent to

$$|R(z_1,\dots,z_n)| \leqslant 1, \quad \forall (z_1,\dots,z_n) \in \mathbb{C}^n \text{ such that } \mathrm{Re}(z_i) \leqslant 0, \ \forall i=1,\dots,n.$$

Although the A-stability has been introduced in a simple scalar case, this property also ensures a contractive behaviour for some (vectorial) problems of type

$$y'(t) = \sum_{i=1}^{n} (L_i(t)y(t) + g_i(t)), \quad i=1,\dots,n,$$
$$y(t_0) = y_0.$$

Concretely in [14] the case $L_i(t) \equiv L_i$ is studied, for all $i=1,\dots,n$. In this case it holds that

$$y_{m+1} = R(\tau L_1,\dots,\tau L_n)y_m + S(\tau \hat{L}_1,\dots,\tau \hat{L}_n, \tau \hat{G}_1^m(\tau),\dots,\tau\, \hat{G}_n^m(\tau)),$$

where

$$S(\tau \hat{L}_1,\dots,\tau \hat{L}_n, \tau \hat{G}_1^m(\tau),\dots,\tau \hat{G}_n^m(\tau))$$
$$= \tau \sum_{i=1}^{n} (\overline{b^i})^{\mathrm{T}} \left(\hat{G}_i^m(\tau) + \hat{L}_i \left(\bar{I} - \tau \sum_{j=1}^{n} \overline{\mathscr{A}^j} \hat{L}_j \right)^{-1} \left(\tau \sum_{k=1}^{n} \overline{\mathscr{A}^k} \hat{G}_k^m(\tau) \right) \right)$$

and where now

$$\hat{L}_i = \begin{pmatrix} L_i & 0 & \dots & 0 \\ 0 & L_i & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & L_i \end{pmatrix};$$

the contractivity of scheme (3) applied to some of these problems is deduced easily from the following result.

Theorem 3.1. (Jorge and Lisbona [14]). *Let $\{\Lambda_i\}_{i\in\{1,\dots,n\}}$ be a commutative system of maximal coercive operators in a Hilbert space H, such that the commutative system of contractions $\{(I-\tau\Lambda_i)(I+\tau\Lambda_i)^{-1}\}_{i\in\{1,\dots,n\}}$ admits a unitary dilation. Then for all A-stable additive Runge–Kutta method it holds that*

$$\|R(\tau\Lambda_1,\dots,\tau\Lambda_n)\| \leqslant 1, \tag{26}$$

where $R(z_1,\dots,z_n)$ is the amplification function of such method.

Classically, when the operator L depends on t, i.e., $L \equiv L(t)$, AN-stability has been imposed to preserve contractivity. It is well known that this property is strongly restrictive since it is only verified by simple low-order methods or by high order fully implicit methods. Nevertheless, in [8], and more recently in [9], it is shown that, under suitable hypotheses of variation in time for operators

$L(t)$, A-stability can be a sufficient condition for a stable integration, at least in finite intervals of time $[t_0, T]$.

For FSRK methods a similar stability result is proved in [2]. Such a result is based on the following recurrence obtained from (9):

$$y_{m+1} = \tilde{R}(\tau\hat{L}_1^m(\tau), \dots, \tau\hat{L}_n^m(\tau))y_m + \tilde{S}(\tau\hat{L}_1^m(\tau), \dots, \tau\hat{L}_n^m(\tau), \tau\hat{G}_1^m(\tau), \dots, \tau\,\hat{G}_n^m(\tau)),$$

where

$$\tilde{R}(\tau\hat{L}_1^m(\tau), \dots, \tau\hat{L}_n^m(\tau)) = \bar{I} + \sum_{i=1}^{n} (\overline{b^i})^{\mathrm{T}} \tau\hat{L}_i^m(\tau) \left(\bar{I} - \sum_{j=1}^{n} \overline{\mathscr{A}^j} \tau\hat{L}_j^m(\tau) \right)^{-1} \bar{e},$$

can be considered as a perturbation of $R(\tau L_1(t_m), \dots, \tau L_n(t_m))$ and

$$\tilde{S}(\tau\hat{L}_1^m(\tau), \dots, \tau\hat{L}_n^m(\tau), \tau\hat{G}_1^m(\tau), \dots, \tau\hat{G}_n^m(\tau))$$

$$= \tau\sum_{i=1}^{n} (\overline{b^i})^{\mathrm{T}} \left(\hat{G}_i^m(\tau) + \hat{L}_i^m(\tau) \left(\bar{I} - \tau\sum_{j=1}^{n} \overline{\mathscr{A}^j} \hat{L}_j^m(\tau) \right)^{-1} \left(\tau\sum_{k=1}^{n} \overline{\mathscr{A}^k} \hat{G}_k^m(\tau) \right) \right)$$

is the contribution of the source term.

Theorem 3.2. (Bujanda and Jorge [2]). *Let (9) be an A-stable FSRK method, and let $\{L_i(t)\}_{i=1}^n$ be a linear maximal coercive system of operators satisfying*:

(a) *for every $t \in [t_0, T]$ the system of operators $\{L_i(t)\}_{i=1}^n$ is commutative and the commutative system of contractions $\{(I - \tau L_i(t))(I + \tau L_i(t))^{-1}\}_{i \in \{1,\dots,n\}}$ admits a unitary dilation,*

(b) *there exist n constants M_i such that*

$$\|L_i(t')y - L_i(t)y\| \leqslant |t - t'| M_i \|L_i(t)y\|, \quad \forall i = 1, \dots, n, \ \forall t, t' \in [t_0, T].$$

Then there exists a constant β, usually positive [3], *independent of τ, such that*

$$\|\tilde{R}(-\tau\hat{L}_1^m(\tau), \dots, -\tau\hat{L}_n^m(\tau))\| \leqslant e^{\beta\tau}. \tag{27}$$

From the consistency result (10) and the stability result (26) or (27) it is easy to prove

$$\|y(t_m) - y_m\| \leqslant \|y(t_m) - \hat{y}_m\| + \|\hat{y}_m - y_m\| \leqslant C\tau^{p+1} + e^{\beta\tau}\|y(t_{m-1}) - y_{m-1}\|.$$

Using this recurrence we obtain the convergence of scheme (3) since

$$\|y(t_m) - y_m\| \leqslant C\sum_{k=0}^{m-1} e^{\beta\tau k}\tau^{p+1} \leqslant CM\tau^p,$$

[3] Some additional A-stability conditions, of type strong A-stability, together with small variations in time for the operators $L_i(t)$ allow to obtain negative values for β and, consequently, preserve a contractive behaviour on the numerical solutions of scheme (3).

where

$$M = \begin{cases} \dfrac{e^{T\beta}-1}{\beta} & \text{if } \beta > 0, \\ T & \text{if } \beta = 0, \\ \tau_0 \dfrac{1-e^{T\beta}}{1-e^{\beta\tau_0}} \ \forall \tau \in (0, \tau_0], & \text{if } \beta < 0. \end{cases}$$

4. A third-order FSRK method and some numerical tests

4.1. A third-order FSRK method

In this subsection, we summarize the main lines that we have followed to construct an L-stable third-order FSRK method of six stages and two levels (more details of such construction can be seen in [1]) with a particular coefficient structure following the ideas of alternating direction methods.

Definition 4.1. An FSRK method of two levels is of type alternating directions if $b_k^1 = a_{ik}^1 = 0$ when k is even and $b_k^2 = a_{ik}^2 = 0$ when k is odd.

In order to use some previous stability studies as well as to simplify the construction process of this method we have firstly imposed that the nonnull diagonal terms are equal, i.e., $a_{ii}^1 = a_{jj}^2 = a$ for $i = 1, 3, 5$ and for $j = 2, 4, 6$, (in [1] it can be seen that to obtain a third-order method at least five stages are necessary and six are convenient).

From (20), we deduce that the number of necessary order conditions to attain order 3 are 26. We have grouped these conditions in such a way that in the first group only the superindex 1 appears, in the second group only the superindex 2 appears and finally in the third group we have put the remaining equations. The seven equations of the first group correspond to the study of the order 3 conditions for the reduced SDIRK

$$\begin{array}{c|ccc} c_1 & a & & \\ c_3 & a_{31}^1 & a & \\ c_5 & a_{51}^1 & a_{53}^1 & a \\ \hline & b_1^1 & b_3^1 & b_5^1 \end{array}$$

(see [10]), and analogously the second group correspond to the study of the method

$$\begin{array}{c|ccc} c_2 & a & & \\ c_4 & a_{42}^2 & a & \\ c_6 & a_{62}^2 & a_{64}^2 & a \\ \hline & b_2^2 & b_4^2 & b_6^2. \end{array}$$

We solve these two groups of equations taking as free parameters a, c_3, c_4, c_5 and c_6.

To finish the search of the family of third-order methods, we solve the remaining 12 order conditions, obtaining a family of methods with the following free parameters: a, c_3, c_4, c_6 and a^1_{43}.

Some of these parameters are used to obtain an L-stable method. In order to reduce the study of absolute stability of these methods, as far as possible, to the study of the absolute stability of the standard RK methods

$$\begin{array}{c|c} Ce & \mathscr{A}^1 \\ \hline & (b^1)^{\mathrm{T}} \end{array}$$

and

$$\begin{array}{c|c} Ce & \mathscr{A}^2 \\ \hline & (b^2)^{\mathrm{T}} \end{array},$$

we decompose their amplification functions as follows:

$$R(z_1,z_2) = R_1(z_1)R_2(z_2) + \text{Rest}, \tag{28}$$

where $R_i(z_i)$ is the amplification function associated with

$$\begin{array}{c|c} Ce & \mathscr{A}^i \\ \hline & (b^i)^{\mathrm{T}} \end{array}$$

for $i = 1,2$ and

$$\text{Rest} = \frac{F_1 z_1 z_2^3 + F_2 z_1^2 z_2^2 + F_3 z_1^3 z_2 + G_2 z_1^2 z_2^3 + G_3 z_1^3 z_2^2 + H_3 z_1^3 z_2^3}{(1 - a z_1)^3 (1 - a z_2)^3},$$

with

$$F_j = E_{j,4-j} - \frac{1}{j!(4-j)!}, \quad j = 1,2,3,$$

$$G_j = E_{j,5-j} - 3a\,(F_{j-1} + F_j) - \frac{1}{12}, \quad j = 1,2,$$

$$H_3 = E_{3,3} + 3a^2(F_1 + F_3) + 9a^2 F_2 - 3a(G_1 + G_2) - \frac{1}{36},$$

$$E_{i_1,i_2} = \sum_{\substack{\bar{k} \equiv (k_1,\dots,k_j) \in \{1,2\}^j \\ n_l(\bar{k}) = i_l, \forall l = 1,2}} (b^{k_1})^{\mathrm{T}} \mathscr{A}^{k_2} \dots \mathscr{A}^{k_j} e$$

and $n_l(\bar{k})$ is the number of times that the index l appears in the multiindex $\bar{k}$.

From (28) it is easy to see that if we make null the Rest and the RK methods

$$\begin{array}{c|c} Ce & \mathscr{A}^i \\ \hline & (b^i)^{\mathrm{T}} \end{array}$$

for $i=1,2$ are A-stable, then the alternating direction implicit method will be A-stable. In [10] it can be seen that if $a=0.435866521508459$ the RK methods

$$\begin{array}{c|c} Ce & \mathscr{A}^i \\ \hline & (b^i)^{\mathrm{T}} \end{array}$$

for $i=1,2$ are not only A-stable but also L-stable.

For this family of third-order methods, we have seen that a unique choice of a^1_{43}, as function of the parameters c_3, c_4 and c_6, make null the terms F_1, F_3, G_2, G_3 and H_3, but then F_2 takes the constant value of -0.474949. Nevertheless, this election of a^1_{43} is sufficient to get an L-stable alternating direction implicit method (for more details see [1]).

The remaining free parameters c_3, c_4 and c_6 have been used to minimize the main term of the local truncation error. From this minimization process, we have obtained the following alternating direction implicit method:

$$Ce=(0.435866521508459,\ 0.435866521508459,\ 0.7,\ 0.3,\ 0.56413347849154,\ 0.75)^{\mathrm{T}}, \tag{29}$$

$$\left(\frac{\mathscr{A}^1}{(b^1)^{\mathrm{T}}}\right)=\left(\begin{array}{llllll} 0.435866521508459 & & & & & \\ 0.435866521508459 & 0 & & & & \\ 0.264133478491540 & 0 & 0.435866521508459 & & & \\ 0.524203567293128 & 0 & -0.22420356729312 & 0 & & \\ 0.054134244066592 & 0 & 0.074132712916489 & 0 & 0.435866521508459 & \\ 2.005981609913539 & 0 & 1.336337252930893 & 0 & -2.59231886284469 & 0 \\ \hline 2.838287230686191 & 0 & 2.207497360663944 & 0 & -4.04578459135012 & 0 \end{array}\right) \tag{30}$$

$$\left(\frac{\mathscr{A}^2}{(b^2)^{\mathrm{T}}}\right)=\left(\begin{array}{llllll} 0 & & & & & \\ 0 & 0.435866521508459 & & & & \\ 0 & 0.170931386851894 & 0 & & & \\ 0 & -0.13586652150846 & 0 & 0.435866521508459 & & \\ 0 & 0.062944816984284 & 0 & -0.09326511998115 & 0 & \\ 0 & -0.54301448024727 & 0 & 0.857147958738813 & 0 & 0.435866521508459 \\ \hline 0 & -0.78100185474576 & 0 & 1.100752954088072 & 0 & 0.680248900657693 \end{array}\right).$$

4.2. Numerical tests

In this section, we will perform some numerical experiences with the alternating direction implicit method (29)–(30) combined with a simple spatial discretization of the following convection–diffusion–reaction problem:

$$\frac{\partial u}{\partial t}-d_1\frac{\partial^2 u}{\partial x^2}-d_2\frac{\partial^2 u}{\partial y^2}+v_1\frac{\partial u}{\partial x}+v_2\frac{\partial u}{\partial y}+k_1u+k_2u=g_1+g_2,$$

$$(x,y,t)\in\Omega\times[0,5]\ \text{ with }\Omega=[0,1]\times[0,1],$$

$$u(0,y,t)=u(1,y,t)=0, \quad y\in[0,1], \ t\in[0,5],$$

$$u(x,0,t)=u(x,1,t)=0, \quad x\in[0,1], \ t\in[0,5],$$

$$u(x,y,0)=e^{(x+y)/2}\sin(\pi x)\sin(\pi y), \ x,y\in\Omega,$$

where $d_1=(1-e^{-t})(1+x)$, $d_2=(2+\cos(\pi t))(1+y)$, $v_1=(2+\sin(\pi t))(1+y)$, $v_2=(2-e^{-t})(1+x)$, $k_1=(1+e^{-t})(2-xy)$, $k_2=(1+e^{-t})(1+x^2)$, $g_1=e^{-t}x(1-x)y(1-y)e^{(x+y)/2}$ and $g_2=e^{-t}x(1-x)y(1-y)e^{(x+y)/2}$.

To numerically integrate this problem we have realized firstly a spatial semidiscretization by using a simple upwind finite difference scheme on a uniform mesh Ω_h with $(N+1)\times(N+1)$ points being $h=1/N$ and obtaining the following family of IVPs:

$$\frac{\mathrm{d}u_{i,j}(t)}{\mathrm{d}t}+A_{1h}u_{i,j}(t)+A_{2h}u_{i,j}(t)=g_1(x_i,y_j,t)+g_2(x_i,y_j,t),$$

$$u_{0,j}(t)=u_{N,j}(t)=0, \quad j=0,\ldots,N, \ t\in[0,5],$$

$$u_{i,0}(t)=u_{i,N}(t)=0, \quad i=0,\ldots,N, \ t\in[0,5],$$

$$u_{i,j}(0)=u(x_i,y_j,0), \quad i,j=0,\ldots,N, \tag{31}$$

where

$$A_{1h}u_{i,j}(t)=-d_1(x_i,y_j,t)\frac{u_{i-1,j}(t)-2u_{i,j}(t)+u_{i+1,j}(t)}{h^2}$$
$$+v_1(x_i,y_j,t)\frac{u_{i,j}(t)-u_{i-1,j}(t)}{h}+k_1(x_i,y_j,t)u_{i,j}(t), \quad i,j=1,\ldots,N-1,$$

$$A_{2h}u_{i,j}(t)=-d_2(x_i,y_j,t)\frac{u_{i,j-1}(t)-2u_{i,j}(t)+u_{i,j+1}(t)}{h^2}$$
$$+v_2(x_i,y_j,t)\frac{u_{i,j}(t)-u_{i,j-1}(t)}{h}+k_2(x_i,y_j,t)u_{i,j}(t), \quad i,j=1,\ldots,N-1,$$

with $x_i=i\,h$, for $i=0,\ldots,N$ and $y_j=j\,h$ for $j=0,\ldots,N$.

Secondly, we have made the time discretization by using the FSRK method given by (29) and (30). In Table 1, we show the obtained maximum nodal errors $E_{N\tau}$ from the time instant $t=0.05$ for different mesh sizes $1/N$ and for different values of τ, which are computed as

$$E_{N\tau}=\max_{\substack{i,j=1,\ldots,N-1,\\ t_m\equiv m\tau\in[0.05,5]}}\|\tilde{u}_{i,j}(t_m)-u^m_{i,j}\|,$$

where $\tilde{u}_{i,j}(t_m)$ is the solution of the semidiscrete scheme in space (31) that we have estimated by using a fourth-order L-stable SDIRK method of five stages (see [10]) with enough small time step. The value of τ_0 is $5E-2$.

In Table 2 we show the numerical orders of convergence $p_{N\tau}$ that have been computed as follows:

$$p_{N\tau}=\log_2\frac{E_{N\tau}}{E_{N\tau/2}}$$

Table 1
Maximum nodal errors $E_{N\tau}$

$E_{N\tau}$	$N=8$	$N=16$	$N=32$	$N=64$	$N=128$
$\tau=\tau_0$	4.5413 E-2	4.6690 E-2	4.6699 E-2	4.6729 E-2	4.6729 E-2
$\tau=\frac{\tau_0}{2}$	4.3538 E-3	4.4508 E-3	4.4696 E-3	4.4860 E-3	4.4928 E-3
$\tau=\frac{\tau_0}{2^2}$	6.0788 E-4	6.1223 E-4	6.1652 E-4	6.1806 E-4	6.1906 E-4
$\tau=\frac{\tau_0}{2^3}$	8.3720 E-5	8.4310 E-5	8.4889 E-5	8.5100 E-5	8.5234 E-5
$\tau=\frac{\tau_0}{2^4}$	1.1149 E-5	1.1229 E-5	1.1307 E-5	1.1335 E-5	1.1352 E-5
$\tau=\frac{\tau_0}{2^5}$	1.4453 E-6	1.4561 E-6	1.4662 E-6	1.4698 E-6	1.4719 E-6
$\tau=\frac{\tau_0}{2^6}$	1.8425 E-7	1.8564 E-7	1.8694 E-7	1.8739 E-7	1.8766 E-7
$\tau=\frac{\tau_0}{2^7}$	2.3268 E-8	2.3443 E-8	2.3608 E-8	2.3665 E-8	2.3699 E-8

Table 2
Numerical orders of convergence $p_{N\tau}$

$p_{N\tau}$	$N=8$	$N=16$	$N=32$	$N=64$	$N=128$
$\tau=\tau_0$	3.3827	3.3909	3.3851	3.3808	3.3786
$\tau=\frac{\tau_0}{2}$	2.8404	2.8619	2.8579	2.8596	2.8595
$\tau=\frac{\tau_0}{2^2}$	2.8601	2.8603	2.8605	2.8605	2.8605
$\tau=\frac{\tau_0}{2^3}$	2.9086	2.9084	2.9083	2.9084	2.9084
$\tau=\frac{\tau_0}{2^4}$	2.9474	2.9472	2.9471	2.9471	2.9471
$\tau=\frac{\tau_0}{2^5}$	2.9716	2.9715	2.9715	2.9715	2.9715
$\tau=\frac{\tau_0}{2^6}$	2.9852	2.9853	2.9852	2.9852	2.9852

References

[1] B. Bujanda, Métodos Runge–Kutta de Pasos Fraccionarios de orden alto para la resolución de problemas evolutivos de convección-difusión-reacción, Tesis, Universidad Pública de Navarra, Secc. 2, n. 15, 1999, pp. 1–174.
[2] B. Bujanda, J.C. Jorge, Stability results for fractional step discretizations of time dependent coefficient evolutionary problems, Appl. Numer. Math. 38 (2001) 69–86.
[3] B. Bujanda, J.C. Jorge, Third order fractional step methods for multidimensional evolutionary convection–diffusion problems, Finite Difference Methods: Theory and Applications, Nova Science, New York 1999, pp. 110–117.
[4] J.C. Butcher, The Numerical Analysis of Ordinary Differential Equations, Wiley, New York, 1987.
[5] C. Clavero, J.C. Jorge, F. Lisbona, G.I. Shishkin, A fractional step method on a special mesh for the resolution of multidimensional evolutionary convection–diffusion problems, Appl. Numer. Math. 27 (1998) 211–231.
[6] G.J. Cooper, A. Sayfy, Aditive methods for the numerical solution of ordinary differential equations, Math. Comp. 35 (1980) 1159–1172.
[7] G.J. Cooper, A. Sayfy, Additive Runge–Kutta methods for stiff ordinary differential equations, Math. Comp. 40 (1983) 207–218.
[8] M. Crouzeix, Sur l'aproximation des èquations differentielles opèrationelles linèaires par des mèthodes de Runge–Kutta, These d'Etat, Univ. de Paris VI, 1975.
[9] C. González, C. Palencia, Stability of Runge–Kutta methods for abstract time-dependent parabolic problems: the Hölder case, Preprint of Departamento de Matemática Aplicada y Computación, Facultad de Ciencias, Universidad de Valladolid, 1996.
[10] E. Hairer, G. Wanner, Solving ordinary differential equations II, Springer, Berlin, 1987.
[11] W. Hundsdorfer, Trapezoidal and midpoint splittings for initial-boundary value problems, Math. Comp. 67 (233) (1998) 1047–1062.
[12] J.C. Jorge, Los métodos de pasos fraccionarios para la integración de problemas parabólicos lineales: formulación general, análisis de la convergencia y diseño de nuevos métodos, Tesis, Universidad de Zaragoza, 1992.
[13] J.C. Jorge, Métodos de Runge–Kutta adtitivos para la integración de ecuaciones diferenciales ordinarias stiff, Publicaciones del seminario matemático García de Galdeano, Universidad de Zaragoza, Secc. 7, n. 3, 1990, pp. 1–73.
[14] J.C. Jorge, F. Lisbona, Contractivity results for alternating direction schemes in Hilbert spaces, Appl. Numer. Math. 15 (1994) 65–75.
[15] D.W. Peaceman, H.H. Rachford, The numerical solution of parabolic and elliptic differential equations, J. SIAM 3 (1955) 28–42.
[16] J.M. Sanz-Serna, The numerical integration of Hamiltonian systems, Proceedings of the Conference on Computational Differential Equations, Imperial College, London, 3rd–7th, July, 1989, pp. 1–14.
[17] P.J. van der Houwen, The development of Runge–Kutta methods for partial differential equations, Appl. Numer. Math. 20 (1996) 261–272.
[18] N.N. Yanenko, The Method of Fractional Steps, Springer, Berlin, 1971.