Available at WWW.MATHEMATICSWEB.ORG
POWERED BY SCIENCE DIRECT

Applied Numerical Mathematics 45 (2003) 99–122

APPLIED NUMERICAL MATHEMATICS

www.elsevier.com/locate/apnum

Fractional Step Runge–Kutta methods for time dependent coefficient 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*

Abstract

We study the uniform convergence for general Fractional Step Runge–Kutta methods in the integration of a class of evolution problems which includes linear parabolic problems whose coefficients depend on time. Such analysis is performed by suitably decomposing the contribution to the global error of this time integration procedure and the contribution of some standard spatial discretization methods.

© 2002 IMACS. Published by Elsevier Science B.V. All rights reserved.

Keywords: Additive Runge–Kutta; Uniform convergence; Fractional Step Runge–Kutta

1. Introduction

This paper is devoted to present and analyze a class of efficient and robust methods which includes the classical Splitting methods (see [15,17]), Alternating Direction schemes (see [13,19]) or, more generally Fractional Step methods (see [9]) used to discretize efficiently some parabolic problems.

From the classical methods (see [26]) until the most recent ones (see [14,16]), every method of this type has been designed and analyzed separately, in close relation to the specific continuous problem whose resolution is required. In [18], a framework to analyze this kind of methods for general parabolic problems, as well as to develop new ones of higher orders, is proposed. The analysis carried out there only covers the case of time independent coefficients. Here we extend such ideas to the case of evolution problems with time dependent coefficients.

☆ This research is partially supported by the DGES research project num. PB97-1013, BFM2000-0803 and research project No. API-00/A24 of the University of La Rioja.

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

0168-9274/02/$30.00 © 2002 IMACS. Published by Elsevier Science B.V. All rights reserved.
doi:10.1016/S0168-9274(02)00191-5

The main idea for analyzing, in a unified way, the Splitting, Alternating Direction or Fractional Step methods is based in the fact that these methods are composed by a standard spatial discretization scheme of type Finite Differences or Finite Elements and a special class of Additive Runge–Kutta methods to advance in time, called Fractional Step Runge–Kutta methods. For these methods we give suitable, easy to check, consistency and linear absolute stability properties which will be sufficient to ensure efficiency and robustness for the integration processes described here.

In order to get results for a wide set of problems, we will consider an abstract linear evolution problem whose coefficients may depend on time and we will obtain some uniform convergence results for the totally discrete schemes derived from the combination of Fractional Step Runge–Kutta methods and some standard discretization methods for elliptic problems of type Finite Differences or Finite Elements. To do this, we will follow one of the techniques which Jorge proposed in [18] for studying the convergence of Fractional Step schemes for linear evolution problems with constant coefficients.

Let $u(t)$ be solution of an evolution problem with time dependent coefficients which admits an operational formulation as follows:

$$\begin{cases} \dfrac{\mathrm{d}u(t)}{\mathrm{d}t} = -L(t)\,u(t) + g(t), \\ u(t_0) = u_0; \end{cases} \tag{1}$$

here $L(t):\mathcal{D}(L(t)) \subseteq H \to H$ are, generally, unbounded differential operators[1] defined on a domain $\mathcal{D}(L(t))$ dense on a Hilbert space H, with scalar product $((\cdot,\cdot))$ and with associated norm $\|\cdot\|$. Some results concerning the existence, uniqueness and smoothness in time for $u(t)$ can be consulted in [5,12, 21]. For carrying out our analysis, we will assume that for every t, that the operators $L(t)$ are coercive and maximal, i.e.,

$$\begin{cases} \big(\big(L(t)v, v\big)\big) \geqslant \alpha\|v\|^2, \quad \text{with } \alpha > 0,\ \forall\, v \in \mathcal{D}\big(L(t)\big) \\ \text{and } \forall\, f \in H, \exists v \in \mathcal{D}\big(L(t)\big), \quad \text{such that } v + L(t)v = f, \end{cases}$$

and also that the operators $L(t)$ are defined on the same domain $\mathcal{D}(L(t)) \equiv \mathcal{D} \subseteq H$ for all $t \in [t_0, T]$. The practical meaning of this restriction is that only time independent boundary conditions are included in the scope of our study.

A discretization of problem (1) by using a Fractional Step scheme departs from a natural decomposition of the operators $L(t)$ in n simpler addends:

$$L(t) = \sum_{i=1}^{n} L_i(t), \tag{2}$$

where the operators $L_i(t):\mathcal{D}_i \subseteq H \to H$ preserve some properties of $L(t)$; concretely, we will assume that:

$$\begin{cases} \forall i = 1,\dots,n, \exists \alpha_i > 0 \quad \text{such that } \big(\big(L_i(t)v, v\big)\big) \geqslant \alpha_i\|v\|^2, \forall v \in \mathcal{D}_i \\ \text{and } \forall\, f \in H, \exists v \in \mathcal{D}_i, \quad \text{such that } v + L_i(t)v = f, \end{cases} \tag{3}$$

and also $\mathcal{D} = \bigcap_{i=1}^{n} \mathcal{D}_i$.

[1] In the case of (1) is a parabolic problem, the operators $L(t)$, which contain the spatial derivative terms, will be elliptic operators for all $t \in [t_0, T]$.

For example, if we consider an n-dimensional convection–diffusion problem of type $\frac{\partial u}{\partial t} = \mathrm{d}\Delta u - \vec{v}\vec{\nabla} u - ku + g$, typically

$$L_i u \equiv -\mathrm{d}\frac{\partial^2 u}{\partial x_i^2} + v_i \frac{\partial u}{\partial x_i} + k_i u,$$

for $i = 1, \ldots, n$ being $k_1 + \cdots + k_n = k$ (where n is the number of the spatial variables); thus, each operator L_i contains only derivatives with respect to one of the spatial variables.

We will also consider a decomposition for the source term: $g(t) = \sum_{i=1}^{n} g_i(t)$, with $g_i : [t_0, T] \to H$. Using this decomposition together with (2) we will rewrite the problem (1) as follows:

$$\begin{cases} \dfrac{\mathrm{d}u(t)}{\mathrm{d}t} = -\displaystyle\sum_{i=1}^{n} \big(L_i(t)u(t) - g_i(t)\big), \\ u(t_0) = u_0 \in \mathcal{D}, \end{cases} \tag{4}$$

a Fractional Step Runge–Kutta method (abbreviated as FSRK) to discretize in time problem (4) is an algorithm structured in the following way:

$$\begin{cases} U^0 = u_0(\in \mathcal{D}), \\ U^{m+1} = U^m - \tau \displaystyle\sum_{j=1}^{s} b_j^{i_j} \big(L_{i_j}(t_{m,j})U^{m,j} - g_{i_j}(t_{m,j})\big), \\ \quad \text{where } U^{m,j} = U^m - \tau \displaystyle\sum_{k=1}^{j} a_{jk}^{i_k} \big(L_{i_k}(t_{m,k})U^{m,k} - g_{i_k}(t_{m,k})\big), \quad \text{for } j = 1, \ldots, s, \end{cases} \tag{5}$$

here, τ denotes the time step which we have chosen constant for simplifying our presentation, U^{m+1} will be approximations to the exact solution $u(t)$ at the instants $t_{m+1} = (m+1)\tau$, for all $m = 0, 1, \ldots$ and $U^{m,j}$ can be considered approximations to the exact solution at the intermediate times, $t_{m,j} = (m + c_j)\tau$, which are usually called stages of the method.

A FSRK method is determined by the coefficients $a_{kj}^{i_j}, b_j^{i_j}$ and c_i with $j, k = 1, \ldots, s$ and $i_j = 1, \ldots, n$. We will assume that a choice of these coefficients determines a unique FSRK method, although different partitions on the derivative function $-L(t)u(t) + g(t)$ can give different numerical approaches U^m. In order to connect this type of method with other one step methods of type Runge–Kutta we want to remark that the FSRK methods form a special subset of semiexplicit Additive Runge–Kutta methods which were introduced by Cooper and Sayfy in [10,11] and generalized by Jorge in [18] in the following form:

$$\begin{cases} U^0 = u_0, \\ U^{m+1} = U^m - \tau \displaystyle\sum_{i=1}^{n}\sum_{j=1}^{s} b_j^{i} \big(L_i(t_{m,j})U^{m,j} - g_i(t_{m,j})\big), \\ \quad \text{where } U^{m,j} = U^m - \tau \displaystyle\sum_{i=1}^{n}\sum_{k=1}^{s} a_{jk}^{i} \big(L_i(t_{m,k})U^{m,k} - g_i(t_{m,k})\big), \quad \text{for } j = 1, \ldots, s. \end{cases} \tag{6}$$

It is clear that if we impose on these methods the restrictions:

$$\begin{cases} a^i_{jk} = 0,\ \forall i = 1, \dots, n, \quad \text{if } k > j, \\ \left|b^i_k\right| + \sum\limits_{j=1}^{s} \left|a^i_{jk}\right| \neq 0 \quad \text{for some } i \in \{1, \dots, n\} \\ \quad \Longrightarrow \sum\limits_{\substack{l=1 \\ l\neq i}}^{s} \left(\left|b^l_k\right| + \sum\limits_{j=1}^{s} \left|a^l_{jk}\right| \right) = 0, \quad \forall k = 1, \dots, s, \end{cases} \tag{7}$$

we obtain a scheme of type (5) where i_j for all $j = 1, \dots, s$, are the indexes which verify the property $\sum_{j=1}^{s} \sum_{\substack{l=1 \\ l\neq i_j}}^{n} (|b^l_j| + \sum_{i=1}^{s} |a^l_{ij}|) = 0$.

It is not difficult to see that an additive Runge–Kutta method involves n overlapped standard Runge–Kutta schemes of type:

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

where $\mathrm{A}^i = (a^i_{jk})$, $\mathrm{b}^i = (b^i_j)$ for $i \in \{1, \dots, n\}$, $\mathrm{C} = \mathrm{diag}(c_1, \dots, c_s)$ and $e = (1, \dots, 1)^{\mathrm{T}} \in \mathbb{R}^s$, in such way that each one of them determines the contribution of the term $-(L_i(t)u - g_i(t))$ in the numerical resolution of (1). In order to use a compact notation of type Butcher table for these methods we will organize the coefficients of them as follows:

$$\begin{array}{c|c|c|c|c} \mathrm{C}e & \mathrm{A}^1 & \mathrm{A}^2 & \dots & \mathrm{A}^n \\ \hline & (\mathrm{b}^1)^{\mathrm{T}} & (\mathrm{b}^2)^{\mathrm{T}} & \dots & (\mathrm{b}^n)^{\mathrm{T}} \end{array} \tag{8}$$

This table can be reduced for an Additive method of type FSRK by avoiding the null columns in the form:

$$\begin{array}{c|c} & \boldsymbol{i}^{\mathrm{T}} \\ \hline \mathrm{C}e & \mathrm{A} \\ \hline & \mathrm{b}^{\mathrm{T}} \end{array}$$

where $\mathrm{A} = \sum_{i=1}^{n} \mathrm{A}^i$, $\mathrm{b}^{\mathrm{T}} = \sum_{i=1}^{n} (\mathrm{b}^i)^{\mathrm{T}}$ and $\boldsymbol{i}^{\mathrm{T}} = (i_1, \dots, i_s)$, being $i_j \in \{1, \dots, n\}$ the index of the unique non-null jth column of the extended matrices

$$\begin{pmatrix} \mathrm{A}^i \\ \hline (\mathrm{b}^i)^{\mathrm{T}} \end{pmatrix} \in \mathbb{R}^{s+1\times s}.$$

In scheme (5) we can observe that the use of FSRK methods for discretizing in time problem (1) reduces it to a family of elliptic problems, one problem per stage, in the form:

$$\big(I + \tau k L_{i_j}(t_{m,j})\big) U^{m,j} = F_j, \tag{9}$$

where F_j only contains evaluations of $g_i(t)$ and terms which are explicitly computed from the previous stages, and where $L_i(t)$ can be simpler than $L(t)$. For example, for convection–diffusion problems, if every operator $L_i(t)$ contains only derivatives with respect to one of the spatial variables, then problem (9) is essentially one-dimensional; consequently, the analytical or numerical resolution of (5) requires simpler methods than the classical time integrators (for example, semiexplicit RK methods) used to discretize (1), because the calculus of the stages with them would involve multidimensional elliptic problems of type $(I + \tau k L(t_{m,j}))U^{m,j} = F_j$.

To analyze the convergence of the discretization in time, we combine consistency properties which we will study here with the stability properties obtained in [6,7]. In this context, the linear absolute stability is a suitable property to reach unconditional convergence.

To finish our study we carry out the discretization in space of the family of elliptic problems resulting from the time discretization process (6) with the restrictions (7). To set the properties of this combination we introduce an abstract formulation for a general discretization method developed by Vainikko (see [25]), which includes Finite Differences and Finite Elements and we deduce some τ-independent convergence results for this spatial discretization stage.

The rest of the paper is structured in five sections. In the following section we show some results which ensure that the time discretization procedure has unique solution (U^m, $m = 0, \dots, T/\tau$) and it is stable. As well, we give some consistency results for the discretization in time and we use such results joint to the stability ones developed in [6,7] to prove that it is convergent.

In Section 3 we introduce a general abstract formulation for spatial discretization of (6), (7) and we study the convergence of the derived totally discrete scheme.

Section 4 contains the proofs of the main theorems of Sections 2 and 3 as well as some previous technical lemmas together with their proofs.

Finally, in Section 5 we show two numerical tests where we check the results obtained in this paper.

2. Convergence of the discretization in time

In this section we will denote by C a generic constant independent of τ.

In order to show our analysis in a shorter form we will use the following tensorial notations:

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

$$\overline{M} \equiv \begin{pmatrix} m_{11} I_H & \dots & m_{1s} I_H \\ \vdots & \ddots & \vdots \\ m_{s1} I_H & \dots & m_{ss} I_H \end{pmatrix} \in H^{s\times s} \quad \text{and} \quad \bar{v} \equiv \begin{pmatrix} v_1 I_H \\ \vdots \\ v_s I_H \end{pmatrix} \in H^s, \tag{10}$$

where I_H is the identity in H,

$$\hat{L}_i^m(\tau) = \begin{pmatrix} L_i(t_{m,1}) & 0 & \dots & 0 \\ 0 & L_i(t_{m,2}) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & L_i(t_{m,s}) \end{pmatrix} \in \mathcal{L}(\mathcal{D}_i, H)^{s\times s}, \quad \forall i = 1, \dots, n,$$

$$\widehat{U}^m = \left(U^{m,1}, \dots, U^{m,s}\right)^{\mathrm{T}} \in H^s, \tag{11}$$

$$\widehat{G}_i^m(\tau) = \left(g_i(t_{m,1}), \dots, g_i(t_{m,s})\right)^{\mathrm{T}} \in H^s, \quad \forall i = 1, \dots, n.$$

By using these notations we can write scheme (6) as follows:

$$\begin{cases} U^{m+1} = U^m - \tau \sum_{i=1}^{n} \overline{(\mathrm{b}^i)}^{\mathrm{T}} \big(\hat{L}_i^m(\tau)\widehat{U}^m - \widehat{G}_i^m\big), \\ \quad \text{where } \widehat{U}^m = \bar{e}U^m - \tau \sum_{i=1}^{n} \overline{\mathrm{A}^i}\big(\hat{L}_i^m(\tau)\widehat{U}^m - \widehat{G}_i^m\big). \end{cases}$$

In [7] it is proven that the operator $(\bar{I} + \tau \sum_{i=1}^{n} \overline{\mathrm{A}^i}\hat{L}_i^m(\tau)) : \mathcal{D}_{k_1} \times \cdots \times \mathcal{D}_{k_s} \to H^s$ is invertible and its inverse operator $(\bar{I} + \tau \sum_{i=1}^{n} \overline{\mathrm{A}^i}\hat{L}_i^m(\tau))^{-1} : H^s \to H^s$ is bounded[2] independently of $\tau \in (0, \tau_0]$ if the FSRK method has all its stages implicit, i.e.,

$$a_{ii}^{k_i} > 0, \quad \text{for all } i = 1, \ldots, s, \tag{12}$$

and if $\{L_i(t)\}_{i=1}^n$ is a system of operators fulfilling (3). As many of the classical Alternating Direction or Fractional Step methods are designed in such way that their formulation as FSRK method has a first explicit stage and the last stage gives directly U^{m+1}, i.e.,

$$\begin{aligned} &a_{11}^{k_1} = 0 \quad \text{and} \quad a_{ii}^{k_i} > 0 \quad \text{for all } i = 2, \ldots, s, \\ &(0, \ldots, 0, 1)\mathrm{A}^i = \big(\mathrm{b}^i\big)^{\mathrm{T}} \quad \text{for all } i = 1, \ldots, n, \end{aligned} \tag{13}$$

together with the additional property $a_{ss}^{k_1} \neq 0$, we have studied also this case. It is proven in [7] that, in this case, the operator $(\bar{I} + \tau \sum_{i=1}^{n} \overline{\mathrm{A}^i}\hat{L}_i^m(\tau)) : \mathcal{D}_{k_1} \times \cdots \times \mathcal{D}_{k_s} \to \mathcal{D}_{k_1} \times H^{s-1}$ is invertible and its inverse is bounded independently of τ.

The inversibility of these operators, together with notations (10), (11) permit us to write the solution of scheme (6), (7) as follows:

$$\begin{aligned} U^{m+1} = {} & \widetilde{R}\big(-\tau\hat{L}_1^m(\tau), \ldots, -\tau\hat{L}_n^m(\tau)\big)U^m \\ & - \tau \sum_{i=1}^{n} \overline{(\mathrm{b}^i)}^{\mathrm{T}} \hat{L}_i^m(\tau) \left(\bar{I} + \tau \sum_{j=1}^{n} \overline{\mathrm{A}^j}\hat{L}_j^m(\tau)\right)^{-1} \left(\tau \sum_{k=1}^{n} \overline{\mathrm{A}^k}\widehat{G}_k^m(\tau)\right) \\ & + \tau \sum_{i=1}^{n} \overline{(\mathrm{b}^i)}^{\mathrm{T}} \widehat{G}_i^m(\tau), \end{aligned} \tag{14}$$

where

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

is called transition operator.

To prove that the solution of scheme (6), (7) is bounded independently of $\tau \in (0, \tau_0]$ in [6,7] the decomposition (14) is used together with some additional results. Firstly it is seen that the operator

$$T_{(12)}^m \equiv \tau \sum_{i=1}^{n} \overline{(\mathrm{b}^i)}^{\mathrm{T}} \hat{L}_i^m(\tau) \left(\bar{I} + \tau \sum_{j=1}^{n} \overline{\mathrm{A}^j}\hat{L}_j^m(\tau)\right)^{-1} : H^s \to H,$$

[2] The norms considered in H^s are the induced norms by the norm $\|\cdot\|$ of H and any norm in $\mathbb{R}^s$.

obtained from a FSRK verifying (12) is bounded (uniformly in τ). A similar result is deduced for the operator

$$T^m_{(13)} \equiv \tau \sum_{i=1}^{n} \overline{(\mathrm{b}^i)}^{\mathrm{T}} \hat{L}^m_i(\tau) \left(\bar{I} + \tau \sum_{j=1}^{n} \overline{\mathrm{A}^j} \hat{L}^m_j(\tau) \right)^{-1} : \mathcal{D}_{k_1} \times H^{s-1} \to H,$$

obtained from a FSRK verifying (13).

Besides, in [6,7] it is shown that if $\{L_i(t)\}_{i=1}^n$ is a commutative system for all $t \in [t_0, T]$, admits unitary dilation and there exist n constants M_i such that

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

the transition operator can be bounded in the following form:

$$\| \widetilde{R}(-\tau \hat{L}^m_1(\tau), \ldots, -\tau \hat{L}^m_n(\tau)) \| \leqslant e^{\beta\tau}, \tag{15}$$

where β is a constant, usually positive, independent of $\tau \in (0, \tau_0]$; this condition ensures the stability of the discretization in time, at least in finite intervals of time. Some additional A-stability conditions, of type strong A-stability together with small variations in time for the operators $L_i(t)$ allow to get negative values for β and, consequently, preserve a contractive behaviour on the numerical solutions of scheme (6), (7) and ensure the stability even in infinite intervals of time.

To study the consistency of the semidiscretization (6), (7), we define, as usual, the local error

$$e^{m+1} = u(t_{m+1}) - \breve{U}^{m+1},$$

being $\breve{U}^{m+1}$ the numerical solution obtained with only one step of scheme (6), (7) starting from $u(t_m)$ and we say that the FSRK method is consistent of order p if, for sufficiently smooth data, it is verified that

$$\| e^{m+1} \| \leqslant C\tau^{p+1}, \quad \forall m \geqslant 0 \text{ and } \forall \tau \in (0, \tau_0]. \tag{16}$$

In the following theorem, which is proven in Section 4, we give sufficient conditions to get consistency of order p.

Theorem 2.1. *Let* (8) *be a FSRK method satisfying the reductions* [3]

$$(\mathrm{C})^k e - k\mathrm{A}^i(\mathrm{C})^{k-1} e = 0, \quad \forall i \in \{1, \ldots, n\}, \ \forall k \in \{1, \ldots, k_0 - 1\}, \tag{17}$$

together with the order conditions

$$(\mathrm{b}^{i_1})^{\mathrm{T}} (\mathrm{C})^{\rho_1} e = \frac{1}{\rho_1 + 1}, \tag{18}$$

$$(\mathrm{b}^{i_1})^{\mathrm{T}} (\mathrm{C})^{\rho_1} \mathrm{A}^{i_2} (\mathrm{C})^{\rho_2} \cdots \mathrm{A}^{i_r} (\mathrm{C})^{\rho_r} e = \prod_{j=1}^{r} \frac{1}{(r - j + 1) + \sum_{k=j}^{r} \rho_k}, \tag{19}$$

$\forall r \in \{2, \ldots, p\}, \ \forall (i_1, \ldots, i_r) \in \{1, \ldots, n\}^r,$

$\forall (\rho_1, \ldots, \rho_{r-1}) \in \{0, \ldots, p-1\}^{r-1}$ *and* $\forall \rho_r \in \{k_0, \ldots, p-1\},$

such that $1 \leqslant r + \sum_{k=1}^{r} \rho_k \leqslant p,$

[3] If there is not any reduction of type (17) we must consider $k_0 = 1$.

and let us apply it to a problem of type (4) *where the operators* $\{L_i(t)\}_{i=1}^n$ *satisfy* (3)*, whose solution satisfies the smoothness requirements*

$$\|u_{i_1}^{(p+1)}(t)\| \leqslant C, \tag{20}$$

$$\|L_{i_1}(t)u_{i_2}^{(p)}(t)\| \leqslant C, \tag{21}$$

$$\|L_{i_1}^{(\varrho_1)}(t)\cdots L_{i_{l-1}}^{(\varrho_{l-1})}(t)u_{i_l}^{(\varrho_l)}(t)\| \leqslant C, \quad \forall t \in [t_0, T], \tag{22}$$

$$\forall l \in \{2, \ldots, p\},\ \forall (i_1, \ldots, i_l) \in \{1, \ldots, n\}^l,$$

$$\forall (\varrho_1, \ldots, \varrho_{l-1}) \in \{0, \ldots, p-1\}^{l-1} \text{ and } \forall \varrho_l \in \{k_0, \ldots, p-1\},$$

$$\text{such that } 2 \leqslant l + \sum_{k=1}^{l} \varrho_k \leqslant p+2,$$

being $u_i'(t) = -L_i(t)u(t) + g_i(t)$.

Then (16) *is verified.*

It is well known that if (1) is an operational formulation of an initial boundary value problem, conditions (22) are generally not satisfied for high values of p, unless severe compatibility conditions between the initial and the boundary conditions are imposed. Such compatibility conditions, which some authors call non-natural, joined with the smoothness of data are sufficient but not necessary to ensure enough smoothness on the solution of (1). If conditions (22) are not satisfied the order-reduction phenomenon will occur. This behavior is typical in discretizations of partial differential problems with a numerical method with its internal stages which can be considered as approximations, of low order, of the solution at intermediate time steps (see, for example, [22,24]).

Recently, several techniques have been developed to avoid the order reduction in classical time discretization methods (see [1,20]). In [2,4] a clever correction of the boundary conditions for the calculus of the internal stages permits avoiding the order reduction phenomenon for Rosembrock and Runge–Kutta methods. In [3] similar ideas are used to avoid the order reduction of Fractional Step Runge–Kutta time discretizations in the case of considering operators L_i with time independent coefficients. Nowadays, we are studying corrections of this type for the case of having time-dependent operators $L_i(t)$.

To end the study of the convergence of the semidiscrete scheme (6), (7) we define the global error associated to the time discretization as $E^\tau \equiv \sup_{m\leqslant T/\tau} \|u(t_m) - U^m\|$, and we say that scheme (6), (7) is convergent of order p if $E^\tau \leqslant C\tau^p$, for sufficiently smooth functions $u(t)$ joining the last consistency results and the stability property (15) it is not difficult to prove the following

Theorem 2.2. *If the semidiscrete scheme* (6), (7)*, satisfies* (15) *and* (16)*, then it is convergent of order* p.

3. Convergence of the totally discrete scheme

In this section we will denote by C a generic constant independent of the time step τ and also of the spatial mesh size h.

To obtain a totally discrete scheme of type Fractional Steps we must discretize in space the stationary problems obtained in (6), (7). To study how this discretization process works we use an abstract formulation, which was proposed by Vainikko in [25] for a general discretization method. To introduce this framework in our context in a simple way we will focus our attention in discretizing elliptic problems of type

$$\big(I + \tau a_{jj}^{i_j} L_{i_j}(t_{m,j})\big)U^{m,j} = f^{m,j}, \tag{23}$$

which come from the time discretization stage (6).

Let us first take a positive parameter $h \in (0, h_0]$, destined to tend to zero. For every h we consider a finite-dimensional space V_h.[4]

A discretization method of (23) reduces it to a linear system of type

$$\big(I_h + \tau a_{jj}^{i_j} L_{i_j h}(t_{m,j})\big)U_h^{m,j} = f_h^{m,j}, \tag{24}$$

here the operators $L_{ih}(t) \in \mathcal{L}(V_h, V_h)$ are discrete approximations of the operators $L_i(t)$ for $i = 1, \dots, n$, which must preserve their essential qualities; as well $f_h^{m,j}$ will be discrete approximations (for example, projections, restrictions to the mesh or interpolations) of $f^{m,j}$.

It is expected that $U_h(t) \in V_h$ is a suitable approximation of $u(t) \in \mathcal{D}$ in a metric that we must precise. To do this, we take for every (h, V_h, t) two connecting applications between $\mathcal{R} \supseteq \bigcup_{i=1}^n \mathcal{D}_i$ or H and V_h:[5]

$$r_h(t) : \mathcal{R} \subset H \to V_h,$$
$$\pi_h : H \to V_h,$$

and a norm $\|\cdot\|_h$ in V_h associated to the scalar product $((\cdot,\cdot))_h$ satisfying the following compatibility properties:

$$\lim_{h\to 0} \big\|r_h(t)u\big\|_h = \|u\|, \quad \forall u \in \mathcal{R} \quad \text{and} \quad \lim_{h\to 0} \|\pi_h g\|_h = \|g\|, \quad \forall g \in H. \tag{25}$$

By using these discretizations we obtain the totally discrete scheme:

$$\begin{cases} U_h^0 = r_h(t_0)(u_0), \\ U_h^{m+1} = U_h^m - \tau \displaystyle\sum_{i=1}^n \sum_{j=1}^s b_j^i \big(L_{ih}(t_{m,j})U_h^{m,j} - g_{ih}(t_{m,j})\big), \\ \quad \text{with } U_h^{m,j} = U_h^m - \tau \displaystyle\sum_{i=1}^n \sum_{k=1}^s a_{jk}^i \big(L_{ih}(t_{m,k})U_h^{m,k} - g_{ih}(t_{m,k})\big), \quad \text{for all } j = 1, \dots, s. \end{cases} \tag{26}$$

Note that the calculation of every stage $U^{m,i}$ of (6), (7) involved the resolution of a problem of type (23) where

$$f^{m,j} = U^m - \tau \sum_{i=1}^n \sum_{k=1}^{j-1} a_{jk}^i L_i(t_{m,k})U^{m,k} + \tau \sum_{i=1}^n \sum_{k=1}^{j} a_{jk}^i g_i(t_{m,k}).$$

---

[4] h will be the thickness of the mesh in Finite Differences and the diameter of the elements in Finite Elements and V_h will be spaces of discrete functions on a mesh in Finite Differences and will be subspaces of H in Finite Elements. Note also that the dimension of these spaces will tend to infinity when h tends to zero.

[5] In Finite Differences $r_h(t)$ and π_h use to be restrictions to the mesh nodes and in Finite Elements can be projections in the subspace V_h or interpolations if sufficiently smooth functions are considered.

Now we are approaching $U^{m,i}$ with $U_h^{m,i}$ by posing a linear system, which admits a similar expression, in the form (24), where

$$f_h^{m,j} = U_h^m - \tau \sum_{i=1}^{n} \sum_{k=1}^{j-1} a_{jk}^i L_{ih}(t_{m,k}) U_h^{m,k} + \tau \sum_{i=1}^{n} \sum_{k=1}^{j} a_{jk}^i g_{ih}(t_{m,k}).$$

Let us now introduce the concepts involved in the study of the convergence of problem (24) to problem (23). We call local truncation error associated to the operator $L_i(t)$, as the following operator

$$\tau_h^{L_i(t)}(v) \equiv L_{ih}(t) r_h(t) v - \pi_h L_i(t) v, \quad \forall v \in \mathcal{D}_i, \tag{27}$$

and we say that $L_{ih}(t)$ is a consistent approximation of $L_i(t)$ of order q if for sufficiently smooth functions v the following bound is verified

$$\left\| \tau_h^{L_i(t)}(v) \right\|_h \leqslant C h^q. \tag{28}$$

In similar way, we say that $g_h(t)$ is a consistent approximation of order q of the source term, if for sufficiently smooth (in space) functions $g(t)$, it holds that

$$\left\| g_h(t) - \pi_h g(t) \right\|_h \leqslant C h^q. \tag{29}$$

In order to obtain the convergence of the solution of a scheme of type (24) to the solution of (23) it is necessary to impose some stability requirements. In this case we will impose that $L_{ih}(t)$ preserve the coercitivity of $L_i(t)$, i.e.,

$$\big(\big(L_{ih}(t) v, v \big) \big)_h \geqslant \gamma_i \|v\|_h^2, \quad \forall v \in V_h,\ \forall i = 1, \ldots, n. \tag{30}$$

This property ensures that

$$\left\| \big(I + \tau a_{ii}^k L_{kh}(t_{m,i}) \big) \right\|_h \leqslant 1 \tag{31}$$

and consequently schemes (26) are stable. A classical reasoning, which is typical in the study of the convergence of Finite Difference methods, permits to combine (31) with (28) and (29) to deduce the convergence of the discretization scheme (24).

In order to study the convergence of the totally discrete scheme (26), let us denote by $E_h^m \equiv \|r_h(t_m) u(t_m) - U_h^m\|_h$ the global error associated to it at the time t_m. We say that the discretization is unconditionally convergent of order p in time and of order q in space if, for sufficiently smooth functions $u(t)$, it holds that

$$E_h^m \leqslant C \big(h^q + \tau^p \big), \quad \forall h \in [0, h_0] \text{ and } \forall m = 1, 2, \ldots, T/\tau. \tag{32}$$

In order to analyze the convergence of the total discretization in a clever way we have split the contributions to the global error, E_h^m, of the temporal and spatial parts of the discretization; the contribution of the space discretization stage will be studied by using an intermediate term which we call local error of the space discretization.

Definition 3.1. We define the local error of the discretization in space as:

$$\hat{e}_h^m = \left\| r_h(t_m) \hat{u}^m - \widehat{U}_h^m \right\|_h,$$

where $\hat{u}^m$ is obtained with one step of the semidiscrete scheme (6), (7) taking as starting value $U^{m-1} = u(t_{m-1})$ and $\widehat{U}_h^m$ is obtained with one step of the totally discrete scheme (26) taking as starting value $U_h^{m-1} = r_h(t_{m-1}) u(t_{m-1})$.

Thus, for every time step of scheme (26) we obtain the following convergence result:

Theorem 3.1. *Let $u(t)$ be the solution of problem* (4) *with $\{L_i(t)u(t)\}_{i=1}^n$ and $\{g_i(t)\}_{i=1}^n$ sufficiently smooth data, let* (8) *be a FSRK method, let $\{L_{ih}(t)\}_{i=1}^n$ be n discretization operators of $\{L_i(t)\}_{i=1}^n$ satisfying* (28) *and* (30) *let $\{g_{ih}(t)\}_{i=1}^n$ be n functions which discretize $\{g_i(t)\}_{i=1}^n$ satisfying* (29) *and let $r_h(t)$ and π_h connecting applications satisfying, for sufficiently smooth functions $f(t)$, the following properties*:

$$\begin{aligned} &\left\|r_h(t)f(t)-r_h(t')f(t)\right\|_h \leqslant C|t-t'|h^q, && \forall t,t'\in[t_0,T],\\ &\left\|\pi_h f(t)-r_h(t)f(t)\right\|_h \leqslant Ch^q, && \forall t\in[t_0,T]. \end{aligned} \tag{33}$$

Then

$$\hat{e}_h^m \leqslant C\tau h^q. \tag{34}$$

In Section 2 we referenced some papers concerning the stability of the FSRK methods and revised the conditions in which the stability bound $\|\widetilde{R}(-\tau\hat{L}_1^m(\tau),\ldots,-\tau\hat{L}_n^m(\tau))\| \leqslant e^{\beta\tau}$ is obtained. The results of these papers are also applicable to obtain the bound

$$\left\|\widetilde{R}\left(-\tau\hat{L}_{1h}^m(\tau),\ldots,-\tau\hat{L}_{nh}^m(\tau)\right)\right\|_h \leqslant e^{\beta\tau}, \tag{35}$$

where β is independent of τ and also of h, if the system $\{L_{ih}(t)\}_{i=1}^n$ preserves some properties of $\{L_i(t)\}_{i=1}^n$. For example, if $\{L_{ih}(t)\}_{i=1}^n$ is coercive, commutative, admits unitary dilation $\forall t\in[t_0,T]$ and

$$\left\|L_{ih}(t')u_h - L_{ih}(t)u_h\right\|_h \leqslant |t-t'|M_i\left\|L_{ih}(t)u_h\right\|_h, \quad \forall i=1,\ldots,n,\ \forall t,t'\in[t_0,T],\ \forall u_h\in V_h,$$

with M_i independent of h, and the FSRK method is A-stable the same reasoning used to prove Theorem 1.1 in [7], permits to deduce (35).

A suitable combination of the stability and consistency properties of the time semidiscretization and the convergence of the spatial discretization stage, permit us to reach the expected result of convergence for the total discretization scheme (26):

Theorem 3.2. *If problem* (4) *has a sufficiently smooth solution and we use an A-stable FSRK method of order p for the discretization in time and a discretization in space such that the connecting applications satisfy* (33) *and preserve properties* (34) *and* (35), *then the global error verifies* (32).

4. Proofs of main Theorems

In order to shorten and clear the proofs of the theorems of Sections 2 and 3 we first introduce and prove two previous technical lemmas:

Lemma 4.1. *Let* (8) *be a FSRK method and let $\{L_i(t)\}_{i=1}^n$ be a system of operators fulfilling* (3). *If $\{L_i(t),g_i(t)\}_{i=1}^n \subseteq \mathcal{C}^p([t_0,T];H)$, then the local error can be written as follows*:

$$e^{m+1} = \varsigma^{m+1} - T^m\,\Xi^m, \tag{36}$$

where $T^m \equiv T_{(12)}^m$ if the FSRK satisfies (12) *or $T^m \equiv T_{(13)}^m$ if the FSRK satisfies* (13) *and where*

$$\varsigma^{m+1} = \sum_{k=1}^{p} \frac{\tau^k}{k!} \sum_{i=1}^{n} \big(1 - k(\mathrm{b}^i)^{\mathrm{T}}(\mathrm{C})^{k-1} e\big) u_i^{(k)}(t_m) + \sum_{i=1}^{n} \int_{t_m}^{\tilde{t}} \left(\frac{(t_{m+1} - \zeta)_+^p}{p!} - \tau \sum_{j=1}^{s} b_j^i \frac{(t_{m,j} - \zeta)_+^{p-1}}{(p-1)!} \right) u_i^{(p+1)}(\zeta)\, \mathrm{d}\zeta, \tag{37}$$

$$\Xi^m = \sum_{k=1}^{p} \frac{\tau^k}{k!} \Delta^{k,m} + \Upsilon^m, \tag{38}$$

with

$$\Delta^{k,m} = \sum_{i=1}^{n} \big((\overline{\mathrm{C}})^k \bar{e} - k \overline{\mathrm{A}^i} (\overline{\mathrm{C}})^{k-1} \bar{e} \big) u_i^{(k)}(t_m), \quad \forall k = 1, \ldots, p, \tag{39}$$

and $\Upsilon^m = (\upsilon^{m,1}, \ldots, \upsilon^{m,s})^{\mathrm{T}} \in H^s$ *is such that*

$$\upsilon^{m,j} = \sum_{i=1}^{n} \int_{t_m}^{\tilde{t}'} \left(\frac{(t_{m,j} - \zeta)_+^p}{p!} - \tau \sum_{l=1}^{i} a_{jl}^i \frac{(t_{m,l} - \zeta)_+^{p-1}}{(p-1)!} \right) u_i^{(p+1)}(\zeta)\, \mathrm{d}\zeta, \quad \forall j = 1, \ldots, s, \tag{40}$$

being $\tilde{t} = \max_{j \in \{1,\ldots,s\}} \{t_{m,j}, t_{m+1}\}$, $\tilde{t}' = \max_{l \in \{1,\ldots,s\}} \{t_{m,l}\}$ *and*

$$(t - \zeta)_+ = \begin{cases} t - \zeta & \text{if } \zeta \leqslant t, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. If we use notation (10), (11) for scheme (6), (7) we can rewrite the local error as follows:

$$e^{m+1} = u(t_{m+1}) - u(t_m) + T^m \bar{e} u(t_m) + T^m \left(\tau \sum_{i=1}^{n} \overline{\mathrm{A}^i} \widehat{G}_i^m(\tau) \right) - \tau \sum_{i=1}^{n} \overline{(\mathrm{b}^i)}^{\mathrm{T}} \widehat{G}_i^m(\tau). \tag{41}$$

In order to get (36), we introduce the contributions of the local error of quadrature formulae

$$\varsigma^{m+1} = u(t_{m+1}) - u(t_m) + \tau \sum_{i=1}^{n} \sum_{j=1}^{s} b_j^i \big(L_i(t_{m,j}) u(t_{m,j}) - g_i(t_{m,j}) \big), \tag{42}$$

as well, we will use the contributions to the local error of similar formulae for every stage

$$\xi^{m,j} = u(t_{m,j}) - u(t_m) + \tau \sum_{i=1}^{n} \sum_{k=1}^{i} a_{jk}^i \big(L_i(t_{m,k}) u(t_{m,k}) - g_i(t_{m,k}) \big), \tag{43}$$

for $j = 1, \ldots, s$, and we group them in

$$\Xi^m = \big(\xi^{m,1}, \ldots, \xi^{m,s} \big)^{\mathrm{T}} \in H^s. \tag{44}$$

If we also group the evaluations of $u(t)$ at the intermediate steps $t_{m,j}$, for $j = 1, \ldots, s$ as $\mathcal{U}^m = (u(t_{m,1}), \ldots, u(t_{m,s}))^{\mathrm{T}} \in H^s$ then we can rewrite (42) and (43) in a compact form as follows:

$$\varsigma^{m+1} = u(t_{m+1}) - u(t_m) + \tau \sum_{i=1}^{n} \overline{(\mathrm{b}^i)}^{\mathrm{T}} \big(\hat{L}_i^m(\tau) \mathcal{U}^m - \widehat{G}_i^m(\tau) \big), \tag{45}$$

and

$$\Xi^m = \left(\bar{I} + \tau \sum_{i=1}^{n} \overline{\mathrm{A}^i}\,\hat{L}_i^m(\tau)\right)\mathcal{U}^m - \bar{e}u(t_m) - \tau \sum_{i=1}^{n} \overline{\mathrm{A}^i}\,\widehat{G}_i^m(\tau), \tag{46}$$

if we remove $\mathcal{U}^m$ in (45) using (46) we obtain

$$\varsigma^{m+1} = u(t_{m+1}) - u(t_m) + T^m\bar{e}u(t_m) + T^m\,\Xi^m + T^m\left(\tau \sum_{i=1}^{n} \overline{\mathrm{A}^i}\,\widehat{G}_i^m(\tau)\right) - \tau \sum_{i=1}^{n} \overline{(\mathrm{b}^i)}^{\mathrm{T}}\widehat{G}_i^m(\tau). \tag{47}$$

Comparing (47) with (41) it is immediate to get (36).

Finally to obtain (37) and (38) for ς^{m+1} and Ξ^m respectively, it only rests to us to use the Taylor's expansions

$$u(t_{m+1}) = \sum_{k=0}^{p} \frac{\tau^k}{k!}u^{(k)}(t_m) + \int_{t_m}^{t_{m+1}} \frac{(t_{m+1}-\zeta)^p}{p!}u^{(p+1)}(\zeta)\,\mathrm{d}\zeta,$$

$$u(t_{m,j}) = \sum_{k=0}^{p} \frac{\tau^k}{k!}(c_j)^k u^{(k)}(t_m) + \int_{t_m}^{t_{m,j}} \frac{(t_{m,j}-\zeta)^p}{p!}u^{(p+1)}(\zeta)\,\mathrm{d}\zeta \quad \text{and}$$

$$u_i'(t_{m,j}) = \sum_{k=0}^{p-1} \frac{\tau^k}{k!}(c_j)^k u_i^{(k+1)}(t_m) + \int_{t_m}^{t_{m,j}} \frac{(t_{m,j}-\zeta)^{p-1}}{(p-1)!}u_i^{(p+1)}(\zeta)\,\mathrm{d}\zeta,$$

in expressions (42), (43) and (44) and regroup terms in function of the powers of τ. □

Lemma 4.2. *Let* (8) *be a FSRK method satisfying the order conditions* (19)*; then, for all* $k = k_0, \ldots, p-1$, *it holds that*

$$\frac{T^m}{\tau} = \sum_{l=0}^{p-k-1} (-\tau)^l \beta^{l,t_m} + (-\tau)^{p-k}\Psi^{p-k,t_m}(\tau)\left(\bar{I} + \tau \sum_{i=1}^{n} \overline{\mathrm{A}^i}\,\hat{L}_i^m(\tau)\right)^{-1} \tag{48}$$

for all $\Delta^{k,m}$ *defined by* (39)*, being*

$$\Psi^{l,t_m}(\tau) \equiv \sum_{i=1}^{n} \overline{(\mathrm{b}^i)}^{\mathrm{T}}\hat{L}_i^m(\tau)\left(\sum_{j=1}^{n} \overline{\mathrm{A}^j}\,\hat{L}_j^m(\tau)\right)^l, \tag{49}$$

and

$$\beta^{l,t_m} = \int_0^{\tau} \frac{(\tau-\zeta)^{p-k-l-1}}{(p-k-l-1)!}\frac{\mathrm{d}^{p-k-l}\Psi^{l,t_m}(\zeta)}{\mathrm{d}\tau^{p-k-l}}\,\mathrm{d}\zeta. \tag{50}$$

Proof. We first use the development

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

to rewrite $\frac{T^m}{\tau}=\sum_{i=1}^{n}\overline{(\mathrm{b}^i)}^{\mathrm{T}}\hat{L}_i^m(\tau)(\bar{I}+\tau\sum_{j=1}^{n}\overline{\mathrm{A}^j}\hat{L}_j^m(\tau))^{-1}$ in the form

$$\sum_{l=0}^{p-k-1}(-\tau)^l\Psi^{l,t_m}(\tau)+(-\tau)^{p-k}\Psi^{p-k,t_m}(\tau)\left(\bar{I}+\tau\sum_{j=1}^{n}\overline{\mathrm{A}^j}\hat{L}_j^m(\tau)\right)^{-1},$$

for all $k=k_0,\ldots,p-1$.

Note that the second addend of the last expression applied to $\Delta^{k,m}$ gives directly the second addend of (48). To end the proof of (48) it only rests to us to check that $\Psi^{l,t_m}(\tau)\Delta^{k,m}=\beta^{l,t_m}\Delta^{k,m}$ for all $k=k_0,\ldots,p-1$ and for all $l=0,\ldots,p-k-1$.

If we consider $\Psi^{l,t_m}(\tau)$ as function of τ, the following Taylor's expansion at the point t_m can be used:

$$\Psi^{l,t_m}(\tau)=\sum_{j=0}^{p-k-l-1}\frac{\tau^j}{j!}\frac{d^j\Psi^{l,t_m}(0)}{d\tau^j}+\beta^{l,t_m}, \tag{51}$$

where $\frac{\mathrm{d}^j\Psi^{l,t_m}(\tau)}{\mathrm{d}\tau^j}$ is

$$\sum_{\substack{(\rho_1,\ldots,\rho_{l+1})\in\{0,\ldots,j\}^{l+1}\\ \rho_1+\cdots+\rho_{l+1}=j}}\frac{j!}{\rho_1!\cdots\rho_{l+1}!}\sum_{i_1=1}^{n}\overline{(\mathrm{b}^{i_1})}^{\mathrm{T}}\big(\overline{\mathrm{C}}\big)^{\rho_1}\hat{L}_{i_1}^{m(\rho_1)}(\tau)$$
$$\times\sum_{i_2=1}^{n}\overline{\mathrm{A}^{i_2}}\big(\overline{\mathrm{C}}\big)^{\rho_2}\hat{L}_{i_2}^{m(\rho_2)}(\tau)\cdots\sum_{i_{l+1}=1}^{n}\overline{\mathrm{A}^{i_{l+1}}}\big(\overline{\mathrm{C}}\big)^{\rho_{l+1}}\hat{L}_{i_{l+1}}^{m(\rho_{l+1})}(\tau), \tag{52}$$

being $\hat{L}_i^{m(j)}(\tau)=\frac{\mathrm{d}^j\hat{L}_i^m(\tau)}{\mathrm{d}\tau^j}$.

As the operators $\hat{L}_i^{m(j)}(0)$ commute with $\overline{\mathrm{A}^k}$ and also with $(\overline{\mathrm{C}})^l$ for all $i,k=1,\ldots,n$, and for all $j,l,m\geqslant 0$, we can rewrite $\frac{\mathrm{d}^j\Psi^{l,t_m}(0)}{\mathrm{d}\tau^j}$ as follows:

$$\sum_{\substack{(\rho_1,\ldots,\rho_{l+1})\in\{0,\ldots,j\}^{l+1}\\ \rho_1+\cdots+\rho_{l+1}=j\\ (i_1,\ldots,i_{l+1})\in\{1,\ldots,n\}^{l+1}}}\frac{j!}{\rho_1!\cdots\rho_{l+1}!}\overline{(\mathrm{b}^{i_1})}^{\mathrm{T}}\big(\overline{\mathrm{C}}\big)^{\rho_1}\overline{\mathrm{A}^{i_2}}\big(\overline{\mathrm{C}}\big)^{\rho_2}\cdots\overline{\mathrm{A}^{i_{l+1}}}\big(\overline{\mathrm{C}}\big)^{\rho_{l+1}}$$
$$\times\hat{L}_{i_1}^{m(\rho_1)}(0)\hat{L}_{i_2}^{m(\rho_2)}(0)\cdots\hat{L}_{i_{l+1}}^{m(\rho_{l+1})}(0). \tag{53}$$

It is easy to check that the order conditions (19) ensure that

$$\big(\mathrm{b}^{i_1}\big)^{\mathrm{T}}(\mathrm{C})^{\rho_1}\mathrm{A}^{i_2}(\mathrm{C})^{\rho_2}\cdots\mathrm{A}^{i_{l+1}}(\mathrm{C})^{\rho_{l+1}}\big((\mathrm{C})^k e-k\mathrm{A}^{i_{l+2}}(\mathrm{C})^{k-1}e\big)=0,$$

for all $k = k_0, \dots, p-1$, for all $l = 0, \dots, p-k-1$, for all $j = 0, \dots, p-k-l-1$, for all $(\rho_1, \dots, \rho_{l+1}) \in \{0, \dots, j\}^{l+1}$ such that $\rho_1 + \cdots + \rho_{l+1} = j$ and for all $(i_1, \dots, i_{l+2}) \in \{1, \dots, n\}^{l+2}$.

Making use of this property, together with (53), it is immediate that it holds $\frac{\mathrm{d}^j \Psi^{l,t_m}(0)}{\mathrm{d}\tau^j} \Delta^{k,m} = 0$, $\forall k = k_0, \dots, p-1$, $\forall l = 0, \dots, p-k-1$, $\forall j = 0, \dots, p-k-l-1$. □

Proof of Theorem 2.1. We will use the decomposition (36)–(40) for e^{m+1}. From expression (37) for ς^{m+1} it is immediate to deduce

$$\left\| \varsigma^{m+1} \right\| \leqslant C\tau^{p+1},$$

taking into account the order conditions (18) and the smoothness hypotheses (20).

Let us see now the obtaining of the same bound for the second addend of e^{m+1} which is $T^m \Xi^m$. If we use in expression (38) for Ξ^m the reduction (17), which implies $\Delta^{k,m} = 0 \in H^s$ for all $k = 1, \dots, k_0 - 1$, then:

$$T^m \Xi^m = \sum_{k=k_0}^{p} \frac{\tau^k}{k!} T^m \Delta^{k,m} + T^m \Upsilon^m. \tag{54}$$

Using the bounds for T^m obtained in [7], together with the expression of Υ^m, given in (40), and the smoothness requirements given by (20), we obtain that the second addend of (54) verifies

$$\left\| T^m \Upsilon^m \right\| \leqslant C\tau^{p+1}.$$

In order to obtain the corresponding bound for the first addend of (54) we use the decomposition given in Lemma 4.2 for $T^m \Delta^{k,m}$ with $k = k_0, \dots, p-1$ to deduce that

$$\sum_{k=k_0}^{p} \frac{\tau^k}{k!} T^m \Delta^{k,m} = \sum_{k=k_0}^{p-1} \frac{\tau^{k+1}}{k!} \sum_{l=0}^{p-k-1} (-\tau)^l \beta^{l,t_m} \Delta^{k,m} \tag{55}$$

$$+ \tau^{p+1} \sum_{k=k_0}^{p-1} \frac{(-1)^{p-k}}{k!} \Psi^{p-k,t_m}(\tau) \left(\bar{I} + \tau \sum_{i=1}^{n} \overline{\mathrm{A}^i} \hat{L}_i^m(\tau) \right)^{-1} \Delta^{k,m} \tag{56}$$

$$+ \frac{\tau^p}{p!} T^m \Delta^{p,m}. \tag{57}$$

From the expressions (50) and (52) for β^{l,t_m} and the smoothness hypotheses (22), a bound of type $C\tau^{p+1}$ is obtained for the addend given in (55).

To obtain a bound of type $C\tau^{p+1}$ for the addend given in (56), we will show that, for all $k = k_0, \dots, p-1$, the following bound is verified:

$$\left\| \Psi^{p-k,t_m}(\tau) \left(\bar{I} + \tau \sum_{j=1}^{n} \overline{\mathrm{A}^j} \hat{L}_j^m(\tau) \right)^{-1} \Delta^{k,m} \right\| \leqslant C. \tag{58}$$

To prove such bound, we first rewrite the operator $\Psi^{p-k,t_m}(\tau)$ given in (49) as follows:

$$\Psi^{p-k,t_m}(\tau) = \overline{\mathrm{b}^{\mathrm{T}}} \left(\overline{\mathrm{A}} \right)^{-1} \left(\sum_{i=1}^{n} \overline{\mathrm{A}^i} \hat{L}_i^m(\tau) \right)^{p-k+1}, \quad \forall k = k_0, \dots, p-1, \tag{59}$$

this can be done thanks to the special coefficient structure of a FSRK which makes that

$$\sum_{i=1}^{n}\overline{\big(\mathrm{b}^i\big)}^{\mathrm{T}}\hat{L}_i^m(\tau)=\big(b_1^{k_1}L_{k_1}(t_{m,1}),\dots,b_s^{k_s}L_{k_s}(t_{m,s})\big);$$

and we can write it in the form

$$\big(b_1^{k_1}I_H,\dots,b_s^{k_s}I_H\big)\begin{pmatrix} L_{k_1}(t_{m,1}) & 0 & \dots & 0\\ 0 & L_{k_2}(t_{m,2}) & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & L_{k_s}(t_{m,s})\end{pmatrix},$$

where the first tensor is $\bar{\mathrm{b}}^{\mathrm{T}}$. If we multiply it by $(\overline{\mathrm{A}})^{-1}\overline{\mathrm{A}}$ and regroup products in the form

$$\big[\bar{\mathrm{b}}^{\mathrm{T}}\big(\overline{\mathrm{A}}\big)^{-1}\big]\left[\overline{\mathrm{A}}\begin{pmatrix} L_{k_1}(t_{m,1}) & 0 & \dots & 0\\ 0 & L_{k_2}(t_{m,2}) & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & L_{k_s}(t_{m,s})\end{pmatrix}\right],$$

it results that the second square bracket is equal to $\sum_{i=1}^{n}\overline{\mathrm{A}^i}\hat{L}_i^m(\tau)$ obtaining directly (59).

Secondly, we use that for all $k=k_0,\dots,p-1$, the following equality is verified

$$\left(\sum_{i=1}^{n}\overline{\mathrm{A}^i}\hat{L}_i^m(\tau)\right)^{p-k+1}\left(\bar{I}+\tau\sum_{j=1}^{n}\overline{\mathrm{A}^j}\hat{L}_j^m(\tau)\right)^{-1}=\left(\bar{I}+\tau\sum_{j=1}^{n}\overline{\mathrm{A}^j}\hat{L}_j^m(\tau)\right)^{-1}\left(\sum_{i=1}^{n}\overline{\mathrm{A}^i}\hat{L}_i^m(\tau)\right)^{p-k+1}. \tag{60}$$

Finally, to obtain (58), we use the bounds for $(\bar{I}+\tau\sum_{j=1}^{n}\overline{\mathrm{A}^j}\hat{L}_j^m(\tau))^{-1}$ which appear in [7] and we take into account that the components of the term

$$\left(\sum_{i=1}^{n}\overline{\mathrm{A}^i}\hat{L}_i^m(\tau)\right)^{p-k+1}\Delta^{k,m},$$

contain linear combinations of elementary differentials which are bounded by the hypotheses (22).

It remains to verify (60) to end the proof of (58). We only prove that (60) is true for $p-k+1=1$, and for other values of $p-k+1$ an inductive reasoning can be applied.

Following the same techniques developed in [7] for the obtaining of bounds of $(\bar{I}+\tau\sum_{i=1}^{n}\overline{\mathrm{A}^i}\hat{L}_i^m(\tau))^{-1}$, it is easy to prove that the operator

$$\left(\bar{I}+\tau\sum_{i=1}^{n}\overline{\mathrm{A}^i}\hat{L}_i^m(\tau)\right)^{-1}\left(\sum_{j=1}^{n}\overline{\mathrm{A}^j}\hat{L}_j^m(\tau)\right):\mathcal{D}_{k_1}\times\cdots\times\mathcal{D}_{k_s}\to\mathcal{D}_{k_1}\times\cdots\times\mathcal{D}_{k_s},$$

is bounded for any norm of H^s. Thus, by using that $\mathcal{D}_{k_i}$ is dense in H for all $k_i\in\{1,\dots,n\}$, this operator can be extended, in a unique way, by preserving the linearity and the boundness to

$$\left(\bar{I}+\tau\sum_{i=1}^{n}\overline{\mathrm{A}^i}\hat{L}_i^m(\tau)\right)^{-1}\left(\sum_{j=1}^{n}\overline{\mathrm{A}^j}\hat{L}_j^m(\tau)\right):H^s\to H^s.$$

If the FSRK satisfies (12), as the non-null diagonal coefficients of the FSRK method satisfy $\sum_{i=1}^{n} a_{jj}^{i} \neq 0$ for all $j = 1, \ldots, s$, then the operator $\sum_{i=1}^{n} \overline{\mathrm{A}^i} \hat{L}_i^m(\tau)$ is invertible; therefore, we can write for all $\mathcal{U} \in \mathcal{D}_{k_1} \times \cdots \times \mathcal{D}_{k_s}$ that

$$\left(\sum_{i=1}^{n} \overline{\mathrm{A}^i} \hat{L}_i^m(\tau)\right)\left(\bar{I} + \tau \sum_{j=1}^{n} \overline{\mathrm{A}^j} \hat{L}_j^m(\tau)\right)^{-1} \mathcal{U} = \left(\sum_{i=1}^{n} \overline{\mathrm{A}^i} \hat{L}_i^m(\tau)\right)$$
$$\times \left[\left(\sum_{j=1}^{n} \overline{\mathrm{A}^j} \hat{L}_j^m(\tau)\right)^{-1} \left(\sum_{j=1}^{n} \overline{\mathrm{A}^j} \hat{L}_j^m(\tau)\right) + \tau \left(\sum_{j=1}^{n} \overline{\mathrm{A}^j} \hat{L}_j^m(\tau)\right)^{-1} \left(\sum_{j=1}^{n} \overline{\mathrm{A}^j} \hat{L}_j^m(\tau)\right)\left(\sum_{j=1}^{n} \overline{\mathrm{A}^j} \hat{L}_j^m(\tau)\right)\right]^{-1} \mathcal{U}$$
$$= \left(\sum_{i=1}^{n} \overline{\mathrm{A}^i} \hat{L}_i^m(\tau)\right)\left(\sum_{j=1}^{n} \overline{\mathrm{A}^j} \hat{L}_j^m(\tau)\right)^{-1} \left(\bar{I} + \tau \sum_{j=1}^{n} \overline{\mathrm{A}^j} \hat{L}_j^m(\tau)\right)^{-1} \left(\sum_{j=1}^{n} \overline{\mathrm{A}^j} \hat{L}_j^m(\tau)\right)\mathcal{U}$$
$$= \left(\bar{I} + \tau \sum_{j=1}^{n} \overline{\mathrm{A}^j} \hat{L}_j^m(\tau)\right)^{-1} \left(\sum_{i=1}^{n} \overline{\mathrm{A}^i} \hat{L}_i^m(\tau)\right)\mathcal{U},$$

and considering the same extension argument by density, this equality is also right for all $\mathcal{U} \in H^s$.

To obtain a bound of the same order for (57), we have carried out a similar process to the one used to obtain (58), but in this case we consider the hypotheses (21) instead of (22).

If the FSRK satisfies (13), instead of (12), to obtain bounds of type $C\tau^{p+1}$ for (56) and (57) we can not use that $\overline{\mathrm{A}} = \sum_{i=1}^{n} \overline{\mathrm{A}^i}$ is invertible. To obtain (58), we will take into account that, in this case, the first component of $\Delta^{k,m} \in H^s$ is null and that the first row of the matrix $\overline{\mathrm{A}}^i$ is also null for all $i = 1, \ldots, n$. These two facts cause that the first components of (56) and (57) are null and also that the first columns of operators $(\bar{I} + \tau \sum_{i=1}^{n} \overline{\mathrm{A}^i} \hat{L}_i^m(\tau))^{-1}$ and $\sum_{i=1}^{n} \overline{\mathrm{A}^i} \hat{L}_i^m(\tau)$ do not play a role in the calculus of the bounds of (56) and (57) resulting that

$$\Psi^{p-k,t_m}(\tau)\left(\bar{I} + \tau \sum_{i=1}^{n} \overline{\mathrm{A}^i} \hat{L}_i^m(\tau)\right)^{-1} \Delta^{k,m}$$
$$= \sum_{i=1}^{n} \overline{(\boldsymbol{b}^i)^{\mathrm{T}}} \hat{\boldsymbol{L}}_i^m(\tau)\left(\sum_{i=1}^{n} \overline{\boldsymbol{A}^i} \hat{\boldsymbol{L}}_i^m(\tau)\right)^{p-k} \left(\bar{\boldsymbol{I}} + \tau \sum_{i=1}^{n} \overline{\boldsymbol{A}^i} \hat{\boldsymbol{L}}_i^m(\tau)\right)^{-1} \boldsymbol{\Delta}^{k,m},$$

where

$$\overline{(\boldsymbol{b}^i)^{\mathrm{T}}} = \left(b_2^i I_H, \ldots, b_s^i I_H\right)^{\mathrm{T}} \in H^{s-1},$$

$$\overline{\boldsymbol{A}^i} = \begin{pmatrix} a_{22}^i I_H & 0 & \ldots & 0 \\ a_{32}^i I_H & a_{33}^i I_H & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{s2}^i I_H & a_{s3}^i I_H & \ldots & a_{ss}^i I_H \end{pmatrix} \in H^{s-1\times s-1},$$

$$\hat{\boldsymbol{L}}_i^m(\tau)=\begin{pmatrix} L_i(t_{m,2}) & 0 & \dots & 0\\ 0 & L_i(t_{m,3}) & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & L_i(t_{m,s})\end{pmatrix}\in\mathcal{L}(\mathcal{D}_i,H)^{s-1\times s-1},$$

$$\boldsymbol{\Delta}^{k,m}=\left(\Delta_2^{k,m},\dots,\Delta_s^{k,m}\right)^{\mathrm{T}}\in H^{s-1}\quad\text{and}\quad\bar{\boldsymbol{I}}=\mathrm{Diag}(I_H,\dots,I_H)\in H^{s-1}.$$

Now, as $\overline{\boldsymbol{A}}=\sum_{i=1}^n\overline{\boldsymbol{A}^i}$ is invertible, we can repeat the reasoning used for the all implicit stage case to conclude (58) and also that $\|\tau^p/p!T^m\Delta^{p,m}\|\leqslant C\tau^{p+1}$. □

Proof of Theorem 2.2. In order to introduce the contribution of the local truncation error, we decompose it in the form

$$\left\|u(t_m)-U^m\right\|\leqslant\left\|u(t_m)-\breve{U}^m\right\|+\left\|\breve{U}^m-U^m\right\|,$$

and using (16) and (15) we get the recurrence relation

$$\left\|u(t_m)-U^m\right\|\leqslant C\tau^{p+1}+e^{\beta\tau}\left\|u(t_{m-1})-U^{m-1}\right\|,$$

which permits us to deduce that:

(i) if $\beta>0$, then $\|u(t_m)-U^m\|\leqslant C\frac{1-e^{m\beta\tau}}{1-e^{\beta\tau}}\tau^{p+1}$ and, using that $1-e^{\beta\tau}\leqslant-\beta\tau$, we deduce $\|u(t_m)-U^m\|\leqslant C\frac{e^{T\beta}-1}{\beta}\tau^p$.
(ii) If $\beta=0$, as $e^{\beta k\tau}=1$, then $\|u(t_m)-U^m\|\leqslant Cm\tau^{p+1}\leqslant CT\tau^p$.
(iii) If $\beta<0$, then $\|u(t_m)-U^m\|\leqslant C\frac{\tau_0}{1-e^{\beta\tau_0}}\tau^p$, for all $\tau\in(0,\tau_0]$ and $\forall m$; therefore, in this case the convergence is also reached for infinite periods of time. □

Proof of Theorem 3.1. As we have that

$$\begin{aligned}\hat{u}^{m+1}&=u(t_m)-\tau\sum_{j=1}^n\sum_{i=1}^s b_i^j\left(L_j(t_{m,i})\hat{u}^{m,i}-g_j(t_{m,i})\right)\\&=u(t_m)-\tau\sum_{i=1}^s b_i^{k_i}\left(L_{k_i}(t_{m,i})\hat{u}^{m,i}-g_{k_i}(t_{m,i})\right),\\\widehat{U}_h^{m+1}&=r_h(t_m)u(t_m)-\tau\sum_{j=1}^n\sum_{i=1}^s b_i^j\left(L_{jh}(t_{m,i})\widehat{U}_h^{m,i}-g_{jh}(t_{m,i})\right)\\&=r_h(t_m)u(t_m)-\tau\sum_{i=1}^s b_i^{k_i}\left(L_{k_ih}(t_{m,i})\widehat{U}_h^{m,i}-g_{k_ih}(t_{m,i})\right),\end{aligned}$$

using hypotheses (29) and (33), the achievement of (34) is immediately reduced to prove that

$$\left\|L_{k_ih}(t_{m,i})\widehat{U}_h^{m,i}-\pi_hL_{k_i}(t_{m,i})\hat{u}^{m,i}\right\|_h\leqslant Ch^q,\quad\forall i=1,\dots,s.\tag{61}$$

We proceed by induction in the number of the stages to get (61).

For the first stage we have that $\hat{u}^{m,1}$ is solution of

$$\hat{u}^{m,1}=u(t_m)-\tau a_{11}^{k_1}\left(L_{k_1}(t_{m,1})\hat{u}^{m,1}-g_{k_1}(t_{m,1})\right);\tag{62}$$

as well, $\widehat{U}_h^{m,1}$ is solution of

$$\widehat{U}_h^{m,1} = r_h(t_m)u(t_m) - \tau a_{11}^{k_1}\big(L_{k_1h}(t_{m,1})\widehat{U}_h^{m,1} - g_{k_1h}(t_{m,1})\big), \tag{63}$$

with $g_{k_1h}(t_{m,1}) = \pi_h g_{k_1}(t_{m,1}) + \mathcal{O}(h^q)$.

By applying π_h to expression (62), and subtracting it to (63) we deduce that

$$\begin{aligned}
&\big\|\tau a_{11}^{k_1}\big(L_{k_1h}(t_{m,1})\widehat{U}_h^{m,1} - \pi_h L_{k_1}(t_{m,1})\hat{u}^{m,1}\big)\big\|_h \\
&\quad\leqslant \big\|\big(r_h(t_m) - \pi_h\big)u(t_m) - \widehat{U}_h^{m,1} + \pi_h\hat{u}^{m,1}\big\|_h + \mathcal{O}\big(\tau h^q\big) \\
&\quad\leqslant \big\|\big(r_h(t_m) - \pi_h\big)\big(u(t_m) - \hat{u}^{m,1}\big)\big\|_h + \big\|r_h(t_m)\hat{u}^{m,1} - \widehat{U}_h^{m,1}\big\|_h + \mathcal{O}\big(\tau h^q\big) \\
&\quad= \big\|\tau a_{11}^{k_1}\big(r_h(t_m) - \pi_h\big)\big(L_{k_1}(t_{m,1})\hat{u}^{m,1} - g_{k_1}(t_{m,1})\big)\big\|_h \\
&\qquad+ \big\|r_h(t_m)\hat{u}^{m,1} - r_h(t_{m,1})\hat{u}^{m,1}\big\|_h + \big\|r_h(t_{m,1})\hat{u}^{m,1} - \widehat{U}_h^{m,1}\big\|_h + \mathcal{O}\big(\tau h^q\big);
\end{aligned} \tag{64}$$

as the first two addends of (64) are bounded by $C\tau h^q$, because of hypotheses (33), it only rests us to obtain that

$$\big\|r_h(t_{m,1})\hat{u}^{m,1} - \widehat{U}_h^{m,1}\big\|_h \leqslant C\tau h^q, \tag{65}$$

to deduce (61) for $i = 1$.

To prove it we again apply π_h to (62) and we use (28) and (29), to deduce

$$\begin{aligned}
&\big(I + \tau a_{11}^{k_1}L_{k_1h}(t_{m,1})\big)r_h(t_{m,1})\hat{u}^{m,1} \\
&\quad= \pi_h\big(u(t_m) - \hat{u}^{m,1}\big) + r_h(t_{m,1})\hat{u}^{m,1} + \tau a_{11}^{k_1}g_{k_1h}(t_{m,1}) + \mathcal{O}\big(\tau h^q\big).
\end{aligned}$$

As well, (63) can be written in similar form as

$$\big(I + \tau a_{11}^{k_1}L_{k_1h}(t_{m,1})\big)\widehat{U}_h^{m,1} = r_h(t_m)u(t_m) + \tau a_{11}^{k_1}g_{k_1h}(t_{m,1}).$$

Subtracting the two last expressions we obtain

$$\begin{aligned}
&\big(I + \tau a_{11}^{k_1}L_{k_1h}(t_{m,1})\big)\big(r_h(t_{m,1})\hat{u}^{m,1} - \widehat{U}_h^{m,1}\big) \\
&\quad= \pi_h u(t_m) - \pi_h\hat{u}^{m,1} + r_h(t_{m,1})\hat{u}^{m,1} - r_h(t_m)u(t_m) + \mathcal{O}\big(\tau h^q\big).
\end{aligned} \tag{66}$$

As we have assumed (30), then $\|(I + \tau a_{11}^{k_1}L_{k_1h}(t_{m,1}))^{-1}\|_h \leqslant 1$ is verified; thus, to prove (65), it remains to prove that $\pi_h u(t_m) - \pi_h\hat{u}^{m,1} + r_h(t_{m,1})\hat{u}^{m,1} - r_h(t_m)u(t_m)$ is bounded by $C\tau h^q$; so, we rewrite this term as follows:

$$\begin{aligned}
&\big(r_h(t_{m,1}) - r_h(t_m)\big)\hat{u}^{m,1} - \big(r_h(t_m) - \pi_h\big)\big(u(t_m) - \hat{u}^{m,1}\big) \\
&\quad= \big(r_h(t_{m,1}) - r_h(t_m)\big)\hat{u}^{m,1} - \tau a_{11}^{k_1}\big(r_h(t_m) - \pi_h\big)\big(L_{k_1}(t_{m,1})\hat{u}^{m,1} - g_{k_1}(t_{m,1})\big)
\end{aligned} \tag{67}$$

and the compatibility properties (33) for $r_h(t)$ and π_h give directly the expected bound.

We must note that the coefficient structure of a FSRK method, together with the bound (61) for $i = 1$, permit us to deduce:

$$\big\|\tau a_{r1}^{k_i}\big(L_{k_ih}(t_{m,i})\widehat{U}_h^{m,1} - \pi_h L_{k_i}(t_{m,i})\hat{u}^{m,1}\big)\big\|_h \begin{cases} = 0, & \text{if } k_i \neq k_1, \\ \leqslant C\tau h^q, & \text{if } k_i = k_1, \end{cases}$$

with $i, r \in \{1, \ldots, s\}$ and besides that

$$\big\|\tau b_1^{k_i}\big(L_{k_ih}(t_{m,i})\widehat{U}_h^{m,1} - \pi_h L_{k_i}(t_{m,i})\hat{u}^{m,1}\big)\big\|_h \begin{cases} = 0, & \text{if } k_i \neq k_1, \\ \leqslant C\tau h^q, & \text{if } k_i = k_1, \end{cases}$$

with $i \in \{1, \ldots, s\}$.

To end this proof by induction let us suppose that for all $l < j$ the following bounds are true:

$$\begin{cases} \left\| \tau a_{rl}^{k_i} \left(L_{k_i h}(t_{m,i}) \widehat{U}_h^{m,l} - \pi_h L_{k_i}(t_{m,i}) \hat{u}^{m,l} \right) \right\|_h \leqslant C\tau h^q, \\ \left\| \tau b_l^{k_i} \left(L_{k_i h}(t_{m,i}) \widehat{U}_h^{m,l} - \pi_h L_{k_i}(t_{m,i}) \hat{u}^{m,l} \right) \right\|_h \leqslant C\tau h^q \quad \text{and} \\ \left\| \widehat{U}_h^{m,l} - r_h(t_{m,l}) \hat{u}^{m,l} \right\|_h \leqslant C\tau h^q, \end{cases} \tag{68}$$

for all $r, i \in \{1, \ldots, s\}$ and we will obtain similar bounds for $\hat{u}^{m,j}$ and $\widehat{U}_h^{m,j}$. Such jth stages are obtained by solving:

$$\hat{u}^{m,j} = u(t_m) - \tau \sum_{l=1}^{j} a_{jl}^{k_l} \left(L_{k_l}(t_{m,l}) \hat{u}^{m,l} - g_{k_l}(t_{m,l}) \right),$$

and

$$\widehat{U}_h^{m,j} = r_h(t_m) u(t_m) - \tau \sum_{l=1}^{j} a_{jl}^{k_l} \left(L_{k_l h}(t_{m,l}) \widehat{U}_h^{m,l} - g_{k_l h}(t_{m,l}) \right),$$

respectively.

The same process used for the first stage can be repeated for the jth stage to arrive at the next relations

$$\begin{aligned} &\left(I + \tau a_{jj}^{k_j} L_{k_j h}(t_{m,j}) \right) \left(\widehat{U}_h^{m,j} - r_h(t_{m,j}) \hat{u}^{m,j} \right) \\ &\quad = -\tau \sum_{l=1}^{j-1} a_{jl}^{k_l} \left(L_{k_l h}(t_{m,l}) \widehat{U}_h^{m,l} - \pi_h L_{k_l}(t_{m,l}) \hat{u}^{m,l} \right) \\ &\qquad + r_h(t_m) u(t_m) - \pi_h u(t_m) + \pi_h \hat{u}^{m,j} - r_h(t_{m,j}) \hat{u}^{m,j} + \mathcal{O}\left(\tau h^q\right). \end{aligned} \tag{69}$$

Note that the induction hypotheses (68) ensures that the first addend of (69) is bounded by $C\tau h^q$. Therefore the last expression can be reduced to one with the form (66), with indexes j instead of 1, and we can repeat the same reasoning used for the first stage to obtain now that

$$\begin{aligned} &\left\| \tau a_{jj}^{k_j} \left(L_{k_j h}(t_{m,j}) \widehat{U}_h^{m,j} - \pi_h L_{k_j}(t_{m,j}) \hat{u}^{m,1} \right) \right\|_h \\ &\quad \leqslant \left\| \tau a_{jj}^{k_j} \left(r_h(t_m) - \pi_h \right) \left(L_{k_j}(t_{m,j}) \hat{u}^{m,j} - g_{k_j}(t_{m,1}) \right) \right\|_h \\ &\qquad + \left\| r_h(t_m) \hat{u}^{m,j} - r_h(t_{m,j}) \hat{u}^{m,j} \right\|_h + \left\| r_h(t_{m,j}) \hat{u}^{m,j} - \widehat{U}_h^{m,j} \right\|_h + \mathcal{O}\left(\tau h^q\right) \end{aligned}$$

and

$$\left\| r_h(t_{m,j}) \hat{u}^{m,j} - \widehat{U}_h^{m,j} \right\|_h \leqslant C\tau h^q.$$

Finally, due to the coefficient structure of a FSRK method which makes that $a_{ij}^{k_l} = 0$ and $b_j^{k_l} = 0$ if $j \neq l$, we can to deduce that (68) is also true for $l = j$. □

Remark 4.3. If the FSRK verifies (13) instead of (12), then the proof of this theorem is similar, simpler because in this case $\widehat{U}_h^{m,1} = r_h(t_m) u(t_m)$ and $\hat{u}^{m,1} = u(t_m)$ and to obtain the bound (61) it is sufficient with using the property (28).

Proof of Theorem 3.2. We decompose the global error as follows:

$$E_h^m \leqslant \left\| r_h(t_m)u(t_m) - r_h(t_m)\hat{u}^m \right\|_h + \left\| r_h(t_m)\hat{u}^m - \widehat{U}_h^m \right\|_h + \left\| \widehat{U}_h^m - U_h^m \right\|_h, \tag{70}$$

the consistency result (16) and the compatibility properties (25) permit us to bound the first addend of (70) as follows:

$$\left\| r_h(t_m)u(t_m) - r_h(t_m)\hat{u}^m \right\|_h \leqslant C\tau^{p+1}.$$

As the second addend of (70) is $\hat{e}_h^m$, it admits the bound $C\tau h^q$ by hypotheses. For the third addend of (70) we can use that

$$\widehat{U}_h^m - U_h^m = \widetilde{R}\left(-\tau \hat{L}_{1h}^m(\tau), \ldots, -\tau \hat{L}_{nh}^m(\tau)\right)\left(r_h(t_{m-1})u(t_{m-1}) - U_h^{m-1}\right)$$

and taking into account (35), we deduce immediately

$$\left\| \widehat{U}_h^m - U_h^m \right\|_h \leqslant e^{\beta\tau} \left\| r_h(t_{m-1})u(t_{m-1}) - U_h^{m-1} \right\|_h \equiv e^{\beta\tau} E_h^{m-1}.$$

Joining the last bounds, we get the following recurrence relation for the global errors

$$E_h^m \leqslant C\left(\tau^{p+1} + \tau h^p\right) + e^{\beta k\tau} E_h^{m-1},$$

which permit us to deduce that

(i) if $\beta > 0$ then $E_h^m \leqslant C\tau \frac{1-e^{m\beta\tau}}{1-e^{\beta\tau}}(\tau^p + h^q) \leqslant C\frac{e^{T\beta}-1}{\beta}(\tau^p + h^q)$.
(ii) If $\beta = 0$ then $E_h^m \leqslant Cm\tau(\tau^p + h^q) \leqslant CT(\tau^p + h^q)$.
(iii) If $\beta < 0$ then $E_h^m \leqslant C\frac{\tau_0}{1-e^{\beta\tau_0}}(\tau^p + h^q)$, $\forall \tau \in (0, \tau_0]$ and $\forall m$ and in this case we obtain again the unconditional convergence even for infinite intervals of time. □

5. Numerical examples

In this section we present two numerical tests which show the numerical behavior of the methods obtained in this paper. Firstly we show a numerical test for the following reaction–diffusion problem:

$$\begin{cases} \dfrac{\partial u}{\partial t} - d_1(x,y,t)\dfrac{\partial^2 u}{\partial x^2} - d_2(x,y,t)\dfrac{\partial^2 u}{\partial y^2} + k_1(x,y,t)u + k_2(x,y,t)u = f(x,y,t), \\ \quad \forall x, y \in \Omega \text{ and } \forall t \in [0,5], \\ u(x,0,t) = u(x,1,t) = 0, \quad \forall x \in [0,1] \text{ and } \forall t \in [0,5], \\ u(0,y,t) = u(1,y,t) = 0, \quad \forall y \in [0,1] \text{ and } \forall t \in [0,5], \\ u(x,y,0) = x^2(1-x)^2 y^2(1-y)^2, \quad \forall x, y \in \Omega, \end{cases}$$

with $d_1(x,y,t) = (1+e^{-t})(1+y)$, $d_2(x,y,t) = (2-e^{-t})(1+xy)$, $k_1(x,y,t) = 1+\sin(\pi x)e^{-t}$, $k_2(x,y,t) = 1+y^2$, and the source term $f(x,y,t) = e^{-t}x(1-x)y(1-y)$ in the spatial range $\Omega = [0,1] \times [0,1]$.

The total discretization has been realized by using A-stable FSRK method of Peaceman and Rachford of order two (see [6,23]) for the time discretization and a standard central-difference scheme on a uniform mesh with $(N+1) \times (N+1)$ for the spatial discretization. As both discretization processes spatial and temporal are of the same order, the results that we show in Tables 1 and 2 have been obtained by taking the relation $\tau N = C = 0.4$ in order to preserve contributions of the same order in the discretizations in space and time. In Table 1, we show the numerical errors that we have estimated by using the double mesh principle:

$$E_{N,\tau} = \max_{x_i, y_j, t_m} \left| U^{N,\tau}(x_i, y_j, t_m) - U^{2N,\tau/2}(x_i, y_j, t_m) \right|,$$

where $U^{N,\tau}(x_i, y_j, t_m)$ is the numerical solution obtained in the spatial node $(i/N, j/N)$ and in the time $t_m = m\tau$ and $U^{2N,\tau/2}(x_i, y_j, t_m)$ is the numerical solution obtained in the same point using a mesh with $(2N+1) \times (2N+1)$ points and time step $\tau/2$. In Table 2 we show their corresponding numerical orders of convergence, which we have computed with the formula

$$p = \log_2 \frac{E_{N,\tau}}{E_{2N,\tau}}.$$

Secondly, we present the following convection–diffusion problem:

$$\begin{cases} \dfrac{\partial u}{\partial t} - d_1(x, y, t)\dfrac{\partial^2 u}{\partial x^2} - d_2(x, y, t)\dfrac{\partial^2 u}{\partial y^2} + v_1(x, y, t)\dfrac{\partial u}{\partial x} + v_2(x, y, t)\dfrac{\partial u}{\partial y} \\ \quad + k_1(x, y, t)u + k_2(x, y, t)u = f(x, y, t), \quad \forall x, y \in \Omega \text{ and } \forall t \in [0, 5], \\ u(x, 0, t) = u(x, 1, t) = 0, \quad \forall x \in [0, 1] \text{ and } \forall t \in [0, 5], \\ u(0, y, t) = u(1, y, t) = 0, \quad \forall y \in [0, 1] \text{ and } \forall t \in [0, 5], \\ u(x, y, 0) = x^3(1-x)^3 y^3 (1-y)^3, \quad \forall x, y \in \Omega, \end{cases}$$

with $d_1(x, y, t) = (2 - e^{-t})(2 - y)$, $d_2(x, y, t) = (2 - e^{-t})(1 + x)$, $v_1(x, y, t) = (2 + \cos(\pi t)e^{-t})$, $v_2(x, y, t) = (2 - \sin(\pi t)e^{-t})(2 + y^2)$, $k_1(x, y, t) = 1 + y^2$, $k_2(x, y, t) = 1 + \sin(\pi x)$, and the source term $f(x, y, t) = e^{-t}x(1-x)y(1-y)$ in the spatial range $\Omega = [0, 1] \times [0, 1]$.

In this case the total discretization has been realized by using the L-stable FSRK of third order which appears in [8] to discretize the time variable and a standard upwind scheme on a uniform mesh with

Table 1

$N = 8$	$N = 16$	$N = 32$	$N = 64$	$N = 128$	$N = 256$	$N = 512$
2.3821E−5	5.8055E−6	1.4460E−6	3.6184E−7	9.0447E−8	2.2613E−8	5.6361E−9

Table 2

$N = 8$	$N = 16$	$N = 32$	$N = 64$	$N = 128$	$N = 256$
2.0368	2.0054	1.9986	2.0002	1.9999	2.0043

Table 3

$N = 8$	$N = 16$	$N = 32$	$N = 64$	$N = 128$	$N = 256$	$N = 512$
7.4403E−5	4.5183E−5	2.4429E−5	1.2920E−5	6.6653E−6	3.4315E−6	1.7590E−6

Table 4

$N = 8$	$N = 16$	$N = 32$	$N = 64$	$N = 128$	$N = 256$
0.7196	0.8872	0.9190	0.9549	0.9578	0.9641

$(N+1) \times (N+1)$ points for the spatial discretization. In this case, the results that we show have been obtained by taking the relation $\tau \sqrt[3]{N} = 0.1$ and with the same formulae used in Tables 3 and 4 to compute $E_{N,\tau}$ and p, respectively.

References

[1] S. Abarbanel, D. Gottielb, M.H. Carpenter, On the removal of boundary errors caused by Runge–Kutta integration of nonlinear partial differential equations, SIAM J. Sci. Comput. 17 (3) (1996) 777–782.
[2] I. Alonso-Mallo, B. Cano, Spectral–Rosenbock discretizations without order reduction for lineal parabolic problems, Report 2000/3, Departamento de Matemática Aplicada y Computación, Universidad de Valladolid, Valladolid, 2000.
[3] I. Alonso-Mallo, B. Cano, J.C. Jorge, Spectral–Fractional step Runge–Kutta discretizations for initial boundary value problems with time dependent boundary conditions, Report 2001/9 Departamento de Matemática Aplicada y Computación, Universidad de Valladolid, Valladolid, 2001.
[4] I. Alonso-Mallo, B. Cano, Avoiding order reduction of Runge–Kutta discretizations for linear time dependent parabolic problems, Report 2001/10 Departamento de Matemática Aplicada y Computación, Universidad de Valladolid, Valladolid, 2001.
[5] H. Brezis, Analyse Fonctionnelle, Thèorie et Applications, Masson, Paris, 1983.
[6] 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, 1999.
[7] B. Bujanda, J.C. Jorge, Stability results for fractional step discretizations of time dependent coefficient evolutionary problems, Appl. Numer. Math. 38 (2001) 69–86.
[8] B. Bujanda, J.C. Jorge, Additive Runge–Kutta methods for the resolution of linear parabolic problems, J. Comput. Appl. Math. 140 (2002) 99–117.
[9] J. Blair Perot, An analysis of the fractional step method, J. Comput. Phys. 108 (1993) 51–58.
[10] G.J. Cooper, A. Sayfy, Aditive methods for the numerical solution of ordinary differential equations, Math. Comp. 35 (1980) 1159–1172.
[11] G.J. Cooper, A. Sayfy, Additive Runge–Kutta methods for Stiff ordinary differential equations, Math. Comp. 40 (1983) 207–218.
[12] M. Crouzeix, Sur l'aproximation des équations differentielles opérationelles linéaires par des méthodes de Runge–Kutta, These d'Etat, Univ. de Paris VI, 1975.
[13] G. Fairweather, A.R. Mitchell, A new computational procedure for A.D.I. methods, SIAM J. Numer. Anal. 4 (1967) 163–170.
[14] P.J. van der Houwen, The development of Runge–Kutta methods for partial differential equations, Appl. Numer. Math. 20 (1996) 261–272.
[15] P.J. van der Houwen, J.G. Verwer, One-step splitting methods for semi-discrete parabolic problems, Computing 22 (1979) 291–309.
[16] W. Hundsdorfer, Unconditional convergence of some Crank–Nicolson LOD methods for initial-boundary value problems, Math. Comp. 58 (1992) 33–53.
[17] W. Hundsdorfer, Trapezoidal and midpoint splittings for initial-boundary value problems, Math. Comp. 67 (233) (1998) 1047–1062.
[18] 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.
[19] J.C. Jorge, F. Lisbona, Contractivity results for alternating direction schemes in Hilbert spaces, Appl. Numer. Math. 15 (1994) 65–75.
[20] S.L. Keeling, Runge–Kutta discretizations for parabolic equations with time dependent coefficients, Math. Comp. 52 (186) (1989) 561–586.
[21] J.L. Lions, Equations Differentielles Operationelles, et Problèmes aux Limites, Springer, Berlin, 1961.
[22] A. Ostermann, M. Roche, Runge–Kutta methods for partial differential equations and fractional order of convergence, Math. Comp. 59 (200) (1992) 403–420.

[23] D.W. Peaceman, H.H. Rachford, The numerical solution of parabolic and elliptic differential equations, J. SIAM 3 (1955) 28–42.
[24] J.M. Sanz-Serna, J.G. Verwer, W.H. Hundsdorfer, Convergence and order reduction of Runge–Kutta schemes applied to evolutionary problems in partial differential equations, Numer. Math. 50 (1987) 405–418.
[25] G. Vainikko, Approximative methods for nonlinear equations (two approaches to the convergence problem), Nonlinear Anal. Theory Meth. Appl. 6 (1978) 647–687.
[26] N.N. Yanenko, The Method of Fractional Steps, Springer, Berlin, 1971.