Applied Numerical Mathematics 38 (2001) 69–86
APPLIED NUMERICAL MATHEMATICS
www.elsevier.nl/locate/apnum

Stability results for fractional step discretizations of time dependent coefficient evolutionary 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 consider a class of additive Runge–Kutta methods, which include most of the classical alternating direction or fractionary step methods, for discretizing the time variable in an evolutionary problem whose coefficients depend on time. Some stability results are proven for these methods which, together with suitable consistency properties, permit us to show the convergence of these discretizations.

Keywords: Fractional steps; Stability

1. Introduction and notation

Let H be a Hilbert space with scalar product $((\cdot,\cdot))$, with associated norm $\|\cdot\|$, and let $u:[0,T]\to H$ be the solution of the evolution problem:

$$\begin{cases} \dfrac{\mathrm{d}u(t)}{\mathrm{d}t} + A(t)\,u(t) = g(t), \\ u(0) = u_0, \end{cases} \tag{1}$$

where $A(t):\mathcal{D}\big(A(t)\big) \subseteq H \to H$ are linear, generally unbounded operators, which we suppose maximal and coercive for all t, i.e.,

$$\begin{cases} \forall\, g(t)\in H,\ \exists\, v\in\mathcal{D}(A(t)),\ \text{such that } v + A(t)v = g(t) \text{ and} \\ \exists\,\alpha > 0 \text{ such that } ((A(t)v, v)) \geqslant \alpha\|v\|^2,\ \forall\, v\in\mathcal{D}(A(t)). \end{cases}$$

☆ This research is partially supported by the DGES research project No. PB97-1013 and a project of University of La Rioja.

* Corresponding author.

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

0168-9274/01/$ – see front matter
PII: S0168-9274(00)00063-5

Typically, multidimensional initial-boundary value problems involving partial differential equations can be formulated in the operational form (1), where $A(t)$ are operators which contain the spatial differential derivatives (for example, elliptic operators if (1) is a parabolic problem).

In [7,13] some existence and uniqueness results, as well as the study of the smoothness in time are exposed.

Usually, in an operational formulation of type (1) of an evolutionary initial-boundary value problem, boundary conditions are included in the election of the domains $\mathcal{D}(A(t))$, which may change in time for different reasons, such us the case of evolutionary boundary conditions. For simplicity, we shall study only the case $\mathcal{D}\big(A(t)\big) \equiv \mathcal{D} \subseteq H$, for all $t \in [0, T]$.

A discretization in time of problem (1), using for example a standard RK method, permits to obtain approximations U^m to $u(m\,\Delta t)$ by a recurrence, which can be written in the following way:

$$\begin{aligned}U^{m+1} =\ & R\big(-\Delta t A(t_{m,1}), \ldots, -\Delta t A(t_{m,s})\big)U^m \\ & + S\big(\Delta t A(t_{m,1}), \ldots, \Delta t A(t_{m,s}), \Delta t g(t_{m,1}), \ldots, \Delta t g(t_{m,s})\big),\end{aligned}$$

where $R(-\Delta t z(t_{m,1}), \ldots, -\Delta t z(t_{m,s}))$ and $S(\Delta t z(t_{m,1}), \ldots, \Delta t z(t_{m,s}), \Delta t w(t_{m,1}), \ldots, \Delta t w(t_{m,s}))$ are rational approximations of $E(t_m, t_{m+1})$ and of $\int_{t_m}^{t_{m+1}} E(t, t_{m+1}) w(t)\,\mathrm{d}t$, respectively, where $E(a,b) \equiv \exp(-\int_a^b z(t)\,\mathrm{d}t)$.

A time discretization process of type alternating directions or fractional steps usually admits a similar formulation

$$\begin{aligned}U^{m+1} =\ & R\big(-\Delta t A_1(t_{m,1}), \ldots, -\Delta t A_1(t_{m,s}), -\Delta t A_2(t_{m,1}), \ldots, -\Delta t A_n(t_{m,s})\big)U^m \\ & + S\big(\Delta t A_1(t_{m,1}), \ldots, \Delta t A_n(t_{m,s}), \Delta t g_1(t_{m,1}), \ldots, \Delta t g_n(t_{m,s})\big),\end{aligned}$$

where $R(-\Delta t z_1(t_{m,1}), \ldots, -\Delta t z_n(t_{m,s}))$ and $S(\Delta t z_1(t_{m,1}), \ldots, \Delta t z_n(t_{m,s}), \Delta t w_1(t_{m,1}), \ldots, \Delta t w_n(t_{m,s}))$ are rational approximations of $E(t_m, t_{m+1})$ and of $\int_{t_m}^{t_{m+1}} E(t, t_{m+1}) w(t)\,\mathrm{d}t$, respectively, where now $z(t) = \sum_{i=1}^n z_i(t)$, $w(t) = \sum_{i=1}^n w_i(t)$, $\sum_{i=1}^n A_i(t)$ is a decomposition of $A(t)$ in n simpler addends and $\sum_{i=1}^n g_i(t) = g(t)$. Concretely (see [14,19]), an alternating direction or fractional step method can be viewed as a time integrator which uses suitable decompositions of $A(t)$ and $g(t)$ to compute approximations to $u(t_m)$ more easily than using standard implicit methods. This is usually carried out by computing some intermediate fractionary steps $U^{m,i}$, between U^m and U^{m+1}, which are implicit only in one of the operators $A_i(t)$. Such discretization processes can be structured in the following way:

$$\begin{cases} U^0 = u_0, \\ U^{m,i} = U^m + \Delta t \displaystyle\sum_{j=1}^{i} a_{ij}^{k_j}\big(-A_{k_j}(t_{m,j})U^{m,j} + g_{k_j}(t_{m,j})\big), \quad \text{for } i = 1, \ldots, s, \\ U^{m+1} = U^m + \Delta t \displaystyle\sum_{i=1}^{s} b_i^{k_i}\big(-A_{k_i}(t_{m,i})U^{m,i} + g_{k_i}(t_{m,i})\big), \quad \text{with } k_i, k_j \in \{1, \ldots, n\}. \end{cases} \tag{2}$$

If we compare this scheme with a classical time discretization scheme, like a semiexplicit Runge–Kutta method, it is clear that we can obtain remarkable cost reductions if the stationary problems

$$\big(I + \alpha A_i(t)\big)v = f \quad (\alpha > 0) \tag{3}$$

are easier than problems of type

$$\big(I + \alpha A(t)\big)v = f \quad (\alpha > 0) \tag{4}$$

in some way. For example, when (1) is a spatial semidiscretization of type central differences in a rectangular grid of the multidimensional heat conduction equation, the operators $A_i(t)$ can contain the discretizations of the terms $-\partial^2 u/\partial x_i^2$, and then the linear systems of type (3) will involve tridiagonal matrices, while the problems of type (4) will involve block tridiagonal matrices and its resolution has a higher order of computational complexity. A second simple example can be considered if (1) is a system of coupled partial differential equations; in this case, some schemes of type prediction–correction can be easily considered by taking

$$A_i = \begin{pmatrix} 0 & \dots & 0 & L_{i1} & 0 & \dots & 0 \\ 0 & \dots & 0 & L_{i2} & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & L_{in} & 0 & \dots & 0 \end{pmatrix} \quad \text{if } A = \begin{pmatrix} L_{11} & L_{12} & \dots & L_{1n} \\ L_{21} & L_{22} & \dots & L_{2n} \\ \vdots & & \ddots & \vdots \\ L_{n1} & L_{n2} & \dots & L_{nn} \end{pmatrix} \text{ and } u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}.$$

Scheme (2) seems a special kind of one-step method of type Runge–Kutta. In fact, if we fill this formulation with some null coefficients b_j^k, a_{ij}^k, we will obtain the scheme

$$\begin{cases} U^0 = u_0, \\ U^{m,i} = U^m + \Delta t \displaystyle\sum_{j=1}^{s}\sum_{k=1}^{n} a_{ij}^k \big(-A_k(t_{m,j})U^{m,j} + g_k(t_{m,j})\big), \quad \text{for } i = 1, \dots, s, \\ U^{m+1} = U^m + \Delta t \displaystyle\sum_{i=1}^{s}\sum_{k=1}^{n} b_i^k \big(-A_k(t_{m,i})U^{m,i} + g_k(t_{m,i})\big), \end{cases} \tag{5}$$

which is called, for the case $n = 2$, additive Runge–Kutta method in [5] and [6].

Most of the classical alternating direction schemes (see [19]) as well as some new ones of high orders (see [1,2,12]) can be reformulated by using the special subset (2) of additive Runge–Kutta methods, which is called fractional step Runge–Kutta methods (see also [18]).

Definition 1.1. A fractional step Runge–Kutta method (abbreviately FSRK), is an additive RK method satisfying:

$$\begin{cases} a_{ii}^k \geqslant 0, \quad \forall i \in \{1, \dots, s\},\ k \in \{1, \dots, n\}, & a_{ij}^k = 0, \quad \forall j > i, \\ |b_j^k| + \displaystyle\sum_{i=1}^{s} |a_{ij}^k| \neq 0 \Rightarrow |b_j^l| + \displaystyle\sum_{i=1}^{s} |a_{ij}^l| = 0, & \forall l \neq k,\ l,k \in \{1, \dots, n\},\ i,j \in \{1, \dots, s\}, \\ a_{ii}^l a_{ii}^k = 0, & \text{if } k \neq l,\ i \in \{1, \dots, s\},\ k,l \in \{1, \dots, n\}. \end{cases} \tag{6}$$

The coefficients of these methods can be organized in a Butcher table of type

$$\begin{array}{c|c|c|c|c} c & \mathcal{A}^1 & \mathcal{A}^2 & \dots & \mathcal{A}^n \\ \hline & (b^1)^{\mathrm{T}} & (b^2)^{\mathrm{T}} & \dots & (b^n)^{\mathrm{T}}, \end{array} \tag{7}$$

where $\mathcal{A}^k = (a_{ij}^k)$, $b^k = (b_i^k)$ and $c = (c_1, \dots, c_s)^{\mathrm{T}}$ with $i, j = 1, \dots, s$ and $k = 1, \dots, n$.

Avoiding the null columns in $\left(\frac{\mathcal{A}^i}{(b^i)^{\mathrm{T}}}\right)$ we can reduce notation (7) to the following:

$$\begin{array}{c|c} & k^{\mathrm{T}} \\ \hline c & \mathcal{A} \\ \hline & b^{\mathrm{T}} \end{array}, \tag{8}$$

with $\mathcal{A} \equiv (a_{ij}^{k_j}) = \sum_{i=1}^n \mathcal{A}^i$, $b \equiv (b_j^{k_j}) = \sum_{i=1}^n b^i$ and $k^{\mathrm{T}} = (k_1, \ldots, k_n)$ where k_j $(\in \{1, \ldots, n\})$ satisfy that

$$\sum_{j=1}^{s} \sum_{\substack{l=1 \\ l \neq k_j}}^{n} \left(|b_j^l| + \sum_{i=1}^{s} |a_{ij}^l| \right) = 0, \quad \text{for } j = 1, \ldots, s. \tag{9}$$

From now on, we will assume that the operators $A_i(t): \mathcal{D}_i \subseteq H \to H$ preserve the maximality and coercitivity of $A(t)$, i.e.,

$$\begin{cases} \forall g(t) \in H, \exists v \in \mathcal{D}_i, \text{ such that } v + A_i(t)v = g(t) \text{ and} \\ \exists \alpha_i > 0 \text{ such that } ((A_i(t)v, v)) \geqslant \alpha_i \|v\|^2, \ \forall v \in \mathcal{D}_i, \end{cases} \tag{10}$$

for $i = 1, \ldots, n$, and $\mathcal{D} = \bigcap_{i=1}^n \mathcal{D}_i$.

In order to abbreviate the formulation of scheme (2) we introduce the following tensorial notation:

$$\begin{cases} \text{given } M \equiv (m_{ij}) \in \mathbb{R}^{s \times s} \text{ and } v \equiv (v_i) \in \mathbb{R}^s, \text{ we denote} \\ \overline{M} \equiv \begin{pmatrix} m_{11} I_H & \ldots & m_{1s} I_H \\ \vdots & \ddots & \vdots \\ m_{s1} I_H & \ldots & m_{ss} I_H \end{pmatrix} \in H^{s \times s} \quad \text{and} \quad \bar{v} \equiv \begin{pmatrix} v_1 I_H \\ \vdots \\ v_s I_H \end{pmatrix} \in H^s, \end{cases} \tag{11}$$

and we group the stages $U^{m,i}$, as well as the evaluations of $g_i(t)$ and $A_i(t)$ for all $i = 1, \ldots, n$, and for all $m = 1, 2, \ldots$, in the form

$$\widetilde{U}^m = \left(U^{m,1}, \ldots, U^{m,s}\right)^{\mathrm{T}} \in H^s, \qquad G_i^m = \left(g_i(t_{m,1}), \ldots, g_i(t_{m,s})\right)^{\mathrm{T}} \in H^s,$$

$$\hat{A}_i^m = \begin{pmatrix} A_i(t_{m,1}) & 0 & \ldots & 0 \\ 0 & A_i(t_{m,2}) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & A_i(t_{m,s}) \end{pmatrix} \in \mathcal{L}(\mathcal{D}_i, H)^{s \times s}. \tag{12}$$

Using (11) and (12), the scheme (2) can be written as follows:

$$\begin{cases} \left(\bar{I} + \Delta t \sum_{i=1}^{n} \overline{\mathcal{A}^i} \hat{A}_i^m \right) \widetilde{U}^m = \bar{e} U^m + \Delta t \sum_{i=1}^{n} \overline{\mathcal{A}^i} G_i^m, \\ U^{m+1} = U^m + \Delta t \sum_{i=1}^{n} \left(\overline{b^i}\right)^{\mathrm{T}} \left(-\hat{A}_i^m \widetilde{U}^m + G_i^m\right), \end{cases}$$

where $e^{\mathrm{T}} = (1, \ldots, 1) \in \mathbb{R}^s$.

With the matrix coefficient structure defined in (6) for scheme (2), it is not difficult to show that this scheme has unique solution under conditions (10). In fact, we will prove in Lemmas 4.1 and 4.5 that the operator $(\bar{I} + \Delta t \sum_{i=1}^{n} \overline{\mathcal{A}^i} \hat{A}_i^m)$ is invertible for some classes of FSRK methods. Consequently, we can rewrite such schemes in a similar format to (5):

$$U^{m+1} = \widetilde{R}\big(-\Delta t \hat{A}_1^m, \ldots, -\Delta t \hat{A}_n^m\big) U^m + \widetilde{S}\big(\Delta t \hat{A}_1^m, \ldots, \Delta t \hat{A}_n^m, \Delta t G_1^m, \ldots, \Delta t G_n^m\big), \tag{13}$$

where

$$\widetilde{R}\big(-\Delta t \hat{A}_1^m, \ldots, -\Delta t \hat{A}_n^m\big) = \bar{I} - \sum_{i=1}^{n} \big(\overline{b^i}\big)^{\mathrm{T}} \Delta t \hat{A}_i^m \left(\bar{I} + \sum_{j=1}^{n} \overline{\mathcal{A}^j} \Delta t \hat{A}_j^m\right)^{-1} \bar{e}, \tag{14}$$

and

$$\begin{aligned}
&\widetilde{S}\big(\Delta t \hat{A}_1^m, \ldots, \Delta t \hat{A}_n^m, \Delta t G_1^m, \ldots, \Delta t G_n^m\big) \\
&\quad = \Delta t \sum_{i=1}^{n} \big(\overline{b^i}\big)^{\mathrm{T}} \left(G_i^m - \hat{A}_i^m \left(\bar{I} + \Delta t \sum_{j=1}^{n} \overline{\mathcal{A}^j} \hat{A}_j^m\right)^{-1} \left(\Delta t \sum_{k=1}^{n} \overline{\mathcal{A}^k} G_k^m\right)\right).
\end{aligned}$$

A way to study the convergence of time discretization methods consists of combining the properties of consistency and contractivity of the discrete transition operator. If we consider FSRK methods, consistency means that for sufficiently smooth data $u(t_m)$, $A_i(t)$, $g_i(t)$ it holds that

$$\big\|u(t_{m+1}) - \hat{u}^{m+1}\big\| \leqslant C \Delta t^{p+1}, \quad m = 0, 1, \ldots,$$

where $\hat{u}^{m+1}$ is the result of giving one step with (2) taking $u(t_m)$ as the starting point u^m.

On the other hand, as any two exact solutions of (1), $u(t)$ and $v(t)$, obtained with different initial conditions, u_0 and v_0, show the following contractive behaviour:

$$\big\|u(t_m + h) - v(t_m + h)\big\| \leqslant \big\|u(t_m) - v(t_m)\big\|, \quad \text{for } h \geqslant 0,$$

it seems natural to search time discretization schemes which preserve this property. So, we shall say that a method of type RK applied to (1) is contractive iff

$$\big\|U^{m+1} - V^{m+1}\big\| \leqslant \big\|U^m - V^m\big\|, \quad m = 0, 1, \ldots,$$

where U^m and V^m are two sequences generated by the algorithm (2) from different initial values U^0 and V^0, respectively; if we use (13), it is clear that the contractivity of a FSRK method is equivalent to

$$\big\|\widetilde{R}\big(-\Delta t \hat{A}_1^m, \ldots, -\Delta t \hat{A}_n^m\big)\big\| \leqslant 1.$$

In the case $A(t) = A$ for all $t \in [0, T]$, there exists a wide range of stability results in different metrics, mainly in the case of univariate approximations $R(-\Delta t A)$ to the semigroup $\mathrm{e}^{-\Delta t A}$. In the case of considering Hilbert spaces, Crouzeix proved in [7] several contractivity results of type $\|R(-\Delta t A)\| \leqslant 1$ for certain classes of maximal monotone operators A combined suitably with $A(\theta)$-acceptable rational functions $R(z)$. More generally, when A is a densely defined closed linear operator which generates a bounded strongly continuous semigroup e^{-tA}, in a Banach space X, some weaker stability results of type $\|R^m(-\Delta t A)\| \leqslant C/\sqrt{\Delta t}$ can be seen in [3,8]. More recently, Palencia proved (see [16]) that the factor $1/\sqrt{\Delta t}$ can be superseded if operator A is θ-sectorial and $R(z)$ is a $A(\theta)$-acceptable rational function. In the case of considering multivariate approximations of e^{-tA}, there exist only some contractivity results of type $\|R(-\Delta t A_1, \ldots, -\Delta t A_n)\| \leqslant 1$, when $(A_i)_{i=1}^n$ is a commutative system of maximal monotone operators in a Hilbert space H. Such results can be found in [15].

It is well known that, in the case of consider standard RK methods for discretizing the time variable and arbitrary dependencies in time of operators $A(t)$, AN-stability (see [11]) is a necessary condition to preserve contractivity. It is also well known that AN-stability is satisfied only by simple low order methods, such as the implicit Euler rule, or by fully implicit methods of high orders, such as some Gaussian methods.

Nevertheless, in [7], and more recently in [10], it is shown that, under suitable hypotheses of variation in time for operators $A(t)$, A-stability can be a sufficient condition for a stable integration, at least in finite intervals of time $[0, T]$.

Similar situations are produced if we use additive RK methods to discretize the time variable, i.e., a natural generalization of the AN-stability condition would lead us to preserve the contractivity of the numerical solutions of problem (1), but this condition could be satisfied only for semiexplicit schemes of low order or for high order fully implicit schemes of Gaussian type. Therefore, the AN-stability, and consequently the contractivity for arbitrary variations of $A_i(t)$, is not present in most of the alternating direction or fractional step methods, since they can be formulated as semiexplicit FSRK methods. We will show that A-stability, together with suitable time variations of $\{A_i(t)\}_{i=1}^n$ ensure a stable behaviour for numerical solutions of scheme (2).

In order to introduce the A-stability for an additive RK method given by (6), in an easy way, we apply it to the test scalar initial value problem

$$\begin{cases} y'(t) = \displaystyle\sum_{i=1}^{n} \lambda_i y(t), & \text{with } \operatorname{Re}(\lambda_i) \leqslant 0, i = 1, \ldots, n, \\ y(t_0) = y_0. \end{cases}$$

This gives us the recurrence

$$y_{m+1} = \left(1 + \sum_{i=1}^{n} \Delta t \lambda_i \left(b^i\right)^{\mathrm{T}} \left(I - \sum_{j=1}^{n} \Delta t \lambda_j \mathcal{A}^j\right)^{-1} e\right) y_m, \tag{15}$$

and substituting $\Delta t \lambda_i$ by z_i, we obtain what we call the amplification function associated to (6), which is a rational complex function of n complex variables $z_1, z_2, \ldots, z_n$, defined by

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

It is clear that the contractivity of solutions y_m given by (15) is equivalent to the A-stability property introduced in the following definition.

Definition 1.2. An additive RK method is said A-stable iff

$$\left|R(\bar{z})\right| \leqslant 1, \quad \forall \bar{z} \in \left\{(z_1, \ldots, z_n) \mid z_i \in \mathbb{C} \text{ and } \operatorname{Re}(z_i) \leqslant 0, \forall i = 1, \ldots, n\right\}.$$

We will see that in some problems, considered, for example, in [7,10], A-stability is a sufficient condition to obtain

$$\left\|\widetilde{R}\left(-\Delta t \hat{A}_1^m, \ldots, -\Delta t \hat{A}_n^m\right)\right\| \leqslant \mathrm{e}^{\beta \Delta t}, \tag{16}$$

where β is a constant, usually positive, but small, under conditions of type

$$\left\|A_i(t')u - A_i(t)u\right\| \leqslant \left|t - t'\right| M_i \left\|A_i(t)u\right\|, \quad \forall i = 1, \ldots, n, \ \forall t, t' \in [0, T],$$

which are related to a Lipschitz variation in the coefficients that define the differential operators $A_i(t)$. Obviously, condition (16) will not ensure the preservation of the contractivity, unless $\beta \leqslant 0$, but in finite periods of time we can weaken the contractivity requirement and preserve a stable integration. In this context we say that a method of type RK is A-stable if the time discretizations U^m and V^m, obtained with such method and with time step Δt, of problem (1) and the perturbed problem

$$\begin{cases} v'(t) + \sum_{i=1}^{n} A_i(t)v(t) = \sum_{i=1}^{n} \widetilde{g}_i(t) \quad (t \in [0, T]), \\ v(0) = \widetilde{u}_0, \end{cases}$$

satisfy

$$\|U^m - V^m\| \leqslant C\left(\|u_0 - \widetilde{u}_0\| + \sum_{i=1}^{n} \max_{t\in[0,T]} \|g_i(t) - \widetilde{g}_i(t)\|\right), \quad \forall m = 1, 2, \ldots, M = \frac{T}{\Delta t}, \tag{17}$$

where C is independent of Δt.

Thus, it is easy to see that the bound

$$\|\widetilde{S}(\Delta t \hat{A}_1^m, \ldots, \Delta t \hat{A}_n^m, \Delta t G_1^m, \ldots, \Delta t G_n^m)\| \leqslant C \Delta t \sum_{i=1}^{n} \|G_i^m\| \tag{18}$$

together with the bound (16) obtained for $\widetilde{R}(-\Delta t \hat{A}_1^m, \ldots, -\Delta t \hat{A}_n^m)$, will guarantee us the stability, in finite intervals of time, of the numerical integration process (2). Results (16) and (18) are proved in Sections 2 and 4, respectively.

To get the bound (16) we will use two main ideas. The first one, developed deeply in [15], is focused in the fact that a null variation in time of operators $A_i(t)$ ($M_i = 0$) implies that $\widetilde{R}(-\Delta t \hat{A}_1^m, \ldots, -\Delta t \hat{A}_n^m) \equiv R(-\Delta t A_1(t_m), \ldots, -\Delta t A_n(t_m))$ can be a contraction in H and, consequently, the FSRK will preserve the contractivity. Secondly, if we consider smooth variations of $A_i(t)$, then $\widetilde{R}(-\Delta t \hat{A}_1^m, \ldots, -\Delta t \hat{A}_n^m)$ can be viewed as a perturbation of $R(-\Delta t A_1(t_m), \ldots, -\Delta t A_n(t_m))$ and it will be bounded in the form

$$\|\widetilde{R}(-\Delta t \hat{A}_1^m, \ldots, -\Delta t \hat{A}_n^m) - R(-\Delta t A_1(t_m), \ldots, -\Delta t A_n(t_m))\| \leqslant C \Delta t.$$

In the next theorem, which we prove in Section 2, we set conditions to obtain (16).

Theorem 1.1. *Let* (7) *be an A-stable FSRK method such that*

$$\sum_{k=1}^{n} a_{ii}^k \neq 0, \quad \text{for } i = 1, \ldots, s \quad (\text{all their stages are implicit}) \tag{19}$$

and let $\{A_i(t)\}_{i=1}^n$ *be a linear maximal coercive system of operators satisfying:*

(a) *for each* $t \in [0, T]$ *the system of operators* $\{A_i(t)\}_{i=1}^n$ *is commutative and the commutative system of contractions* $\{(I - \Delta t A_i(t))(I + \Delta t A_i(t))^{-1}\}_{i\in\{1,\ldots,n\}}$ *admits a unitary dilation,*

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

$$\|A_i(t')u - A_i(t)u\| \leqslant |t - t'| M_i \|A_i(t)u\|, \\ \forall i = 1, \ldots, n, \ \forall t, t' \in [0, T] \text{ and } \forall u \in \mathcal{D}_i. \tag{20}$$

Then there exists a constant β, *independent of* Δt, *such that* (16) *is verified.*

When the operators $\{A_i(t)\}_{i=1}^n$ are also self-adjoint we can weaken the A-stability requirements to obtain (16).

Definition 1.3. An additive RK method is said A(0)-stable iff

$$\left|R(\bar{x})\right| \leqslant 1, \quad \forall \bar{x} \in \left\{(x_1, \dots, x_n) \mid x_i \leqslant 0, i = 1, \dots, n\right\}.$$

In similar way of the previous theorem we have proven

Theorem 1.2. *Let* (7) *be an A(0)-stable FSRK method such that satisfies* (19) *and let* $\{A_i(t)\}_{i=1}^n$ *be a commutative system of self-adjoint linear maximal coercive operators fulfilling* (20). *Then there exists a constant* β *such that* (16) *is verified with* β *independent of* Δt.

For problems with a small time variation in operators $A_i(t)$ it is possible to obtain the stability result (17) even in infinite-length intervals of time (i.e., for $M = \infty$), if we impose some additional A-stability conditions to the additive RK which we use to integrate them. Concretely in Section 2 we will show that conditions (20) together with M_i small permit us to prove a stronger contractivity result of type

$$\left\|\widetilde{R}\left(-\Delta t \hat{A}_1^m, \dots, -\Delta t \hat{A}_n^m\right)\right\| \leqslant \mathrm{e}^{-\beta \Delta t} \tag{21}$$

under some additional stability requirements on functions $R(z_1, \dots, z_n)$ which we introduce in the following two definitions:

Definition 1.4. An additive RK method is said strongly A-stable if it is A-stable and there exist $c < 1$ and M such that $R(\bar{z}) < c$ for all $\bar{z} \in \{(z_1, \dots, z_n) \mid z_i \in \mathbb{C}$ and $\mathrm{Re}(z_i) \leqslant 0,\ i = 1, \dots, n,$ and $|z_1| + \cdots + |z_n| \geqslant M\}$.

Definition 1.5. An additive RK method is said strongly A(0)-stable if

(1) $|R(\bar{x})| < 1,\ \forall \bar{x} \in \{(x_1, \dots, x_n) \mid x_i \leqslant 0,\ i = 1, \dots, n,$ and $x_1 + \cdots + x_n < 0\}$.
(2) There exist $c < 1$ and M such that $R(\bar{x}) < c,\ \forall \bar{x} \in \{(x_1, \dots, x_n) \mid x_i \leqslant 0,\ i = 1, \dots, n,$ and $|x_1| + \cdots + |x_n| \geqslant M\}$.
(3) $(\partial R/\partial x_i)(0, 0, \dots, 0) = -1, i = 1, \dots, n$.

The rest of this paper is structured in three sections. Section 2 is mainly devoted to the proof of the Theorem 1.1; the proof of Theorem 1.2 and the obtaining of (21) are light variations of this proof which we also explain in this section. In Section 3 we obtain similar stability results for FSRK methods with an explicit first stage and with additional restrictions on the coefficients of the last stage of them. Finally, in Section 4 we will give some technical results which we have used in the previous two sections for proving Theorems 1.1, 1.2, 3.1 and 3.2.

Henceforth we will use $C, C_1, C_2, C_3,$ as arbitrary constants independent of Δt.

2. Proof of Theorem 1.1

Let us take in H^s any norm induced by the norm of H and any norm of $\mathbb{R}^s$, for example: $\|u\|_{\infty,H^s} = \max_{1 \leqslant i \leqslant s} \|u_i\|$ and $\|u\|_{1,H^s} = \sum_{i=1}^s \|u_i\|$. Note that any two of these norms in H^s are equivalent.

For each FSRK we denote

$$\hat{A}^m = \begin{pmatrix} A_{k_1}(t_{m,1}) & & \\ & \ddots & \\ & & A_{k_s}(t_{m,s}) \end{pmatrix} \quad \text{and} \quad \check{A}^m = \begin{pmatrix} A_{k_1}(t_m) & & \\ & \ddots & \\ & & A_{k_s}(t_m) \end{pmatrix}, \tag{22}$$

where $k_1, \ldots, k_s \in \{1, \ldots, n\}$ are defined in (9).

Conditions (6) imposed to the FSRK coefficients permit us to use notation (22) to reduce the expression (14) since the following equalities are true:

$$\sum_{i=1}^{n} \overline{\mathcal{A}^i} \hat{A}_i^m = \overline{\mathcal{A}} \hat{A}^m, \qquad \sum_{i=1}^{n} \left(\overline{b^i}\right)^{\mathrm{T}} \hat{A}_i^m = \overline{b}^{\mathrm{T}} \hat{A}^m,$$

and, analogously, if we consider

$$\check{A}_i^m = \begin{pmatrix} A_i(t_m) & 0 & \ldots & 0 \\ 0 & A_i(t_m) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & A_i(t_m) \end{pmatrix} \in \mathcal{L}(\mathcal{D}_i, H)^{s \times s}, \tag{23}$$

it is also true that

$$\sum_{i=1}^{n} \overline{\mathcal{A}^i} \check{A}_i^m = \overline{\mathcal{A}} \check{A}^m \quad \text{and} \quad \sum_{i=1}^{n} \left(\overline{b^i}\right)^{\mathrm{T}} \check{A}_i^m = \overline{b}^{\mathrm{T}} \check{A}^m.$$

Using these notations, we will consider now the following decomposition for the transition operator

$$\widetilde{R}\left(-\Delta t \hat{A}_1^m, \ldots, -\Delta t \hat{A}_n^m\right) = R\left(-\Delta t A_1(t_m), \ldots, -\Delta t A_n(t_m)\right) + \Delta t P \tag{24}$$

with

$$P = \overline{b}^{\mathrm{T}} \left(\overline{\mathcal{A}}\right)^{-1} \left(\bar{I} + \Delta t \overline{\mathcal{A}} \check{A}^m\right)^{-1} \left(\overline{\mathcal{A}} \check{A}^m - \overline{\mathcal{A}} \hat{A}^m\right) \left(\bar{I} + \Delta t \overline{\mathcal{A}} \hat{A}^m\right)^{-1} \bar{e}; \tag{25}$$

in Lemma 4.4 we will prove that this decomposition is possible under conditions of Theorem 1.1.

As we have supposed that the FSRK method is A-stable and that the commutative system of operators $\{(I - \Delta t A_i(t))(I + \Delta t A_i(t))^{-1}\}_{i=1}^{n}$ admits unitary dilation, Theorem 2.3 of [15] ensures that

$$\left\| R\left(-\Delta t A_1(t_m), \ldots, -\Delta t A_n(t_m)\right) \right\| \leqslant 1. \tag{26}$$

Thus, to get (16), it is still to be proven that the operator P given in (25) is bounded independently of Δt. In order to get this bound, we decompose the operator P in the form $P = P_3 P_2 P_1$ where:

$$P_3 \equiv \overline{b}^{\mathrm{T}} \left(\overline{\mathcal{A}}\right)^{-1} \left(\bar{I} + \Delta t \overline{\mathcal{A}} \check{A}^m\right)^{-1} \in L\left(H^s, H\right),$$
$$P_2 \equiv \left(\overline{\mathcal{A}} \check{A}^m - \overline{\mathcal{A}} \hat{A}^m\right) \left(\overline{\mathcal{A}} \hat{A}^m\right)^{-1} \in L\left(H^s, H^s\right),$$
$$P_1 \equiv \left(\overline{\mathcal{A}} \hat{A}^m\right) \left(\bar{I} + \Delta t \overline{\mathcal{A}} \hat{A}^m\right)^{-1} \bar{e} \in L\left(H, H^s\right),$$

and we bound P_3, P_2 and P_1 separately.

Lemma 4.3 ensures that

$$\|P_3\| \leqslant C_3, \tag{27}$$

where C_3 is a constant independent of Δt.

For bounding P_2 we rewrite it in the form $P_2 = \overline{\mathcal{A}}(\check{A}^m - \hat{A}^m)(\hat{A}^m)^{-1}(\overline{\mathcal{A}})^{-1}$; so

$$\|P_2\widetilde{V}\| \leqslant \|\bar{\mathcal{A}}\|\|(\check{A}^m - \hat{A}^m)(\hat{A}^m)^{-1}\widetilde{W}\| \leqslant C\|(\check{A}^m - \hat{A}^m)(\hat{A}^m)^{-1}\widetilde{W}\|,$$

where $\widetilde{W} = (\overline{\mathcal{A}})^{-1}\widetilde{V}$.

Let us prove now that for all $\widetilde{W} = (W^1, \ldots, W^s)^{\mathrm{T}} \in H^s$ it holds that

$$\|\widetilde{W^m}\| \equiv \|(\check{A}^m - \hat{A}^m)(\hat{A}^m)^{-1}\widetilde{W}\| \leqslant M\Delta t\|\widetilde{W}\|. \tag{28}$$

Note that every component $W^{m,i}$ of $\widetilde{W^m}$ satisfies

$$W^{m,i} = \left(A_{k_i}(t_m) - A_{k_i}(t_{m,i})\right)\left(A_{k_i}(t_{m,i})\right)^{-1}W^i,$$

and using (20) it is deduced that

$$\|W^{m,i}\| \leqslant |t_m - t_{m,i}|M_{k_i}\|A_{k_i}(t_{m,i})\left(A_{k_i}(t_{m,i})\right)^{-1}W^i\| = |c_i|\Delta t M_{k_i}\|W^i\|;$$

therefore (28) is verified with $M = \max_{i=1,\ldots,s}\{|c_i|M_{k_i}\}$; thus, we can conclude that

$$\|P_2\widetilde{V}\| \leqslant \|\mathcal{A}\|M\Delta t\|\widetilde{W}\| \leqslant C_2 M\Delta t\|\widetilde{V}\|, \tag{29}$$

where $C_2 = \|\mathcal{A}\|\,\|\mathcal{A}^{-1}\|$.

To check that $\|P_1\| \leqslant C_1/\Delta t$, with C_1 independent of Δt, it suffices to use Lemma 4.1 for proving that

$$\|L\widetilde{U}\| = \|\Delta t\overline{\mathcal{A}}\hat{A}^m(\bar{I} + \Delta t\overline{\mathcal{A}}\hat{A}^m)^{-1}\widetilde{U}\| \leqslant \|\widetilde{U}\| + \|(\bar{I} + \Delta t\overline{\mathcal{A}}\hat{A}^m)^{-1}\widetilde{U}\| \leqslant C\|\widetilde{U}\|.$$

As $\Delta t P_1 u = \Delta t L\bar{e}u$ it is clear that

$$\|\Delta t P_1\| \leqslant C_1. \tag{30}$$

Joining (26), (27), (29) and (30) we deduce

$$\|\widetilde{R}(-\Delta t\hat{A}_1^m, \ldots, -\Delta t\hat{A}_n^m)\| \leqslant 1 + C_1C_2C_3M\Delta t \leqslant \mathrm{e}^{\beta\Delta t}$$

with $\beta = C_1C_2C_3M = CM$. □

Remark 2.1. To prove Theorem 1.2 we carry out a process identical to this one, except that to obtain the bound (26) we apply Theorem 3.2 of [15].

Remark 2.2. If the FSRK method is strongly A-stable (or similarly if the FSRK method is strongly A(0)-stable and the operators $\{A_i(t)\}_{i=1}^n$ are self-adjoint) it is proven in Theorem 4.3 (Theorem 4.6 for the self-adjoint case) of [15] that $\|R(-\Delta t A_1(t_m), \ldots, -\Delta t A_n(t_m))\| \leqslant \mathrm{e}^{-\beta'\Delta t}$ under a coercitivity property of type $(A_i(t)v, v) \geqslant \alpha\|v\|^2$ (with $\alpha > 0$) for any $i = 1, \ldots, n$. Using the same reasoning of the proof of Theorem 1.1, for $\Delta t \in (0, \Delta t_0]$, it is deduced in these cases that

$$\|\widetilde{R}(-\Delta t\hat{A}_1^m, \ldots, -\Delta t\hat{A}_n^m)\| \leqslant \mathrm{e}^{-\beta'\Delta t} + CM\Delta t \leqslant \mathrm{e}^{\beta\Delta t},$$

where $\beta = -\beta' + CM\mathrm{e}^{\beta'\Delta t_0}$. Therefore, for $M < \beta'/(C\mathrm{e}^{\beta'\Delta t_0})$, i.e., for small time variations of coefficients of $A_i(t)$, we can obtain a negative β, and consequently, as we pointed out before, the stability of scheme (2) also in infinite intervals of time.

3. The case of an explicit first stage

Most of the classical alternating direction or fractional step methods are designed in such a way that their formulation as FSRK method has an explicit first stage. In order to include such methods in our analysis it is convenient to give some stability results in this case. Besides, in the development of new FSRK methods of high order (see [1]) it is interesting to consider reductions in the number of the order conditions by imposing (classical) restrictions of the form $\mathcal{A}^i e = c$, $\forall i = 1, \ldots, n$, and such restrictions require that the first stage will be explicit.

A FSRK has only the first stage explicit if it satisfies that

$$\sum_{i=1}^{n} a_{11}^{i} = 0 \quad \text{and} \quad \sum_{i=1}^{n} a_{jj}^{i} \neq 0, \quad \forall j = 2, \ldots, s. \tag{31}$$

Let us call k_1 ($\in \{1, \ldots, n\}$) the integer such that any of the coefficients $a_{i1}^{k_1} \neq 0$ or $b_1^{k_1} \neq 0$. [1]

In this section we take for $u \in \mathcal{D}_i$ the norm $\|u\|_{i,t} = \|u\| + \Delta t \|A_i(t)u\|$ and in $\mathcal{D}_i \times H^{s-1}$ any norm induced by the norm $\|\cdot\|_{i,t}$ for the first component, the norm of H for the remaining ones and any norm of $\mathbb{R}^s$; we also denote these equivalent norms by $\|\cdot\|_{i,t}$. In $\mathcal{D}_{k_1} \times \cdots \times \mathcal{D}_{k_s}$ we take the norms induced by the norm $\|\cdot\|_{k_i, t_{m,i}}$ of $\mathcal{D}_{k_i}$, for $i = 1, \ldots, s$, and any norm of $\mathbb{R}^s$ and we denote these equivalent norms by $\|\cdot\|_{\bar{k}, \bar{t}_m}$ where $\bar{k} = (k_1, \ldots, k_s)$ and $\bar{t}_m = (t_{m,1}, \ldots, t_{m,s})$. In the following results we use these metrics.

Using Lemmas 4.5–4.8 of Section 4 we can prove the following:

Theorem 3.1. *Let us consider an A-stable FSRK method such that satisfies* (31), (49) *and* (50) *and let* $\{A_i(t)\}_{i=1}^n$ *be a linear maximal coercive system of operators satisfying hypotheses (a) and (b) of Theorem* 1.1. *Then there exists a constant* β, *independent of* Δt, *such that* (16) *is verified.*

Proof. To bound the perturbation P we proceed in a different way with respect to the proof of Theorem 1.1 since $\mathcal{A}$ is not invertible. In this case, we decompose this operator in the form $P = P_1 P_2 P_3$ where:

$$\begin{aligned}
P_3 &\equiv \bar{\tau}^{\mathrm{T}} \big(\bar{I} + \Delta t \overline{\mathcal{A}} \check{A}^m\big)^{-1} : \mathcal{D}_{k_1} \times H^{s-1} \to \mathcal{D}_{k_1}, \\
P_2 &\equiv \overline{\mathcal{A}} \big(\check{A}^m - \hat{A}^m\big) : \mathcal{D}_{k_1} \times \cdots \times \mathcal{D}_{k_s} \to \mathcal{D}_{k_1} \times H^{s-1}, \\
P_1 &\equiv \big(\bar{I} + \Delta t \overline{\mathcal{A}} \hat{A}^m\big)^{-1} \bar{e} : \mathcal{D}_{k_1} \to \mathcal{D}_{k_1} \times \cdots \times \mathcal{D}_{k_s}.
\end{aligned}$$

Lemma 4.7 ensures that

$$\big\|\big(\bar{I} + \Delta t \overline{\mathcal{A}} \check{A}^m\big)^{-1} \widetilde{W}\big\|_{\bar{k}, \bar{t}_m} \leqslant C \big\|\widetilde{W}\big\|_{k_1, t_m}. \tag{32}$$

Applying (49) and (50) as well as $\tau^{\mathrm{T}} = (0, \ldots, 0, 1)$ we can deduce for $\widetilde{Z}^{\mathrm{T}} = (Z^1, \ldots, Z^s) \in \mathcal{D}_{k_1} \times \cdots \times \mathcal{D}_{k_s}$ that

$$\big\|\bar{\tau}^{\mathrm{T}} \widetilde{Z}\big\|_{k_1, t_{m+1}} = \big\|Z^s\big\|_{k_1, t_{m+1}} \leqslant C \big\|\widetilde{Z}\big\|_{\bar{k}, \bar{t}_m}. \tag{33}$$

Joining (32) and (33) we obtain that $P_3 \widetilde{W} \in \mathcal{D}_{k_1}$ satisfies

$$\big\|P_3 \widetilde{W}\big\|_{k_1, t_{m+1}} \leqslant C_3 \big\|\widetilde{W}\big\|_{k_1, t_m}, \tag{34}$$

[1] If k_1 does not exist then we will reduce the FSRK method to a FSRK method where all of its stages are implicit, by eliminating the first stage.

where C_3 is a constant independent of Δt.

To prove that P_2 is bounded, we use that the ith component of

$$\left(\check{A}^m - \hat{A}^m\right)\begin{pmatrix} V_1 \\ \vdots \\ V_s \end{pmatrix} \quad \text{is} \quad \left(A_{k_i}(t_m) - A_{k_i}(t_{m,i})\right)V^i, \quad \text{for } i = 1, \ldots, s,$$

and applying (20) we deduce that

$$\left\|\left(A_{k_i}(t_m) - A_{k_i}(t_{m,i})\right)V^i\right\| \leqslant |t_m - t_{m,i}| M_{k_i} \left\|A_{k_i}(t_{m,i})V^i\right\| \leqslant |c_i| M_{k_i} \left\|V^i\right\|_{k_i, t_{m,i}};$$

therefore taking also into account that the first row of $\mathcal{A}$ is null, it is deduced, for all $\widetilde{V} \in \mathcal{D}_{k_1} \times \cdots \times \mathcal{D}_{k_s}$, that

$$\left\|P_2 \widetilde{V}\right\|_{k_1, t_m} \leqslant C_2 M \left\|\widetilde{V}\right\|_{\bar{k}, \bar{t}_m}, \tag{35}$$

where $C_2 = \|\mathcal{A}\|$ and $M = \max_{i=2,\ldots,s}\{|c_i| M_{k_i}\}$.

Finally, to bound P_1 we use that Lemma 4.5 ensures that

$$\left\|\left(\bar{I} + \Delta t \overline{\mathcal{A}} \hat{A}^m\right)^{-1} \widetilde{U}\right\|_{\bar{k}, \bar{t}_m} \leqslant C \left\|\widetilde{U}\right\|_{k_1, t_m}, \quad \forall \widetilde{U} \in \mathcal{D}_{k_1} \times H^{s-1}. \tag{36}$$

As $\forall u \in \mathcal{D}_{k_1}$ it is obvious that $\bar{e}u \in \mathcal{D}_{k_1} \times H^{s-1}$ and

$$\|\bar{e}u\|_{k_1, t_m} \leqslant C \|u\|_{k_1, t_m}; \tag{37}$$

from (36) and (37) we conclude that

$$\left\|P_1 u\right\|_{\bar{k}, \bar{t}_m} \leqslant C_1 \|u\|_{k_1, t_m}. \tag{38}$$

Joining (34), (35) and (38) we deduce

$$\left\|Pu\right\|_{k_1, t_{m+1}} \leqslant C_1 C_2 C_3 M \|u\|_{k_1, t_m}. \qquad \square$$

In a similar way, we can prove for the self-adjoint operator case the following:

Theorem 3.2. *Let us consider an A(0)-stable FSRK method such that satisfies* (31), (49) *and* (50) *and let* $\{A_i(t)\}_{i=1}^n$ *be a commutative system of self-adjoint linear maximal coercive operators satisfying* (20). *Then there exists a constant* β *such that* (16) *is verified with* β *independent of* Δt.

A strong A-stability property for the FSRK together with a small variation in $A_i(t)$ permits to deduce the contractivity result (21) using the same reasoning of Section 2, for the case in which all the stages are implicit.

4. Technical lemmas

Lemma 4.1. *Let us consider a FSRK method satisfying* (19) *and let* $\{A_i(t)\}_{i=1}^n$ *be a system of linear maximal coercive operators in H. Then the operator*

$$\left(\bar{I} + \Delta t \sum_{i=1}^{n} \overline{\mathcal{A}^i} \hat{A}_i^m\right) : \mathcal{D}_{k_1} \times \cdots \times \mathcal{D}_{k_s} \to H^s,$$

associated to such FSRK, is invertible, and the inverse operator is bounded in H^s*, independently of* Δt.

Proof. We will use an induction principle in the components of H^s to prove that for each $\widetilde{g} \equiv (g_1, \ldots, g_s)^{\mathrm{T}} \in H^s$, there exists a unique $\widetilde{U} \equiv (U^{m,1}, \ldots, U^{m,s})^{\mathrm{T}} \in H^s$ such that

$$\left(\bar{I} + \Delta t \sum_{i=1}^{n} \overline{\mathcal{A}^i}\, \hat{A}_i^m\right) \widetilde{U} = \widetilde{g},$$

and it holds that $\|\widetilde{U}\|_{\infty, H^s} \leqslant C \|\widetilde{g}\|_{1,H^s}$.

Starting with the first stage, conditions (8) and (19) imposed on the coefficients of a FSRK method imply that $U^{m,1}$ is solution of a (stationary) problem of type

$$U^{m,1} + \Delta t a_{11}^{k_1} A_{k_1}(t_{m,1}) U^{m,1} = g_1, \tag{39}$$

where $a_{11}^{k_1} > 0$. As we are supposing that $\{A_i(t)\}_{i=1}^n$ are maximal and coercive operators in H, it holds (see [4]) that

$$\left\| \left(I + \Delta t a_{11}^{k_1} A_{k_1}(t_{m,1})\right)^{-1} \right\| \leqslant 1;$$

thus

$$\left\|U^{m,1}\right\| \leqslant \|g_1\|. \tag{40}$$

If we rewrite (39) in the form

$$\Delta t a_{11}^{k_1} A_{k_1}(t_{m,1}) U^{m,1} = g_1 - U^{m,1},$$

and we use (40) as well as $a_{11}^{k_1} > 0$, then it is clear that

$$\left\|\Delta t A_{k_1}(t_{m,1}) U^{m,1}\right\| \leqslant C \left\|U^{m,1}\right\|, \quad \text{with } C = \frac{2}{a_{11}^{k_1}}.$$

To apply the induction principle we assume that

$$\begin{cases} \left\|U^{m,l}\right\| \leqslant C \displaystyle\sum_{i=1}^{l} \|g_i\|, & \forall l = 1, \ldots, j-1, \\ \left\|\Delta t A_{k_l}(t_{m,l}) U^{m,l}\right\| \leqslant C \displaystyle\sum_{i=1}^{l} \|g_i\|, & \forall l = 1, \ldots, j-1, \end{cases} \tag{41}$$

and we prove that (41) is also true for $l = j$.

By using again the restrictions (8) and (19) on the FSRK method coefficients, the jth component of $\widetilde{U}$ is solution of the problem

$$U^{m,j} + \Delta t a_{jj}^{i_j} A_{k_j}(t_{m,j}) U^{m,j} = f_j, \quad \text{with } f_j = g_j - \Delta t \sum_{l=1}^{j-1} a_{jl}^{k_l} A_{k_l}(t_{m,l}) U^{m,l}. \tag{42}$$

Again the maximality and coercitivity of the operator $A_{k_j}(t_{m,j})$ ensure that

$$\left\|U^{m,j}\right\| \leqslant \|f_j\|, \tag{43}$$

and using the induction hypotheses (41) we have that

$$\|f_j\| \leqslant \|g_j\| + C \sum_{l=1}^{j-1} \left\|\Delta t A_{k_l}(t_{m,l}) U^{m,l}\right\| \leqslant C \sum_{l=1}^{j} \|g_l\|, \tag{44}$$

therefore

$$\|U^{m,j}\| \leqslant C\sum_{i=1}^{j}\|g_i\|.$$

Using (42)–(44) we can also deduce easily the bound

$$\|\Delta t A_{k_j}(t_{m,j})U^{m,j}\| \leqslant C\sum_{i=1}^{j}\|g_i\|. \qquad \square$$

Lemma 4.2. *Under the same hypotheses of Lemma* 4.1*, the operator*

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

is bounded independently of Δt.

Proof. We consider now $U^m = T^m\widetilde{g}$, where $\widetilde{g} = (g_1,\dots,g_s)^{\mathrm T}\in H^s$; T^m can be staged as

$$\begin{cases} U^m = \Delta t\displaystyle\sum_{i=1}^{s} b_i^{k_i}A_{k_i}(t_{m,i})U^{m,i}, \quad \text{where} \\ \left(I+\Delta t a_{ii}^{k_i}A_{k_i}(t_{m,i})\right)U^{m,i} = g_i - \Delta t\displaystyle\sum_{j=1}^{i-1} a_{ij}^{k_j}A_{k_j}(t_{m,j})U^{m,j}, \quad \text{for } i=1,\dots,s, \\ \quad \text{with } k_i,k_j\in\{1,\dots,n\}. \end{cases} \tag{45}$$

Thus, as we proved (41) for $l=1,\dots,s$, it is immediately deduced that

$$\|U^m\| = \|T^m\widetilde{g}\| \leqslant C\|\widetilde{g}\|_{1,H^s}. \tag{46}$$

$\square$

Remark 4.1. It is clear that (46) permits us to deduce (18) since

$$\widetilde{S}\left(\Delta t\hat{A}_1^m,\dots,\Delta t\hat{A}_n^m,\Delta t G_1^m,\dots,\Delta t G_n^m\right) = \Delta t\sum_{i=1}^{n}\left(\overline{b^i}\right)^{\mathrm T}G_i^m - \Delta t T^m\widetilde{g},$$

where $\widetilde{g} = \sum_{k=1}^{n}\overline{\mathcal{A}^k}G_k^m$.

Using the notations (22) and (23), introduced in Section 2, we will prove the following lemmas:

Lemma 4.3. *Let us consider a FSRK method satisfying* (19) *and let* $\{A_i(t)\}_{i=1}^n$ *be a system of linear maximal coercive operators in H. Then the operator*

$$\left(\bar{I}+\Delta t\sum_{i=1}^{n}\overline{\mathcal{A}^i}\check{A}_i^m\right): \mathcal{D}_{k_1}\times\cdots\times\mathcal{D}_{k_s}\to H^s,$$

associated to such FSRK, is invertible, and the inverse operator is bounded in H^s*, independently of* Δt.

Proof. It is identical to the proof of Lemma 4.1. $\square$

Lemma 4.4. *Let us consider a FSRK method satisfying* (19) *and let* $\{A_i(t)\}_{i=1}^n$ *be a system of linear maximal coercive operators. Then the decomposition* (24), (25) *for the operator* $\widetilde{R}(-\Delta t\hat{A}_1^m, \ldots, -\Delta t\hat{A}_n^m)$ *is valid.*

Proof. As the FSRK method satisfies (19), $\mathcal{A} = \sum_{i=1}^n \mathcal{A}^i$ is invertible; using notation (22) we can rewrite (14) as follows:

$$\widetilde{R}(-\Delta t\hat{A}_1^m, \ldots, -\Delta t\hat{A}_n^m) = \bar{I} - \Delta t\bar{b}^{\mathrm{T}}(\overline{\mathcal{A}})^{-1}\overline{\mathcal{A}}\hat{A}^m(\bar{I} + \Delta t\overline{\mathcal{A}}\hat{A}^m)^{-1}\bar{e},$$

and using the same notation we can decompose $R(-\Delta tA_1(t_m), \ldots, -\Delta t\ A_n(t_m))$ in a similar format:

$$\bar{I} - \Delta t\bar{b}^{\mathrm{T}}(\overline{\mathcal{A}})^{-1}\overline{\mathcal{A}}\check{A}^m(\bar{I} + \Delta t\overline{\mathcal{A}}\check{A}^m)^{-1}\bar{e} = \bar{I} - \bar{b}^{\mathrm{T}}(\overline{\mathcal{A}})^{-1}(\bar{I} - (\bar{I} + \Delta t\overline{\mathcal{A}}\check{A}^m)^{-1})\bar{e}.$$

In this way, it is immediately seen that

$$\widetilde{R}(-\Delta t\hat{A}_1^m, \ldots, -\Delta t\hat{A}_n^m) = \bar{I} - \bar{b}^{\mathrm{T}}(\overline{\mathcal{A}})^{-1}\bar{e} + \bar{b}^{\mathrm{T}}(\overline{\mathcal{A}})^{-1}(\bar{I} + \Delta t\overline{\mathcal{A}}\hat{A}^m)^{-1}\bar{e}, \tag{47}$$

and, analogously, it is deduced that

$$R(-\Delta tA_1(t_m), \ldots, -\Delta tA_n(t_m)) = \bar{I} - \bar{b}^{\mathrm{T}}(\overline{\mathcal{A}})^{-1}\bar{e} + \bar{b}^{\mathrm{T}}(\overline{\mathcal{A}})^{-1}(\bar{I} + \Delta t\overline{\mathcal{A}}\check{A}^m)^{-1}\bar{e}. \tag{48}$$

Using (47) and (48) we obtain

$$\begin{aligned}&\widetilde{R}(-\Delta t\hat{A}_1^m, \ldots, -\Delta t\hat{A}_n^m)\\ &\quad= R(-\Delta tA_1(t_m), \ldots, -\Delta tA_n(t_m)) + \bar{b}^{\mathrm{T}}(\overline{\mathcal{A}})^{-1}\big[(\bar{I} + \Delta t\overline{\mathcal{A}}\hat{A}^m)^{-1} - (\bar{I} + \Delta t\overline{\mathcal{A}}\check{A}^m)^{-1}\big]\bar{e};\end{aligned}$$

from this, it is immediately deduced (24), (25). □

Now we will expose briefly some similar results that we use to prove Theorems 3.1 and 3.2 for the case of an explicit first stage.

Lemma 4.5. *Let us consider a FSRK method satisfying* (31) *and let* $\{A_i(t)\}_{i=1}^n$ *be a system of linear maximal coercive operators in H. Then the operator*

$$\left(\bar{I} + \Delta t\sum_{i=1}^n \overline{\mathcal{A}^i}\hat{A}_i^m\right) : \mathcal{D}_{k_1} \times \cdots \times \mathcal{D}_{k_s} \to \mathcal{D}_{k_1} \times H^{s-1},$$

associated to such FSRK method, is invertible, and the inverse operator verifies

$$\left\|\left(\bar{I} + \Delta t\sum_{i=1}^n \overline{\mathcal{A}^i}\hat{A}_i^m\right)^{-1}\widetilde{U}\right\|_{\bar{k},\bar{t}_m} \leqslant C\|\widetilde{U}\|_{k_1,t_m},$$

where C is independent of Δt.

Proof. The proof of this lemma is similar to the corresponding to Lemma 4.1, excepting that, in this case, the first stage is

$$U^{m,1} = g_1,$$

thus, it is obvious that

$$\|U^{m,1}\|_{k_1,t_m} = \|g_1\|_{k_1,t_m},$$

and

$$\left\|\Delta t A_{k_1}(t_m)U^{m,1}\right\| \leqslant \|g_1\|_{k_1,t_m}. \qquad \square$$

Applying an identical reasoning to the one used in Lemmas 4.2 and 4.3 except for the first stage, we have proved the following two results:

Lemma 4.6. *Under the same hypotheses of Lemma* 4.5*, the operator*

$$T^m \equiv \Delta t \sum_{i=1}^{n} \left(\overline{b^i}\right)^{\mathrm{T}} \hat{A}_i^m \left(\bar{I} + \Delta t \sum_{j=1}^{n} \overline{\mathcal{A}^j}\,\hat{A}_j^m\right)^{-1} : \mathcal{D}_{k_1} \times H^{s-1} \to H$$

satisfies $\|T^m \widetilde{g}\| \leqslant C\|\widetilde{g}\|_{k_1,t_m}$ *for all* $\widetilde{g} \equiv (g_1, \ldots, g_s)^{\mathrm{T}} \in D_{k_1} \times H^{s-1}$.

If besides

$$g_s \in \mathcal{D}_{k_1}, \qquad a_{ss}^{k_1} \neq 0 \tag{49}$$

and [2]

$$(0, \ldots, 0, 1)\mathcal{A}^i = \left(b^i\right)^{\mathrm{T}}, \quad \forall i = 1, \ldots, n, \tag{50}$$

then $U^m \equiv T^m \widetilde{g} \in \mathcal{D}_{k_1}$ *and the operator* $T^m : \mathcal{D}_{k_1} \times H^{s-2} \times \mathcal{D}_{k_1} \to \mathcal{D}_{k_1}$ *is bounded independently of* Δt*, i.e.,*

$$\left\|T^m \widetilde{g}\right\|_{k_1,t_{m+1}} \leqslant C\left(\|g_1\|_{k_1,t_m} + \sum_{i=2}^{s-1} \|g_i\| + \|g_s\|_{k_1,t_{m+1}}\right).$$

Proof. It is sufficient to write now the operator T^m in the form (45) and repeat the calculations of the proof of Lemma 4.1, together with the indications of the Lemma 4.5 for the first stage.

When the additional hypotheses (49) and (50) are verified, we use $U^m = g_s - U^{m,s}$, $g_s \in \mathcal{D}_{k_1}$ and $U^{m,s} \in \mathcal{D}_{k_1}$, to deduce that $U^m \in \mathcal{D}_{k_s} \equiv \mathcal{D}_{k_1}$. $\square$

Note that the last result permits us to deduce the bound (18) in the norm $\|\cdot\|_{k_1,t_{m+1}}$.

As in the case without explicit stages, we will give an intermediate technical result, which is proved like Lemma 4.5.

Lemma 4.7. *Let us consider a FSRK method satisfying* (31) *and let* $\{A_i(t)\}_{i=1}^n$ *be a system of linear maximal coercive operators in H. Then the operator*

$$\left(\bar{I} + \Delta t \sum_{i=1}^{n} \overline{\mathcal{A}^i}\,\check{A}_i^m\right) : \mathcal{D}_{k_1} \times \cdots \times \mathcal{D}_{k_s} \to \mathcal{D}_{k_1} \times H^{s-1},$$

with $\check{A}_i^m$ *defined in* (23)*, associated to such FSRK, is invertible, and the inverse operator is bounded in* $\mathcal{D}_{k_1} \times H^{s-1}$*, independently of* Δt.

[2] The classical alternating direction methods formulated as FSRK methods with the first stage explicit satisfy (50) or, in other words, the last fractionary step calculates U^{m+1}. This condition is closely related to the concept of $\bar{\mathrm{A}}$-stability studied in [7].

Lemma 4.8. *Let us consider a FSRK method satisfying* (31) *and* (50) *and let* $\{A_i(t)\}_{i=1}^n$ *be a system of linear maximal coercive operators. Then it holds*

$$\widetilde{R}\big(-\Delta t\hat{A}_1^m,\ldots,-\Delta t\hat{A}_n^m\big)=R\big(-\Delta t A_1(t_m),\ldots,-\Delta t A_n(t_m)\big)+\Delta t P$$

with

$$P=\overline{\tau}^{\mathrm{T}}\big(\bar{I}+\Delta t\overline{\mathcal{A}}\check{A}^m\big)^{-1}\overline{\mathcal{A}}\big(\check{A}^m-\hat{A}^m\big)\big(\bar{I}+\Delta t\overline{\mathcal{A}}\hat{A}^m\big)^{-1}\bar{e}:\mathcal{D}_{k_1}\to\mathcal{D}_{k_1},\quad u\to Pu$$

and with $\tau^{\mathrm{T}}=(0,\ldots,0,1)$.

Proof. To prove this lemma we have to use only that the hypotheses (50) implies that $b^{\mathrm{T}}=\tau^{\mathrm{T}}\mathcal{A}$ and repeat the reasoning employed to prove Lemma 4.4, excepting that now $\overline{\tau}^{\mathrm{T}}=(0,\ldots,0,I_H)$ plays the role of $\overline{b^{\mathrm{T}}}(\overline{\mathcal{A}})^{-1}$. □

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, 1999.
[2] B. Bujanda, J.C. Jorge, Third order fractional step methods for multi-dimensional evolutionary convection-diffusion problems, Finite Difference Methods: Theory and Applications, Nova Science (1999) 110–117.
[3] P. Brenner, V. Thomèe, On rational approximations of semigroups, SIAM J. Numer. Anal. 16 (1979) 694.
[4] H. Brezis, Analyse fonctionnelle, thèorie et applications, Masson, 1983.
[5] G.J. Cooper, A. Sayfy, Additive methods for the numerical solution of ordinary differential equations, Math. Comp. 35 (1980) 1159–1172.
[6] G.J. Cooper, A. Sayfy, Additive Runge–Kutta methods for stiff ordinary differential equations, Math. Comp. 40 (1983) 207–218.
[7] M. Crouzeix, Sur l'aproximation des èquations differentielles opèrationelles linèaires par des mèthodes de Runge–Kutta, These d'Etat, Univ. de Paris VI, 1975.
[8] M. Crouzeix, S. Larsson, S. Piskarev, V. Thomèe, The stability of rational approximations of analytic semigroups, BIT 33 (1993) 74–84.
[9] J. Douglas, H.H. Rachford, On the numerical solution of heat conduction problems in two- and three-space variables, Trans. Amer. Math. Soc. 82 (1956) 421–439.
[10] C. González, C. Palencia, Stability of Runge–Kutta methods for abstract time-dependent parabolic problems: the Hölder case, Math. Comp. 68 (1999) 73–89.
[11] E. Hairer, G. Wanner, Solving Ordinary Differential Equations II, Springer-Verlag, 1987.
[12] W. Hundsdorfer, Trapezoidal and midpoint splittings for initial-boundary value problems, Math. Comp. 67(233) (1998) 1047–1062.
[13] J.L. Lions, Equations differentielles operationelles, et problemes aux limites, Springer-Verlag, 1961.
[14] 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.
[15] J.C. Jorge, F. Lisbona, Contractivity results for alternating direction schemes in Hilbert spaces, Appl. Numer. Math. 15 (1994) 65–75.
[16] C. Palencia, A stability result for sectorial operators in Banach spaces, SIAM J. Numer. Anal. 30 (1993) 1373–1384.

[17] B. Sz-Nagy, Unitary dilations of Hilbert space operators and related topics, C.B.M.S. Regional Conference Series in Math. 19 (1974).
[18] P.J. van der Houwen, The development of Runge–Kutta methods for partial differential equations, Appl. Numer. Math. 20 (1996) 261–272.
[19] N.N. Yanenko, The Method of Fractional Steps, Springer, 1971.