

Blanca Bujanda, Juan Carlos Jorge

NEW EFFICIENT TIME INTEGRATORS FOR NON-LINEAR PARABOLIC PROBLEMS

Abstract. In this work a new numerical method is constructed for time-integrating multi-dimensional parabolic semilinear problems in a very efficient way. The method reaches the fourth order in time and it can be combined with standard spatial discretizations of any order to obtain unconditionally convergent numerical algorithms. The main theoretical results which guarantee this property are explained here, as well as the method characteristics which guarantee a very strong reduction of computational cost in comparison with classical discretization methods.

Keywords: fractional step methods, non-linear parabolic problems, convergence.

Mathematics Subject Classification: 65M06, 65M12.

1. INTRODUCTION

In this paper we deal with the construction of numerical methods which efficiently integrate the following non-linear parabolic problems:

$$\begin{cases} \text{find } y(t) : [t_0, T] \rightarrow H, \text{ a solution of} \\ y'(t) = L(t)y(t) + f(t) + g(t, y(t)), \\ y(t_0) = y_0. \end{cases} \quad (1)$$

Here H is a Hilbert space of functions defined on a certain domain $\Omega \subseteq \mathbb{R}^n$, $L(t)$ is a linear second order elliptic differential operator, $f(t)$ is the source term, $g(t, y(t))$ is a non-linear function which meets some additional requirements and $y_0 \equiv y_0(\bar{x})$ is the initial condition.

The numerical integration of this problem can be viewed as a double process of discretization which adequately combines a time integrator (typically an adapted ODE Solver) with a discretization of the spatial variables. It is well known that the use of classical implicit schemes for the numerical integration of this type of problems poses mainly two problems. On the one hand, the convergence of the numerical scheme is usually obtained by imposing excessively restrictive stability requirements which force

us to use low order methods (like the implicit Euler rule) or complicated fully implicit methods for reaching higher orders of convergence in time. On the other hand, we must use slow iterative methods to resolve the internal stage equations because they are defined as the solution of large non-linear systems. Also, in this type of schemes, the convergence is deduced by imposing strong stability requirements (for example, B-stability).

In [5] a semiexplicit method, which is used to determinate the contribution of the linear term, is combined in an additive way with an explicit method used for the contribution of the non-linear term. Thus, linear systems only appear in the computation of the internal stages. In [7], splitting methods are applied in the numerical integration of some non-linear advection-diffusion-reaction problems; here the advection terms are integrated with an explicit method.

In order to avoid these problems, in [3] a new class of linearly implicit methods is developed; they are built by combining a standard spatial discretization of Finite Differences or Finite Elements type with a suitable time discretization, which is deduced from a Fractional Step Runge–Kutta (FSRK) method to deal with the linear terms and an explicit Runge–Kutta method to define the contribution of the non-linear term. By applying this new linearly implicit methods (see [1, 2, 3]) we obtain numerical algorithms which are convergent under a single requirement of the linear absolute stability type. It is well known (see [8]) that the Fractional Step methods are very efficient in integrating multidimensional linear parabolic problems if we suitably decompose the linear elliptic operator and the source term; the advantages of these methods manifest themselves also in these new linearly implicit methods. Thus in the resolution of the internal stages only we must solve linear systems with very simple matrices and it is not necessary to apply classical iterative methods to solve them. For example, if the elliptic linear operator $L(t)$ is $\sum_{i=1}^n d_i(t) \frac{\partial^2 y}{\partial x_i^2} + \sum_{i=1}^n v_i(t) \frac{\partial y}{\partial x_i} + \sum_{i=1}^n k_i(t)$, and we realize a spatial discretization by applying classical Finite Differences, the matrices which appear at each stage are tridiagonal; consequently, the computational cost of the final method is of the same order as that of an explicit method.

2. NUMERICAL SCHEME

By applying firstly a spatial discretization process of type Finite Difference, Finite Elements, . . . to problem (1) we obtain a family of Initial Value Problems which depend on the parameter $h \in (0, h_0]$ used to discretize in space as follows:

$$\begin{aligned} &\text{Find } Y_h(t) : [t_0, T] \rightarrow V_h, \text{ a solution of} \\ &\begin{cases} Y_h'(t) = L_h(t)Y_h(t) + f_h(t) + g_h(t, Y_h(t)), \\ Y_h(t_0) = y_{h0}. \end{cases} \end{aligned} \quad (2)$$

For each h it is common to consider a finite dimensional space V_h ; such space is the space of discrete functions on a mesh in Finite Differences, the subspace of H of piecewise polynomial functions in a classical Finite Element discretization. We

suppose that such spatial semidiscretization process is uniformly convergent of order q , i.e., for sufficiently smooth functions $y(t)$ there is

$$\|\pi_h y(t) - Y_h(t)\|_h \leq C h^q, \quad \forall t \in [t_0, T]. \tag{3}$$

Here, the operator π_h denotes the restriction to the mesh nodes if Finite Differences used and if we use the standard Finite Elements, then π_h are suitable projections in the space V_h .

In order to integrate efficiently this problem by applying the linearly implicit numerical methods which are first appeared in [3], we must adequately decompose the spatial discretization of the elliptic operator and the source term in n addends as follows: $L_h(t) = \sum_{i=1}^n L_{ih}(t)$ and $f_h(t) = \sum_{i=1}^n f_{ih}(t)$; the FSRK method will act on these terms. The contribution of the non-linear term $g_h(t, U_h(t))$ is given by the explicit RK method. From this process, we obtain the following scheme:

$$\left\{ \begin{array}{l} Y_h^{m+1} = Y_h^m + \tau \sum_{i=1}^s b_i^{k_i} \left(L_{k_i h}(t_{m,i}) Y_h^{m,i} + f_{k_i h}(t_{m,i}) \right) + \tau \sum_{i=1}^s b_i^{n+1} g_h(t_{m,i}, Y_h^{m,i}), \\ \text{where} \\ Y_h^{m,i} = Y_h^m + \tau \sum_{j=1}^i a_{ij}^{k_j} \left(L_{k_j h}(t_{m,j}) Y_h^{m,j} + f_{k_j h}(t_{m,j}) \right) + \tau \sum_{j=1}^{i-1} a_{ij}^{n+1} g_h(t_{m,j}, Y_h^{m,j}). \end{array} \right. \tag{4}$$

In this algorithm, we denote with Y_h^m the numerical approximations to the exact solution $y(t_m)$ at the times $t_m = t_0 + m \tau$ (τ is the time step), $U_h^0 = U_h(t_0)$, $k_i, k_j \in \{1, \dots, n\}$, n is the number of the levels of the method, $t_{m,i} = t_0 + (m + c_i) \tau$ and the intermediate approximations $U_h^{m,i}$ for $i = 1, \dots, s$ are the internal stages of the method. By setting some coefficients to zero in the latest formula as follows

$$\mathcal{A}^k = (a_{ij}^k) \text{ where } a_{ij}^k = \begin{cases} a_{ij}^{n+1} & \text{if } k = n + 1 \text{ y } i > j, \\ a_{ij}^{k_j} & \text{if } k = k_j \text{ and } i \geq j, \\ 0 & \text{otherwise,} \end{cases}$$

$$b^k = (b_j^k) \text{ where } b_j^k = \begin{cases} b_j^{n+1} & \text{if } k = n + 1, \\ b_j^{k_j} & \text{if } k = k_j, \\ 0 & \text{otherwise,} \end{cases}$$

we can organize these coefficients in a table which can be considered as an extension of the tables introduced by Butcher for standard RK methods, in the following way

$\mathcal{C} e$	\mathcal{A}^1	\mathcal{A}^2	\dots	\mathcal{A}^n	\mathcal{A}^{n+1}
	$(b^1)^T$	$(b^2)^T$	\dots	$(b^n)^T$	$(b^{n+1})^T$

where $\mathcal{C} = \text{diag}(c_1, \dots, c_s)$ and $e = (1, \dots, 1)$.

We will use $(\mathcal{C}, \mathcal{A}^i, \mathbf{b}^i)_{i=1}^{n+1}$ to refer to this method shortly and $(\mathcal{C}, \mathcal{A}^i, \mathbf{b}^i)$ to denote every standard RK method involved in (4).

To study the convergence of numerical scheme (4) we introduce, as usual, the global error at the time $t = t_m$ by $E_h^m = \|\pi_h y(t_m) - Y_h^m\|_h$; thus we say that (4) is uniformly convergent of order p in time and of order q in space, if E_h^m satisfies

$$E_h^m \leq C(h^q + \tau^p). \quad (5)$$

To get this bound, we begin by studying the stability of numerical scheme (4); thus we rewrite it in a more compact way by using the following tensorial notation: given $M \equiv (m_{ij}) \in \mathbb{R}^{s \times s}$ and $v \equiv (v_i) \in \mathbb{R}^s$, we denote $\bar{M} \equiv (m_{ij} I_{V_h}) \in V_h^{s \times s}$ and $\bar{v} \equiv (v_i I_{V_h}) \in V_h^s$; we group the evaluations of the source terms $f_{ih}(t)$ and of the linear operators $L_{ih}(t)$ for $i = 1, \dots, n$ and for all $m = 1, 2, \dots$, $F_{ih}^m = (f_{ih}(t_{m,1}), \dots, f_{ih}(t_{m,s}))^T \in V_h^s$ and $\hat{L}_{ih}^m = \text{diag}(L_{ih}(t_{m,1}), \dots, L_{ih}(t_{m,s})) \in \mathcal{L}(V_h, V_h)$. We also group the contribution of the stages $\tilde{Y}_h^m = (Y_h^{m,1}, \dots, Y_h^{m,s})^T \in V_h^s$ and the nonlinear term $\hat{G}_h^m(\tilde{Y}_h^m) = (g_h(t_{m,1}, Y_h^{m,1}), \dots, g_h(t_{m,s}, Y_h^{m,s}))^T \in V_h^s$. In this way we obtain

$$\begin{cases} \left(\bar{I} - \tau \sum_{i=1}^n \bar{\mathcal{A}}^i \hat{L}_{ih}^m \right) \tilde{Y}_h^m = \bar{e} Y_h^m + \tau \sum_{i=1}^n \bar{\mathcal{A}}^i F_{ih}^m + \tau \bar{\mathcal{A}}^{n+1} \hat{G}_h^m(\tilde{Y}_h^m), \\ Y_h^{m+1} = Y_h^m + \tau \sum_{i=1}^n (\bar{\mathbf{b}}^i)^T (\hat{L}_{ih}^m \tilde{Y}_h^m + F_{ih}^m) + \tau (\bar{\mathbf{b}}^{n+1})^T \hat{G}_h^m(\tilde{Y}_h^m). \end{cases}$$

When the operator $(\bar{I} - \sum_{j=1}^n \bar{\mathcal{A}}^j \hat{L}_{jh}^m)$ is invertible¹⁾, the numerical solution can be written as

$$\begin{aligned} Y_h^{m+1} &= \tilde{R}(\tau \hat{L}_{1h}^m, \dots, \tau \hat{L}_{nh}^m) Y_h^m + \tilde{S}(\tau \hat{L}_{1h}^m, \dots, \tau \hat{L}_{nh}^m, \tau F_{1h}^m, \dots, \tau F_{nh}^m) \\ &+ \tilde{T}(\tau \hat{L}_{1h}^m, \dots, \tau \hat{L}_{nh}^m, \tau \hat{G}_h^m(\tilde{Y}_h^m)), \end{aligned} \quad (6)$$

where

$$\tilde{R}(\tau \hat{L}_{1h}^m, \dots, \tau \hat{L}_{nh}^m) = \bar{I} + \sum_{i=1}^n (\bar{\mathbf{b}}^i)^T \tau \hat{L}_{ih}^m \left(\bar{I} - \sum_{j=1}^n \bar{\mathcal{A}}^j \tau \hat{L}_{jh}^m \right)^{-1} \bar{e}, \quad (7)$$

$$\begin{aligned} \tilde{S}(\tau \hat{L}_{1h}^m, \dots, \tau \hat{L}_{nh}^m, \tau F_{1h}^m, \dots, \tau F_{nh}^m) &= \\ &= \tau \sum_{i=1}^n (\bar{\mathbf{b}}^i)^T \left(F_{ih}^m + \hat{L}_{ih}^m \left(\bar{I} - \tau \sum_{j=1}^n \bar{\mathcal{A}}^j \hat{L}_{jh}^m \right)^{-1} \left(\tau \sum_{k=1}^n \bar{\mathcal{A}}^k F_{kh}^m \right) \right) \end{aligned} \quad (8)$$

and

$$\begin{aligned} \tilde{T}(\tau \hat{L}_{1h}^m, \dots, \tau \hat{L}_{nh}^m, \tau \hat{G}_h^m(\tilde{Y}_h^m)) &= \\ &= \tau \sum_{i=1}^n (\bar{\mathbf{b}}^i)^T \hat{L}_{ih}^m \left(\bar{I} - \sum_{j=1}^n \bar{\mathcal{A}}^j \hat{L}_{jh}^m \right)^{-1} \tau \bar{\mathcal{A}}^{n+1} \hat{G}_h^m(\tilde{Y}_h^m) + \tau (\bar{\mathbf{b}}^{n+1})^T \hat{G}_h^m(\tilde{Y}_h^m). \end{aligned} \quad (9)$$

¹⁾ In [4] it is proved that this operator is invertible and that its inverse operator is bounded independently of h and τ .

Note that first term (7) is the transition operator if we consider linear homogeneous problems; the contribution of the source term appears in second term (8) and, finally, the contribution of the nonlinear term $g_h(t, Y_h(t))$ is defined by third addend (9).

To study the stability of the numerical scheme we must bound these three addends (see [1]). The contribution of the linear terms, given by (7), can be bounded by applying the following result (see [4]):

Theorem 2.1. *Let an A-stable FSRK method be given such that all their stages are implicit (i.e., $\sum_{k=1}^n a_{ii}^k \neq 0$, for $i = 1, \dots, s$) or its first stage only is explicit (i.e., $\sum_{k=1}^n a_{ii}^k \neq 0$, for $i = 2, \dots, s$ and $a_{11}^{k_1} = 0$ for all $i = 1, \dots, s$) and it also satisfies $(0, \dots, 0, 1)^T \mathcal{A}^i = (b^i)^T$ and $a_{ss}^{k_1} \neq 0$, and let $\{L_{ih}(t)\}_{i=1}^n$ be a linear maximal coercive system of operators such that:*

- a) for each $t \in [0, T]$ the system of operators $\{L_{ih}(t)\}_{i=1}^n$ is commutative,
- b) there exist n constants M_i such that $\|L_{ih}(t')Y_h - L_{ih}(t)Y_h\|_h \leq |t-t'| C \|L_{ih}(t)Y_h\|_h \forall t, t' \in [t_0, T]$.

Then there exists a constant γ , independent of τ , such that

$$\|\tilde{R}(\tau \hat{L}_{1h}^m, \dots, \tau \hat{L}_{nh}^m)\|_h \leq e^{\gamma\tau} \tag{10}$$

holds.

In [4], the following bound for second addend (8) is also proved to hold:

$$\|\tilde{S}(\tau \hat{L}_{1h}^m, \dots, \tau \hat{L}_{nh}^m, \tau F_{1h}^m, \dots, \tau F_{nh}^m)\|_h \leq C \tau \sum_{i=1}^n \|F_{ih}^m\|_h.$$

To finish the study of the stability, we must bound third term (9); in [1], the following result is proved, rewritten for our purposes:

Theorem 2.2. *Let (4) be a linearly implicit method meeting the conditions of Theorem 2.1. If we use such method to integrate nonlinear parabolic problem (2) in time, where the nonlinear part $g_h(t, Y_h)$ satisfies $\|g_h(t, X_h) - g_h(t, Y_h)\|_h \leq L \|X_h - Y_h\|_h$, $\forall t \in [t_0, T]$, then:*

$$\begin{aligned} \|\tilde{T}(\tau \hat{L}_{1h}^m, \dots, \tau \hat{L}_{nh}^m, \tau \hat{G}_h^m(\tilde{X}_h^m)) - \tilde{T}(\tau \hat{L}_{1h}^m, \dots, \tau \hat{L}_{nh}^m, \tau \hat{G}_h^m(\tilde{Y}_h^m))\|_h \\ \leq C \tau \|X_h^m - Y_h^m\|_h. \end{aligned} \tag{11}$$

To study the consistency of the numerical scheme, we define the local error of the time discretization method as $e_h^m = \|Y_h(t_m) - \hat{Y}_h^m\|_h$, where \hat{Y}_h^m is obtained with one step of (4) taking as starting value $Y_h^{m-1} = Y_h(t_{m-1})$.

We say that (4) is uniformly consistent of order p if

$$e_h^m \leq C \tau^{p+1}, \quad \forall m \geq 0. \tag{12}$$

In [2] a theorem which can be rewritten for this case as follows is proved:

Theorem 2.3. *Let us assume that problem (2) satisfies the smoothness requirements*

$$\|Y_{i_1 h}^{(p+1)}(t)\| \leq C, \quad \|L_{i_1 h}^{(\varrho_1)}(t) \cdots L_{i_{\ell-1} h}^{(\varrho_{\ell-1})}(t) Y_{i_\ell h}^{(\varrho_\ell)}(t)\| \leq C, \quad \forall t \in [t_0, T],$$

$$\forall \ell \in \{2, \dots, p\}, \quad \forall (i_1, \dots, i_\ell) \in \{1, \dots, n+1\}^\ell, \quad \forall (\varrho_1, \dots, \varrho_{\ell-1}) \in \{0, \dots, p-1\}^{\ell-1}$$

$$\text{and } \forall \varrho_\ell \in \{r+1, \dots, p\} \text{ such that } 2 \leq \ell + \sum_{k=1}^{\ell} \varrho_k \leq p+2,$$

where $Y'_{ih}(t) = L_{ih}(t)Y_h(t) + f_{ih}(t)$, $\forall i = 1, \dots, n+1$ with $L_{(n+1)h}(t) = \frac{\partial g}{\partial Y_n}(t, Y_h(t))$ and $f_{(n+1)h}(t) = \frac{\partial g}{\partial Y_n}(t, Y_h(t))Y_h(t) + g_h(t, Y_h(t))$.

Let $(C, \mathcal{A}^i, b^i)_{i=1}^{n+1}$ be a linearly implicit FSRK method satisfying the reductions

$$(C)^k e - k \mathcal{A}^j (C)^{k-1} e = 0, \quad \forall j = 1, \dots, n+1, \quad k = 1, \dots, r$$

with $r = E \left[\frac{p-1}{2} \right]$ together with the order conditions

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

$$(b^{i_1})^T (C)^{\rho_1} \mathcal{A}^{i_2} (C)^{\rho_2} \cdots \mathcal{A}^{i_\ell} (C)^{\rho_\ell} e = \prod_{j=1}^{\ell} \frac{1}{(\ell - j + 1) + \sum_{k=j}^{\ell} \rho_k},$$

$$\forall \ell \in \{2, \dots, p\}, \quad \forall (i_1, \dots, i_\ell) \in \{1, \dots, n+1\}^\ell,$$

$$\forall (\rho_1, \dots, \rho_{\ell-1}) \in \{0, \dots, p-1\}^{\ell-1}$$

$$\text{and } \forall \rho_\ell \in \{r+1, \dots, p-1\} \text{ such that } 1 \leq \ell + \sum_{k=1}^{\ell} \rho_k \leq p.$$

Then (12) is true.

To obtain the convergence of the scheme (bound (5)), we decompose the global error as follows:

$$E_h^m \leq \|\pi_h y(t_m) - Y_h(t_m)\|_h + \|Y_h(t_m) - \hat{Y}_h^m\|_h + \|\hat{Y}_h^m - Y_h^m\|_h$$

where we have used the intermediate approximations $Y_h(t_m)$ and \hat{Y}_h^m . The space discretization is convergent of order q , thus by applying (3) we obtain

$$E_h^m \leq C h^q + e_h^m + \|\hat{Y}_h^m - Y_h^m\|_h.$$

Now we use the fact that the time discretization is consistent of order p (i.e. it satisfies (12))

$$E_h^m \leq C h^q + C \tau^{p+1} + \|\hat{Y}_h^m - Y_h^m\|_h.$$

Finally, we bound the third addend it by using (6)–(9) as follows

$$\begin{aligned} \|\hat{Y}_h^m - Y_h^m\|_h &= \|\tilde{R}(\tau \hat{L}_{1h}^m, \dots, \tau \hat{L}_{nh}^m) (\pi_h y(t_{m-1}) - Y_h(t_{m-1})) + \\ &\quad + \tilde{T}(\tau \hat{L}_{1h}^m, \dots, \tau \hat{L}_{nh}^m, \tau \hat{G}_h^m(\pi_h y(t_{m-1}))) - \\ &\quad - \tilde{T}(\tau \hat{L}_{1h}^m, \dots, \tau \hat{L}_{nh}^m, \tau \hat{G}_h^m(Y_h^{m-1}))\|_h. \end{aligned}$$

Thus, by applying (10) and (11), we obtain

$$E_h^m \leq C h^q + C \tau^{p+1} + (e^{\gamma \tau} + C \tau) E_h^{m-1}.$$

From this result is easy to conclude that scheme (4) satisfies (5).

3. A NEW FOURTH ORDER LINEARLY IMPLICIT METHOD

Now we present the main ideas of the construction process for a linearly implicit three-level FSRK method; such method attains fourth order in time when it is applied to the numerical integration of the semilinear problems of type (1) combined with spatial discretizations. To the best of our knowledge, this is the only fourth order method of this class which appears in the literature. The construction process for this kind of methods is very complicated due to two main drawbacks; the first one is related to the high number of order conditions that we must impose, which makes us cope with large and complicated non-linear systems. The second one is related to the additional requirements that we must impose in order to guarantee the stability of the final method.

Precisely, the order conditions that we must impose are (see [2]):

(r_i)	$\mathcal{A}^i e = C e$	for $i = 1, 2, 3$	
(α_1)	$(b^1)^T e = 1$	(α_2)	$(b^1)^T C e = \frac{1}{2}$
(α_4)	$(b^1)^T C C e = \frac{1}{3}$	(α_5)	$(b^1)^T C C C e = \frac{1}{4}$
(α_7)	$(b^1)^T \mathcal{A}^1 C C e = \frac{1}{12}$	(α_8)	$(b^1)^T \mathcal{A}^1 \mathcal{A}^1 C e = \frac{1}{24}$
(β_2)	$(b^2)^T C e = \frac{1}{2}$	(β_3)	$(b^2)^T \mathcal{A}^2 C e = \frac{1}{6}$
(β_5)	$(b^2)^T C C C e = \frac{1}{4}$	(β_6)	$(b^2)^T C \mathcal{A}^2 C e = \frac{1}{8}$
(β_8)	$(b^2)^T \mathcal{A}^2 \mathcal{A}^2 C e = \frac{1}{24}$	(γ_1)	$(b^3)^T e = 1$
(γ_3)	$(b^3)^T \mathcal{A}^3 C e = \frac{1}{6}$	(γ_4)	$(b^3)^T C C e = \frac{1}{3}$
(γ_6)	$(b^3)^T C \mathcal{A}^3 C e = \frac{1}{8}$	(γ_7)	$(b^3)^T \mathcal{A}^3 C C e = \frac{1}{12}$
(\times_{ij})	$(b^i)^T \mathcal{A}^j C e = \frac{1}{6}$	for $\{i, j\} \in \{1, 2, 3\}$	
(\diamond_{ij})	$(b^i)^T C \mathcal{A}^j C e = \frac{1}{8}$	for $\{i, j\} \in \{1, 2, 3\}$	
(\star_{ij})	$(b^i)^T \mathcal{A}^j C C e = \frac{1}{12}$	for $\{i, j\} \in \{1, 2, 3\}$	
(\odot_{ijk})	$(b^i)^T \mathcal{A}^{1j} \mathcal{A}^k C e = \frac{1}{24}$	for $\{i, j, k\} \in \{1, 2, 3\}$	
(α_3)	$(b^1)^T \mathcal{A}^1 C e = \frac{1}{6}$	(α_6)	$(b^1)^T C \mathcal{A}^1 C e = \frac{1}{8}$
(β_1)	$(b^2)^T e = 1$	(β_4)	$(b^2)^T C C e = \frac{1}{3}$
(β_7)	$(b^2)^T \mathcal{A}^2 C C e = \frac{1}{12}$	(γ_2)	$(b^3)^T C e = \frac{1}{2}$
		(γ_5)	$(b^3)^T C C C e = \frac{1}{4}$
		(γ_8)	$(b^3)^T \mathcal{A}^3 \mathcal{A}^3 C e = \frac{1}{24}$

One may check that we need at least nine stages to obtain a linearly implicit method which fulfils these order conditions. Restrictions (r_1) , (r_2) and (r_3) require the assumption that the first stage of the method (C, \mathcal{A}^1, b^1) is explicit; we must note that in this way the computational cost of the final method will be similar to the cost of an eight stage method with all stages implicit. To obtain a stable method, it is convenient to impose the restrictions $(0, \dots, 0, 1)^T \mathcal{A}^i = (b^i)^T$ for $i = 1, 2, 3$ (see [3]); in order to get the numerical solution directly from the calculation of the last stage $Y_h^{m+1}(= Y_h^{m,9})$ and, consequently, to reduce lightly the computational cost of the final method, we impose the same restriction on the third level $(0, \dots, 0, 1)^T \mathcal{A}^3 = (b^3)^T$.

The computational cost of the final method can be reduced (when the coefficients of the linear operator do not depend on time) by assuming that $a_{ii}^1 = a_{jj}^2 = a$ for $i = 3, 5, 7, 9$ and for $j = 2, 4, 6, 8$; besides, using these restrictions we simplify the study of the stability of the FSRK method. Thus, the table of the deduced method will have the following structure:

$$\begin{array}{c|c|c}
 \mathcal{C}e & \mathcal{A}^1 & \mathcal{A}^2 \\
 \hline
 & (b^1)^T & (b^2)^T \\
 \hline
 \begin{array}{l} 0 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \\ c_7 \\ c_8 \\ c_9 \end{array} & \begin{array}{cccccccc} 0 & & & & & & & & \\ a_{21}^1 & 0 & & & & & & & \\ a_{31}^1 & 0 & a & & & & & & \\ a_{41}^1 & 0 & a_{43}^1 & 0 & & & & & \\ a_{51}^1 & 0 & a_{53}^1 & 0 & a & & & & \\ a_{61}^1 & 0 & a_{63}^1 & 0 & a_{65}^1 & 0 & & & \\ a_{71}^1 & 0 & a_{73}^1 & 0 & a_{75}^1 & 0 & a & & \\ a_{81}^1 & 0 & a_{83}^1 & 0 & a_{85}^1 & 0 & a_{87}^1 & 0 & \\ b_1^1 & 0 & b_3^1 & 0 & b_5^1 & 0 & b_7^1 & 0 & a \end{array} & \begin{array}{cccccccccccc} 0 & & & & & & & & & & & & & & \\ 0 & a & & & & & & & & & & & & & \\ 0 & a_{32}^2 & 0 & & & & & & & & & & & & \\ 0 & a_{42}^2 & 0 & a & & & & & & & & & & & \\ 0 & a_{52}^2 & 0 & a_{54}^2 & 0 & & & & & & & & & & \\ 0 & a_{62}^2 & 0 & a_{64}^2 & 0 & a & & & & & & & & & \\ 0 & a_{72}^2 & 0 & a_{74}^2 & 0 & a_{74}^3 & 0 & & & & & & & & \\ 0 & a_{82}^2 & 0 & a_{84}^2 & 0 & a_{84}^3 & 0 & a & & & & & & & \\ 0 & b_2^2 & 0 & b_4^2 & 0 & b_6^2 & 0 & b_8^2 & 0 & & & & & & \\ \hline 0 & b_2^2 & 0 & b_4^2 & 0 & b_6^2 & 0 & b_8^2 & 0 & & & & & & \end{array}
 \end{array}$$

$$\begin{array}{c|c}
 \mathcal{A}^3 & \\
 \hline
 (b^3)^T & \begin{array}{ccccccccccc} 0 & & & & & & & & & & \\ a_{21}^3 & 0 & & & & & & & & & \\ a_{31}^3 & a_{32}^3 & 0 & & & & & & & & \\ a_{41}^3 & a_{42}^3 & a_{43}^3 & 0 & & & & & & & \\ a_{51}^3 & a_{52}^3 & a_{53}^3 & a_{54}^3 & 0 & & & & & & \\ a_{61}^3 & a_{62}^3 & a_{63}^3 & a_{64}^3 & a_{65}^3 & 0 & & & & & \\ a_{71}^3 & a_{72}^3 & a_{73}^3 & a_{74}^3 & a_{75}^3 & a_{76}^3 & 0 & & & & \\ a_{81}^3 & a_{82}^3 & a_{83}^3 & a_{84}^3 & a_{85}^3 & a_{86}^3 & a_{87}^3 & 0 & & & \\ b_1^3 & b_2^3 & b_3^3 & b_4^3 & b_5^3 & b_6^3 & b_7^3 & b_8^3 & 0 & & \\ \hline b_1^3 & b_2^3 & b_3^3 & b_4^3 & b_5^3 & b_6^3 & b_7^3 & b_8^3 & 0 & & \end{array}
 \end{array}$$

We begin the construction process for the method by resolving the following order conditions: $(r_1), (r_2), (\alpha_1), \dots, (\alpha_8), (\beta_1), \dots, (\beta_8), (\times_{12}), (\times_{21}), (\diamond_{21}), (\diamond_{12}), (\odot_{112}), (\odot_{121}), (\odot_{211}), (\odot_{221}), (\odot_{212})$ and (\odot_{122}) ; which involve the coefficients of the FSRK methods $(\mathcal{C}, \mathcal{A}^i, b^i)_{i=1}^2$ only. Thus we obtain a four order family of FSRK methods depending on the free parameters $a, c_4, c_5, c_6, c_7, a_{65}^1, a_{83}^1$ and a_{87}^1 .

To perform, in an efficient way, the numerical integration of the linear stiff problems, it is necessary to impose additional restrictions of type A-stability. In order to introduce this concept in the simplest way we apply a two-level FSRK method to the numerical integration of the following test problem

$$y'(t) = (\lambda_1 + \lambda_2) y(t) \text{ with } Re(\lambda_i) \leq 0, \text{ for } i = 1, 2;$$

obtaining the recurrence:

$$y_{m+1} = \left(1 + \sum_{i=1}^2 \tau \lambda_i (b^i)^T (I - \sum_{j=1}^2 \tau \lambda_j \mathcal{A}^j)^{-1} e\right) y_m.$$

By substituting in this expression z_i for $\tau \lambda_i$ for $i = 1, 2$ we obtain a rational function of two complex variables; which we call it the Amplification Function associated with the FSRK method

$$R(z_1, z_2) = 1 + \sum_{i=1}^2 z_i (b^i)^T \left(I - \sum_{j=1}^2 z_j \mathcal{A}^j\right)^{-1} e.$$

We say that an FSRK method is A-stable iff $|R(z_1, z_2)| \leq 1, \forall z_i \in \mathbb{C}$ being $Re(z_i) \leq 0, i = 1, 2$.

As the FSRK method has nine stages, with the first one of them explicit, and attains fourth order, the amplification function of the method can be written as follows:

$$R(z_1, z_2) = R_1(z_1)R_2(z_2) + Rest$$

where

$$R_i(z_i) = \frac{1 + (1 - 4a)z_i + \frac{1}{2}(1 - 68a + 12a^2)z_i^2 + (\frac{1}{6} - 2a + 6a^2 - 4a^3)z_i^3}{(1 - az_i)^4} + \frac{(\frac{1}{24} - \frac{2a}{3} + 3a^2 - 4a^3 + a^4)z_i^4}{(1 - az_i)^4}$$

are the amplification functions associated with the each standard RK method $(\mathcal{C}, \mathcal{A}^i, b^i)$ for $i = 1, 2$ and

$$Rest = \frac{e_{41}z_1^4z_2 + e_{32}z_1^3z_2^2 + e_{23}z_1^2z_2^3 + e_{14}z_1z_2^4 + e_{42}z_1^4z_2^2 + e_{33}z_1^3z_2^3}{(1 - az_1)^4(1 - az_2)^4} + \frac{e_{24}z_1^2z_2^4 + e_{43}z_1^4z_2^3 + e_{34}z_1^3z_2^4 + e_{44}z_1^4z_2^4}{(1 - az_1)^4(1 - az_2)^4},$$

where:

$$\begin{aligned} e_{41} &= E_{4,1} - \frac{1}{24}, \\ e_{14} &= E_{1,4} - \frac{1}{24}, \\ e_{23} &= E_{2,3} - \frac{1}{12}, \\ e_{32} &= E_{3,2} - \frac{1}{12}, \\ e_{42} &= E_{4,2} - 4ae_{32} - 4ae_{41} - \frac{1}{48}, \\ e_{24} &= E_{2,4} - 4ae_{23} - 4ae_{14} - \frac{1}{48}, \\ e_{33} &= E_{3,3} - 4ae_{23} - 4ae_{32} - \frac{1}{36}, \\ e_{43} &= E_{4,3} - 4ae_{42} - 10a^2e_{41} - 4ae_{33} - 16a^2e_{32} - 10a^2e_{23} - \frac{1}{144}, \\ e_{34} &= E_{3,4} - 4ae_{24} - 10a^2e_{14} - 4ae_{33} - 16a^2e_{23} - 10a^2e_{32} - \frac{1}{144}, \\ e_{44} &= E_{4,4} - 4ae_{43} - 4ae_{34} - 10a^2e_{42} - 20a^3e_{41} - 16a^2e_{33} - 40a^3e_{32} - 10a^2e_{24} \\ &\quad - 40a^3e_{23} - 20a^3e_{14} - \frac{1}{576}, \end{aligned}$$

where

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

with $n_l(\bar{k})$ the number of the times that the index l appears in the integer vector \bar{k} .

Given this FSRK method of Alternating Directions type attains fourth order, and the standard RK method $(\mathcal{C}, \mathcal{A}^1, b^1)$ may be reduced to a DIRK method of five stages with the first stage explicit, the standard RK method $(\mathcal{C}, \mathcal{A}^2, b^2)$ may be reduced to a four-stage SDIRK method, we can see in [6] that such methods $(\mathcal{C}, \mathcal{A}^i, b^i)$ for $i = 1, 2$ are L-stable for $a = 0.572816062482134746$. Taking into account that the strong stability and, particularly, the L-stability are very interesting properties when we carry out the numerical integration of the parabolic problems with irregular or incompatible data, we have fixed $a = 0.572816062482134746$.

Now, if we were able to make null the *Rest* we would obtain an L-stable FSRK method. We have checked that it is impossible to cancel the *Rest* because by substituting the previous values we observe that e_{44} has the constant value 0.011591. Nevertheless, we can determine the remaining free parameters in order to minimize to some extent the contribution of the *Rest* to the amplification function. Due to the complexity, size and nonlinearity of the last expressions, we have taken $a_{87}^1 = 0$ in order to shorten them; now, we can check that $e_{14} = e_{41}$, $e_{23} = e_{32}$, $e_{24} = e_{42}$ and $e_{43} = e_{34}$, simplifying the subsequent study. By cancelling $e_{43} = e_{34}$ and e_{33} we can fix the parameters a_{65}^1 and a_{83}^1 . By using this option, the remaining coefficients of the *Rest* have the following values: $e_{14} = 0.0165332$, $e_{23} = 0.0171413$ and $e_{24} = -0.0423441$.

With the imposed restrictions, have obtained a family of FSRK methods which depend on the parameters c_4, c_5, c_6, c_7 , attain fourth order and are A-stable.

In the last part of the construction of this method we solve the remaining conditions. In such order conditions, some coefficients of the last explicit level $(\mathcal{C}, \mathcal{A}^3, b^3)$ appear. Namely we must solve: $(\mathcal{C}, \mathcal{A}^3, b^3)$: $(r_3), (\gamma_1), \dots, (\gamma_8), (\times_{13}), (\times_{23}), (\times_{31}), (\times_{32}), (\diamond_{13}), (\diamond_{23}), (\diamond_{31}), (\diamond_{32}), (\star_{13}), (\star_{23}), (\star_{31}), (\star_{32}), (\odot_{113}), (\odot_{131}), (\odot_{311}), (\odot_{331}), (\odot_{313}), (\odot_{133}), (\odot_{223}), (\odot_{232}), (\odot_{322}), (\odot_{332}), (\odot_{323}), (\odot_{233}), (\odot_{123}), (\odot_{132}), (\odot_{213}), (\odot_{231}), (\odot_{321})$ and (\odot_{312}) . By resolving these order conditions we obtain a family of methods which depend on the parameters: $a_{63}^3, a_{73}^3, a_{74}^3, a_{76}^3, a_{82}^3, a_{83}^3, a_{84}^3, a_{85}^3, a_{86}^3, a_{87}^3, c_4, c_5, c_6$ and c_7 . Such coefficients are fixed in order to minimize to some extent the main term of the local error and also to simplify coefficients. Thus, the final method is given by:

$$\left(\begin{array}{c} \frac{\mathcal{A}^1}{(b^1)^T} \\ 0 \\ \left(\begin{array}{cccccccc} 0.5728160624821349 & 0 & & & & & & \\ 0.5728160624821349 & 0 & 0.5728160624821349 & & & & & \\ 0.23306240731962216 & 0 & 0.10027092601371115 & 0 & & & & \\ 0.14320401562053764 & 0 & -0.14320401562053764 & 0 & 0.5728160624821349 & & & \\ 0.2475254450891478 & 0 & -0.1146139158037628 & 0 & 0.5337551373812817 & 0 & & \\ 0.16584521264111465 & 0 & 0.16898133629871276 & 0 & -0.574309278088629 & 0 & 0.5728160624821349 & \\ -0.0519764686433798 & 0 & -0.0051822334268295 & 0 & 0.48434263958807455 & 0 & 0 & \\ 0.06578033900854208 & 0 & -0.2245756812245128 & 0 & -0.0453534758107523 & 0 & 0.6313327555445881 & 0 & 0.5728160624821349 \\ 0.0657803390085421 & 0 & -0.2245756812245128 & 0 & -0.0453534758107523 & 0 & 0.6313327555445881 & 0 & 0.5728160624821349 \end{array} \right) \end{array} \right)$$

$$\left(\frac{-A^2}{(b^2)^T} \right) = \begin{pmatrix} 0 \\ 0 & 0.5728160624821349 \\ 0 & 1.1456321249642698 & 0 \\ 0 & -0.2394827291488016 & 0 & 0.5728160624821349 \\ 0 & -0.1122401054590243 & 0 & 0.6850561679411592 & 0 \\ 0 & -0.5036376505875728 & 0 & 0.5974882547721045 & 0 & 0.5728160624821349 \\ 0 & 0.6845495903831673 & 0 & 0.3072638220447982 & 0 & -0.6584800790946324 & 0 \\ 0 & -1.1893898642709955 & 0 & 0.7073896907436742 & 0 & 0.3363680485630513 & 0 & 0.5728160624821349 \\ 0 & -1.2359082397450338 & 0 & 1.7359082397450343 & 0 & 1.7359082397450338 & 0 & -1.2359082397450343 & 0 \\ 0 & -1.2359082397450338 & 0 & 1.7359082397450343 & 0 & 1.7359082397450338 & 0 & -1.2359082397450343 & 0 \end{pmatrix}$$

$$\left(\frac{-A^3}{(b^3)^T} \right) = \begin{pmatrix} 0 \\ 0.5728160624821349 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.145632124964014 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.36458308229640946 & -0.02059633495004886 & -0.0106534140130273 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1.5973005621929421 & -2.335377030662664 & 1.3108925309518593 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.4571508517861478 & 0.020596334950028255 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.236346605491203 & 0.23975256771974143 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.4271839375178651 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1.2359082397450865 & 0 & 0 & 1.7359082397450232 & 0 & 0 & 0 & 0 \\ 0 & -1.2359082397450865 & 0 & 0 & 1.7359082397450232 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.18891947993049055 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.14276583987761113 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.735908239745029 & 0 & -1.2359082397450358 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.735908239745029 & 0 & -1.2359082397450358 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$C = (0, 0.5728160624821349, 1.1456321249642698, \frac{1}{3}, 0.5728160624821349, \frac{2}{3}, \frac{1}{3}, 0.4271839375178651)^T.$$

4. NUMERICAL EXPERIMENT

In this section we show a numerical test where we have integrated the following non-linear convection-diffusion problem with a non-linear reaction term $r(y) = k_1 y + k_2 y + \frac{y^3}{(1+y^2)^2}$

$$\begin{cases} \frac{\partial y}{\partial t} = \Delta y - v_1 y_{x_1} - v_2 y_{x_2} - r(y) + f, & \forall (x_1, x_2, t) \in \Omega \times [0, 5], \\ y(x_1, 0, t) = y(x_1, 1, t) = 0, & \forall x_1 \in [0, 1] \text{ and } \forall t \in [0, 5], \\ y(0, x_2, t) = y(1, x_2, t) = 0, & \forall x_2 \in [0, 1] \text{ and } \forall t \in [0, 5], \\ y(x_1, x_2, 0) = x_1^3(1-x_1)^3 x_2^3(1-x_2)^3, & \forall (x_1, x_2) \in \bar{\Omega}, \end{cases}$$

with $\Omega = (0, 1) \times (0, 1)$, $v_1 = 1 + x_1 x_2 e^{-t}$, $v_2 = 1 + x_1 t$, $k_1 = k_2 = \frac{1+(x_1+x_2)^2 e^{-t}}{2}$ and $f = 10^3 e^{-t} x_1(1-x_1)x_2(1-x_2)$.

We have carried out the time discretization using the method described in the previous section and we have discretized in space on a rectangular mesh by using a central difference scheme. The combination of methods of finite differences type of order q in the spatial discretization stage and p th-order time integrators of the type described in this paper provides a totally discrete scheme whose global error is $[y(t_m)]_h - Y_h^m = \mathcal{O}(h^q + \tau^p)$; here $[v]_h$ denotes the restriction of the function v to the mesh node.

In order to apply the integration method described here, we have taken for $i = 1, 2$ and $j = 1, \dots, 9$

$$L_{ih}(t_{m,j})Y_h^{m,j} = \delta_{\bar{x}_i x_i} Y_h^{m,j} - [v_i]_h \delta_{\hat{x}_i} Y_h^{m,j} - [k_i]_h Y_h^{m,j} - \frac{3(Y_h^m)^2 - (Y_h^m)^4}{2(1 + (Y_h^m)^2)^3} Y_h^{m,j},$$

$$g_h(t_{m,j}, Y_h^{m,j}) = \frac{3(Y_h^m)^2 - (Y_h^m)^4}{(1 + (Y_h^m)^2)^3} Y_h^{m,j} - \frac{(Y_h^{m,j})^3}{(1 + (Y_h^{m,j})^2)^2}$$

where $\delta_{\bar{x}_i x_i}$ and $\delta_{\hat{x}_i}$ denote the classical central differences.

Note that, by using this decomposition in each stage of (4) we must solve only one linear system whose matrix is tridiagonal; thus the computational complexity of computing the internal stages of this method is of the same order as in the cases of an explicit method. Note that these operators $L_{ih}(t)$ are not commutative: nevertheless, the numerical results obtained are correct.

The numerical maximum global errors have been estimated as follows

$$E_{N,\tau} = \max_{x_{1i}, x_{2i}, t_m} |Y^{N,\tau}(x_{1i}, x_{2i}, t_m) - Y^*|$$

where $Y^{N,\tau}(x_{1i}, x_{2i}, t_m)$ are the numerical solutions obtained at the mesh point (x_{1i}, x_{2i}) and the time point $t_m = m\tau$, on a rectangular mesh with $(N+1) \times (N+1)$ nodes and with time step τ , while Y^* is the numerical solution calculated at the same mesh points and time steps, but by using a spatial mesh with $(2N+1) \times (2N+1)$ nodes, halving the mesh size, and with time step $\frac{\tau}{2}$. We compute the numerical orders of convergence as $\log_2 \frac{E_{N,\tau}}{E_{2N,\tau/2}}$.

Table 1. Numerical Errors ($E_{N,\tau}$)

$N = 8$	$N = 16$	$N = 32$	$N = 64$	$N = 128$	$N = 256$	$N = 512$
2.2183E-2	5.6902E-3	1.4231E-3	3.5569E-4	8.8921E-5	2.2289E-5	5.6054E-6

In Table 1, we show the numerical errors and, in Table 2, their corresponding numerical orders of convergence. In order to guarantee that the contribution to the error of the spatial and temporal part are of the same size, we have assumed the relation $\sqrt{N}\tau \equiv C = 0.1$ between the time step τ and the mesh size $\frac{1}{N}$.

Table 2. Numerical orders of convergence

$N = 8$	$N = 16$	$N = 32$	$N = 64$	$N = 128$	$N = 256$
1.9629	1.9995	2.0003	2.0000	1.9962	1.9915

REFERENCES

- [1] B. Bujanda, J.C. Jorge, *Stability results for linearly implicit Fractional Step discretizations of non-linear time dependent parabolic problems*, to appear in Appl. Num. Math.
- [2] B. Bujanda, J.C. Jorge, *Order conditions for linearly implicit Fractional Step Runge–Kutta methods (Preprint)*, Universidad Pública de Navarra.
- [3] B. Bujanda, J.C. Jorge, *Efficient linearly implicit methods for nonlinear multidimensional parabolic problems*, J. Comput. Appl. Math. **164/165** (2004), 159–174.
- [4] B. Bujanda, J.C. Jorge, *Stability results for fractional step discretizations of time dependent coefficient evolutionary problems*, Appl. Numer. Math. **38**, (2001), 69–86.
- [5] G.J. Cooper, A. Sayfy, *Additive methods for the numerical solution of ordinary differential equations*, Math. of Comp. **35**, (1980), 1159–1172.
- [6] E. Hairer, G. Wanner, *Solving ordinary differential equations II*, Springer-Verlag, 1987.
- [7] W. Hundsdorfer, J.G. Verwer, *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*, Springer, 2003.
- [8] N.N. Yanenko, *The method of fractional steps*, Springer, 1971.

Blanca Bujanda
blanca.bujanda@unavarra.es

Universidad Pública de Navarra
Dpto. Matemática e Informática

Juan Carlos Jorge
jcjorge@unavarra.es

Universidad Pública de Navarra
Dpto. Matemática e Informática

Received: October 3, 2005.