MATHEMATICS OF COMPUTATION
Volume 69, Number 230, Pages 609–628
S 0025-5718(99)01156-4
Article electronically published on May 20, 1999

STABILITY OF RUNGE–KUTTA METHODS FOR QUASILINEAR PARABOLIC PROBLEMS

C. GONZÁLEZ AND C. PALENCIA

ABSTRACT. We consider a quasilinear parabolic problem

$$u'(t) = Q\big(u(t)\big)u(t), \qquad u(t_0) = u_0 \in \mathcal{D},$$

where $Q(w) : \mathcal{D} \subset X \to X$, $w \in W \subset X$, is a family of sectorial operators in a Banach space X with fixed domain $\mathcal{D}$. This problem is discretized in time by means of a strongly $\mathrm{A}(\theta)$-stable, $0 < \theta \leq \pi/2$, Runge–Kutta method. We prove that the resulting discretization is stable, under some natural assumptions on the dependence of $Q(w)$ with respect to w. Our results are useful for studying in L^p norms, $1 \leq p \leq +\infty$, many problems arising in applications. Some auxiliary results for time-dependent parabolic problems are also provided.

1. INTRODUCTION

The present paper is devoted to the study of the stability and convergence of semidiscretizations in time, based on Runge–Kutta methods, of quasilinear parabolic problems. We believe that the issues discussed in this paper are a decisive step towards a rigorous analysis of the fully discrete (i.e. discrete in space and time) methods.

Quasilinear parabolic problems arise in the study of diffusion phenomena with state-dependent diffusivity. They also appear in the equations of fluids in porous media, the study of polymers, models for cartilages (see e.g. [5, 8, 10, 11] and the references therein), etc.

The continuous problem is considered in the abstract setting of Banach spaces. This point of view is not only general but also very convenient, since the proofs rely on two clear and simple abstract hypotheses. Let $(X, \|\cdot\|)$ be a complex Banach space, let $W \subset X$, and let $\mathcal{D} \subset X$ be a dense subspace of X. For each $w \in W$, let $Q(w) : \mathcal{D} \subset X \to X$ be a linear operator. We are interested in the semidiscretizations in time of the abstract quasilinear problem

$$\left\{ \begin{array}{lcll} u'(t) & = & Q\big(u(t)\big)u(t), & \quad t > t_0, \\ u(t_0) & = & u_0 \in \mathcal{D}. & \end{array} \right. \tag{1}$$

Thus, our attention is restricted to the simplest case for which the domains of the operators $Q(w)$ are independent of $w \in W$. The practical meaning of this limitation (see Section 6) is that only boundary conditions of Dirichlet type can be included within the scope of our study. However, by using the theory of extrapolation spaces

Received by the editor March 12, 1997 and, in revised form, February 23, 1998 and June 9, 1998.

1991 *Mathematics Subject Classification.* Primary 65M12, 65M15, 65M20, 65L06, 65J10, 65J15.

(see [6]), it seems possible to reduce the study of quasilinear problems with variable domains to the case of fixed domains. The investigation of this possibility will be carried out elsewhere.

Our hypotheses concerning problem (1) are standard. Roughly speaking, we will assume that (i) the "frozen" linear evolution problems $u'(t) = Q(w)u(t)$, where $w \in W$ is fixed, are parabolic, and (ii) the dependence of $Q(w)$ on w is Lipschitz in a suitable sense. Before we state our hypotheses, we introduce some notation. For each angle $\theta \in (0, \pi/2)$, we set

$$S_\theta := \{0\} \cup \{\, z \in \mathbf{C} \,:\, z \neq 0,\ |\arg(-z)| \leq \theta \,\}.$$

The Banach space X is also denoted by X_0. The domain $\mathcal{D}$ is assumed to be endowed with a norm $\|\cdot\|_1$ such that $X_1 := (\mathcal{D}, \|\cdot\|_1)$ is a Banach space and such that X_1 is continuously embedded in X_0. With no loss of generality, we will assume that $\|x\| \leq \|x\|_1$ for $x \in X_1$. For $0 < \eta < 1$, X_η stands for the Calderón interpolation space $[X_0, X_1]_\eta$ (see e.g. [9, 25]). The norm in X_η is denoted by $\|\cdot\|_\eta$. Notice that, for $0 \leq \eta \leq \nu \leq 1$, we have $\|x\|_\eta \leq \|x\|_\nu$ for $x \in X_\nu$. Let us point out that the interpolation spaces obtained by means of the real interpolation method (see e.g. [25]) could be used instead. We wish to emphasize that only basic facts about interpolation theory will be required.

Our hypotheses **H**1-2 concerning (1) are as follows.

H1. There exist $M \geq 1$, $\omega_0 \in \mathbf{R}$ and $\theta \in (0, \pi/2)$ such that, for any complex $z \notin \omega_0 + S_\theta$ and for any $w \in W$, the resolvent $\big(zI - Q(w)\big)^{-1} : X \to X$ exists and satisfies

$$\big\|\big(z - Q(w)\big)^{-1}\big\| \leq \frac{M}{|z - \omega_0|}.$$

As it is well known (see e.g.[22]), this assumption implies that for fixed $w \in W$ the semigroup $e^{tQ(w)}$, $t \geq 0$, is analytic. It is this condition that renders the abstract problem (1) parabolic.

H2. There exists $\mu \in [0, 1)$ such that W is an open subset of X_μ, and there exists $L > 0$ such that the bound

$$\|Q(w)x - Q(v)x\| \leq L\|x\|_1\, \|w - v\|_\mu$$

holds for arbitrary v, $w \in W$ and $x \in X_1$.

Fix $w \in W$. Since $Q(w) : \mathcal{D} \subset X \to X$ is a closed operator, the space $\mathcal{D}$ endowed with the graph norm $\|x\|_{\mathcal{D}} = \|x\| + \|Q(w)x\|$, $x \in \mathcal{D}$, is a Banach space. Therefore, the norm $\|\cdot\|_1$ in X_1 is equivalent to such a graph norm. It turns out that X_1, and hence the intermediate Calderón spaces X_η, $0 < \eta < 1$, can be built up from any of the operators $Q(w)$, $w \in W$. However, due to the nonapplicability of Heinz's theorem, the domains of the fractional powers $D\big((\omega_0 I - Q(w))^\eta\big)$ may depend on the point $w \in W$. This is the reason why we use the Calderón spaces.

It is known that **H**1-2 guarantee the existence and uniqueness of the solution of the initial value problem (1) (see e.g. [3, 4, 23]). In [4] it is proved that (1) defines a semiflow in $W \bigcap X_\nu$, for $\mu < \nu < 1$. Notice that the limiting value $\nu = \mu$ is not covered by this result.

Problem (1) is discretized in time by means of a Runge–Kutta method, defined by its Butcher array

$$\left(\begin{array}{c|c} \mathbf{c} & \mathcal{A} \\ \hline & \mathbf{b}^T \end{array}\right), \tag{2}$$

where $\mathbf{b} = [b_1, \dots, b_s]^T \in \mathbf{R}^s$, $\mathbf{c} = [c_1, \dots, c_s]^T \in \mathbf{R}^s$ and $\mathcal{A} = [a_{ij}]_{i,j=1}^s \in \mathbf{R}^{s\times s}$. We assume that $0 \le c_i \le 1$ for $1 \le i \le s$, and that $\mathcal{A}$ is invertible. Let us recall that the stability function of the method is the rational function $r(z) = 1 + \mathbf{b}^T(\mathcal{I} - z\mathcal{A})^{-1}\mathbf{e}$, where $\mathbf{e} = [1, \dots, 1]^T \in \mathbf{R}^s$ and $\mathcal{I}$ is the identity matrix in $\mathbf{R}^{s\times s}$. We define the method to be $A(\theta)$-stable, $0 < \theta \le \pi/2$, when (i) the spectrum of the matrix $\mathcal{A}$ is contained in the complement in $\mathbf{C}$ of the sector S_θ, and (ii) $|r(z)| \le 1$, for $z \in S_\theta$. If in addition $|r(\infty)| < 1$, then the method is called strongly A(θ)-stable. In the sequel only strongly $A(\theta)$-stable methods will be considered. Though this excludes, among others, the Gaussian methods, there is a wide range of methods lying in the class of strongly A(θ)-stable methods (see e.g. [15]).

Let t_n, $0 \le n \le N$, be a finite sequence in $\mathbf{R}$, with constant step-size $h > 0$. The method, applied to problem (1), starts with the value u_0 and produces the numerical approximations u_n to the values $u(t_n)$, $1 \le n \le N$, by means of the recurrence

$$u_{n+1} = u_n + h\sum_{i=1}^{s} b_i Q\big(U_n^i\big)U_n^i, \qquad 0 \le n \le N-1, \tag{3}$$

where, for $0 \le n \le N-1$, the intermediate stages $U_n^i \in W$ are implicitly defined by the system

$$U_n^i = u_n + h\sum_{j=1}^{s} a_{ij} Q\big(U_n^j\big)U_n^j, \qquad i = 1, \dots, s. \tag{4}$$

For $R > 0$, $0 \le \eta \le 1$ and $v \in X_\eta$, $D_\eta(v, R)$ stands for the closed ball in X_η centered at v and with radius R. In Lemma 2.2 it is proved that, assuming that $u_0 \in W \bigcap X_\nu$ for some $\nu \in (\mu, 1)$, there exist R and h_0 such that, for $0 < h \le h_0$ and $v \in D_\nu(u_0, 2R/3)$, the stage system

$$V^i = v + h\sum_{j=1}^{s} a_{ij} Q\big(V^j\big)V^j, \qquad i = 1, \dots, s, \tag{5}$$

possesses a unique solution $\sigma_h(v) := [V^1, V^2, \dots, V^s] \in D_\mu(v, R/3)^s$. Thus, for $0 < h \le h_0$, we can construct a non-linear operator $\mathcal{N}_h : D_\nu(u_0, 2R/3) \to X_\nu$ in such a way that (3) is equivalent to

$$u_{n+1} = \mathcal{N}_h(u_n), \qquad 0 \le n \le N-1, \tag{6}$$

provided that $u_n \in D_\nu(u_0, 2R/3)$. Two important remarks are in order. First, notice that while v in (5) lies in the finer space X_ν, the vector of stages $\sigma_h(v)$ lies in X_μ^s. Second, it is possible to see that

$$\|\mathcal{N}_h(v) - v\|_\nu = \mathcal{O}(h^{\nu-\mu}) \tag{7}$$

and that this estimate is sharp. Thus, if we start at u_0 and assume that we progress with (6) up to u_n, then (7) yields

$$\|u_n - u_0\|_\nu = \mathcal{O}(nh^{\nu-\mu}) = \mathcal{O}((t_n - t_0)/h^{1-\nu+\mu}).$$

However, this does not guarantee that the terms in (6) remain in $D_\nu(u_0, 2R/3)$, when $h \to 0+$. This means that the applicability of the method for $Nh = T$, $T > 0$ fixed, must be established by another approach. The possibility of this solvability for $T > 0$ fixed, even when source terms are incorporated into (6), is provided by the next theorem, which is the main contribution of the present paper. The continuous dependence the initial data and the source terms is also considered. The inclusion of such source terms is required in studying the convergence of the method.

Theorem 1.1. *Assume that the quasilinear parabolic problem* (1) *satisfies conditions* ***H****1-2, for some* $M \geq 0$, $\theta \in (0, \pi/2)$, $L > 0$ *and* $\mu \in [0, 1)$. *Assume also that the Runge–Kutta method given by* (3) *and* (4) *is strongly* $A(\theta)$*-stable. Let* $u_0 \in W \bigcap X_\nu$, *for some* $\nu \in (\mu, 1)$, *and fix* $S > 0$. *Then there exist* $R > 0$, $h_0 > 0$, $T > 0$ *and* $C > 0$ *such that for* $0 < h \leq h_0$ *and for* $N = [T/h]$ *the following results hold.*

(i) For each $v \in D_\nu(u_0, R/3)$ *and each finite sequence* τ_n, $1 \leq n \leq N$, *in* $D_\nu(0, S)$ *the recurrence*

$$v_{n+1} = \mathcal{N}_h(v_n) + h\tau_{n+1}, \qquad 0 \leq n \leq N-1; \quad v_0 = v,$$

is solvable in $D_\nu(u_0, R)$.

(ii) Let v, $v^* \in D_\nu(u_0, R/3)$ *and let* τ_n, τ_n^*, $1 \leq n \leq N$, *be two finite sequences in* $D_\nu(0, S)$. *Consider the two sequences* v_n, v_n^*, $0 \leq n \leq N$, *in* $D_\nu(u_0, R)$ *defined by* $v_0 = v$, $v_0^* = v^*$ *and*

$$v_{n+1} = \mathcal{N}_h(v_n) + h\tau_{n+1}, \qquad v_{n+1}^* = \mathcal{N}_h(v_n^*) + h\tau_{n+1}^*, \qquad 0 \leq n \leq N-1.$$

Then

$$\|v_n^* - v_n\|_\nu \leq C\big(\|v^* - v\|_\nu + h\sum_{j=1}^{n} \|\tau_j^* - \tau_j\|_\nu\big).$$

Notice that, as in the result in [4] mentioned above for the continuous problem, the limiting value $\nu = \mu$ is not covered by Theorem 1.1.

We only know two references [7, 16] for the semidiscretization in time of abstract quasilinear parabolic problems. The functional setting in [16], given by the Lions–Gelfand triplets in Hilbert spaces, is more restrictive than ours. Notice that in a Hilbert space setting we can use Fourier transforms and energy methods, tools that are not valid in our general framework. In [16], only convergence is considered, not stability.

As in the results in [4] for the continuous problem, the proof of Theorem 1.1 relies on a fix-point argument. However, the parallelism with the continuous problem is broken by the fact that the stage vectors in (5) are in a space different from that containing the nodal values of the numerical solution. This forces a tricky choice of the base space for the fix-point argument. The necessary estimates are consequences of some bounds for linear time-dependent parabolic problems. In addition to the stability results in [14] for these problems, we need some estimates that are provided in the final Section 7.

Theorem 1.1 is of local nature, as expected in general for a nonlinear problem. Nevertheless, it is possible to trace the constants in Theorem 1.1. An important remark is that they do not depend on the Banach space X, but only on the Runge–Kutta method, on the constants in **H**1-2, and on the distance from u_0 to the boundary of W. It turns out that the constant C in Theorem 1.1.(ii) is of the form

$$C = \mathcal{O}(T^m \exp\big((\omega_0 + C^* L^{1/\alpha})T\big)),$$

where $m \geq 1$, $0 < \alpha < 1$ and $C^* > 0$ are independent of T. In view of this estimate we envisage the existence of some global result for $T = +\infty$, at least for asymptotically stable cases, when $\omega_0 < 0$ and L is small enough. This interesting issue will be considered elsewhere.

Section 2 is devoted to two auxiliary lemmas. The proof of the main result, Theorem 1.1, is postponed to Section 3. Convergence is considered in Section 4. In Section 5 we give the extension to time-dependent problems and to problems with semilinear terms. In Section 6 we show the applicability of our results to some basic examples.

We will use the following notation. The product spaces X_η^m, m an integer ≥ 1 and $0 \leq \eta \leq 1$, are endowed with the maximum norm component-wise. The norm in X_η^m is also denoted by $\|\cdot\|_\eta$. Given l, m integers ≥ 1 and $\delta, \eta \in [0,1]$, the operator norm correspoding to a bounded operator $F : \big(X_\delta^l, \|\cdot\|_\delta\big) \to \big(X_\eta^m, \|\cdot\|_\eta\big)$ is denoted by $\|F\|_{\delta\to\eta}$. Any matrix $\mathcal{M} \in \mathbf{C}^{l\times m}$ is identified with the operator $\mathcal{M} \otimes I : X_\eta^m \to X_\eta^l$, for $0 \leq \eta \leq 1$. The letter a stands for an upper bound on the norms of the operators defined by the matrices $\mathbf{b}^T$, $\mathbf{c}$, $\mathcal{A}$ and $\mathcal{A}^{-1}$ with respect to any pair of fractional spaces. Finally, given a family F_n, $1 \leq n \leq N$, of linear operators acting in some common space, we set

$$\prod_{j=m+1}^{n} F_j = F_n \cdot F_{n-1} \cdots F_{m+1}$$

for $0 \leq m < n \leq N$, and $\prod_{j=m+1}^{n} F_j = I$ for $0 \leq n \leq m \leq N$.

2. Some lemmas

In this section we study the local solvability of the stage system (5) and give some basic estimates.

For $V = [V^1, \dots, V^s]^T \in W^s$, we take $B(V) : X_1^s \subset X^s \to X^s$ to be the operator defined by $\mathrm{diag}(Q(V^1), \dots, Q(V^s))$.

Lemma 2.1. *Let $u_0 \in W \subset X_\mu$. There exist $R > 0$, $h_1 > 0$ and $K_1 > 0$ such that $D_\mu(u_0, R) \subset W$ and, for $0 < h \leq h_1$ and $V \in D_\mu(u_0, R)^s \subset W^s$, the operator $\mathcal{I} - h\mathcal{A}B(V) : X_1^s \subset X^s \to X^s$ is boundedly invertible and for $0 \leq \eta \leq \delta \leq 1$*

$$\big\| \big(\mathcal{I} - h\mathcal{A}B(V)\big)^{-1} \big\|_{\eta\to\delta} \leq K_1 h^{\eta-\delta}, \tag{8}$$

$$\big\| \big(\mathcal{I} - h\mathcal{A}B(V)\big)^{-1} - \mathcal{I} \big\|_{\delta\to\eta} \leq K_1 h^{\delta-\eta}, \tag{9}$$

and

$$\big\| \big(\mathcal{I} - h\mathcal{A}B(V_2)\big)^{-1} - \big(\mathcal{I} - h\mathcal{A}B(V_1)\big)^{-1} \big\|_{\delta\to\eta} \leq K_1 h^{\delta-\eta} \|V_2 - V_1\|_\mu. \tag{10}$$

Proof. Let $B_0 : X_1^s \subset X^s \to X^s$ be the operator given by

$$B_0 = \mathrm{diag}\big(Q(u_0), \dots, Q(u_0)\big).$$

As in the proof of Theorem 4.1 of [13], due to the fact that the diagonal elements of B_0 are all equal to $Q(u_0)$, we can see that there exist $h_1 > 0$ and $K > 0$ such that, for $0 < h \leq h_1$, the operator $\mathcal{I} - h\mathcal{A}B_0 : X_1^s \subset X^s \to X^s$ is boundedly invertible and

$$\big\| \big(\mathcal{I} - h\mathcal{A}B_0\big)^{-1} \big\|_{\eta\to\delta} \leq K h^{\eta-\delta}, \qquad 0 \leq \eta \leq \delta \leq 1. \tag{11}$$

Select $R_1 > 0$ with $D_\mu(u_0, R_1) \subset W$. For $V \in D_\mu(u_0, R_1)^s$ and $0 < h \leq h_1$, we define a new operator $\Delta(V, h) : X_1^s \subset X^s \to X^s$ by

$$\Delta(V, h) := h\mathcal{A}\big(B(V) - B_0\big)\big(\mathcal{I} - h\mathcal{A}B_0\big)^{-1}.$$

Take $R = \min\{R_1, 1/(2aLK)\}$ and fix $V \in D_\mu(u_0, R)^s$. For $0 \le \eta \le 1$, from (11) and **H**2, we get

$$\begin{aligned} \|\Delta(V,h)\|_{\eta\to 0} &\le h\|\mathcal{A}\|_{0\to 0}\big\| B(V) - B_0 \big\|_{1\to 0}\big\| \big(\mathcal{I} - h\mathcal{A}B_0\big)^{-1} \big\|_{\eta\to 1} \\ &\le aL\|V - \mathbf{e}\otimes u_0\|_\mu K h^\eta \\ &\le h^\eta/2. \end{aligned} \tag{12}$$

By the estimate (12), with $\eta = 0$, we have

$$\|\Delta(V,h)\|_{0\to 0} \le 1/2. \tag{13}$$

Thus, noticing that

$$\mathcal{I} - h\mathcal{A}B(V) = \mathcal{I} - h\mathcal{A}B_0 - h\mathcal{A}\big(B(V) - B_0\big) = \big(\mathcal{I} - \Delta(V,h)\big)\big(\mathcal{I} - h\mathcal{A}B_0\big),$$

we can establish the existence of the inverse of $\mathcal{I} - h\mathcal{A}B(V)$, for $V \in D_\mu(u_0, R)^s$ by means of the Neumann series, which yields

$$\big(\mathcal{I} - h\mathcal{A}B(V)\big)^{-1} = \big(\mathcal{I} - h\mathcal{A}B_0\big)^{-1}\sum_{k=0}^{\infty}\Delta^k(V,h). \tag{14}$$

Fix $0 \le \eta \le \delta \le 1$. By taking norms in (14) and using (11), (12) and (13), we deduce that

$$\begin{aligned} \big\| \big(\mathcal{I} - h\mathcal{A}B(V)\big)^{-1} \big\|_{\eta\to\delta} &\le \big\| \big(\mathcal{I} - h\mathcal{A}B_0\big)^{-1} \big\|_{\eta\to\delta} + \big\| \big(\mathcal{I} - h\mathcal{A}B_0\big)^{-1} \big\|_{0\to\delta} \\ &\quad \times \left(\sum_{k=1}^{\infty}\|\Delta^{k-1}(V,h)\|_{0\to 0}\right)\|\Delta(V,h)\|_{\eta\to 0} \\ &\le 2Kh^{\eta-\delta}. \end{aligned}$$

This proves (8).

Next we prove (9). As we mentioned in the introduction, because of **H**2, the norm $\|\cdot\|_1$ is equivalent to the graph norm of any operator $Q(w)$, $w \in W$. In fact, it is clear that there exists $b > 0$ such that

$$\|B(V)\|_{1\to 0} \le b, \qquad V \in D_\mu(u_0, R)^s. \tag{15}$$

Fix $0 < h \le h_0$, $V \in D_\mu(u_0, R)^s$ and $0 \le \eta \le \delta \le 1$. By (8),

$$\big\| \big(\mathcal{I} - h\mathcal{A}B(V)\big)^{-1} - \mathcal{I} \big\|_{\delta\to\delta} \le 1 + \big\| \big(\mathcal{I} - h\mathcal{A}B(V)\big)^{-1} \big\|_{\delta\to\delta} \le 1 + K. \tag{16}$$

Moreover, the identity

$$\big(\mathcal{I} - h\mathcal{A}B(V)\big)^{-1} - \mathcal{I} = h\mathcal{A}B(V)\big(\mathcal{I} - h\mathcal{A}B(V)\big)^{-1}$$

holds. Hence, using (8) and (15) again, we conclude that

$$\begin{aligned} \big\| \big(\mathcal{I} - h\mathcal{A}B(V)\big)^{-1} - \mathcal{I} \big\|_{\delta\to 0} &\le h\|\mathcal{A}\|_{0\to 0}\|B(V)\|_{1\to 0}\big\| \big(\mathcal{I} - h\mathcal{A}B(V)\big)^{-1} \big\|_{\delta\to 1} \\ &\le abKh^\delta. \end{aligned}$$

Now (9) is obtained by interpolating between the last estimate and (16).

It remains to prove (10). Given V_1, $V_2 \in D_\mu(u_0, R)^s$, we have

$$\begin{aligned} &\big(\mathcal{I} - h\mathcal{A}B(V_2)\big)^{-1} - \big(\mathcal{I} - h\mathcal{A}B(V_1)\big)^{-1} \\ &\qquad = \big(\mathcal{I} - h\mathcal{A}B(V_2)\big)^{-1}h\mathcal{A}\big(B(V_2) - B(V_1)\big)\big(\mathcal{I} - h\mathcal{A}B(V_1)\big)^{-1}. \end{aligned}$$

Hence, by (8) and **H**2,

$$\begin{aligned}\big\| \big(\mathcal{I} - h\mathcal{A}B(V_2)\big)^{-1} - \big(\mathcal{I} - h\mathcal{A}B(V_1)\big)^{-1} \big\|_{\delta\to\eta} \\ \leq \quad \big\| \big(\mathcal{I} - h\mathcal{A}B(V_2)\big)^{-1} \big\|_{0\to\eta} h\|\mathcal{A}\|_{0\to0} \\ \times \|B(V_2) - B(V_1)\|_{1\to0} \big\| \big(\mathcal{I} - h\mathcal{A}B(V_1)\big)^{-1} \big\|_{\delta\to1} \\ \leq \quad aK^2 L h^{\delta-\eta} \|V_2 - V_1\|_\mu. \qquad \square\end{aligned}$$

Lemma 2.2. *Let $\mu < \lambda \leq 1$ and $u_0 \in X_\lambda \bigcap W$. There exist $h_0 > 0$, $R > 0$ and $K > 0$ such that $D_\mu(u_0, R) \subset W$ and such that, for $0 < h \leq h_0$ and $v \in D_\lambda(u_0, 2R/3)$, the stage system*

$$V = \mathbf{e} \otimes v + h\mathcal{A}B(V)V \tag{17}$$

possesses a unique solution, denoted by $\sigma_h(v) := V$, in $D_\mu(v, R/3)^s \bigcap X_1^s$. Moreover,

$$\|\sigma_h(v^*) - \sigma_h(v)\|_\mu \leq K\|v^* - v\|_\lambda, \qquad v,\, v^* \in D_\lambda(u_0, 2R/3). \tag{18}$$

Proof. Let $R > 0$, $K_1 > 0$ and $h_1 > 0$ be the radius, constant and threshold provided by Lemma 2.1 for u_0. Fix $v \in D_\lambda(u_0, 2R/3)$ and $0 < h \leq h_1$. According to Lemma 2.1, for $V \in D_\mu(v, R/3)^s \subset D_\mu(u_0, R)^s$ the inverse $\big(\mathcal{I} - h\mathcal{A}B(V)\big)^{-1}$ exists. Therefore, (17) is equivalent to the system

$$V = \big(\mathcal{I} - h\mathcal{A}B(V)\big)^{-1}(\mathbf{e} \otimes v), \qquad V \in D_\mu(v, R/3)^s.$$

Let $F_{h,v} : D_\mu(v, R/3)^s \to X_\mu^s$ be the operator defined by

$$F_{h,v}(V) = \big(\mathcal{I} - h\mathcal{A}B(V)\big)^{-1}(\mathbf{e} \otimes v), \qquad V \in X_1^s.$$

Because of the above considerations, it turns out that the solutions $V \in D_\mu(v, R/3)^s$ of system (17) are the fixed points of $F_{h,v}$ and conversely. Next we are going to see that, possibly after a reduction of h_1, $F_{h,v}$ maps $D_\mu(v, R/3)^s$ to $D_\mu(v, R/3)^s$ and that $F_{h,v}$ is a contraction on $D_\mu(v, R/3)^s$, with respect to the metric induced by X_μ^s. Then, by the Banach fixed-point theorem, we will conclude that (17) possesses a unique solution in $D_\mu(v, R/3)^s$.

For $V \in D_\mu(v, R/3)^s$ we have

$$F_{h,v}(V) - \mathbf{e} \otimes v = \Big(\big(\mathcal{I} - h\mathcal{A}B(V)\big)^{-1} - \mathcal{I} \Big)(\mathbf{e} \otimes v),$$

so that, by (9), we get

$$\begin{aligned}\|F_{h,v}(V) - \mathbf{e} \otimes v\|_\mu \quad &\leq \quad \big\| \big(\mathcal{I} - h\mathcal{A}B(V)\big)^{-1} - \mathcal{I} \big\|_{\lambda\to\mu} \|\mathbf{e} \otimes v\|_\lambda \\ &\leq \quad K_1 h^{\lambda-\mu}(\|u_0\|_\lambda + 2R/3).\end{aligned}$$

Select $h_2 > 0$ such that $K_1 h_2^{\lambda-\mu}(\|u_0\|_\lambda + 2R/3) \leq R/3$. It is clear now that, for $0 < h \leq \min\{h_1, h_2\}$, $F_{h,v}$ maps $D_\mu(v, R/3)^s$ to itself. Moreover, for $V_1, V_2 \in D_\mu(v, R/3)^s$ we have

$$F_{h,v}(V_2) - F_{h,v}(V_1) = \Big(\big(\mathcal{I} - h\mathcal{A}B(V_2)\big)^{-1} - \big(\mathcal{I} - h\mathcal{A}B(V_1)\big)^{-1}\Big)(\mathbf{e} \otimes v).$$

Then, by (10),

$$\begin{aligned}&\|F_{h,v}(V_2) - F_{h,v}(V_1)\|_\mu \\ &\quad \leq \quad \big\| \big(\mathcal{I} - h\mathcal{A}B(V_2)\big)^{-1} - \big(\mathcal{I} - h\mathcal{A}B(V_1)\big)^{-1} \big\|_{\lambda\to\mu} \|\mathbf{e} \otimes v\|_\lambda \\ &\quad \leq \quad K_1(\|u_0\|_\lambda + 2R/3) h^{\lambda-\mu} \|V_2 - V_1\|_\mu.\end{aligned}$$

Take $h_3 > 0$ such that $K_1(\|u_0\|_\lambda + 2R/3)h_3^{\lambda-\mu} \le 1/2$ and set $h_0 = \min\{h_1, h_2, h_3\}$. Then, for $0 < h \le h_0$, the operator $F_{h,v}$ is a contraction on $D_\mu(v, R/3)^s$.

Finally, let v, $v^* \in D_\lambda(u_0, 2R/3)$ and set $V^* = \sigma_h(v^*)$ and $V = \sigma_h(v)$. For $0 < h \le h_0$, we have the identity

$$\begin{aligned} V^* - V &= F_{h,v^*}(V^*) - F_{h,v}(V) \\ &= \big(\mathcal{I} - h\mathcal{A}B(V^*)\big)^{-1}(\mathbf{e}\otimes v^*) - \big(\mathcal{I} - h\mathcal{A}B(V)\big)^{-1}(\mathbf{e}\otimes v) \\ &= \big(\mathcal{I} - h\mathcal{A}B(V^*)\big)^{-1}(\mathbf{e}\otimes v^* - \mathbf{e}\otimes v) \\ &\qquad + \Big(\big(\mathcal{I} - h\mathcal{A}B(V^*)\big)^{-1} - \big(\mathcal{I} - h\mathcal{A}B(V)\big)^{-1}\Big)(\mathbf{e}\otimes v). \end{aligned}$$

On the one hand, by (8), we get

$$\big\|\big(\mathcal{I} - h\mathcal{A}B(V^*)\big)^{-1}(\mathbf{e}\otimes v^* - \mathbf{e}\otimes v)\big\|_\mu \le K_1\|v^* - v\|_\mu,$$

and, on the other hand, (10) yields

$$\begin{aligned} &\Big\|\Big(\big(\mathcal{I} - h\mathcal{A}B(V^*)\big)^{-1} - \big(\mathcal{I} - h\mathcal{A}B(V)\big)^{-1}\Big)(\mathbf{e}\otimes v)\Big\|_\mu \\ &\qquad\le K_1 h^{\lambda-\mu}\|V^* - V\|_\mu\|v\|_\lambda \\ &\qquad\le K_1 h^{\lambda-\mu}(\|u_0\|_\lambda + 2R/3)\|V^* - V\|_\mu \\ &\qquad\le \tfrac{1}{2}\|V^* - V\|_\mu. \end{aligned}$$

Therefore, we have proved that

$$\|V^* - V\|_\mu \le K_1\|v^* - v\|_\mu + \tfrac{1}{2}\|V^* - V\|_\mu$$

and hence we obtain

$$\|\sigma_h(v^*) - \sigma_h(v)\|_\mu = \|V^* - V\|_\mu \le 2K_1\|v^* - v\|_\mu \le 2K_1\|v^* - v\|_\lambda. \quad \square$$

3. Proof of the main result

Choose $\lambda > 0$ and $\rho > 0$ such that $\mu < \lambda < \nu$ and $\rho < \nu - \lambda$. Let us apply Lemma 2.2 to $u_0 \in W \bigcap X_\lambda$. Let $R > 0$, K and h_0 be the radius, constant and threshold provided by that lemma. Fix $L^* > 0$ and $T > 0$ satisfying $L^*T^\rho < R/3$. For $0 < h < h_0$ we set $N = [T/h]$, and we define

$$\begin{aligned} H = \big\{\, \mathbf{v} = \{v_n\}_{n=0}^N \in D_\nu(u_0, R/3) \times X_\lambda^N \,:\, &\|v_n - v_m\|_\lambda \le L^*(t_n - t_m)^\rho, \\ &0 \le m < n \le N \,\big\}. \end{aligned}$$

and

$$\Sigma = D_\nu(0, S)^N = \big\{\, \boldsymbol{\tau} = \{\tau_n\}_{n=1}^N \in X_\nu^N \,:\, \|\tau_n\|_\nu \le S, \quad 1 \le n \le N \,\big\}.$$

Let $\mathbf{v} = \{v_n\}_{n=0}^N \in H$. Due to our choice of L^* and T, for $0 \le n \le N$ we have

$$\|v_n - u_0\|_\lambda \le \|v_n - v_0\|_\lambda + \|v_0 - u_0\|_\nu \le L^*T^\rho + R/3 \le 2R/3,$$

and hence $v_n \in D_\lambda(u_0, 2R/3)$. Thus, the stage vectors $V_n = \sigma_h(v_n) \in D_\mu(v_n, R/3)^s \subset D_\mu(u_0, R)^s$, $0 \le n \le N$, are well defined. Therefore, it makes sense to consider the linear operators in X, depending on $\mathbf{v}$, given by

$$p_n(\mathbf{v}) = I + h\mathbf{b}^T B(V_{n-1})\big(\mathcal{I} - h\mathcal{A}B(V_{n-1})\big)^{-1}(\mathbf{e}\otimes I), \qquad 1 \le n \le N,$$

and by

$$P_{n,m}(\mathbf{v}) = \prod_{j=m+1}^{n} p_j(\mathbf{v}), \qquad 0 \le m \le n \le N.$$

With this notation we are going to define a non-linear operator $\mathbf{F} : H \times \Sigma \to X_\nu \times X_\mu^N$. For $\mathbf{v} \in H$ and $\boldsymbol{\tau} = \{\tau_n\}_{n=1}^N \in \Sigma$, we set

$$\mathbf{F}(\mathbf{v}, \boldsymbol{\tau}) = \{f_n(\mathbf{v}, \boldsymbol{\tau})\}_{n=0}^N,$$

where $f_0(\mathbf{v}, \boldsymbol{\tau}) = v_0$ and, for $1 \le n \le N$,

$$f_n(\mathbf{v}, \boldsymbol{\tau}) = P_{n,0}(\mathbf{v})v_0 + h\sum_{j=1}^{n} P_{n,j}(\mathbf{v})\tau_j.$$

Notice that, after Lemma 2.1, the operator $\mathbf{F}$ takes values in X_ν^{1+N}.

The proof of the present theorem is based on the following remark. Let $\mathbf{v} = \{v_n\}_{n=0}^N \in H$, $\boldsymbol{\tau} = \{\tau_n\}_{n=1}^N \in \Sigma$, and assume that

$$\mathbf{v} = \mathbf{F}(\mathbf{v}, \boldsymbol{\tau}). \tag{19}$$

Then, by the discrete variation-of-constants formula, it is clear that the recurrence

$$v_{n+1} = \mathcal{N}_h(v_n) + \tau_{n+1}, \qquad 0 \le n \le N-1,$$

is satisfied. In conclusion, for $0 < h \le h_0$ the Runge-Kutta method, starting from v_0 and with sources $\boldsymbol{\tau}$, is applicable up to time T. For $v_0 \in D_\nu(u_0, R/3)$, let H_{v_0} be the subset of H of those elements whose first component is v_0. The idea of the proof is to fix $v_0 \in D_\nu(u_0, R/3)$ and $\boldsymbol{\tau} \in \Sigma$, and then to solve the fixed-point equation (19) in H_{v_0}. The proofs of the existence of the fixed point in (19) and of its continuous dependence on v_0 and $\boldsymbol{\tau}$ are both based on the estimates (20) and (21) below. The proofs of (20) and (21) require the results in Section 7 and are given at the end of the present section. Thus, assume that there exist continuous mappings $C_k : [0, +\infty) \to (0, +\infty)$, $1 \le k \le 3$, such that, for $0 \le m \le n \le N$, $\mathbf{v} \in H$ and $\boldsymbol{\tau} \in \Sigma$,

$$\|f_n(\mathbf{v}, \boldsymbol{\tau}) - f_m(\mathbf{v}, \boldsymbol{\tau})\|_\lambda \le C_1(T)\big(T^{\nu-\lambda-\rho}\|v_0\|_\nu + ST^{1-\rho}\big)(t_n - t_m)^\rho, \tag{20}$$

and for $0 \le \eta \le \nu$

$$\begin{aligned} &\|f_n(\mathbf{v}^*, \boldsymbol{\tau}^*) - f_n(\mathbf{v}, \boldsymbol{\tau})\|_\eta \\ &\quad \le C_2(T)\|\mathbf{v}^* - \mathbf{v}\|_\lambda\Big(T^{\nu-\eta}\|v_0^*\|_\nu + h\sum_{j=1}^{n}\|\tau_j\|_\nu\Big) \\ &\qquad + C_3(T)\Big(\|v_0^* - v_0\|_\nu + h\sum_{j=1}^{n}\|\tau_j^* - \tau_j\|_\nu\Big). \end{aligned} \tag{21}$$

Reducing T if necessary, in the rest of the proof we can suppose that

$$C_1(T)\big(T^{\nu-\lambda-\rho}(\|u_0\|_\nu + R/3) + ST^{1-\rho}\big) \le L^*$$

and

$$C_2(T)\big(T^{\nu-\lambda}(\|u_0\|_\nu + R/3) + TS\big) \le 1/2.$$

Fix $v_0 \in D_\nu(u_0, R/3)$ and $\boldsymbol{\tau} \in \Sigma$. Notice that H_{v_0} is a complete metric space with respect to the distance induced by the one in X_λ^N. By (20), we deduce that the restriction mapping $\mathbf{v} \to \mathbf{F}(\mathbf{v}, \boldsymbol{\tau})$ maps H_{v_0} into H_{v_0}. Moreover, by (21), this restriction mapping contracts the distance in H_{v_0}. Therefore, by the Banach contraction principle, this mapping possesses a unique fixed point $\mathbf{v} = \{v_n\}_{j=0}^N \in H_{v_0}$. As we mentioned before, this proves part (i) of the theorem.

Next we prove the stability estimate (ii). It is enough to show (ii) for $n = N$. Thus, assume that $\mathbf{v}^*$, $\mathbf{v} \in H$ and $\boldsymbol{\tau}^*$, $\boldsymbol{\tau} \in H$ satisfy

$$\mathbf{v}^* = \mathbf{F}(\mathbf{v}^*, \boldsymbol{\tau}^*), \qquad \mathbf{v} = \mathbf{F}(\mathbf{v}, \boldsymbol{\tau}).$$

From (21), with $\eta = \lambda$, and the choice of T we deduce that

$$\begin{aligned} \|\mathbf{v}^* - \mathbf{v}\|_\lambda &= \|\mathbf{F}(\mathbf{v}^*, \boldsymbol{\tau}^*) - \mathbf{F}(\mathbf{v}, \boldsymbol{\tau})\|_\lambda \\ &\leq (1/2)\|\mathbf{v}^* - \mathbf{v}\|_\lambda + C_3(T)\big(\|v_0^* - v_0\|_\nu + h\sum_{j=1}^N \|\tau_j^* - \tau_j\|_\nu\big), \end{aligned}$$

which yields

$$\|\mathbf{v}^* - \mathbf{v}\|_\lambda \leq 2C_3(T)\big(\|v_0^* - v_0\|_\nu + h\sum_{j=1}^N \|\tau_j^* - \tau_j\|_\nu\big).$$

Using the last estimate in (21) again, but now with $\eta = \nu$, we get

$$\begin{aligned} \|\mathbf{v}^* - \mathbf{v}\|_\nu &= \|\mathbf{F}(\mathbf{v}^*, \boldsymbol{\tau}^*) - \mathbf{F}(\mathbf{v}, \boldsymbol{\tau})\|_\nu \\ &\leq C_2^*(T)\|\mathbf{v}^* - \mathbf{v}\|_\lambda + C_3(T)\big(\|v_0^* - v_0\|_\nu + h\sum_{j=1}^N \|\tau_j^* - \tau_j\|_\nu\big) \\ &\leq C_3(T)(2C_2^*(T) + 1)\big(\|v_0^* - v_0\|_\nu + h\sum_{j=1}^N \|\tau_j^* - \tau_j\|_\nu\big), \end{aligned}$$

where $C_2^*(T) = C_2(T)(\|u_0\| + R/3 + TS)$. This is precisely (ii) for $n = N$.

It remains to prove the basic estimates (20) and (21). Let $\mathbf{v} = \{v_n\}_{n=0}^N \in H$, and consider the associated sequence of internal stages $\{V_n^i\}_{i=1}^s = \sigma_h(v_n) \in D_\mu(v_n, R/3)^s$, $0 \leq n \leq N$. Let $\varphi_{\mathbf{v}} : [t_0, t_0 + T] \to D_\mu(u_0, R) \subset W \subset X_\mu$ be the unique piecewise linear interpolant with nodal values

$$\varphi_{\mathbf{v}}(t_n + c_i h) = V_n^i, \qquad 0 \leq n \leq N, \quad 1 \leq i \leq s.$$

It is clear that, due to the Lemma 2.2, there exists a constant $K > 0$ satisfying

$$\|\varphi_{\mathbf{v}}(t) - \varphi_{\mathbf{v}}(s)\|_\mu \leq KL^*|t - s|^\rho, \qquad s,\, t \in [t_0, t_0 + T],\ \mathbf{v} \in H. \tag{22}$$

For $\mathbf{v} \in H$, let $A_{\mathbf{v}}(t) : X_1 \subset X \to X$ be the family of linear operators

$$A_{\mathbf{v}}(t) = Q(\varphi_{\mathbf{v}}(t)), \qquad t_0 \leq t_0 + T.$$

Notice that, because of $\mathbf{H}1$ and (22), we have for $x \in X_1$ and t, $s \in [t_0, t_0 + T]$

$$\|(A_{\mathbf{v}}(t) - A_{\mathbf{v}}(s))x\| \leq L\|\varphi_{\mathbf{v}}(t) - \varphi_{\mathbf{v}}(s)\|_\mu \|x\|_1 \leq KLL^*|t - s|^\rho \|x\|_1.$$

Furthermore, it is plain that the operators $P_{n,m}(\mathbf{v})$, $0 \leq m \leq n \leq N$, are the discrete transition operators corresponding to the Runge–Kutta method applied to the non-autonomous problem

$$u'(t) = A_{\mathbf{v}}(t)u(t),$$

Therefore, by Theorem 7.1 we get (20). Finally, let us prove (21). Let $\mathbf{v}$, $\mathbf{v}^* \in H$. By Lemma 2.2, we have

$$\|\varphi_{\mathbf{v}^*}(t) - \varphi_{\mathbf{v}}(t)\|_\mu \leq KL^*\|\mathbf{v}^* - \mathbf{v}\|_\lambda,$$

so that

$$\big\| \big(A_{\mathbf{v}^*}(t) - A_{\mathbf{v}}(t)\big)x \big\| \leq L\|\varphi_{\mathbf{v}^*}(t) - \varphi_{\mathbf{v}}(t)\|_\mu \|x\|_1 \leq KLL^*\|\mathbf{v}^* - \mathbf{v}\|_\lambda \|x\|_1,$$

and (21) is readily obtained from Theorem 7.2 with $\varepsilon = KLL^*\|\mathbf{v}^* - \mathbf{v}\|_\lambda$. □

4. Convergence

As we will see, the stability result in Theorem 1.1 is well suited to the analysis of the convergence. The order of convergence of the method turns out to be essentially the so-called stage order q^*, rather than the classical order p. Some extra fractional order may be present, depending on the norm in which the error is measured.

Let us recall that the stage order q^* of the Runge–Kutta method is defined as the maximum integer $q^* \geq 1$ with the property that, for $1 \leq i \leq s$, the quadrature formula in $[0, 1]$ with nodes c_j and weights a_{ij}, $1 \leq j \leq s$, is of order q^*. We set $q = \min\{q^*, p\}$ (in practice $q = q^*$ for all the methods).

Let $u : J \to W$ be a solution of (1) defined on a compact interval $J = [t_0, t_0 + T]$, $t_0 \in \mathbf{R}$, $T > 0$. We suppose that $u \in C(J, X_1) \bigcap C^{p+1}(J, X_{\nu^*})$, where $\mu < \nu^* < 1$. The error can be measured in the norms of any X_ν, with $\mu < \nu \leq \nu^*$. In the sequel we fix $\nu \in (\mu, \nu^*]$ and set $\beta = \nu^* - \nu \geq 0$ and $\bar{q} = \min\{q + \beta, p\}$. As we will show, the order of convergence, in the norm of X_ν, is $\bar{q}$.

As a first step we study the local error. Choose $R > 0$ and $h_0 > 0$ in such a way that Lemmas 2.1 and 2.2 remains valid uniformly along the values $u(t)$, $t \in J$, for $0 < h \leq h_0$. We set $J_h = [t_0, t_0 + T - h]$, for $0 < h \leq h_0$. After reducing h_0 if necessary, we can also suppose that $\|u(t + h) - u(t)\|_\mu \leq R$, for $0 < h \leq h_0$ and $t \in J_h$. Thus, for $0 < h \leq h_0$, the Runge–Kutta method turns out to be well defined at all the points $u(t)$, $t \in J_h$. The local error $\epsilon_h(t)$ is defined by

$$u(t + h) = \mathcal{N}_h\big(u(t)\big) + \epsilon_h(t), \qquad t \in J_h.$$

Suppose we have proved an estimate of the form

$$\|\epsilon_h(t)\|_\nu = \mathcal{O}(h^{\bar{q}+1}), \tag{23}$$

where the leading constant on the right hand side, though depending on u, is uniform in $t \in J_h$. Let t_j, $0 \leq j \leq N$, be a finite sequence of time levels in J, with constant step-size. For $0 \leq n \leq N - 1$ we have

$$\begin{aligned} u(t_{n+1}) &= \mathcal{N}_h(u(t_n)) + \epsilon_h(t_n), \\ u_{n+1} &= \mathcal{N}_h(u_n), \end{aligned}$$

as long as the method is defined. Therefore, after a possible reduction of h_0 and T, Theorem 1.1 applied with $\tau_{n+1} := h^{-1}\epsilon_h(t_n)$, $0 \leq n \leq N$, yields the bound

$$\|u(t_n) - u_n\|_\nu = \mathcal{O}(h^{\bar{q}}), \qquad 0 \leq n \leq N - 1, \tag{24}$$

which constitutes the estimate for the global error.

Thus, it remains to prove (23). For $t \in J_h$, set

$$U_h(t) = [u(t + c_1 h), u(t + c_2 h), \dots, u(t + c_s h)]^T \in X_\nu^s.$$

Define the residuals $\Delta_h(t) \in X^s$ and $\delta_h(t)$ (cf. [20]) by means of the expressions

$$u(t + h) = u(t) + h\mathbf{b}^T U_h'(t) + \delta_h(t) = u(t) + h\mathbf{b}^T B\big(U_h(t)\big)U_h(t) + \delta_h(t)$$

and

$$U_h(t) = \mathbf{e} \otimes u(t) + h\mathcal{A}U_h'(t) + \Delta_h(t) = \mathbf{e} \otimes u(t) + h\mathcal{A}B\big(U_h(t)\big)U_h(t) + \Delta_h(t)$$

It is well known that the quadrature formula with weights $\mathbf{b}$ and nodes $\mathbf{c}$ is of order p. This gives $\|\delta_h(t)\|_{\nu^*} = \mathcal{O}(h^{p+1})$, since $u \in C^{p+1}(J, X_{\nu^*})$. Therefore,

$$\|\delta_h(t)\|_{\nu^*} = \mathcal{O}(h^{\bar{q}+1}).$$

Analogously, by the definition of q^*, we also have

$$\|\Delta_h(t)\|_{\nu^*} = \mathcal{O}(h^{q+1}).$$

Notice that the constants in the previous two upper bounds depend on the size of $\|u^{(q+1)}\|_{\nu^*}$, so that they are uniform in $t \in J_h$.

On the other hand, $\|U_h(t) - \mathbf{e} \otimes u(t)\|_\mu \leq R$, because of our choice of h_0. Then we can apply Lemma 2.1 and conclude that the inverse of $\mathcal{I} - h\mathcal{A}B\big(U_h(t)\big)$ exists as a bounded operator in X^s. Hence, we get the representation

$$U_h(t) = \big(\mathcal{I} - h\mathcal{A}B\big(U_h(t)\big)\big)^{-1}\big(\mathbf{e} \otimes u(t) + \Delta_h(t)\big), \qquad t \in J_h.$$

Let us apply the Runge–Kutta method at $u(t)$, $t \in J_h$. The stages $V_h(t) = \sigma_h\big(u(t)\big)$ are uniquely defined as the solutions of

$$V_h(t) = \mathbf{e} \otimes u(t) + h\mathcal{A}B\big(V_h(t)\big)V_h(t)$$

or, by Lemma 2.2, by the expression

$$V_h(t) = \big(\mathcal{I} - h\mathcal{A}B\big(V_h(t)\big)\big)^{-1}\big(\mathbf{e} \otimes u(t)\big).$$

Denoting $E_h(t) = U_h(t) - V_h(t)$, we get

$$\begin{aligned} E_h(t) \quad = \quad & \big(\mathcal{I} - h\mathcal{A}B\big(U_h(t)\big)\big)^{-1}\Delta_h(t) \\ & + \Big(\big(\mathcal{I} - h\mathcal{A}B\big(U_h(t)\big)\big)^{-1} - \big(\mathcal{I} - h\mathcal{A}B\big(V_h(t)\big)\big)^{-1}\Big)\big(\mathbf{e} \otimes u(t)\big). \end{aligned}$$

By (10), it is clear that

$$\|E_h(t)\|_\mu \leq K_1\|\Delta_h(t)\|_{\nu^*} + K_1 h^{\nu^*-\mu}\|E_h(t)\|_\mu \|u(t)\|_{\nu^*}.$$

Thus, by reducing h_0 so as to have $h_0^{\nu^*-\mu} K_1 \max_{s\in J} \|u(s)\|_{\nu^*} \leq 1/2$, we deduce that

$$\|E_h(t)\|_\mu \leq 2K_1\|\Delta_h(t)\|_{\nu^*} = \mathcal{O}(h^{q+1}). \tag{25}$$

In the same way, we have

$$\begin{aligned} E_h(t) - \Delta_h(t) = & \Big(\big(\mathcal{I} - h\mathcal{A}B\big(U_h(t)\big)\big)^{-1} - \mathcal{I}\Big)\Delta_h(t) \\ & + \Big(\big(\mathcal{I} - h\mathcal{A}B\big(U_h(t)\big)\big)^{-1} - \big(\mathcal{I} - h\mathcal{A}B\big(V_h(t)\big)\big)^{-1}\Big)(\mathbf{e} \otimes u(t)). \end{aligned}$$

Therefore, from (9), (10) and (25), we obtain

$$\|E_h(t) - \Delta_h(t)\|_\nu \leq K_1 h^\beta \|\Delta_h(t)\|_{\nu^*} + K_1 h^\beta \|E_h(t)\|_\mu \|u(t)\|_{\nu^*} = \mathcal{O}(h^{q+1+\beta}).$$

Hence, since

$$hB\big(U_h(t)\big)U_h(t) - hB\big(V_h(t)\big)V_h(t) = \mathcal{A}^{-1}\big(E_h(t) - \Delta_h(t)\big),$$

we conclude that

$$h\|\mathbf{b}^T\big(B\big(U_h(t)\big)U_h(t) - B\big(V_h(t)\big)V_h(t)\big)\|_\nu = \mathcal{O}(h^{q+1+\beta}).$$

Now the error estimate (23) is readily proved, by taking norms in the expression

$$\epsilon_h(t) = u(t+h) - \mathcal{N}_h(u(t)) = h\mathbf{b}^T\big(B\big(U_h(t)\big)U_h(t) - B\big(V_h(t)\big)V_h(t)\big) + \delta_h(t).$$

5. Some extensions

First, let us comment on how to extend our results to abstract time-dependent quasilinear problems of the form

$$\text{(26)}\qquad \begin{cases} u'(t) &= Q\big(t,u(t)\big)u(t), \\ u(t_0) &= u_0, \end{cases}$$

where $Q(t,w) : X_1 \subset X \to X$ is a family of linear operators, defined for t in some open interval J and for $w \in W \subset X$. Now, along with the obvious modification of **H**1, we require that there exist $L > 0$, $\mu \in [0,1)$ and $\delta \in (0,1)$ such that W is open in X_μ and such that

H2′. $\|Q(t,w)-Q(s,v)\|_{1\to 0} \le L\big(|t-s|^\delta + \|w-v\|_\mu\big)$, for t, $s \in J$ and v, $w \in W$.

The existence and uniqueness of the solution of (26) are known [3, 4, 23].

Lemmas 2.1 and 2.2 are easily adapted to the context of problem (26). With similar arguments we can prove that, for each initial condition $(t_0,u_0) \in J \times X_\lambda$ with $\mu < \lambda \le 1$, there exist $T > 0$, $R > 0$ and a nonlinear operator $\mathcal{N}_h : [t_0,t_0+T]\times D_\lambda(u_0,2R/3) \to D_\lambda(u_0,R)$ in such a way that the Runge–Kutta scheme, applied to (26), is given by the recurrence

$$u_{n+1} = \mathcal{N}_h(t_n,u_n), \qquad 0 \le n \le N-1,$$

as long as $t_N \le t_0 + T$, and that $u_n \in D_\lambda(u_0,2R/3)$, $0 \le n \le N-1$. The proof of Theorem 1.1 for problem (26) remains the same, except for obvious modifications. The only specific detail is that the parameter ρ in the definition of H must be $< \delta$.

Second, let us consider a semilinear term in (26). Then (26) becomes

$$\text{(27)}\qquad \begin{cases} u'(t) &= Q\big(t,u(t)\big)u(t) + f(t,u(t)), \\ u(t_0) &= u_0, \end{cases}$$

where $f : J \times W \to X$. We keep the hypotheses **H**1-2, and we also suppose that

H3. $\|f(t,w) - f(s,v)\| \le L\big(|t-s|^\delta + \|w-v\|_\mu\big)$, for t, $s \in J$ and v, $w \in W$.

Instead of adapting the proofs of our results to cover (27), it is better to reduce (27) to a problem of the form (26). To this end, we set $\widetilde{X} = \mathbf{R}\times X$, $\widetilde{X}_1 = \mathbf{R}\times X_1$, $\widetilde{W} = J \times W$, $\widetilde{u}_0 = [1,u_0]^T \in \widetilde{X}_1$, and we consider the operator $\widetilde{Q} : J \times \widetilde{X}_1 \subset J \times \widetilde{X} \to J \times \widetilde{W}$ defined by

$$\widetilde{Q}(t,\widetilde{w}) = \begin{bmatrix} 0 & 0 \\ f(t,w) & Q(t,w) \end{bmatrix}, \qquad \widetilde{w} = [\tau,w]^T \in \widetilde{W}.$$

It is clear that (27) is equivalent to the time-dependent quasilinear problem

$$\begin{cases} \widetilde{u}'(t) &= \widetilde{Q}\big(t,\widetilde{u}(t)\big)\widetilde{u}(t), \\ \widetilde{u}(t_0) &= \widetilde{u}_0, \end{cases}$$

which satisfies the standard conditions **H**1-2. Moreover, Runge–Kutta methods are compatible with this reduction.

6. Applications

The abstract formulation of the standard quasilinear parabolic problems arising in the applications is as follows. Let $\Omega \subset \mathbf{R}^d$ be an open and bounded domain with regular boundary Γ, let $J \subset \mathbf{R}$ be an open interval and let $\Lambda \subset \mathbf{R}\times\mathbf{R}^d$ be an open domain. Assume that we are given coefficients $a_{ij} : \Omega \times J \times \Lambda \to \mathbf{R}$,

$1 \leq i,\, j \leq d$. For $w \in C^1(\Omega \times J)$, let $Q(t,w)$ be the second-order linear differential operator defined by

$$Q(t,w)u(x) = \sum_{i,j=1}^{d} a_{ij}(x,t,w(x,t),\nabla w(x,t))\partial_{ij}u(x), \qquad u \in C^2(\Omega), \quad x \in \Omega.$$

For $[t_0, t_0+T] \subset J$, let us consider the initial value problem

$$\text{(28)} \qquad \begin{cases} u_t(x,t) &=& Q(t,u(x,t))u(x,t), \qquad t_0 \leq t \leq t_0+T, \quad x \in \Omega, \\ u(x,t_0) &=& u_0(x), \qquad x \in \Omega, \\ u(x,t) &=& 0, \qquad t_0 \leq t \leq t_0+T, \quad x \in \Gamma. \end{cases}$$

Together with suitable hypotheses on the regularity and boundedness of the coefficients, we impose an ellipticity condition,

$$\sum_{i,j=1}^{d} a_{ij}(x,t,u,\mathbf{p})\xi_i\xi_j \geq \omega \sum_{i=1}^{d} \xi_i^2, \qquad (x,t,u,\mathbf{p}) \in \Omega \times J \times \Lambda, \quad \{\xi_i\}_{i=1}^{d} \in \mathbf{R}^d,$$

for some $\omega > 0$. Take $d < p < +\infty$. Set $X = L^p(\Omega)$, $X_1 = W_0^{1,p}(\Omega) \bigcap W^{2,p}(\Omega)$ and $W = X_\mu = [X, X_1]_\mu$, with $\frac{1}{2} + \frac{d}{2p} < \mu < 1$. Then $X_\mu \subset C^1(\overline{\Omega})$, and $Q(t,w)$ makes sense as a linear operator from X_1 to X, even for $w \in W$. In this way, (28) is written in the abstract setting of (26). The validity of **H**1-2 has been considered e.g. in [3, 23]. Thus, classical quasilinear parabolic problems (28) can be studied in the abstract setting of Banach spaces. The same comment applies to systems of equations. After the previous section, an extra semilinear term $f(x,t,u(x,t),\nabla u(x,t))$ can be incorporated in the right hand side in (28). However, within the present framework, we cannot consider Neumann boundary conditions, because the conormal derivative may depend on the solution and a variable domain $\mathcal{D}_w$ for $Q(t,w)$ may result.

The cases $p = 1, +\infty$ are more delicate. This is due to the absence of the Agmon–Douglis–Nirenberg estimates. For these values of p only simple problems (28) have been shown to be included in the abstract framework (see [18]).

The results of the present paper apply to semidiscretizations in time of (28), by means of Runge–Kutta methods. However, this is only a first step in the study of the discretizations of such problems. In practice, the main interest is in studying the discretizations in both space and time of (28). In general, we first discretize in space and then we apply a time-stepping method, let us say a Runge–Kutta method. In this way we are led to the consideration of a family of problems of the form

$$u'_{\Delta x}(t) = Q_{\Delta x}\big(u_{\Delta x}(t)\big)u_{\Delta x}(t),$$

where $\Delta x > 0$ stands for the parameter of the spatial discretization and $Q_{\Delta x} : X_{\Delta x} \to X_{\Delta x}$, $\Delta x > 0$, are approximations to the operator Q, defined in some discrete spaces $X_{\Delta x}$ approximating X. For the rigorous analysis of these procedures using the results of the present paper, we must ascertain whether **H**1-2 hold uniformly in Δx or, at least, with constants $M_{\Delta x}$ and $L_{\Delta x}$ exhibiting only weak singularities as $\Delta x \to 0+$. For non-Hilbert norms, these issues are far from being satisfactorily solved and constitute an interesting line of investigation. We mention that the results for finite differences in [2] seem likely to be extendable to quasilinear problems, although such an extension has not been carried out. The same comment applies to discretizations based on finite elements (see [12, 21]).

7. Non-autonomous case

In this section we state several results on the semidiscretization in time of linear, non-autonomous parabolic problems. These results are interesting on their own, and are needed for the proof of the basic estimates (20) and (21) in the main Theorem 1.1.

Let $J \subset \mathbf{R}$ be an interval and let $A(t) : X_1 \subset X \to X$, $t \in J$, be a family of linear, densely defined operators. Let us consider the initial value problem

$$\left\{ \begin{array}{rcll} u'(t) & = & A(t)u(t), & t \in J, \\ u(t_0) & = & u_0 \in X_1, & t_0 \in J. \end{array} \right. \tag{29}$$

We assume that the following two standard hypotheses [14] hold:

NA1. There exist $M \geq 1$, $\omega_0 \in \mathbf{R}$ and $\theta \in (0, \pi/2)$ such that, for a complex $z \notin \omega_0 + S_\theta$ and $t \in J$, the resolvent $\big(zI - A(t)\big)^{-1} : X \to X$ exists and

$$\big\|\big(zI - A(t)\big)^{-1}\big\| \leq \frac{M}{|z - \omega_0|}.$$

NA2. There exist $\tilde{L} > 0$ and $\alpha \in (0, 1]$ such that

$$\|A(t)x - A(s)x\| \leq \tilde{L}|t - s|^\alpha \|x\|_1, \qquad x \in X_1, \quad t,\, s \in J.$$

It is well known that **N**A1-2 guarantee the existence and uniqueness of the solution of the problem (29) (see e.g. [1, 19, 22, 23, 24]).

Let $u : J \to X$ be the solution of problem (29) and let t_n, $0 \leq n \leq N$, be a finite sequence of time levels in J, with constant step-size $h > 0$. The Runge–Kutta method given by (2) applied to problem (29) leads to the recurrence

$$u_{n+1} = u_n + h \sum_{i=1}^{s} b_i A(t_n + c_i h) U_n^i, \qquad 0 \leq n \leq N - 1. \tag{30}$$

Here u_n is the approximation to $u(t_n)$, $0 \leq n \leq N$, and the internal stages U_n^i, $0 \leq n \leq N-1$, $1 \leq i \leq s$, are defined by means of the system of equations

$$U_n^i = u_n + h \sum_{j=1}^{s} a_{ij} A(t_n + c_j h) U_n^j. \tag{31}$$

For $t \in J$ such that $t + h \in J$, we let $B(t) : X_1^s \subset X^s \to X^s$ be the operator defined by $B(t) = \operatorname{diag}(A(t + c_1 h), \dots, A(t + c_s h))$. In Lemma 2.3 of [14] it is proved that, for $h > 0$ small enough, the operator $\mathcal{I} - h\mathcal{A}B(t) : X_1^s \subset X^s \to X^s$ possesses a bounded inverse $\big(\mathcal{I} - h\mathcal{A}B(t)\big)^{-1} : X^s \to X^s$. Thus, system (31) is uniquely solvable. Moreover, for $0 \leq n \leq N-1$, let $r(t_{n+1}, t_n) : X \to X$ be the continuous linear mapping defined by

$$r(t_{n+1}, t_n) = I + h\mathbf{b}^T B(t_n)\big(\mathcal{I} - h\mathcal{A}B(t_n)\big)^{-1}(\mathbf{e} \otimes I). \tag{32}$$

Then the recurrence (30) can be written in concise form as $u_{n+1} = r(t_{n+1}, t_n)u_n$, $0 \leq n \leq N$.

For the convenience of the reader, we now state the stability result obtained in [14]: There exist a threshold $\overline{h} > 0$, and constants $K > 0$, $\kappa > 1$, $\Omega > 0$ such that, for $0 < h \le \overline{h}$, $0 \le m < n \le N$ and $0 \le \eta \le \nu \le 1$,

$$\begin{aligned}(33)\qquad &\Big\| \prod_{j=m}^{n-1} r(t_{j+1}, t_j) - \gamma^{n-m} \Big\|_{\eta\to\nu} \\ &\qquad \le K e^{(\omega_0 + \Omega \tilde{L}^{1/\alpha})(t_n - t_m)} (1 + \kappa \tilde{L}(t_n - t_m)^\alpha)^{10} (t_n - t_m)^{\eta-\nu},\end{aligned}$$

where $\gamma = r(\infty)$.

In the sequel, the letter K denotes a positive constant that depends only on M, θ, ω_0 and the Runge-Kutta method. Of course, K may take different values at different places. Furthermore, $\overline{h}$, Ω and κ are the constants in (33). We also set, for a fixed $h > 0$,

$$C^l_{m,n} = e^{(\omega_0 + \Omega \tilde{L}^{1/\alpha})(t_n - t_m)} \big(1 + \kappa \tilde{L}(t_n - t_m)^\alpha\big)^l, \qquad 0 \le m \le n \le N, \quad l \ge 0.$$

Theorem 7.1. *Assume that the parabolic problem* (29) *satisfies hypotheses* ***N****A1-2 and assume that the Runge–Kutta method given by* (30) *and* (31) *is strongly A(θ)-stable. Then there exists a constant* $K > 0$ *such that for each finite sequence of time levels* t_j*,* $0 \le j \le N$*, in* J *with constant step-size* $0 < h \le \overline{h}$ *and for* $0 \le \eta < \nu \le 1$

$$(34)\qquad \| \prod_{j=0}^{N-1} r(t_{j+1}, t_j) - I \|_{\nu\to\eta} \le (K/\nu) C^{11}_{N,0} (t_N - t_0)^{\nu-\eta}.$$

Proof. Fix $0 < h \le \overline{h}$ and set $T = t_N - t_0$. We begin by proving that

$$(35)\qquad \|r(t_{j+1}, t_j) - I\|_{\nu\to 0} \le K\big(1 + \tilde{L}T^\alpha\big)^\nu h^\nu, \qquad 0 \le j \le N-1.$$

Fix $0 \le j \le N-1$, Lemma 2.3 and (20) in [14] give

$$\|r(t_{j+1}, t_j) - I\|_{0\to 0} \le 1 + \|r(t_{j+1}, t_j)\|_{0\to 0} \le K,$$

and

$$\|r(t_{j+1}, t_j) - I\|_{1\to 0} \le h \|\mathbf{b}^T\| \, \|\mathcal{A}^{-1}\| \, \|(\mathcal{I} - h\mathcal{A}B(t_j))^{-1} \mathcal{A}B(t_j)\|_{1\to 0} \le K(1 + \tilde{L}T^\alpha)h.$$

Hence, (35) is obtained by interpolation.

Set

$$G_{n,m} = \prod_{j=m}^{n-1} r(t_{j+1}, t_j), \qquad 0 \le m \le n \le N.$$

We have

$$\begin{aligned} G_{N,0} - I &= \sum_{j=0}^{N-1} (r(t_{j+1}, t_j) - I) G_{j,0} \\ &= \sum_{j=0}^{N-1} (r(t_{j+1}, t_j) - I)(G_{j,0} - \gamma^j) + \sum_{j=0}^{N-1} (r(t_{j+1}, t_j) - I)\gamma^j. \end{aligned}$$

Therefore, by (33) and (35), we get

$$\begin{aligned}
&\|G_{N,0} - I\|_{\nu\to 0} \\
&\quad\le \sum_{j=0}^{N-1} \|r(t_{j+1},t_j) - I\|_{1\to 0}\|G_{j,0} - \gamma^j\|_{\nu\to 1} + \sum_{j=0}^{N-1} \|r(t_{j+1},t_j) - I\|_{\nu\to 0}\gamma^j \\
&\quad\le \left(1+\tilde{L}T^\alpha\right)\left(Kh\sum_{j=0}^{N-1}\|G_{j,0} - \gamma^j\|_{\nu\to 1} + Kh^\nu \sum_{j=0}^{N-1}\gamma^j\right) \\
&\quad\le \left(1+\tilde{L}T^\alpha\right)\left(Kh\sum_{j=0}^{N-1}(t_j - t_0)^{\nu-1}C^{10}_{j,0} + Kh^\nu/(1-\gamma)\right) \\
&\quad\le \left(1+\tilde{L}T^\alpha\right)\left((K/\nu)C^{10}_{N,0} + K/(1-\gamma)\right)(t_N - t_0)^\nu .
\end{aligned}$$

On the other hand, by Theorem 1.1 of [14] we have

$$\|G_{N,0} - I\|_{\nu\to\nu} \le KC^5_{N,0}.$$

Now (34) is proved by interpolating between the two previous estimates. □

In relation with the estimate (21) it is necessary to study the dependence of the numerical solution on the coefficients $A(t)$, $t \in J$. To this end, we consider another family $A^*(t) : X_1 \subset X \to X$, $t \in J$, of linear operators satisfying conditions **N**A1-2. The corresponding discrete operators defined by the Runge-Kutta method are now denoted by $r^*(t_{n+1}, t_n) : X \to X$, $0 \le n \le N-1$.

Next we state a lemma. The first estimate (36) is readily proved by comparing with an integral. The second one (37) was obtained in the proof of Lemma 2.1 of [14].

Lemma 7.1. *Let $\nu \in (0,1)$. Then there exist constants $b_1(\nu)$, $b_2(\nu) > 0$ such that, for $0 \le \eta \le \nu$,*

(36)
$$\sum_{j=m+1}^{n-1} h(t_n - t_j)^{-\eta}(t_j - t_m)^{\nu-1} \le b_1(\nu)(t_n - t_m)^{\nu-\eta}, \qquad 0 \le m < n \le N,$$

(37)
$$\sum_{j=m}^{n-1} h^\nu (t_n - t_j)^{-\eta}\gamma^{j-m} \le b_2(\nu)(t_n - t_m)^{\nu-\eta}, \qquad 0 \le m < n \le N,$$

where t_n, $0 \le n \le N$, is a finite sequence of time levels with step-size $h > 0$.

Theorem 7.2. *Let $A(t)$, $A^*(t) : X_1 \subset X \to X$, $t \in J$, be two families of operators. Assume that they satisfy hypotheses **N**A1-2, with the same constants, and that the Runge–Kutta method is strongly $A(\theta)$-stable. Then there exists $K > 0$ such that, for an arbitrary finite sequence of time levels t_n, $0 \le n \le N$, in J with step-size $0 < h \le \overline{h}$, and for $0 < \nu < 1$ and $0 \le \eta \le \nu$,*

(38)
$$\Big\| \prod_{j=0}^{N-1} r(t_{j+1},t_j) - \prod_{j=0}^{N-1} r^*(t_{j+1},t_j) \Big\|_{\nu\to\eta} \le K\big(b_1(\nu) + b_2(\nu)\big)\varepsilon C^{22}_{N,0}(t_N - t_0)^{\nu-\eta},$$

where

$$\varepsilon = \sup\{\|A(t) - A^*(t)\|_{1\to 0} \,:\, t_0 \le t \le t_N\}.$$

Proof. Fix $0 < h \le \overline{h}$ and set $T = t_N - t_0$. For $0 \le j \le N$, let $B^*(t_j) : X_1^s \subset X^s \to X^s$ be the operator defined by

$$B^*(t_j) = \mathrm{diag}(A^*(t_j + c_1 h), A^*(t_j + c_2 h), \dots, A^*(t_j + c_s h)).$$

Notice that

(39)
$$\begin{aligned} &hB(t_j)\big(\mathcal{I} - h\mathcal{A}B(t_j)\big)^{-1} - hB^*(t_j)\big(\mathcal{I} - h\mathcal{A}B^*(t_j)\big)^{-1} \\ &\quad = \mathcal{A}^{-1}\Big(\big(\mathcal{I} - h\mathcal{A}B(t_j)\big)^{-1} - \big(\mathcal{I} - h\mathcal{A}B^*(t_j)\big)^{-1}\Big) \\ &\quad = h\mathcal{A}^{-1}\big(\mathcal{I} - h\mathcal{A}B(t_j)\big)^{-1}\mathcal{A}(B(t_j) - B^*(t_j))\big(\mathcal{I} - h\mathcal{A}B^*(t_j)\big)^{-1}. \end{aligned}$$

By Lemma 2.4 of [14], there exists $K > 0$ such that, for $\sigma, \beta \in [0,1]$ and for $0 \le j \le N-1$, we have

$$\|\big(\mathcal{I} - h\mathcal{A}B(t_j)\big)^{-1}\|_{0\to\beta} \le K\big(1 + \tilde{L}T^\alpha\big)h^{-\beta}$$

and

$$\|(\mathcal{I} - h\mathcal{A}B^*(t_j))^{-1}\|_{\sigma\to 1} \le K(1 + \tilde{L}T^\alpha)h^{\sigma-1}.$$

Then, by using the identities (32) and (39) and the above estimates, we obtain

(40)
$$\begin{aligned} &\|r(t_{j+1}, t_j) - r^*(t_{j+1}, t_j)\|_{\sigma\to\beta} \\ &\quad = h\Big\|\mathbf{b}^T\Big(B(t_j)\big(\mathcal{I} - h\mathcal{A}B(t_j)\big)^{-1} - B^*(t_j)\big(\mathcal{I} - h\mathcal{A}B^*(t_j)\big)^{-1}\Big)(\mathbf{e}\otimes I)\Big\|_{\sigma\to\beta} \\ &\quad \le Kh\big\|\big(\mathcal{I} - h\mathcal{A}B(t_j)\big)^{-1}\big\|_{0\to\beta}\|B(t_j) - B^*(t_j)\|_{1\to 0}\big\|\big(\mathcal{I} - h\mathcal{A}B^*(t_j)\big)^{-1}\big\|_{\sigma\to 1} \\ &\quad \le K(1 + \tilde{L}T^\alpha)^2\varepsilon h^{\sigma-\beta} \end{aligned}$$

for some suitable $K > 0$.

Set, for $0 \le m \le n \le N$,

$$G_{n,m} = \prod_{j=m}^{n-1} r(t_{j+1}, t_j), \qquad G^*_{n,m} = \prod_{j=m}^{n-1} r^*(t_{j+1}, t_j),$$

and $F_{n,m} = G_{n,m} - \gamma^{n-m}$, $F^*_{n,m} = G^*_{n,m} - \gamma^{n-m}$. Then

(41)
$$\begin{aligned} G_{N,0} - G^*_{N,0} &= \sum_{j=0}^{N-1} G_{N,j+1}(G_{j+1,j} - G^*_{j+1,j})G^*_{j,0} \\ &= \sum_{j=0}^{N-1}\Big(F_{N,j+1}(G_{j+1,j} - G^*_{j+1,j})F^*_{j,0} + F_{N,j+1}(G_{j+1,j} - G^*_{j+1,j})\gamma^j \\ &\quad + \gamma^{N-(j+1)}(G_{j+1,j} - G^*_{j+1,j})F^*_{j,0} + \gamma^{N-(j+1)}(G_{j+1,j} - G^*_{j+1,j})\gamma^j\Big). \end{aligned}$$

Let η and ν as in the theorem. By (33) and (40), for $0 < 1 < j+1 < N$, we have

$$\begin{aligned} &\|F_{N,j+1}(G_{j+1,j} - G^*_{j+1,j})F^*_{j,0}\|_{\nu\to\eta} \\ &\quad \le \|F_{N,j+1}\|_{0\to\eta}\|r(t_{j+1}, t_j) - r^*(t_{j+1}, t_j)\|_{1\to 0}\|F^*_{j,0}\|_{\nu\to 1} \\ &\quad \le K(1 + LT^\alpha)^2 C^{10}_{N,j+1}(t_N - t_{j+1})^{-\eta}(\varepsilon h)C^{10}_{j,0}(t_j - t_0)^{\nu-1} \\ &\quad \le K(1 + LT^\alpha)^2 C^{20}_{N,0}(t_N - t_j)^{-\eta}\varepsilon h(t_j - t_0)^{\nu-1}. \end{aligned}$$

In a similar way, we see that

$$\begin{aligned}&\|F_{N,j+1}(G_{j+1,j}-G^*_{j+1,j})\gamma^j\|_{\nu\to\eta}\\ &\qquad\leq \quad \|F_{N,j+1}\|_{0\to\eta}\|r(t_{j+1},t_j)-r^*(t_{j+1},t_j)\|_{\nu\to 0}\gamma^j\\ &\qquad\leq \quad K(1+LT^\alpha)^2C^{10}_{N,j+1}(t_N-t_j)^{-\eta}\varepsilon h^\nu\gamma^j,\end{aligned}$$

and that

$$\begin{aligned}&\|\gamma^{N-(j+1)}(G_{j+1,j}-G^*_{j+1,j})F^*_{j,0}\|_{\nu\to\eta}\\ &\qquad\leq \quad \gamma^{N-(j+1)}\|r(t_{j+1},t_j)-r^*(t_{j+1},t_j)\|_{1\to\eta}\|F^*_{j,0}\|_{\nu\to 1}\\ &\qquad\leq \quad K(1+LT^\alpha)^2\gamma^{N-(j+1)}\varepsilon h^{1-\eta}C^{10}_{j,0}(t_j-t_0)^{\nu-1}.\end{aligned}$$

Finally, we have

$$\|\gamma^{N-(j+1)}(G_{j+1,j}-G^*_{j+1,j})\gamma^j\|_{\nu\to\eta}\leq\gamma^{N-1}K(1+LT^\alpha)^2\varepsilon h^{\nu-\eta}.$$

Hence, as $F_{j,j}=F^*_{j,j}=0$, the above estimates in (41) yield

$$\begin{aligned}&\|G_{N,0}-G^*_{N,0}\|_{\nu\to\eta}\\ &\quad\leq K(1+LT^\alpha)^2\varepsilon C^{20}_{N,0}\Bigg(h\sum_{j=1}^{N-2}(t_N-t_j)^{-\eta}(t_j-t_0)^{\nu-1}+\sum_{j=0}^{N-2}(t_N-t_j)^{-\eta}h^\nu\gamma^j\\ &\qquad\qquad+\sum_{j=1}^{N-1}\gamma^{N-(j+1)}h^{1-\eta}(t_j-t_0)^{\nu-1}+(N-1)h^{\nu-\eta}\gamma^{N-1}\Bigg).\end{aligned}$$

Now the proof of (38) is clear, by using Lemma 7.1. □

Acknowledgments

The authors have been supported by grant DGICYT PB95-705. The authors are also thankful to the anonymous referee for several valuable comments.

References

[1] P. Acquistapace, *Abstract linear nonautonomous parabolic equations: A survey*, in Differential Equations in Banach Spaces, Proceedings 1991, Ed. G. Dore, A. Favini, E. Obrecht and A. Venni, Lectures Notes in Pure and Applied Mathematics 148, Dekker (1993), pp. 1-19. MR **94f:**34113

[2] Kh. Alibekov and P. E. Sobolevskii, *Stability and convergence of difference schemes of high order of approximation for parabolic equations*, Ukr. Mat. Zh., 31 (1979), pp. 627-634; English transl., Ukrainian Math. J., 31 (1979), pp. 483–489. MR **81i:**65040

[3] H. Amann, *Quasilinear evolution equations and parabolic systems*, Trans. Amer. Math. Soc., 293 (1986), pp. 191-227. MR **87d:**35070

[4] H. Amann, *Dynamic theory of quasilinear parabolic equations - I. Abstract evolution equations*, Nonlin. Anal. Th. Meth. and Appl., 12 (1988), pp. 895-919. MR **89j:**35072

[5] H. Amann, *Highly degenerate quasilinear parabolic systems*, Ann. Scuola Norm. Sup. Pisa, Ser. IV, 18 (1991), pp. 135-166. MR **92m:**35145

[6] H. Amann, *Linear and Quasilinear Parabolic Problems. Vol. I. Abstract Theory*, Monographs in Mathematics, Vol. 89, Birkhäuser, Basel, 1995. MR **96g:**34088

[7] N. Bakaev, *On the stability of nonlinear difference problems constructed by the Runge–Kutta method*, deposited in VINITI, No. 1507-V91, 1991.

[8] J. Bell, A. Friedman and A. A. Lacey, *On solutions of quasi-linear diffusion problem from the study of soft tissue*, SIAM J. Appl. Math., 51 (1991), pp. 484-493. MR **91k:**35131

[9] J. Bergh and J. Löfström, *Interpolation Spaces. An Introduction*, Springer-Verlag, Berlin, 1976. MR **58:**2349

[10] Ph. Clément, C. J. Van Duijn and S. Li, *On a nonlinear elliptic-parabolic partial differential equation system in a two-dimensional groundwater flow problem*, SIAM J. Numer. Anal., 23 (1992), pp. 836-851. MR **93h:**35089

[11] D. S. Cohen and A. B. White, *Sharp fronts due to diffusion and viscoelastic relaxation in polymers*, SIAM J. Appl. Math., 51 (1991), pp. 472-483. MR **91k:**35058

[12] M. Crouzeix, S. Larsson and V. Thomée, *Resolvent estimates for finite elements operators in one dimension*, Math. Comp., 63 (1994), pp. 121-140. MR **95b:**65134

[13] C. González and C. Palencia, *Stability of time-stepping methods for abstract time-dependent parabolic problems*, SIAM J. Numer. Anal., 35 (1998), 979–989. MR **99b:**65072

[14] C. González and C. Palencia, *Stability of time-stepping methods for time-dependent parabolic problems: The Hölder case*, Math. Comp., 68 (1999), 73–89. MR **99c:**65108

[15] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems*, Second Edition, Springer-Verlag, Berlin, 1996. MR **97m:**65007

[16] Ch. Lubich and A. Ostermann, *Runge-Kutta approximation of quasilinear parabolic equations*, Math. Comp., 64 (1995), pp. 601-627. MR **95g:**65122

[17] Ch. Lubich and A. Ostermann, *Runge-Kutta time discretization of reaction-diffusion and Navier–Stokes equations: nonsmooth-data error estimates and applications to long-time behavior*, Appl. Numer. Math., 22 (1996), pp. 279-292. MR **97m:**65148

[18] A. Lunardi, *Global solutions of abstract quasilinear parabolic equations*, J. Diff. Eq., 58 (1985), pp. 228-242. MR **86j:**34071

[19] S. G. Krein, *Linear Differential Equations in Banach Space*, Transl. of Math. Monographs, vol. 29, Amer. Math. Society, 1972. MR **49:**7548

[20] A. Ostermann and M. Roche, *Runge-Kutta methods for partial differential equations and fractional order of convergence*, Math. Comp., 59 (1992), pp. 403-420. MR **93a:**65125

[21] C. Palencia, *Maximum norm analysis of completely discrete finite element methods for parabolic problems*, SIAM J. Numer. Anal., 33 (1996), pp. 1654-1668. MR **97e:**65099

[22] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, New York, 1983. MR **85g:**47061

[23] P. E. Sobolevskii, *Equations of parabolic type in Banach space*, Amer. Math. Soc. Transl., 49 (1966), pp. 1-62. MR **25:**5297

[24] H. Tanabe, *Equations of Evolution*, Pitman, London, 1979. MR **82g:**47032

[25] H. Triebel, *Interpolation Theory, Function Spaces, Differential Operators*, North-Holland, Amsterdam, 1978. MR **80i:**46032

Departamento de Matemática Aplicada y Computación, Universidad de Valladolid, Valladolid, Spain

E-mail address: cesareo@mac.cie.uva.es

E-mail address: palencia@mac.cie.uva.es