MATHEMATICS OF COMPUTATION
VOLUME 60, NUMBER 201
JANUARY 1993, PAGES 105–131

RUNGE-KUTTA METHODS FOR PARABOLIC EQUATIONS AND CONVOLUTION QUADRATURE

CH. LUBICH AND A. OSTERMANN

ABSTRACT. We study the approximation properties of Runge-Kutta time discretizations of linear and semilinear parabolic equations, including incompressible Navier-Stokes equations. We derive asymptotically sharp error bounds and relate the temporal order of convergence, which is generally noninteger, to spatial regularity and the type of boundary conditions. The analysis relies on an interpretation of Runge-Kutta methods as convolution quadratures. In a different context, these can be used as efficient computational methods for the approximation of convolution integrals and integral equations. They use the Laplace transform of the convolution kernel via a discrete operational calculus.

1. INTRODUCTION

The objectives of the present paper are twofold:

- It gives an error analysis of Runge-Kutta time discretizations of parabolic equations.
- It introduces a new class of computational methods for the approximation of convolution integrals, based on Runge-Kutta methods.

The reason for treating these apparently unrelated topics in a single paper is that they both rely on a discrete operational calculus of Runge-Kutta methods in such a way that separation of the second topic would only minimally reduce the length of the article while implying a substantial loss of perspective.

Section 2 introduces and studies Runge-Kutta based convolution quadrature methods approximating integrals of the form

$$\int_0^t k(t-\tau)g(\tau)\,d\tau, \qquad t>0.$$

Like the multistep methods in [16], they require (only) the Laplace transform of the kernel $k(t)$ to be known. The kernel may be weakly singular, may have components with different time scales, and need not itself be known explicitly. Such an approximation of convolution integrals is important in integral equations arising, e.g., as feedback systems in control engineering, as boundary integral equations for various types of initial-boundary value problems, and in viscoelasticity. The methods proposed here can be used for the numerical

Received by the editor November 20, 1991.
1991 *Mathematics Subject Classification.* Primary 65M15, 65M20, 65D30.
Key words and phrases. Parabolic equations, nonstationary Navier-Stokes equation, Runge-Kutta time discretization, convolution integrals, numerical quadrature.

solution of such integral equations in much the same way as those of [16]. A potential advantage of the present methods is that they can combine good stability properties with high order[1] (cf. the discussion of Runge-Kutta vs. multistep methods in the stiff ODE case [9]).

In §3 we study Runge-Kutta time discretizations of *linear* parabolic equations. Of special interest here is the way in which *spatial* regularity and boundary conditions determine the *temporal* approximation properties of the method. Such a relationship, which is not present in multistep methods, has first been observed in the pioneering thesis of Crouzeix [3]. The approximation properties of Runge-Kutta methods for inhomogeneous linear differential equations with unbounded operators have further been studied in [4, 1]. More recently, it has been shown in [19] that noninteger orders of convergence are usually attained. In the present paper, asymptotically sharp error bounds are given both in terms of the data and of the solution. For example, for the inhomogeneous heat equation, the temporal approximation order turns out to be in general higher for homogeneous than for time-dependent nonhomogeneous boundary conditions (by 1), higher for Neumann than for Dirichlet boundary conditions (by $\frac{1}{2}$), and higher when the error is measured in space in the L^2 norm rather than the maximum norm (by $\frac{1}{4}$). In the worst of these cases, the order of convergence equals the stage order plus 1, given sufficient temporal smoothness of the solution.

In §4 we give error bounds of Runge-Kutta methods applied to *semilinear* parabolic equations, under assumptions which include the incompressible Navier-Stokes equations. The results obtained are similar to those for the linear case studied in §3. As in that section, the ideas and techniques of §2 play an important role throughout our analysis. We are not aware of previous work giving convergence estimates of Runge-Kutta time discretizations of the Navier-Stokes equations (except for backward Euler and Crank-Nicolson, see [11] and references therein). A convergence analysis of multistep methods for nonlinear parabolic equations under the present assumptions has been given in [17]. Runge-Kutta time discretizations under assumptions different from ours have been studied in [2, 21] (see further [12, 13] for modified Runge-Kutta methods).

Section 5 contains proofs related to error estimates in terms of the data.

We conclude this section by introducing some terminology. A Runge-Kutta method applied to an initial value problem

$$y' = f(t, y), \qquad y(0) = y_0,$$

with a step size $h > 0$ yields at $t_n = nh$ an approximation y_n, given recursively by

$$\begin{aligned} Y_{ni} &= y_n + h \sum_{j=1}^{m} a_{ij} f(t_n + c_j h, Y_{nj}) \qquad (i = 1, \dots, m), \\ y_{n+1} &= y_n + h \sum_{j=1}^{m} b_j f(t_n + c_j h, Y_{nj}). \end{aligned} \tag{1.1}$$

[1] For example, the l^2 circle condition theorem (Theorem 2.3 in [5]) for nonlinear convolution equations can be shown to hold for algebraically stable Runge-Kutta methods. There is no order barrier for such methods, in contrast to A-stable multistep methods.

The Runge-Kutta method has *order* p if the error of the method after one step, when applied to ordinary differential equations with sufficiently differentiable right-hand side, satisfies

$$y_1 - y(h) = O(h^{p+1}) \quad \text{as } h \to 0. \tag{1.2}$$

The method has *stage order* q if the error of the internal stages is, for all $i = 1, \dots, m$,

$$Y_{0i} - y(c_i h) = O(h^{q+1}), \tag{1.3}$$

or equivalently, if

$$\sum_{j=1}^{m} a_{ij} c_j^{k-1} = \frac{c_i^k}{k} \quad \text{for } k = 1, \dots, q. \tag{1.4}$$

In the following we use the notation

$$\mathscr{A} = (a_{ij})_{i,j=1}^{m}, \qquad b^T = (b_1, \dots, b_m), \qquad \mathbb{1} = (1, \dots, 1)^T. \tag{1.5}$$

A Runge-Kutta method is called $A(\theta)$-*stable* if $I - z\mathscr{A}$ is nonsingular in the sector $|\arg(-z)| \le \theta$, and if the *stability function*

$$R(z) = 1 + z b^T (I - z\mathscr{A})^{-1} \mathbb{1} \tag{1.6}$$

is bounded by

$$|R(z)| \le 1 \quad \text{for } |\arg(-z)| \le \theta. \tag{1.7}$$

The method is called *strongly* $A(\theta)$-*stable* if it is $A(\theta)$-stable and in addition has an invertible Runge-Kutta matrix $\mathscr{A}$, and the limit of the stability function at infinity,

$$R(\infty) = 1 - b^T \mathscr{A}^{-1} \mathbb{1}, \tag{1.8}$$

has absolute value strictly smaller than 1.[2] Well-known examples are the Radau IIA methods (see [9]). These methods have further

$$b_i = a_{mi} \qquad (i = 1, \dots, m), \tag{1.9}$$

which implies $R(\infty) = 0$.

2. Runge-Kutta-based convolution quadrature methods

We are interested in approximating the convolution

$$u(t) = \int_0^t k(t-\tau) g(\tau)\, d\tau, \qquad t > 0. \tag{2.1}$$

In many applications, it is the Laplace transform $K(s)$ of the kernel (the *transfer function* in control engineering terminology) rather than the kernel itself which is known a priori, or is of a simple form. The quadrature methods to be described below use only $K(s)$ and are given by a discrete block convolution, which allows for the use of fast Fourier transforms for an efficient approximation of (2.1) for

[2] In the literature, there exist different variants of strong $A(\theta)$-stability [3, 4, 2]. The above definition is slightly stronger than all these and has been chosen to allow for a briefer statement of our results.

many values of t. We assume that the following *sectorial condition* is satisfied by $K(s)$ (cf. [16]):

(2.2) $K(s)$ is analytic in a sector $|\arg(s-c)| < \pi-\phi$ with $\phi < \frac{\pi}{2}$, $c \in \mathbf{R}$, and is bounded there by $|K(s)| \le M \cdot |s|^{-\mu}$ for some $\mu > 0$.

Then the kernel $k(t)$ can be represented by the Laplace inversion formula as

$$k(t) = \frac{1}{2\pi i} \int_\Gamma K(\lambda) e^{\lambda t}\, d\lambda, \tag{2.3}$$

where Γ is a contour parallel to the boundary of the sector, oriented with increasing imaginary part. It follows from (2.2) that $k(t)$ is analytic in some sector containing the positive real axis, exponentially bounded for $t \to \infty$, and with $k(t) = O(t^{\mu-1})$ as $t \to 0$, so that $k(t)$ may be weakly singular. Conversely, every such kernel has a Laplace transform satisfying (2.2) for some $\phi < \frac{\pi}{2}$.

Inserting (2.3) into (2.1) and interchanging integrals, we get

$$u(t) = \frac{1}{2\pi i} \int_\Gamma K(\lambda) y_\lambda(t)\, d\lambda, \tag{2.4}$$

where $y_\lambda(t) = \int_0^t e^{\lambda(t-\tau)} g(\tau)\, d\tau$ is the solution of the initial value problem

$$y' = \lambda y + g, \qquad y(0) = 0. \tag{2.5}$$

The basic idea now is to discretize this differential equation by a Runge-Kutta method, insert the approximate solution into (2.4), and use Cauchy's integral formula to simplify the resulting expression. We then get the following description of the convolution quadrature method.

Proposition 2.1. *Consider an m-stage Runge-Kutta method satisfying* (1.9), *and suppose that all eigenvalues of the matrix $\mathscr{A}$ are in the open sector $|\arg\lambda| < \pi - \phi$, with ϕ of* (2.2). *When $y_\lambda(t)$ in* (2.4) *is replaced by the Runge-Kutta approximation* (1.1) *of* (2.5), *then one gets at $t = (n+1)h$ the approximation*

$$u_{n+1} = h \sum_{j=0}^{m} \sum_{\nu=0}^{n} w_{n-\nu,j}\, g(t_\nu + c_j h), \tag{2.6}$$

where the quadrature weights $(w_{n1}, \dots, w_{nm})$ form the last row of the $m \times m$ matrix W_n given as the nth coefficient of the generating function

$$h \sum_{n=0}^{\infty} W_n \zeta^n = K\left(\frac{\Delta(\zeta)}{h}\right), \tag{2.7}$$

with

$$\Delta(\zeta) = \left(\mathscr{A} + \frac{\zeta}{1-\zeta} \mathbf{1} b^T\right)^{-1}. \tag{2.8}$$

As will be described below, the quadrature weights can be computed efficiently to arbitrary precision using fast Fourier transform techniques.

Proof. We consider the generating functions

$$y(\zeta) = \sum_{n=0}^{\infty} y_n \zeta^n, \quad Y(\zeta) = \sum_{n=0}^{\infty} Y_n \zeta^n, \quad G(\zeta) = \sum_{n=0}^{\infty} G_n \zeta^n,$$

etc., with $Y_n = (Y_{nj})_{j=1}^m$ and $G_n = (g(t_n + c_j h))_{j=1}^m$. From (1.1) with $f(t, y) = \lambda y + g(t)$ we have

$$Y(\zeta) = \mathbb{1} y(\zeta) + h\lambda \mathscr{Q} Y(\zeta) + h\mathscr{Q} G(\zeta),$$
$$(\zeta^{-1} - 1) y(\zeta) = h\lambda b^T Y(\zeta) + h b^T G(\zeta).$$

Inserting $y(\zeta)$ from the second equation into the first one gives

$$\left(I - h\lambda \mathbb{1} b^T \frac{\zeta}{1-\zeta} - h\lambda \mathscr{Q}\right) Y(\zeta) = h\left(\mathbb{1} b^T \frac{\zeta}{1-\zeta} + \mathscr{Q}\right) G(\zeta),$$

that is,

$$Y(\zeta) = \left(\frac{\Delta(\zeta)}{h} - \lambda\right)^{-1} G(\zeta). \tag{2.9}$$

If (1.9) is satisfied, then $y_{n+1} = Y_{nm}$, and hence $y(\zeta)$ is the last component of $Y(\zeta)$. Otherwise, we get

$$y(\zeta) = \frac{\zeta}{1-\zeta} h b^T \frac{\Delta(\zeta)}{h} \left(\frac{\Delta(\zeta)}{h} - \lambda\right)^{-1} G(\zeta). \tag{2.10}$$

Substituting the numerical solution into (2.4), we get for the approximation the generating function

$$U(\zeta) = \frac{1}{2\pi i} \int_\Gamma K(\lambda) \left(\frac{\Delta(\zeta)}{h} - \lambda\right)^{-1} G(\zeta)\, d\lambda,$$

which by Cauchy's integral formula reduces to

$$U(\zeta) = K\left(\frac{\Delta(\zeta)}{h}\right) G(\zeta). \tag{2.11}$$

It follows from (2.7) that

$$U_n = (U_{ni})_{i=1}^m = h \sum_{\nu=0}^{n} W_{n-\nu} G_\nu. \tag{2.12}$$

If (1.9) holds, then u_{n+1} is the last component of U_n, and so u_{n+1} is given by (2.6). In general, we get from (2.10)

$$u(\zeta) = \frac{\zeta}{1-\zeta} h b^T \frac{\Delta(\zeta)}{h} K\left(\frac{\Delta(\zeta)}{h}\right) G(\zeta). \quad \square \tag{2.13}$$

The above derivation indicates how to obtain error estimates for the approximation u_n of $u(t_n)$: One studies the error of the Runge-Kutta method applied to the linear differential equation (2.5) with λ varying on the unbounded contour Γ, and then integrates along Γ over the error multiplied by $K(\lambda)$. This leads to the following result.

Theorem 2.2 (Convergence of convolution quadrature methods). *Assume* (2.2), *and consider a Runge-Kutta method of order* p *and stage order* q *which is strongly* $A(\theta)$*-stable with* $\theta > \phi$. *Then the error of the convolution quadrature*

(2.6) *(or that given by* (2.13)*, if condition* (1.9) *is not satisfied) is bounded for* $h \le h_0$ *(sufficiently small) by*

$$|u_n - u(t_n)| \le Ch^p \cdot \sum_{l=0}^{q}(1 + t_n^{\mu+l-p}) \cdot |g^{(l)}(0)| + C(h^p + h^{q+1+\mu}|\log h|) \cdot \left(\sum_{l=q+1}^{p-1} |g^{(l)}(0)| + \max_{0\le\tau\le t_n} |g^{(p)}(\tau)| \right). \tag{2.14}$$

If $p \ge q+1$, *the error of the internal stages* U_{ni} *given by* (2.12) *satisfies for* $i = 1, \dots, m$

$$|U_{ni} - u(t_n + c_i h)| \le Ch^{q+1} \cdot \left(\sum_{l=0}^{q}(1 + t_n^{\mu+l-q-1}) \cdot |g^{(l)}(0)| + \max_{0\le\tau\le t_n} |g^{(q+1)}(\tau)| \right).$$

The constants C *and* h_0 *depend only on the Runge-Kutta method, on the constants in* (2.2), *and on the length of the time interval. In particular, they are independent of* n *and* h *with* $nh \le T$, *and of* $g \in C^p[0, T]$. *If* $c < 0$ *in* (2.2), *then the bounds hold uniformly over the whole half-line* $[0, \infty)$, *and* h_0 *can be chosen arbitrarily.*

The proof of (2.14) is deferred to §5. We omit the proof of the error estimate for the internal stages.

The error is thus $O(h^p) + O(h^{q+1+\mu}|\log h|)$ on any finite interval bounded away from 0. One may thus have a noninteger order of convergence. If one wishes to get high convergence order also near 0, then one might add to u_n a correction term of the form $h\sum_{j=1}^{m} \tilde{w}_{nj} g(c_j h)$, where the weights $\tilde{w}_{nj}$ are constructed such that the quadrature becomes exact for polynomials of degree q. Then the error is $O(h^p) + O(h^{q+1+\mu}|\log h|)$ uniformly on every finite interval $[0, T]$ (cf. Corollary 3.2 in [16] for an analogous construction in the multistep case).

For the m-stage Radau IIA method the above assumptions are satisfied with $p = 2m - 1$, $q = m$, and with $\theta = \frac{\pi}{2}$ (see [9]).

Implementation and numerical example. To compute the quadrature weight matrices W_n, one approximates the Cauchy integral

$$hW_n = \frac{1}{2\pi i} \int_{|\zeta|=\rho} \zeta^{-n-1} K(\Delta(\zeta)/h)\, d\zeta$$

by the trapezoidal rule:

$$hW_n \approx \frac{\rho^{-n}}{L} \sum_{l=0}^{L-1} K(\Delta(\zeta_l)/h) e^{-2\pi i n l/L},$$

with $\zeta_l = \rho e^{2\pi i l/L}$. Assuming that the values of K are computed with precision ε, one gets as in [16, §7], that the error in hW_n $(0 \le n \le N)$ is $O(\sqrt{\varepsilon})$ if $L = N$ and $\rho^N = \sqrt{\varepsilon}$, and the error is $O(\varepsilon)$ if $L \ge N|\log\varepsilon|$ and $\rho = e^{-\gamma h}$, with $\gamma > c$ of (2.2). Using fast Fourier transforms, the weights are computed in $O(L \log L)$ operations.

We have used the 1-, 2-, and 3-stage Radau IIA methods to approximate the convolution integral

$$\int_0^t \frac{1}{\sqrt{\pi(t-\tau)}} e^\tau \, d\tau = e^t \cdot \operatorname{erf}\sqrt{t}.$$

The Laplace transform of the kernel $k(t) = (\pi t)^{-1/2}$ is $K(s) = s^{-1/2}$. For the 1- and 2-stage methods, Theorem 2.2 predicts the full convergence order 1 and 3, respectively, while for the 3-stage method the error is bounded by $O(h^{4.5}|\log h|)$. These asymptotic bounds are confirmed by numerical experiments. At $t = 4$ we obtained the following results.

- 2-stage Radau IIA method:

h	relative error	relative error/ h^3
1	$6.4_{10} - 3$	$6.4_{10} - 3$
1/2	$9.6_{10} - 4$	$7.7_{10} - 3$
1/4	$1.4_{10} - 4$	$8.6_{10} - 3$
1/8	$1.8_{10} - 5$	$9.4_{10} - 3$
1/16	$2.4_{10} - 6$	$9.9_{10} - 3$

- 3-stage Radau IIA method:

h	relative error	relative error/ $h^{4.5}$
1	$1.4_{10} - 4$	$1.4_{10} - 4$
1/2	$8.4_{10} - 6$	$1.9_{10} - 4$
1/4	$4.5_{10} - 7$	$2.3_{10} - 4$
1/8	$2.3_{10} - 8$	$2.6_{10} - 4$
1/16	$1.1_{10} - 9$	$3.0_{10} - 4$

For stability investigations, bounds for the coefficients W_n of (2.7) are all-important. The following result has a multistep analogue in Lemma 7 of [17], or Lemma V.7.11 in [9].

Lemma 2.3. *Under the assumptions of Theorem* 2.2 (*now also when* $\mu = 0$ *in* (2.2)), *the coefficients of* $K(\Delta(\zeta)/h) = h\sum_{n=0}^{\infty} W_n \zeta^n$ *satisfy for* $h \le h_0$ (*sufficiently small*)

$$\|W_n\| \le C \cdot (nh)^{\mu-1} \cdot e^{\gamma nh} \quad \text{for } n \ge 1,$$

and for $n = 0$ *the same bound holds as for* $n = 1$. *The constants* C, γ, *and* h_0 *depend only on the method and the constants in* (2.2). *They can be chosen*

independent of $0 \le \mu \le \text{Const}$. *Moreover, if* $c < 0$, *then also* $\gamma < 0$, *and the result holds for arbitrary* h_0.

The proof uses Cauchy's integral formula

$$K\left(\frac{\Delta(\zeta)}{h}\right) = \frac{1}{2\pi i}\int_\Gamma K(\lambda)\left(\frac{\Delta(\zeta)}{h} - \lambda\right)^{-1} d\lambda, \tag{2.15}$$

the following lemma, and bounds for the stability function $R(z)$.

Lemma 2.4. *Under the foregoing assumptions we have*

$$(\Delta(\zeta) - z)^{-1} = \mathscr{Q}(I - z\mathscr{Q})^{-1} + \sum_{n=1}^{\infty} R(z)^{n-1}(I - z\mathscr{Q})^{-1}\mathbb{1}b^T(I - z\mathscr{Q})^{-1}\zeta^n.$$

Proof. By (2.8) we have

$$(\Delta(\zeta) - z)^{-1} = \left(\mathscr{Q} + \frac{\zeta}{1-\zeta}\mathbb{1}b^T\right)\cdot\left(I - z\left(\mathscr{Q} + \frac{\zeta}{1-\zeta}\mathbb{1}b^T\right)\right)^{-1}.$$

With $E = I + (I - z\mathscr{Q})^{-1}z\mathbb{1}b^T$ we get

$$(\Delta(\zeta) - z)^{-1} = \left(\mathscr{Q} + \frac{\zeta}{1-\zeta}\mathbb{1}b^T\right)\cdot(1-\zeta)(I - \zeta E)^{-1}\cdot(I - z\mathscr{Q})^{-1},$$

and we note that

$$(1-\zeta)(I-\zeta E)^{-1} = (1-\zeta)\sum_{n=0}^{\infty} E^n\zeta^n = I + \sum_{n=1}^{\infty}(E^n - E^{n-1})\zeta^n.$$

We have

$$E^2 = I + 2(I - z\mathscr{Q})^{-1}z\mathbb{1}b^T + (I - z\mathscr{Q})^{-1}z\mathbb{1}\cdot[b^T(I - z\mathscr{Q})^{-1}z\mathbb{1}]\cdot b^T.$$

By (1.6), the term in square brackets reduces to $R(z) - 1$, and hence

$$E^2 = I + (I - z\mathscr{Q})^{-1}z\mathbb{1}b^T(1 + R(z)).$$

By induction, we get

$$E^n = I + (I - z\mathscr{Q})^{-1}z\mathbb{1}b^T\cdot(1 + R(z) + \cdots + R(z)^{n-1}).$$

This gives further

$$\mathbb{1}b^TE^n = \mathbb{1}b^T + \mathbb{1}(R(z) - 1)b^T(1 + \cdots + R(z)^{n-1}) = R(z)^n\mathbb{1}b^T$$

and

$$\mathscr{Q}(E^n - E^{n-1}) = R(z)^{n-1}\cdot z\mathscr{Q}(I - z\mathscr{Q})^{-1}\mathbb{1}b^T.$$

Inserting this into the above expressions for $(\Delta(\zeta) - z)^{-1}$ gives the stated result. □

Proof of Lemma 2.3. With Lemma 2.4 available, the proof is now similar to that of Lemma 7 in [17]. We note that there exists $0 < \rho < 1$ such that

$$|R(z)| \le \max(|e^{z/2}|, \rho) \quad \text{for } |\arg(-z)| \le \theta' < \theta, \tag{2.16}$$

and

$$|R(z)| \le e^{2|z|} \quad \text{for } |z| \le r, \ r \text{ sufficiently small}. \tag{2.17}$$

From (2.15) and Lemma 2.4 we have for $n \geq 1$

$$W_n = \frac{1}{2\pi i} \int_\Gamma K(\lambda) \cdot R(h\lambda)^{n-1} (I - h\lambda \mathscr{A})^{-1} \mathbb{1} b^T (I - h\lambda \mathscr{A})^{-1} \cdot d\lambda .$$

For $nh \leq 1$ we substitute $w = nh\lambda$ and replace the resulting integration contour $nh\Gamma$ by an equivalent contour Γ_1 which is independent of nh, has a positive distance to the origin, and, apart from a compact subset near the origin, is again contained in a sector $|\arg(-w)| \leq \theta' < \theta$. We then have, using (2.2) in the second inequality,

$$\begin{aligned} \|W_n\| &\leq \frac{1}{2\pi} \int_{\Gamma_1} \left| K\left(\frac{w}{nh}\right) \right| \cdot \left| R\left(\frac{w}{n}\right)^{n-1} \right| \cdot \frac{C}{(1+|w|/n)^2} \cdot \frac{|dw|}{nh} \\ &\leq (nh)^{\mu-1} \frac{1}{2\pi} \int_{\Gamma_1} M|w|^{-\mu} \cdot \left| R\left(\frac{w}{n}\right)^{n-1} \right| \cdot \frac{C}{(1+|w|/n)^2} \cdot |dw| . \end{aligned}$$

The above bounds of $R(z)$ show that the remaining integral is bounded independently of n. This gives the result for $nh \leq 1$. The result for $nh \geq 1$ follows directly with a contour Γ which is parallel to the boundary of the sector, using the bounds (2.16) and (2.17). □

We give a further lemma similar to Lemma 2.4. It is related to the explicit representation of the coefficients in formula (2.13).

Lemma 2.5. *If the Runge-Kutta matrix $\mathscr{A} = (a_{ij})$ is invertible, then*

$$\frac{b^T \Delta(\zeta)}{1-\zeta} = \sum_{n=0}^{\infty} R(\infty)^n \cdot b^T \mathscr{A}^{-1} \cdot \zeta^n .$$

Proof. We write

$$\frac{b^T \Delta(\zeta)}{1-\zeta} = b^T \mathscr{A}^{-1} (I - \zeta(I - \mathbb{1} b^T \mathscr{A}^{-1}))^{-1} = b^T \mathscr{A}^{-1} \sum_{n=0}^{\infty} (I - \mathbb{1} b^T \mathscr{A}^{-1})^n \zeta^n .$$

The formula

$$b^T \mathscr{A}^{-1} (I - \mathbb{1} b^T \mathscr{A}^{-1}) = R(\infty) \cdot b^T \mathscr{A}^{-1}$$

then implies the desired result. □

In the present section we have tacitly assumed that $K(s)$ is a complex-valued function. Obviously, all the results remain valid if $K(s)$ is operator-valued from one Banach space into another, and $|\cdot|$ is interpreted as a suitable norm. This generalization will actually be used in the following sections.

3. Linear parabolic equations

In this section we derive error bounds for Runge-Kutta time discretizations of linear parabolic equations. Our analysis will be based on an abstract formulation of parabolic equations within the framework of analytic semigroups.

For this, let X be a Banach space. We denote by $\|\cdot\|$ its norm as well as the induced operator norm. We consider a linear initial value problem on X:

$$y' + Ay = g(t), \qquad y(0) = y_0 . \tag{3.1}$$

Here, $-A$ is the generator of an analytic semigroup, i.e., A is a densely defined closed operator on X satisfying

$$\|(s+A)^{-1}\| \le \frac{M}{1+|s|} \quad \text{for } |\arg(s-c)| \le \pi - \phi, \text{ with some } \phi < \frac{\pi}{2}. \tag{3.2}$$

The real number c may be arbitrary, but it simplifies the presentation to assume (3.2) with $c = 0$. Then fractional powers of A are well defined, and the results below are valid uniformly over the whole half-line. All results of this section remain valid for arbitrary $c > 0$ if fractional powers of A are replaced by those of $A + cI$, and if the constants are allowed to grow exponentially with the length of the time interval.

It is well known [20, §7.3] that if $P(x, \partial)$ is a strongly elliptic differential operator, then the partial differential equation

$$\frac{\partial u}{\partial t} + P(x, \partial)u = g(x, t), \qquad x \in \Omega, \ t > 0,$$

with appropriate initial and boundary conditions, can be interpreted in the form (3.1)–(3.2) on $X = L^p(\Omega)$, for $1 \le p < \infty$. Also, finite element or finite difference spatial discretizations of such equations often satisfy (3.2), uniformly in the meshwidth Δx.[3]

Convergence results for one-step time discretizations of equations (3.1) (under varying assumptions on the unbounded operator A) were derived in [3, 4, 1, 15], and more recently in [19]. Closely related to our results are those of [1], where it is assumed that $-A$ generates (only) a C_0-semigroup. The stronger assumption (3.2) of an *analytic* semigroup permits us to obtain stronger estimates. Time discretization of the homogeneous equation ((3.1) with $g \equiv 0$) has been studied under the assumption (3.2) in [15]. An important stability estimate of [15] is

$$\|R(-hA)^n\| \le \text{Const}, \qquad n \ge 0, \tag{3.3}$$

valid for the stability function $R(z)$ of strongly $A(\theta)$-stable methods with $\theta > \phi$ (see also Theorem 3.5 in [18]). This bound will be used repeatedly in the following.

We first give a numerical analogue of the "variation of constants" formula

$$y(t) = e^{-tA}y_0 + \int_0^t e^{-(t-\tau)A}g(\tau)\,d\tau. \tag{3.4}$$

We recall that the analytic semigroup $\{e^{-tA}\}$ is *defined* by formula (2.3) with $K(s) = (s+A)^{-1}$. The following lemma states that the Runge-Kutta solution of (3.1) is identical to the convolution quadrature discretization of (3.4). This establishes the connection to the previous section and will serve as a basic tool in the sequel.

Lemma 3.1 (Discrete variation of constants formula). *Consider a Runge-Kutta method for the solution of* (3.1) *and suppose that all eigenvalues of* $\mathscr{C}$ *are in the*

[3]See, e.g., [22] for the case $p = 2$. We do not know of references showing (3.2) in L^p, $p \ne 2$, for finite element discretizations. When taking the maximum norm, the constant M may grow like $|\log \Delta x|$ in two space dimensions (cf. [22, Chapter 5]): There, Theorem 1 is actually valid in an open sector containing the positive half-axis $t > 0$, implying (3.2) with M growing logarithmically with Δx. The constants appearing in our error bounds below are proportional to M.

sector $|\arg\lambda| < \pi - \phi$, *with* ϕ *of* (3.2). *Then the numerical solution* y_n *and the internal stages* $Y_n = (Y_{ni})_{i=1}^m$ *are given by*

$$Y_n = V_n y_0 + h \sum_{\nu=0}^{n} W_{n-\nu} G_\nu ,$$

$$y_{n+1} = R(-hA)^{n+1} y_0 + b^T \mathscr{A}^{-1} \sum_{k=0}^{n} R(\infty)^{n-k} h \sum_{\nu=0}^{k} W_{k-\nu} G_\nu ,$$

with $G_n = (g(t_n + c_i h))_{i=1}^m$. *Here,* W_n *are the coefficients of* (2.7) *with* $K(s) = (s+A)^{-1}$, *viz.,* $W_n = \Psi_n(-hA)$, *with* $\Psi_n(z)$ *denoting the nth coefficient on the right-hand side of Lemma* 2.4, *and* $V_n = \Phi_n(-hA)$, *with* $\Phi_n(z) = (I - z\mathscr{A})^{-1}\mathbb{1} \cdot R(z)^n$.

Proof. Consider first the homogeneous equation (3.1) with $g \equiv 0$. Then $y_n = R(-hA)^n y_0$, and $Y_n = (I + \mathscr{A} \otimes hA)^{-1}(\mathbb{1} \otimes I) y_n = V_n y_0$ follows from the definition of the method.

Next, for zero initial values the result for the internal stages follows from the first part of the proof of Proposition 2.1 (down to formula (2.9)), with λ formally replaced by $-A$. Lemma 2.4 (with $-hA$ substituted for z) gives the above expressions of W_n. The stated formula for y_{n+1} follows from (2.10) with Lemma 2.5.

Linearity finally gives the general result. □

With the above interpretation of the Runge-Kutta solution, we can apply Theorem 2.2 to get an asymptotic error bound in terms of the data. This almost gives the following result.

Theorem 3.2 (Error estimates in terms of the data). *For an initial value problem* (3.1)–(3.2), *consider a Runge-Kutta method of stage order* q *and order* $p \geq q+1$, *which is strongly* $A(\theta)$*-stable with* $\theta > \phi$. *Let the real numbers* $\alpha \geq 0$ *and* $\beta > 0$ *be such that* $y_0 \in D(A^\alpha)$ *and* $A^{-1} g^{(l)}(t) \in D(A^\beta)$ *for* $0 \leq l \leq p$ *and all* $t \geq 0$. *Then the error is bounded for* $h \leq h_0$ *by*

$$\begin{aligned} \|y_n - y(t_n)\| \leq\ & C(1 + t_n^{\alpha-p}) \cdot h^p \cdot \|A^\alpha y_0\| \\ & + Ch^p \cdot \sum_{l=0}^{q} (1 + t_n^{\beta-p+l}) \cdot \|A^{\beta-1} g^{(l)}(0)\| \\ & + C(h^p + h^{q+1+\beta} |\log h|) \\ & \cdot \left(\sum_{l=q+1}^{p-1} \|A^{\beta-1} g^{(l)}(0)\| + \max_{0 \leq \tau \leq t_n} \|A^{\beta-1} g^{(p)}(\tau)\| \right) . \end{aligned}$$

Remark. In the above formula, β might also depend on the order l of the time derivative. Theorem 3.2 shows mainly the influence of spatially rough data on the temporal approximation. We have not taken compatibility conditions into account, which make the solution smooth from $t = 0$ onwards (see Theorem 3 in [1] for a result of that type in the context of C_0-semigroups). The formulation of Theorem 3.2 has been chosen such that the result is also applicable to nonhomogeneous boundary conditions (see the discussion following Theorem 3.3 below).

Concerning the proof of Theorem 3.2, we note that for $\alpha = 0$ the result follows directly from Theorem 1.1 in [15] (for the homogeneous equation), and for $0 < \beta \leq 1$ it follows with Theorem 2.2 of the present paper (for zero initial conditions), applied with $K(s) = (s + A)^{-1}A^{1-\beta}$, which is bounded by (see, e.g., [10, §1.4])

$$\|(s + A)^{-1}A^{1-\beta}\| \leq C \cdot |s|^{-\beta}, \qquad |\arg s| \leq \pi - \phi.$$

The case of $\beta > 1$ requires a modification in the proof of Theorem 2.2 and is given in §5.

The following result improves slightly on the bounds of [19], where the same fractional order of convergence was obtained under stronger regularity assumptions.

Theorem 3.3 (Error estimates in terms of the solution). *For an initial value problem* (3.1)–(3.2), *consider a Runge-Kutta method of stage order* q *and order* $p \geq q + 1$, *which is strongly* $A(\theta)$*-stable with* $\theta > \phi$. *Let* β *be a nonnegative real number such that* $q + 1 + \beta < p$ *and* $y^{(q+1+l)}(t) \in D(A^{\beta-l})$ *for* $0 \leq l \leq \beta$ *and all* t. *With integer* $\overline{p}$ *defined by* $\overline{p} - 1 \leq q + 1 + \beta < \overline{p}$, *the error is bounded by*

$$\|y_n - y(t_n)\| \leq C \cdot h^{q+1+\beta} \cdot \sum_{l=0}^{\overline{p}-q-2} \max_{0 \leq t \leq t_n} \|A^{\beta-l} y^{(q+1+l)}(t)\|$$
$$+ C \cdot h^{\overline{p}} \cdot \left(\|y^{(\overline{p})}(0)\| + \int_0^{t_n} \|y^{(\overline{p}+1)}(t)\| \, dt \right).$$

For $q + 1 + \beta = p$ *the bound remains valid with* $\overline{p} = p$ *and with an additional factor* $\log n$ *in front of the first term, or alternatively with* $\|A^{\beta-l} y^{(q+1+l)}(0)\| + \int_0^{t_n} \|A^{\beta-l} y^{(q+2+l)}(t)\| \, dt$ *instead of the above maximum.*

Consider the heat equation, where $A = -\Delta$ on a smooth region Ω, equipped with appropriate boundary conditions, is taken as an unbounded operator on $L^2(\Omega)$. Given sufficient smoothness of the solution, it is in general the condition $y^{(q+1)}(t) \in D(A^{\beta})$ which restricts the actual order of convergence. The attainable exponent β depends on the type of boundary conditions. Its determination relies on the characterization of the domains of fractional powers of second-order elliptic operators given in [7] (and [8]).

(i) *Homogeneous Dirichlet boundary conditions.* In the case of a smooth solution, the homogeneous boundary conditions are also satisfied by $y^{(q+1)}(t)$, and one has $y^{(q+1)}(t) \in D(A^{5/4-\varepsilon})$ for arbitrary $\varepsilon > 0$, but it is not contained in the domain of higher powers of A, unless unnatural boundary conditions $\Delta y = 0$ on $\partial\Omega$ are satisfied. Theorem 3.3 is thus applicable with $\beta = \frac{5}{4} - \varepsilon$.

(ii) *Homogeneous Neumann boundary conditions.* Here one can take $\beta = \frac{7}{4} - \varepsilon$ for any $\varepsilon > 0$.

(iii) *Nonhomogeneous Dirichlet boundary conditions.* A standard technique in the analysis of nonhomogeneous boundary conditions consists in subtracting the effect of the boundary term and then considering the corresponding inhomogeneous equation with homogeneous boundary conditions. Time discretization by a Runge-Kutta method is however not invariant under this transformation of the problem, because the resulting inhomogeneity involves a time derivative

(unless the boundary values are constant in time). An alternative way is to restate the heat equation with boundary data $b(t)$ on $\partial\Omega$ in the form considered in [14]:

$$A^{-1}y' + y = Gb(t), \qquad y(0) = y_0, \tag{3.5}$$

where $A = -\Delta$, equipped with *homogeneous* Dirichlet boundary conditions, and G is the solution operator of the nonhomogeneous boundary value problem

$$\Delta v = 0 \quad \text{in } \Omega, \qquad v = b \quad \text{on } \partial\Omega.$$

The solution of (3.5) is given by the semigroup formula

$$y(t) = e^{-tA}y_0 + \int_0^t Ae^{-(t-\tau)A}Gb(\tau)\,d\tau. \tag{3.6}$$

The standard Runge-Kutta time discretization of the heat equation with nonhomogeneous Dirichlet boundary conditions yields exactly the Runge-Kutta convolution quadrature discretization of formula (3.6) according to Lemma 3.1. Since $y(t)$ does not satisfy the homogeneous boundary conditions, the solution derivatives are in general not in the domain of A. Assuming sufficient smoothness, one has $y^{(q+1)}(t) \in D(A^{1/4-\varepsilon})$ for $\varepsilon > 0$, but in general it is not in $D(A^{1/4})$. So we can use Theorem 3.3 with $\beta = \frac{1}{4} - \varepsilon$.

(iv) *Nonhomogeneous Neumann boundary conditions.* By a similar reasoning, one gets here $\beta = \frac{3}{4} - \varepsilon$ for $\varepsilon > 0$.

The order reduction is thus considerably more severe for nonhomogeneous than for homogeneous boundary conditions, and more severe for Dirichlet than for Neumann boundary conditions. Numerical experiments (cf. [23, 19]) have confirmed the convergence rates given by Theorem 3.3 with the above values of β. We remark that the same exponents β are obtained for general second-order, uniformly strongly elliptic differential operators on smooth domains when they are considered as unbounded operators on $L^2(\Omega)$ [7]. However, when taken as operators on $L^r(\Omega)$ $(1 < r < \infty)$, then β varies with r: Theorem 3.3 holds with $\beta < \frac{1}{2r}$ and $\beta < \frac{1}{2} + \frac{1}{2r}$ in the case of time-dependent nonhomogeneous Dirichlet and Neumann boundary conditions, respectively, and with a β higher by 1 for the corresponding homogeneous boundary conditions. (See [8, Théorème 8.1′], and also the discussion and numerical experiments for the one-dimensional heat equation in [19].)

Proof of Theorem 3.3. By possibly reducing p, we may assume $p = \overline{p}$ in the following.

(a) We consider first the defect when the exact solution values are inserted into the Runge-Kutta scheme:

$$d_{ni} = y(t_n + c_i h) - y(t_n) - h\sum_{j=1}^{m} a_{ij}y'(t_n + c_j h),$$

$$d_{n+1} = y(t_{n+1}) - y(t_n) - h\sum_{j=1}^{m} b_j y'(t_n + c_j h).$$

By Taylor expansion and the definition of the stage order q, we have (assuming here for simplicity that all $c_i \in [0, 1]$)

$$d_{ni} = \sum_{k=q+1}^{p} \delta_i^{(k)} \cdot h^k \cdot y^{(k)}(t_n) + h^p \cdot \int_{t_n}^{t_{n+1}} \kappa_i\left(\frac{t - t_n}{h}\right) y^{(p+1)}(t)\,dt,$$

where

$$\delta_i^{(k)} = \left(\sum_{j=1}^{m} a_{ij} c_j^{k-1} - \frac{c_i^k}{k} \right) / (k-1)!,$$

and where κ_i is a bounded Peano kernel. Similarly, we have

$$d_{n+1} = h^p \int_{t_n}^{t_{n+1}} \kappa \left(\frac{t-t_n}{h} \right) y^{(p+1)}(t)\, dt.$$

(b) For the error $e_n = y_n - y(t_n)$ we thus get the recursion

$$e_{n+1} = R(-hA)e_n - hb^T \otimes A(I + h\mathscr{A} \otimes A)^{-1} D_n + d_{n+1},$$

with $D_n = (d_{ni})_{i=1}^m$. By (3.3), the powers of $R(-hA)$ are uniformly bounded. This gives immediately that the contribution from the integral remainder terms in d_{ni} and d_{n+1} is $O(h^p)$ as required. It remains to study the expressions

$$\begin{aligned} e_{n+1}^{(k)} &= h^k \cdot \sum_{\nu=0}^{n} r_{n-\nu}^{(k)}(-hA) \cdot y^{(k)}(t_\nu) \\ &= h^{q+1+\beta} \cdot \sum_{\nu=0}^{n} r_{n-\nu}^{(k)}(-hA)(hA)^{k-q-1-\beta} \cdot A^{q+1+\beta-k} y^{(k)}(t_\nu), \end{aligned}$$

where

$$r_n^{(k)}(z) = R(z)^n \cdot zb^T (I - z\mathscr{A})^{-1} \delta^{(k)}, \tag{3.7}$$

with $\delta^{(k)} = (\delta_i^{(k)})_{i=1}^m$.

(c) For $k \le p-1 \le q+1+\beta$ we have, provided that $\alpha := q+1+\beta-k > 0$,

$$r_n^{(k)}(-hA) \cdot (hA)^{-\alpha} = \frac{1}{2\pi i} \int_\Gamma (z+hA)^{-1} \cdot R(z)^n \cdot zb^T (I - z\mathscr{A})^{-1} \delta^{(k)} \cdot (-z)^{-\alpha} \cdot dz$$

with contour Γ: $|\arg z| = \pi - \phi$. The order conditions for order p,

$$b^T \mathscr{A}^l c^{k-1} = b^T \mathscr{A}^{l-1} c^k / k, \qquad k + l \le p$$

(with $c^k = (c_i^k)_{i=1}^m$), show that

$$zb^T (I - z\mathscr{A})^{-1} \delta^{(k)} = O(z^{p-k+1}), \qquad z \to 0.$$

With the arguments of the proof of Lemma 2.3 (with $(s+A)^{-1}$ satisfying (3.2) in the role of $K(s)$), the presence of the factor $|z|^{p-k+1} \cdot |z|^{-\alpha} = |z|^{p-q-\beta}$ for z near 0 gives the same power of $1/n$ in

$$\| r_n^{(k)}(-hA) \cdot (hA)^{-\alpha} \| \le \frac{C}{n^{p-q-\beta}}.$$

For $q+1+\beta < p$, it is thus majorized by an absolutely summable sequence, and so we get

$$\| e_n^{(k)} \| \le C \cdot h^{q+1+\beta} \cdot \max_{0 \le t \le t_n} \| A^{q+1+\beta-k} y^{(k)}(t) \|,$$

which is the desired estimate.

If $q+1+\beta = p$, then we only have a bound by C/n, which yields an additional factor $\log n$ in the bound of $e_n^{(k)}$.

If $\alpha = 0$, i.e., $k = p - 1 = q + 1 + \beta$, then the above integral is no longer convergent. The remedy is to consider instead the sequence $\hat{r}_n^{(k)}(-hA)$ with $\hat{r}_n^{(k)}(z) = [R(z)^n - R(\infty)^n] \cdot zb^T(I - z\mathscr{A})^{-1}\delta^{(k)}$, which differs from $r_n^{(k)}(-hA)$ by an absolutely summable sequence. One then gets a convergent integral representation, with $R(z)^n$ replaced by $R(z)^n - R(\infty)^n$, and the result follows as before.

(d) For $k = p$ we use a trick of Brenner et al. [1]: With

$$\sigma(z) = zb^T(I - z\mathscr{A})^{-1}\delta^{(p)}/[1 - R(z)],$$

we write

$$\begin{aligned} e_{n+1}^{(p)} &= h^p \cdot \sum_{\nu=0}^{n} R(-hA)^{n-\nu}[I - R(-hA)] \cdot \sigma(-hA) \cdot y^{(p)}(t_\nu) \\ &= h^p \cdot \sigma(-hA) \\ &\quad \cdot \left[y^{(p)}(t_n) + \sum_{\nu=1}^{n} R(-hA)^{n-\nu} \int_{t_{\nu-1}}^{t_\nu} y^{(p+1)}(t)\, dt - R(-hA)^{n+1} y^{(p)}(0) \right]. \end{aligned}$$

Since $\sigma(z)$ is a rational function which is bounded on the sector $|\arg(-z)| \le \theta$, we have $\|\sigma(-hA)\| \le \text{Const}$. This estimate and the uniform power boundedness of $R(-hA)$ now show that $e_n^{(p)} = O(h^p)$ as required. □

In the following section we will need an error estimate also for the internal stages:

Lemma 3.4. *For an initial value problem* (3.1)–(3.2), *consider a Runge-Kutta method of stage order* q *and order* $p \ge q+1$, *which is strongly* $A(\theta)$*-stable with* $\theta > \phi$. *The error of the internal stages is bounded for* $i = 1, \dots, m$ *by*

$$\|Y_{ni} - y(t_n + c_i h)\| \le C \cdot h^{q+1} \cdot \left(\|y^{(q+1)}(0)\| + \int_0^{t_n} \|y^{(q+2)}(t)\|\, dt \right).$$

Proof. With the notation of the previous proof, the internal error vector $E_n = (Y_{ni} - y(t_n + c_i h))_{i=1}^m$ is given by the equation

$$(I + h\mathscr{A} \otimes A)E_n = D_n + \mathbb{1} \otimes e_n.$$

The matrix multiplying E_n has a uniformly bounded inverse, and both D_n and e_n are bounded by expressions of the form of the right-hand side of the desired inequality (take $p = q + 1$ in Theorem 3.3 to see this for e_n). This implies the stated result. □

When equation (3.1) comes from a spatial semidiscretization of a parabolic PDE, then one would like to compare the completely discrete solution to the PDE solution rather than that of the semidiscretization. We are thus led to consider a perturbed equation

$$u' + Au = g(t) + d(t), \qquad u(0) = u_0. \tag{3.8}$$

Here, u may be a projection of the PDE solution onto the finite-dimensional approximation space, in which case d represents the spatial truncation error. To obtain an estimate of $y_n - u(t_n)$, we split

$$y_n - u(t_n) = (y_n - v_n) + (v_n - u_n) + (u_n - u(t_n)),$$

where u_n denotes the Runge-Kutta solution of (3.8), and v_n that of the same perturbed differential equation, but with initial value $v_0 = y_0$. Then $u_n - u(t_n)$ can be bounded using Theorem 3.3, and $v_n - u_n = R(-hA)^n(y_0 - u_0)$, which by (3.3) is bounded by a constant times the initial error. By linearity, $e_n := v_n - y_n$ is the Runge-Kutta solution of

$$e' + Ae = d(t), \qquad e(0) = 0. \tag{3.9}$$

This is bounded in the following stability lemma.

Lemma 3.5. *Assume that A satisfies* (3.2), *and that the Runge-Kutta method is strongly $A(\theta)$-stable with $\theta > \phi$. Then*

(i) $\|e_n\| \le C \cdot \log n \cdot \max_{0\le t\le t_n} \|A^{-1}d(t)\|$;

(ii) $\|e_n\| \le \frac{C}{\varepsilon} \cdot \max_{0\le t\le t_n} \|A^{-1+\varepsilon}d(t)\|$, $0 < \varepsilon \le 1$;

(iii) $\|e_n\| \le C \cdot (\|A^{-1}d(0)\| + \int_0^{t_n} \|A^{-1}d'(t)\|\,dt)$.

The same bounds are valid for the internal stages E_{ni}.

Proof. By Lemma 3.1, the internal stages $E_n = (E_{ni})_{i=1}^m$ satisfy, with $D_n = (d(t_n + c_i h))_{i=1}^m$,

$$E_n = h\sum_{\nu=0}^{n} W_{n-\nu}D_\nu .$$

We note that $W_n A^{1-\varepsilon}$ are the coefficients of (2.7) with $K(s) = (s + A)^{-1}A^{1-\varepsilon}$, which is bounded by (see, e.g., [10, §1.4])

$$\|(s + A)^{-1}A^{1-\varepsilon}\| \le C \cdot |s|^{-\varepsilon}, \qquad |\arg s| \le \pi - \phi, \tag{3.10}$$

with a constant independent of $\varepsilon \in [0, 1]$. It follows with Lemma 2.3 that

$$\|W_n A^{1-\varepsilon}\| \le C \cdot [(n + 1)h]^{\varepsilon-1}, \tag{3.11}$$

and hence the bounds (i) and (ii) follow for E_n, and with Lemma 3.1 also for e_n.

The proof of (iii) is obtained with a variant of the trick of [1] used already at the end of the proof of Theorem 3.3. We have

$$hW_0 = h(\mathscr{A} \otimes I)(I + \mathscr{A} \otimes hA)^{-1} = \Phi(-hA) \cdot (I \otimes A^{-1}),$$

where $\Phi(z) = z(I - z\mathscr{A})^{-1}$ is bounded on the sector $|\arg(-z)| \le \theta$, so that $\|\Phi(-hA)\|$ is uniformly bounded. It follows that

$$\|hW_0 D_n\| \le C \cdot \|(I \otimes A^{-1})D_n\| \le C \cdot \max_{t_n\le t\le t_{n+1}} \|A^{-1}d(t)\|.$$

For $n \ge 1$ we have $W_n = \Psi_n(-hA)$ with

$$\Psi_n(z) = R(z)^{n-1}(I - z\mathscr{A})^{-1}\mathbb{1}b^T(I - z\mathscr{A})^{-1} = \Psi(z) \cdot R(z)^{n-1} \cdot \frac{1 - R(z)}{z},$$

where the rational matrix-valued function

$$\Psi(z) = \frac{z}{1 - R(z)}(I - z\mathscr{A})^{-1}\mathbb{1}b^T(1 - z\mathscr{A})^{-1}$$

is bounded on the sector, so that $\|\Psi(-hA)\| \le \text{Const}$. We thus have

$$h\sum_{\nu=0}^{n-1} W_{n-\nu}D_\nu = \Psi(-hA) \cdot \sum_{\nu=0}^{n-1} R(-hA)^{n-\nu-1}(I - R(-hA)) \cdot (I \otimes A^{-1})D_\nu ,$$

and rearranging the last sum leads to the desired bound. □

4. Nonlinear parabolic equations

In this section we study error estimates of Runge-Kutta methods applied to semilinear parabolic equations. In particular, incompressible Navier-Stokes equations are covered by our assumptions.

On a Banach space X where $\|\cdot\|$ denotes the norm as well as the induced operator norm, we consider the initial value problem

$$y' + Ay = g(t, y), \qquad y(0) = y_0. \tag{4.1}$$

Here, $-A$ is the generator of an analytic semigroup, with spectrum to the left of the origin:

$$\|(s + A)^{-1}\| \leq \frac{M}{1 + |s|} \quad \text{for } |\arg s| \leq \pi - \phi, \text{ with some } \phi < \frac{\pi}{2}. \tag{4.2}$$

For the nonlinearity g, it is assumed that there exist real numbers α and $\gamma < 1$ such that the following local Lipschitz condition is satisfied for every $r > 0$:

$$\|g(t, u) - g(t, v)\|_{\alpha-\gamma} \leq l(r) \cdot \|u - v\|_\alpha \quad \text{for } \|u\|_\alpha \leq r, \ \|v\|_\alpha \leq r, \ 0 \leq t \leq T, \tag{4.3}$$

where $\|\cdot\|_\alpha$ denotes the norm

$$\|v\|_\alpha = \|A^\alpha v\|. \tag{4.4}$$

Studies of existence, uniqueness, and regularity of solutions of semilinear problems can be found in [20], and references therein.

We now turn briefly to the most prominent example of such equations. For further examples and theory, we refer to [10].

Incompressible Navier-Stokes equations [6]. On a bounded two- or three-dimensional smooth domain Ω and for positive times $t > 0$, we consider

$$\begin{aligned} \frac{\partial u}{\partial t} + (u \cdot \nabla)u - \Delta u + \nabla p &= f, \\ \nabla \cdot u &= 0, \end{aligned} \tag{4.5}$$

with Dirichlet boundary conditions $u = 0$ on $\partial\Omega$, and initial conditions $u = u_0$ at $t = 0$. This is to be solved for the velocity $u = (u_1, u_2)(x, t)$ (or $u = (u_1, u_2, u_3)$ in three dimensions) and the pressure $p = p(x, t)$, given the force term $f = f(x, t)$. Let X be the divergence-free subspace of $L^2(\Omega)^d$ ($d = 2$ or 3) defined as the closure of the set of smooth functions on Ω with vanishing divergence and support contained in Ω. Let further P denote the orthogonal projection of $L^2(\Omega)^d$ onto X. Applying P to the differential equation eliminates the pressure p, and one has the differential equation on X,

$$\frac{\partial u}{\partial t} + P(u \cdot \nabla)u - P\Delta u = Pf. \tag{4.6}$$

The Stokes operator $A = -P\Delta$ with Dirichlet boundary conditions satisfies (4.2). The nonlinearity satisfies the Lipschitz condition (4.3) for any $\gamma > \frac{1}{2}$ in the case of two space dimensions, and for any $\gamma > \frac{3}{4}$ in the three-dimensional case, for both $\alpha = \gamma$ and $\alpha = \frac{1}{2}$. Note that for $\alpha = \frac{1}{2}$ one has the Dirichlet norm $\|v\|_{1/2} = \|\nabla v\|_{L^2}$.

Under time discretization by a Runge-Kutta method, equations (4.5) and (4.6) give identical semidiscrete solutions. This invariance is still valid for the analogous transformation of finite element spatial discretizations of (4.5). □

The numerical result obtained by a Runge-Kutta time discretization will be compared to the solution of a perturbed equation

$$u' + Au = g(t, u) + d(t). \tag{4.7}$$

This is again motivated by the situation when (4.1) is a spatial semidiscretization of a parabolic PDE: There, $u(t)$ is taken as a projection of the solution of the PDE into the finite-dimensional approximation space, and $d(t)$ represents the spatial truncation error. Another instance of (4.7) is when (4.1) does not have a smooth solution, but there exists a smooth function u with small defect d.

There is the following error bound in the case that (4.7) has a solution on an interval $[0, T]$ which is sufficiently smooth in time.

Theorem 4.1 (Convergence of Runge-Kutta methods for semilinear parabolic equations). *For an initial value problem* (4.1)–(4.3), *consider a Runge-Kutta method of stage order* q *and order* $p \geq q+1$, *which is strongly* $A(\theta)$*-stable with* $\theta > \phi$. *Then there exists a unique numerical solution* y_n $(0 \leq nh \leq T)$ *for sufficiently small step sizes* $h \leq h_0$, *whose error is bounded by*

$$\|y_n - u(t_n)\|_\alpha \leq C(u) \cdot h^{q+1} + C \cdot \|y_0 - u(0)\|_\alpha + C \cdot \log n \cdot \max_{0 \leq t \leq t_n} \|d(t)\|_{\alpha-1} \tag{4.8}$$

with $C(u) = C \cdot (\|u^{(q+1)}(0)\|_\alpha + \int_0^{t_n} \|u^{(q+2)}(t)\|_\alpha \, dt)$. *This estimate is satisfied for* $t_n = nh \leq T$ *provided that the expression on the right-hand side is bounded by a sufficiently small constant* c. *The constants* h_0, C, *and* c *depend only on the constants in* (4.2) *and* (4.3), *on* T, *and on* $\max_{0 \leq t \leq T} \|u(t)\|_\alpha$. *They are otherwise independent of* A, *and independent of* n *and* h *with* $nh \leq T$.

We remark that the term involving $d(t)$ could be replaced by any of the expressions of Lemma 3.5, taken in the norm $\|\cdot\|_\alpha$.

Proof. (a) Let us first assume that the numerical solution y_n and the internal stages $Y_n = (Y_{ni})_{i=1}^m$ exist for $0 \leq nh \leq T$, with $\|Y_{ni}\|_\alpha \leq 2 \max_{0 \leq t \leq T} \|u(t)\|_\alpha$. For sufficiently small step sizes, this will be verified at the end of the proof.

By the discrete variation of constants formula (Lemma 3.1) we then have

$$Y_n = V_n y_0 + h \sum_{\nu=0}^{n} W_{n-\nu} g(Y_\nu), \tag{4.9}$$

where we have denoted $g(Y_n) = (g(t_n + c_i h, Y_{ni}))_{i=1}^m$, and V_n and W_n are defined as in Lemma 3.1.

(b) For $U_n = (u(t_n + c_i h))_{i=1}^m$ we get the representation

$$U_n = V_n u(0) + h \sum_{\nu=0}^{n} W_{n-\nu}(g(U_\nu) + D_\nu) + E_n, \tag{4.10}$$

with $g(U_n) = (g(t_n + c_i h, U_{ni}))_{i=1}^m$ and $D_n = (d(t_n + c_i h))_{i=1}^m$. Here, $E_n = (E_{ni})_{i=1}^m$ is the error of the internal stages of the Runge-Kutta method when applied to the *linear* equation

$$v' + Av = \varphi(t), \quad \text{with initial value } u(0), \tag{4.11}$$

where the inhomogeneity is chosen such that the exact solution is $u(t)$. This identity follows by applying the discrete variation of constants formula to the perturbed Runge-Kutta schemes obtained by inserting $u(t)$ into the Runge-Kutta formulas for both equations (4.1) and (4.11). Note here that the defect, from inserting the exact solution values into the Runge-Kutta method, depends only on the solution u, but not on the form of the equation.

(c) We observe that $W_n A^\gamma$ are the coefficients of the generating function (2.7) with $K(s) = (s + A)^{-1} A^\gamma$, which is bounded by

$$\|(s + A)^{-1} A^\gamma\| \leq \text{Const} \cdot |s|^{\gamma - 1}, \qquad |\arg s| \leq \pi - \phi. \tag{4.12}$$

It thus follows from Lemma 2.3 that

$$\|W_n A^\gamma\| \leq C \cdot [(n + 1)h]^{-\gamma}, \qquad n \geq 0. \tag{4.13}$$

With the Lipschitz condition (4.3) we get therefore (with $r = 2 \max_{0 \leq t \leq T} \|u(t)\|_\alpha$)

$$\|Y_n - U_n\|_\alpha \leq Ch \cdot l(r) \sum_{\nu=0}^{n} [(n + 1 - \nu)h]^{-\gamma} \cdot \|Y_\nu - U_\nu\|_\alpha + \|F_n\|_\alpha,$$

where

$$F_n = V_n \cdot (y_0 - u(0)) + h \sum_{\nu=0}^{n} W_{n-\nu} D_\nu + E_n.$$

Since $\gamma < 1$, a discrete Gronwall inequality gives us

$$\|Y_n - U_n\|_\alpha \leq C \cdot \max_{0 \leq \nu \leq n} \|F_\nu\|_\alpha. \tag{4.14}$$

(d) We thus have to estimate F_n. Since $R(-hA)$ is uniformly power-bounded, we have

$$\|V_n(y_0 - u(0))\|_\alpha \leq C \cdot \|y_0 - u(0)\|_\alpha.$$

The error of the internal stages E_n of the linear problem (4.11) is estimated using Lemma 3.4 (with $\|\cdot\|_\alpha$ in the role of the norm used there):

$$\|E_n\|_\alpha \leq C(u) \cdot h^{q+1},$$

with $C(u)$ of the same form as in (4.8). For the remaining term in F_n we note that Lemma 3.5 gives

$$\left\| h \sum_{\nu=0}^{n} W_{n-\nu} D_\nu \right\|_\alpha \leq C \cdot \log n \cdot \max_{0 \leq t \leq t_n} \|d(t)\|_{\alpha - 1}.$$

With these bounds for F_n, the inequality (4.14) gives the bound (4.8) for the internal stages Y_n.

(e) To get the estimate also for the numerical solution y_n itself, we use again the variation of constants formula of Lemma 3.1 to get

$$y_{n+1} = R(-hA)^{n+1} y_0 + b^T \mathscr{C}^{-1} h \sum_{k=0}^{n} R(\infty)^{n-k} \sum_{\nu=0}^{k} W_{k-\nu} g(Y_\nu)$$

and

$$u(t_{n+1}) = R(-hA)^{n+1} u(0) + b^T \mathscr{C}^{-1} \sum_{k=0}^{n} R(\infty)^{n-k} h \sum_{\nu=0}^{k} W_{k-\nu} (g(U_\nu) + D_\nu) + e_{n+1},$$

where e_{n+1} is the error of the method applied to the linear equation (4.11), which by Theorem 3.3 is again bounded as E_n above. We now use the obtained error estimates of the internal stages, treat the inner convolution in (4.20) and (4.15) in the same way as in parts (c) and (d) above, and recall that $|R(\infty)| < 1$. This gives the desired error bound (4.8) for y_n.

(f) It remains to prove the unique existence of the numerical solution in a neighborhood (with respect to the $\|\cdot\|_\alpha$ norm pointwise in time) of the exact solution u. Using the arguments of part (c) of the proof, we see that for small step sizes h the mapping

$$(Y_n)_{n=0}^N \mapsto \left(h \sum_{\nu=0}^{n} W_{n-\nu} g(Y_\nu) \right)_{n=0}^N \qquad (Nh \le T)$$

is for a sufficiently large exponent a (independent of h) a contraction in the exponentially weighted maximum norm

$$|||Y||| := \max_{0\le n\le N} \max_{i=1,\dots,m} \|\exp(-a\cdot(t_n + c_i h))Y_{ni}\|_\alpha ,$$

with Lipschitz constant $\rho < 1$ uniformly in h in the ball

$$|||Y||| \le 2 \max_{0\le t\le T} \|e^{-at}u(t)\|_\alpha .$$

For $|||(F_n)_{n=0}^N|||$ sufficiently small, fixed point iteration in (4.9) with initial iterate $(U_n)_{n=0}^N$ gives therefore the unique existence of a numerical solution with the properties assumed in part (a) of the proof. □

Under slightly stronger assumptions the order of convergence can still be raised. In addition to (4.3) we require for every $r > 0$:

$$\begin{aligned} &\left\| g(t,u) - g(t,v) - \frac{\partial g}{\partial y}(t,y)\cdot(u-v) \right\|_{\alpha-1} \\ &\qquad \le\ l'(r)\cdot(\|y-u\|_\alpha + \|y-v\|_\alpha)\cdot\|u-v\|_\alpha \end{aligned} \tag{4.15}$$

for $\max(\|u\|_\alpha, \|v\|_\alpha, \|y\|_\alpha) \le r$. This condition is again satisfied for the Navier-Stokes problem.

Theorem 4.2 (Refined error estimate). *In addition to the conditions of Theorem* 4.1, *we assume* (4.15) *and* $p \ge q+2$. *If* $\beta < 1$ *is such that* $u^{(q+1)}(t) \in D(A^{\alpha+\beta})$, *then the error is bounded for* $t_n \le T$ *by*

$$\begin{aligned} \|y_n - u(t_n)\|_\alpha \le\ & C_\beta(u)\cdot h^{q+1+\beta} + C_1(u)\cdot h^{q+2} \\ & + C\cdot\|y_0 - u(0)\|_\alpha + C\cdot\log n\cdot \max_{0\le t\le t_n}\|d(t)\|_{\alpha-1} \end{aligned} \tag{4.16}$$

with $C_\beta(u) = C\cdot\max_{0\le t\le T}\|u^{(q+1)}(t)\|_{\alpha+\beta}$ *and* $C_1(u) = C\cdot(\|u^{(q+2)}(0)\|_\alpha + \int_0^{t_n}\|u^{(q+3)}(t)\|_\alpha\,dt)$. *For* $\beta = 1$ *the estimate is valid with an additional factor* $\log n$ *in front of the first term. The estimate holds under the assumption that the right-hand side of the inequality and* $h^q\cdot(\max_{t_n\le t\le t_{n+1}}\|u^{(q+1)}(t)\|_\alpha)^2$ *are bounded by a small constant.*

Remark. For parabolic PDEs with homogeneous boundary conditions that have a sufficiently smooth solution, the inclusion $u(t) \in D(A)$ implies $u^{(q+1)}(t) \in D(A)$, and the above estimate is thus applicable at least with $\beta = 1-\alpha$. Returning to the Navier-Stokes equation with homogeneous Dirichlet boundary

conditions, we have a sufficiently smooth solution $u^{(q+1)}(t) \in D(A^{5/4-\varepsilon})$ for any $\varepsilon > 0$ (cf. [7, 8]). Theorem 4.2 can thus be applied with $\alpha = \frac{1}{2}$ and $\beta = \frac{3}{4} - \varepsilon$.

Proof of Theorem 4.2. We concentrate on those aspects which go beyond the proof of Theorem 4.1. We therefore restrict our attention to the case $d(t) \equiv 0$, $u(0) = y_0$ in (4.7), and we ignore the dependence on t in $g(t, y)$.

(a) Let us denote the defect

$$d_{ni} = u(t_n + c_i h) - u(t_n) - h \sum_{j=1}^{m} a_{ij} u'(t_n + c_j h),$$

$$d_{n+1} = u(t_{n+1}) - u(t_n) - h \sum_{j=1}^{m} b_j u'(t_n + c_j h).$$

We have

$$d_{ni} = d_{ni}^{*} + d_{ni}^{**} \equiv \delta_i \cdot h^{q+1} u^{(q+1)}(t_n)/q! + h^{q+1} \int_{t_n}^{t_{n+1}} \kappa_i \left(\frac{t - t_n}{h} \right) u^{(q+2)}(t)\, dt,$$

with $\delta_i = \sum_{j=1}^{m} a_{ij} c_j^q - c_i^{q+1}/(q+1)$, and with a bounded Peano kernel κ_i. A similar representation is valid for d_n, with $q+2$ instead of q.

(b) We study the defect when

$$\hat{Y}_{ni} = u(t_n + c_i h) - d_{ni}, \qquad \hat{y}_n = u(t_n)$$

are inserted into the Runge-Kutta method. With $f(y) = -Ay + g(y)$ we have

$$\begin{aligned} &\hat{Y}_{ni} - \hat{y}_n - h \sum_{j=1}^{m} a_{ij} f(\hat{Y}_{nj}) \\ &\quad = h \sum_{j=1}^{m} a_{ij} A d_{nj} + h \sum_{j=1}^{m} a_{ij} [g(u(t_n + c_j h)) - g(\hat{Y}_{nj})], \\ &\hat{y}_{n+1} - \hat{y}_n - h \sum_{j=1}^{m} b_j f(\hat{Y}_{nj}) \\ &\quad = h \sum_{j=1}^{m} b_j A d_{nj} + h \sum_{j=1}^{m} b_j [g(u(t_n + c_j h)) - g(\hat{Y}_{nj})] + d_{n+1}. \end{aligned}$$

By the discrete variation of constants formula, we get

$$\hat{Y}_n = V_n y_0 + h \sum_{\nu=0}^{n} W_{n-\nu} g(\hat{Y}_\nu) + F_n$$

with

$$F_n = h \sum_{\nu=0}^{n} W_{n-\nu} A D_\nu + h \sum_{\nu=0}^{n} W_{n-\nu} [g(U_\nu) - g(\hat{Y}_\nu)] + \sum_{\nu=1}^{n} V_{n-\nu} d_\nu,$$

where $\hat{Y}_n = (\hat{Y}_{ni})_{i=1}^{m}$ and $D_n = (d_{ni})_{i=1}^{m}$, and the notation is otherwise that employed in the proof of Theorem 4.1. As in part (c) of that proof, it follows that the difference to the numerical solution is bounded by

$$\| Y_n - \hat{Y}_n \|_\alpha \le C \cdot \max_{0 \le \nu \le n} \| F_\nu \|_\alpha.$$

We will show that this is $O(h^{q+1+\beta})$ as in the bound stated in the theorem.

(c) The difficulty in the estimation is contained in the first and second term in F_n, because D_n is only $O(h^{q+1})$. However, with $\delta = (\delta_i)_{i=1}^m$ we have by Lemma 2.4

$$W_n(\delta \otimes I) = \Psi_n(-hA)(\delta \otimes I),$$

where again for $n \geq 1$

$$\Psi_n(z) = R(z)^{n-1}(I - z\mathscr{Q})^{-1}\mathbb{1}b^T(I - z\mathscr{Q})^{-1}.$$

Since $p \geq q+2$, the order conditions give $b^T\delta = 0$, and therefore

$$\|\Psi_n(z)\delta\| \leq \frac{C}{1+|z|^3}|R(z)^{n-1}| \cdot |z|.$$

From the additional factor z we get with the proof of Lemma 2.3 (with $(s+A)^{-1}A^{1-\beta}$ in the role of $K(s)$, and using (4.12)) the factor $1/n$ in

$$\|hW_n(\delta \otimes A^{1-\beta})\| \leq \frac{Ch}{n}(nh)^{\beta-1} = \frac{C}{n^{2-\beta}}h^\beta.$$

For $\beta < 1$ we have therefore, with $D_n^* = (d_{ni}^*)_{i=1}^m$ and d_{ni}^* defined in part (a) of the proof,

$$\left\| h\sum_{\nu=0}^{n} W_{n-\nu}AD_\nu^* \right\|_\alpha \leq C \cdot h^{q+1+\beta} \cdot \max_{0\leq t\leq t_n} \|u^{(q+1)}(t)\|_{\alpha+\beta},$$

whereas for $\beta = 1$ there is an additional factor $\log n$.

(d) By (4.3) we have

$$\left\| \frac{\partial g}{\partial y}(u(t_n))v \right\|_{\alpha-\gamma} \leq C \cdot \|v\|_\alpha$$

for all v in the domain of A^α, and so we get as in part (c)

$$\left\| h\sum_{\nu=0}^{n} W_{n-\nu}\frac{\partial g}{\partial y}(u(t_\nu))D_\nu^* \right\|_\alpha \leq C \cdot h^{q+1+\beta} \max_{0\leq t\leq t_n} \|u^{(q+1)}(t)\|_{\alpha+\beta}.$$

By condition (4.15), the Taylor remainder is bounded by

$$\left\| g(U_n) - g(\hat{Y}_n) - \frac{\partial g}{\partial y}(u(t_n)) \cdot D_n \right\|_{\alpha-1} \leq C \cdot \left(h^{q+1} \cdot \max_{t_n\leq t\leq t_{n+1}} \|u^{(q+1)}(t)\|_\alpha \right)^2,$$

and it follows from (4.13), used with $\gamma = 1$, that the discrete convolution of the sequence $\{hW_n\}$ with this remainder term is bounded in the $\|\cdot\|_\alpha$ norm by $C \cdot \log n$ times the right-hand side of the above inequality.

(e) The remaining terms in F_n are shown to be bounded as needed by the arguments used in the proof of Theorem 4.1. As in part (e) of that proof, the estimate of $y_n - u(t_n) = y_n - \hat{y}_n$ follows with the variation of constants formula, once a bound of the desired type has been established for $Y_n - \hat{Y}_n$. □

Remark. The same order of convergence as in Theorem 3.3 (possibly beyond h^{q+2}, if the regularity of the solution allows it) can be obtained also in the semilinear case, if one assumes conditions like (4.15) also for higher derivatives of g. This can be shown by studying the coefficients in the asymptotic

h-expansion of the internal stages of the method, of which d^*_{ni} in the foregoing proof is actually the first term. We omit the details, which become very technical.

5. Proof of error estimates in terms of the data

We begin with the proof of Theorem 2.2. The proof of Theorem 3.2 will then follow by a slight modification at the end of the section.

We recall the convolution quadrature

$$\begin{aligned} U_n &= h\sum_{\nu=0}^{n} W_{n-\nu}G_\nu\,, \\ u_{n+1} &= b^T\mathscr{Q}^{-1}\sum_{k=0}^{n} R(\infty)^{n-k}h\sum_{\nu=0}^{k} W_{k-\nu}G_\nu\,, \end{aligned} \tag{5.1}$$

with $G_n = (g(t_n + c_i h))_{i=1}^m$, and quadrature weights W_n defined by (2.7).

By linearity, we may treat each of the terms in the Taylor expansion of g at 0 separately:

$$g(t) = \sum_{l=0}^{p-1} \frac{t^l}{l!} g^{(l)}(0) + \int_0^t \frac{\tau^{p-1}}{(p-1)!} g^{(p)}(t-\tau)\,d\tau\,. \tag{5.2}$$

For the Taylor polynomial we will show below:

Lemma 5.1. *Theorem* 2.2 *is valid if* $g(t)$ *is a polynomial of degree at most* $p-1$.

To treat the remainder term in (5.2), we will again use Lemma 5.1, combined with the fact that the application of the quadrature method commutes with convolution:

Lemma 5.2 (Peano kernel representation). *If* $g(0) = g'(0) = \cdots = g^{(p-1)}(0) = 0$, *then the convolution quadrature* (5.1) *can be written in the form*

$$\begin{aligned} U_n &= \int_{-\infty}^{t_n} V_h(\tau) g^{(p)}(t_n - \tau)\,d\tau\,, \\ u_{n+1} &= \int_{-\infty}^{t_n} v_h(\tau) g^{(p)}(t_n - \tau)\,d\tau\,, \end{aligned}$$

where the "Peano kernels" are given by

$$\begin{aligned} V_h(t) &= h\sum_{n\geq 0} W_n P(t - nh)\,, \\ v_h(t) &= hb^T\mathscr{Q}^{-1}\sum_{n\geq 0}\left(\sum_{k=0}^{n} R(\infty)^{n-k}W_k\right)P(t-nh)\,, \end{aligned}$$

with

$$P(t) = \left(\frac{(t + c_i h)_+^{p-1}}{(p-1)!}\right)_{i=1}^m$$

for all real t. *Here,* $\tau_+^{p-1} = (\max(\tau, 0))^{p-1}$.

Proof. The result is obtained by using the associativity of convolution. □

We note that $V_h(t_n)$ and $v_h(t_n)$ are just the results of the convolution quadrature method used with $g(t) = t^{p-1}/(p-1)!$. Hence, Lemma 5.1 gives with $v(t) = \int_0^t k(t-\tau)\tau^{p-1}/(p-1)!\,d\tau$ the error bound

$$|v_h(t) - v(t+h)| \le C(h^p + h^{q+1+\mu}|\log h|) + Ch^p t^{\mu+1}.$$

This holds in fact for all t, not only gridpoints: For t between gridpoints, $v_h(t)$ can be interpreted as the result of the convolution quadrature used with a shifted polynomial $g(t) = (t-ch)^{p-1}/(p-1)!$, and Lemma 5.1 applies again. Since the exact solution can be written as

$$u(t) = \int_0^t v(\tau) g^{(p)}(t-\tau)\,d\tau,$$

the desired estimate of Theorem 2.2 follows. It remains to prove Lemma 5.1.

Proof of Lemma 5.1. (a) A computation similar to that in part (b) of the proof of Theorem 3.3 shows that the error of the Runge-Kutta method, when applied to $y' = \lambda y + t^l/l!$, $y(0) = 0$, is given by

$$\begin{aligned} e_n &= \lambda^{-l-1}(R(h\lambda)^n - e^{nh\lambda}) \\ &\quad - \sum_{k=q+1}^{p} h^k \sum_{\nu=1}^{n-1} r_{n-1-\nu}^{(k)}(h\lambda)\lambda^{k-l-1} \sum_{\kappa=0}^{l-k} \frac{(\lambda t_\nu)^\kappa}{\kappa!}, \end{aligned} \tag{5.3}$$

with $r_n^{(k)}(z)$ of (3.7). We now treat the two terms in the above formula separately.

(b) We show that

$$\left| \int_\Gamma K(\lambda)\lambda^{-l-1}(R(h\lambda)^n - e^{nh\lambda})\,d\lambda \right| \le Ch^p(1 + t_n^{\mu+l-p}). \tag{5.4}$$

Here the integration contour Γ is that of (2.3), possibly replaced by an equivalent one that is bounded away from the origin. We split this contour into three parts:

(i) $|\lambda| \le$ Const: In this situation, $|R(h\lambda)^n|$ as well as $|e^{nh\lambda}|$ are uniformly bounded for $nh \le T$. Using

$$R(z)^n - e^{nz} = (R(z) - e^z) \sum_{\nu=0}^{n-1} R(z)^{n-1-\nu} e^{\nu z},$$

we see that the integral over this part of Γ is $O(h^p)$.

(ii) $h \le |h\lambda| \le 1$ and $|\arg \lambda| = \pi - \phi$: Here we have with $z \equiv h\lambda$ that $|R(z)| \le |e^{\kappa z}|$ for some positive κ. This implies

$$|R(z)^n - e^{nz}| \le C|z^p e^{\kappa n z}|.$$

Hence the integral is bounded by a constant times

$$\int_{\lambda\in\Gamma:\ |\lambda|\ge 1} |\lambda|^{-\mu-l-1}|h\lambda|^p |e^{\kappa n h\lambda}|\,|d\lambda| \le Ch^p(1 + t_n^{\mu+l-p}).$$

(iii) $|h\lambda| \ge 1$ and $|\arg \lambda| = \pi - \phi$: In this case both $R(z)$ and e^z have absolute value bounded strictly below 1, hence

$$|R(z)^n - e^{nz}| \le C\rho^n$$

for some $\rho < 1$. Thus, the integral is bounded by

$$C \int_{\lambda\in\Gamma:\ |\lambda|\geq 1/h} |\lambda|^{-\mu-l-1}\rho^n\,|d\lambda| = O(h^{\mu+l}\rho^n),$$

which is $O(h^p t_n^{\mu+l-p})$.

(c) We finally show that

$$\left|\sum_{k=q+1}^{p} h^k \int_\Gamma K(\lambda) \sum_{\nu=0}^{n-1} r_{n-1-\nu}^{(k)}(h\lambda)\lambda^{k-l-1} \sum_{\kappa=0}^{l-k} \frac{(\lambda t_\nu)^\kappa}{\kappa!}\right| \leq C(h^p + h^{q+1+\mu}|\log h|), \tag{5.5}$$

where $r_n^{(k)}(z)$ is given by (3.7). First we note that $\lambda^{k-l}\sum_{\kappa=0}^{l-k}(\lambda t_\nu)^\kappa/\kappa!$ as well as

$$\sum_{\nu=1}^{n-1} |r_{n-1-\nu}^{(k)}(z)| = \frac{|R(z)^n|-1}{|R(z)|-1} \cdot |zb^T(I-z\mathscr{Q})^{-1}(k\mathscr{Q}c^{k-1}-c^k)/k!| \tag{5.6}$$

(once more with $z = h\lambda$) are bounded on the integration contour. By the order conditions, we have

$$b^T(I-z\mathscr{Q})^{-1}(k\mathscr{Q}c^{k-1}-c^k) = O(z^{p-k}) \quad \text{for } z\to 0.$$

Again we split the contour into three parts:

(i) $|\lambda| \leq$ Const: From the foregoing formula we obtain an estimate $O(h^p)$.

(ii) $h \leq |h\lambda| \leq 1$ and $|\arg\lambda| = \pi - \phi$: Here $|K(\lambda)| = O(|\lambda|^{-\mu})$, and (5.6) is $O(z^{p-q})$. Therefore, we get the estimate

$$C\sum_{k=q+1}^{l} h^k \int_{\lambda\in\Gamma:\ 1\leq|\lambda|\leq 1/h} |\lambda|^{-\mu-1}|h\lambda|^{p-k}\,|d\lambda|, \qquad q+1\leq k\leq l,$$

which for $p-k-\mu-1<-1$ is $O(h^p)$. If $p-k-\mu\geq 0$, we have

$$h^k \int_{\lambda\in\Gamma:\ 1\leq|\lambda|\leq 1/h} |\lambda|^{-\mu-1}|h\lambda|^{p-k}\,|d\lambda| \leq h^{k+\mu}\int_{\lambda\in\Gamma:\ 1\leq|\lambda|\leq 1/h} |\lambda|^{-1}\,|d\lambda|$$

and thus the required bound is $O(h^{q+1+\mu}|\log h|)$.

(iii) $|h\lambda| \geq 1$ and $|\arg\lambda| = \pi - \phi$: Here $|K(\lambda)| = O(|\lambda|^{-\mu})$, and therefore the integral is bounded by a constant times

$$\sum_{k=q+1}^{p} h^k \int_{\lambda\in\Gamma:\ |\lambda|\geq 1/h} |\lambda|^{-\mu-1}\,|d\lambda|,$$

which is $O(h^{q+1+\mu})$. □

Proof of Theorem 3.2 (continued). Here again, it is sufficient to show the desired result for the homogeneous equation, and for the equations with zero initial values and polynomial inhomogeneity of degree not exceeding $p-1$, with coefficients $g^{(l)}(0)$ such that $A^{-1}g^{(l)}(0) \in D(A^\beta)$.

(a) The error for the homogeneous equation is $[(R(-hA)^n - e^{-nhA})A^{-\alpha}]\cdot A^\alpha y_0$. The operator in square brackets has an integral representation (5.4) with $K(\lambda) = (\lambda + A)^{-1}$ (so that $\mu = 1$ by (3.2)), and with α in the role of $l+1$. The bound in (5.4) then gives the result.

(b) The error of the method applied to the equation with inhomogeneity $g^{(l)}(0) \cdot t^l/l!$ is

$$e_n = \varphi_n(h, -A)A^{1-\beta} \cdot A^{\beta-1} g^{(l)}(0),$$

where $\varphi_n(h, \lambda)$ is given by the right-hand side of formula (5.3). The operator $\varphi_n(h, -A)A^{1-\beta}$ is the sum of two terms which have an integral representation as in (5.4) and (5.5), with $K(\lambda) = (\lambda + A)^{-1}\lambda^{1-\beta}$. Hence the bounds of the previous proof apply, with $\mu = \beta$. □

BIBLIOGRAPHY

1. P. Brenner, M. Crouzeix, and V. Thomée, *Single step methods for inhomogeneous linear differential equations in Banach space*, RAIRO Modél. Math. Anal. Numér. **16** (1982), 5–26.
2. K. Burrage, W. H. Hundsdorfer, and J. G. Verwer, *A study of B-convergence of Runge-Kutta methods*, Computing **36** (1986), 17–34.
3. M. Crouzeix, *Sur l'approximation des équations différentielles opérationelles linéaires par des méthodes de Runge-Kutta*, Thèse d'Etat, Université Paris VI, 1975.
4. M. Crouzeix and P. A. Raviart, *Approximation des problèmes d'évolution*, unpublished lecture notes, Université de Rennes, 1980.
5. P. P. B. Eggermont and Ch. Lubich, *Uniform error estimates of operational quadrature methods for nonlinear convolution equations on the half-line*, Math. Comp. **56** (1991), 149–176.
6. H. Fujita and T. Kato, *On the Navier-Stokes initial value problem*. I, Arch. Rational Mech. Anal. **16** (1964), 269–315.
7. D. Fujiwara, *Concrete characterization of the domains of fractional powers of some elliptic differential operators of the second order*, Proc. Japan Acad. **43** (1967), 82–86.
8. P. Grisvard, *Caractérisation de quelques espaces d'interpolation*, Arch. Rational Mech. Anal. **25** (1967), 40–63.
9. E. Hairer and G. Wanner, *Solving ordinary differential equations* II. *Stiff and differential-algebraic problems*, Springer-Verlag, Berlin and Heidelberg, 1991.
10. D. Henry, *Geometric theory of semilinear parabolic equations*, Lecture Notes in Math., vol. 840, Springer-Verlag, 1981.
11. J. G. Heywood and R. Rannacher, *Finite-element approximation of the nonstationary Navier-Stokes problem. Part* IV: *Error analysis for second-order time discretization*, SIAM J. Numer. Anal. **27** (1990), 353–384.
12. O. A. Karakashian, *On Runge-Kutta methods for parabolic problems with time-dependent coefficients*, Math. Comp. **47** (1986), 77–101.
13. S. L. Keeling, *Galerkin/Runge-Kutta discretizations for semilinear parabolic equations*, SIAM J. Numer. Anal. **27** (1990), 394–418.
14. I. Lasiecka, *Galerkin approximations of abstract parabolic boundary value problems with rough boundary data— L_p theory*, Math. Comp. **47** (1986), 55–75.
15. M. N. Le Roux, *Semidiscretization in time for parabolic problems*, Math. Comp. **33** (1979), 919–931.
16. Ch. Lubich, *Convolution quadrature and discretized operational calculus*. I, II, Numer. Math. **52** (1988), 129–145, 413–425.
17. ——, *On the convergence of multistep methods for nonlinear stiff differential equations*, Numer. Math. **58** (1991), 839–853; Erratum, **61** (1992), 277–279.
18. Ch. Lubich and O. Nevanlinna, *On resolvent conditions and stability estimates*, BIT **31** (1991), 293–313.
19. A. Ostermann and M. Roche, *Runge-Kutta methods for partial differential equations and fractional orders of convergence*, Math. Comp. **59** (1992), 403–420.
20. A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Springer-Verlag, New York, 1983.

21. B. A. Schmitt, *Stability of implicit Runge-Kutta methods for nonlinear stiff differential equations*, BIT **28** (1988), 884–897.
22. V. Thomée, *Galerkin finite element methods for parabolic problems*, Lecture Notes in Math., vol. 1054, Springer-Verlag, 1984.
23. J. G. Verwer, *Convergence and order reduction of diagonally implicit Runge-Kutta schemes in the method of lines*, Numerical Analysis (D. F. Griffiths and G. A. Watson, eds.), Pitman Res. Notes Math. Ser., vol. 140, Longman Sci. Tech., Harlow, 1986, pp. 220–237.

INSTITUT FÜR MATHEMATIK UND GEOMETRIE, UNIVERSITÄT INNSBRUCK, TECHNIKERSTR. 13, A-6020 INNSBRUCK, AUSTRIA

Current address, Ch. Lubich: Seminar für Angewandte Mathematik, ETH Zürich, CH-8092 Zürich, Switzerland

E-mail address: lubich@math.ethz.ch