IMA Journal of Numerical Analysis (2000) **20**, 167–184

Non-smooth data error estimates for linearly implicit Runge–Kutta methods

ALEXANDER OSTERMANN AND MECHTHILD THALHAMMER
Institut für Mathematik und Geometrie, Universität Innsbruck, Technikerstraße 13, A-6020 Innsbruck, Austria

[Received 14 January 1999 and in revised form 7 May 1999]

Linearly implicit time discretizations of semilinear parabolic equations with non-smooth initial data are studied. The analysis uses the framework of analytic semigroups which includes reaction–diffusion equations and the incompressible Navier–Stokes equations. It is shown that the order of convergence on finite time intervals is essentially one. Applications to the long-term behaviour of linearly implicit Runge–Kutta methods are given.

1. Introduction

When analysing discretizations of parabolic initial boundary value problems, it is not sufficient to consider only smooth initial data. This is partly because such initial data give solutions that keep their smoothness up to the boundaries. They thus require compatibility conditions which are often unrealistic in practical applications. Apart from that, non-smooth data error estimates are an important tool for obtaining long-term error bounds. This has been emphasized by Larsson (1992) and is also reflected in Assumption 3.2 in Stuart's survey article (Stuart 1995). The long-term behaviour of numerical solutions is closely related to the question of whether the continuous dynamics of the problem is correctly represented in its discretization. Suppose, for example, that the continuous problem has an asymptotically stable periodic orbit. Does the discrete dynamical system then possess an asymptotically stable invariant closed curve that lies close to the continuous orbit? The construction of such discrete invariant objects is usually based on fixed-point iteration, see e.g. Alouges & Debussche (1993), van Dorsselaer (1998), van Dorsselaer & Lubich (1999), Lubich & Ostermann (1996). Although the final result itself might be smooth, the single iterates are, in general, not. The whole construction thus relies on non-smooth data error estimates.

In spite of their importance, surprisingly few such estimates can be found in the literature. For time discretizations of linear parabolic problems, non-smooth data error estimates are first given by Le Roux (1979). But only until recently have these estimates been extended to more general problem classes. For semilinear parabolic problems, optimal results for implicit Runge–Kutta methods are given in Lubich & Ostermann (1996); see also the references therein. The corresponding results for multistep methods can be found in van Dorsselaer (1998).

In this paper we derive optimal error bounds for *linearly implicit Runge–Kutta methods*, applied to semilinear parabolic problems with non-smooth initial values. We work in an abstract Banach space setting of analytic semigroups, given in Henry (1981) and in Pazy (1983). This framework includes reaction–diffusion equations and the incompressible Navier–Stokes equations. The method class is formulated in sufficiently general terms such that it comprises classical Rosenbrock methods as well as extrapolation methods based on the linearly implicit Euler scheme. The latter have proven successful for the time integration of parabolic problems, see Bornemann (1990), Lang (1995), and Nowak (1993).

The present paper is structured as follows. In Section 2 we formulate the analytical framework, and we introduce the numerical method. The main result is stated in Section 3. There we prove that linearly implicit Runge–Kutta methods, when applied to semilinear parabolic problems with non-smooth initial data, converge with order one essentially. Low-order convergence is sufficient for applications to long-term error estimates. We illustrate this in Section 4 where we show that exponentially stable solutions of parabolic problems are uniformly approximated by linearly implicit methods over arbitrarily long time intervals. This result implies stability bounds for certain splitting methods. Under natural assumptions on the nonlinearity, it is possible to improve the convergence result of Section 3. This will be elaborated in Section 5. To keep the paper independent from other work, we have formulated all auxiliary results with an outline of the proofs in Section 6.

Compared to previous work, our convergence proofs are conceptionally simple. We consider the numerical approximation u_n of a linearly implicit Runge–Kutta method to the exact solution $u(t_n)$ as a perturbation of a suitably chosen Runge–Kutta solution $\widetilde{u}_n$. Using the triangular inequality

$$\|u_n - u(t_n)\| \leqslant \|u_n - \widetilde{u}_n\| + \|\widetilde{u}_n - u(t_n)\|,$$

we have to estimate $\|u_n - \widetilde{u}_n\|$. Together with the bounds for $\|\widetilde{u}_n - u(t_n)\|$ from Lubich & Ostermann (1996), we get the desired result. For the reader's convenience, we have collected all the necessary Runge–Kutta bounds in an appendix.

We finally remark that the above approach is not restricted to non-smooth data error estimates. It can equally be used, for example, to derive the conditions for high-order convergence of linearly implicit methods at smooth solutions.

2. Analytical framework and numerical method

In this section we state the assumptions on the evolution equation. Moreover we introduce the numerical method.

2.1 *Evolution equation*

We consider a semilinear parabolic equation of the form

$$u' + Au = f(t, u), \qquad 0 < t \leqslant T \tag{2.1a}$$

$$u(0) = u_0. \tag{2.1b}$$

This abstract evolution equation is given on a Banach space $(X, |\cdot|)$. The domain of the linear operator A on X is denoted by $\mathcal{D}(A)$, and the initial value $u_0 \in V$ is chosen in an interpolation space $\mathcal{D}(A) \subset V \subset X$ which will be specified below. Our basic assumptions on the initial value problem (2.1) are that of Henry (1981).

ASSUMPTION 2.1 Let $A : \mathcal{D}(A) \subset X \to X$ be sectorial, i.e. A is a densely defined and closed linear operator on X satisfying the resolvent condition

$$\left|(\lambda I + A)^{-1}\right|_{X \leftarrow X} \leqslant \frac{M}{|\lambda - \omega|} \tag{2.2}$$

on the sector $\{\lambda \in \mathbb{C}\,;\, |\arg(\lambda - \omega)| \leqslant \pi - \varphi\}$ for $M \geqslant 1$, $\omega \in \mathbb{R}$, and $0 \leqslant \varphi < \frac{\pi}{2}$.

Under this hypothesis, the operator $-A$ is the infinitesimal generator of an analytic semigroup $\{\mathrm{e}^{-tA}\}_{t \geqslant 0}$ which renders (2.1) parabolic. In the sequel we set

$$A_a = A + aI \qquad \text{for some } a > \omega.$$

For this operator, the fractional powers are well defined. We choose $0 \leqslant \alpha < 1$ and define $V = \mathcal{D}(A_a^\alpha)$ which is a Banach space with norm $\|v\| = |A_a^\alpha v|$. Note that this definition does not depend on a, since different choices of a lead to equivalent norms.

We are now ready to give our hypothesis on the nonlinear function f.

ASSUMPTION 2.2 Let $f : [0, T] \times V \to X$ be locally Lipschitz-continuous. Thus there exists a real number $L(R, T)$ such that

$$|f(t_1, v_1) - f(t_2, v_2)| \leqslant L(|t_1 - t_2| + \|v_1 - v_2\|) \tag{2.3}$$

for all $t_i \in [0, T]$ and $\|v_i\| \leqslant R$, $i = 1, 2$.

Reaction–diffusion equations and the incompressible Navier–Stokes equations can be cast into this abstract framework. This is verified in Section 3 of Henry (1981) and in Lubich & Ostermann (1996). For a more general class of reaction–diffusion equations that is included in our framework, we refer to Section 8.4 of Pazy (1983).

We do not distinguish between a norm and its corresponding operator norm. For elements $x = (x_1, \ldots, x_s)$ in a product space, we set $|x| = \max(|x_1|, \ldots, |x_s|)$ and $\|x\| = \max(\|x_1\|, \ldots, \|x_s\|)$, respectively. The norm of linear operators from X^s to V^s is denoted by $\|\cdot\|_{V \leftarrow X}$.

2.2 *Numerical method*

In this paper linearly implicit Runge–Kutta discretizations of parabolic problems are studied. In the sequel we will review these methods in brief. For detailed descriptions, refer to the monographs by Deuflhard & Bornemann (1994), Hairer & Wanner (1996), and Strehmel & Weiner (1992).

A *linearly implicit Runge–Kutta method* with constant stepsize $h > 0$, applied to the initial value problem (2.1), yields an approximation u_n to the value of the solution u at $t_n = nh$ and is given by the internal stages

$$\begin{aligned} U'_{ni} + AU_{ni} &= f(t_n + \alpha_i h, U_{ni}) + hJ_n \sum_{j=1}^{i} \gamma_{ij} U'_{nj} + h\gamma_i g_n \\ U_{ni} &= u_n + h \sum_{j=1}^{i-1} \alpha_{ij} U'_{nj}, \qquad 1 \leqslant i \leqslant s \end{aligned} \tag{2.4a}$$

and the one-step recursion

$$u_{n+1} = u_n + h \sum_{j=1}^{s} b_j U'_{nj}. \tag{2.4b}$$

Here J_n and g_n are approximations to the derivatives of $-Au + f(t, u)$ with respect to the variables u and t

$$J_n \approx -A + D_u f(t_n, u_n), \qquad g_n \approx D_t f(t_n, u_n).$$

The real numbers $\alpha_{ij}, \gamma_{ij}, b_i, \alpha_i, \gamma_i$ are the coefficients of the method. We always assume that $\gamma_{ii} > 0$ for all i.

In contrast to fully implicit Runge–Kutta methods, where the numerical approximation is given as the solution of nonlinear equations, u_{n+1} is obtained from u_n by solving only linear equations.

In order to write the numerical method more compactly, we introduce the following matrix and vector notations

$$\Gamma = (\gamma_{ij})_{1\leqslant i,j\leqslant s}, \qquad \mathcal{O} = (a_{ij})_{1\leqslant i,j\leqslant s}, \qquad 1\!\!1 = (1, \dots, 1)^T \in \mathbb{R}^s, \tag{2.5a}$$

where $a_{ij} = \alpha_{ij} + \gamma_{ij}$ with $\alpha_{ij} = 0$ for $i \leqslant j$ and $\gamma_{ij} = 0$ for $i < j$. Further we set

$$\alpha = (\alpha_1, \dots, \alpha_s)^T, \qquad \gamma = (\gamma_1, \dots, \gamma_s)^T \tag{2.5b}$$

$$b = (b_1, \dots, b_s)^T, \qquad c = (c_1, \dots, c_s)^T = \mathcal{O} 1\!\!1. \tag{2.5c}$$

The numerical scheme has *order* p if the error of the method, applied to ordinary differential equations with sufficiently differentiable right-hand side, fulfils the relation $u_n - u(t_n) = \mathcal{O}(h^p)$ for $h \to 0$, uniformly on bounded time intervals.

A linearly implicit Runge–Kutta method is $A(\vartheta)$*-stable* if the absolute value of the *stability function*,

$$\mathcal{R}(z) = 1 + zb^T(I - z\mathcal{O})^{-1} 1\!\!1, \tag{2.6}$$

is bounded by one for all $z \in M_\vartheta = \{z \in \mathbb{C}\,; |\arg(-z)| \leqslant \vartheta\}$. Note that $(I - z\mathcal{O})$ is invertible in M_ϑ since all γ_{ii} are positive. The numerical method is called *strongly* $A(\vartheta)$*-stable* if in addition the absolute value of $\mathcal{R}$ at infinity, $\mathcal{R}(\infty) = 1 - b^T\mathcal{O}^{-1} 1\!\!1$, is strictly smaller than one.

Two types of linearly implicit Runge–Kutta methods are of particular interest. *Rosenbrock methods* satisfy the conditions

$$\alpha_i = \sum_{j=1}^{i-1} \alpha_{ij}, \qquad \gamma_i = \sum_{j=1}^{i} \gamma_{ij} \tag{2.7}$$

and use the exact Jacobians

$$J_n = -A + D_u f(t_n, u_n), \qquad g_n = D_t f(t_n, u_n). \tag{2.8}$$

As a prominent example, we mention the fourth-order method RODAS from Hairer & Wanner (1996). It is strongly $A(\pi/2)$-stable and satisfies $\mathcal{R}(\infty) = 0$.

A second important class of linearly implicit methods is determined by the requirements

$$\alpha_i = \sum_{j=1}^{i-1} \alpha_{ij} + \sum_{j=1}^{i} \gamma_{ij}, \qquad \text{and} \qquad J_n = -A, \quad g_n = 0.$$

In this paper such methods are called *W-methods*. This differs from the common diction in the literature where this term is often used as a synonym for linearly implicit Runge–Kutta methods. The operator A and the nonlinearity f are not determined uniquely, since bounded parts of A, e.g., can be included into f. Therefore the assumption $J_n = -A$ is not as restrictive as it may seem at first. As an example of W-methods, we mention the extrapolated linearly implicit Euler method which is described briefly in Section 3 of Lubich & Ostermann (1995), see also Hairer & Wanner (1996). It is strongly $A(\vartheta)$-stable with $\vartheta \approx \pi/2$ and satisfies $\mathcal{R}(\infty) = 0$.

3. Non-smooth data error estimates

In Theorem 3.1 below we state the main result of this paper. We give a non-smooth data error estimate for a general class of linearly implicit methods. For smooth initial data, their convergence is studied in Lubich & Ostermann (1995), Ostermann & Roche (1993), and Schwitzer (1995).

THEOREM 3.1 Let (2.1) satisfy Assumptions 2.1 and 2.2, and let $u_0 \in V$ be such that the solution u remains bounded in V for $0 \leqslant t \leqslant T$. Apply a strongly $A(\vartheta)$-stable linearly implicit Runge–Kutta method of order at least one with $\vartheta > \varphi$ to this initial value problem, and assume that $\|J_n + A\|_{X \leftarrow V}$ as well as $|g_n|$ are uniformly bounded for $0 \leqslant t_n \leqslant T$. Then there exist constants h_0 and C such that for all stepsizes $0 < h \leqslant h_0$ the numerical solution u_n satisfies the estimate

$$\|u_n - u(t_n)\| \leqslant C\Big(t_n^{-1} h + t_n^{-\alpha} h\, |\log h|\Big) \qquad \text{for } 0 < t_n \leqslant T.$$

The constants h_0 and C depend on T and the bound of u, on the quantities appearing in Assumptions 2.1 and 2.2, and moreover on the numerical method.

This result can be applied directly to W-methods and Rosenbrock methods. For W-methods this is obvious since $J_n = -A$ and $g_n = 0$. For Rosenbrock methods we have to suppose that the first derivatives of the nonlinearity f are locally bounded. Then, due to (2.8), Theorem 3.1 is applicable.

To study the long-term dynamics of the evolution equation (2.1), apart from a non-smooth data error estimate for finite times, an error estimate for the derivative of the solution with respect to the initial value is often needed, see Stuart (1995). This derivative, evaluated at the point u_0, is a linear operator on V and is denoted here by $v(t) = Du(t; u_0)$. Consequently (u, v) satisfies the system

$$\begin{pmatrix} u' \\ v' \end{pmatrix} + \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} f(t, u) \\ D_u f(t, u) v \end{pmatrix} \tag{3.1}$$

with initial value $(u_0, v_0)^T$, where v_0 is the identity on V. The derivative of the numerical solution u_n with respect to the initial value is denoted by $v_n = Du_n(u_0)$. It is just the second component of the linearly implicit Runge–Kutta solution of (3.1) at $t_n = nh$.

We are now in a position to state the following result.

COROLLARY 3.1 In addition to the assumptions of Theorem 3.1, let the Fréchet derivative $D_u f(t, u)$ be locally Lipschitz-continuous with respect to the variables t and u, and bounded as a linear operator from V to X, uniformly in t and u. Then there exist constants h_0 and C such that for $0 < h \leqslant h_0$ the estimate

$$\|v_n - v(t_n)\| \leqslant C\Big(t_n^{-1} h + t_n^{-\alpha} h\,|\log h|\Big), \qquad 0 < t_n \leqslant T$$

is satisfied. Apart from the quantities given in Theorem 3.1, the maximum stepsize h_0 and the constant C depend on the Lipschitz constants of $D_u f$.

Proof of Corollary 3.1. Obviously (3.1) satisfies Assumptions 2.1 and 2.2. In order to apply Theorem 3.1, it remains to show that $v(t)$ is bounded by a constant, uniformly for $0 \leqslant t \leqslant T$. By means of the variation-of-constants formula, v can be represented as

$$v(t) = \mathrm{e}^{-tA} + \int_0^t \mathrm{e}^{-(t-\tau)A} D_u f(\tau, u(\tau))\, v(\tau)\, \mathrm{d}\tau,$$

see Henry (1981, Lemma 3.3.2). Applying the estimates given in Lemma 6.3 (see later), the boundeness of v follows from a Gronwall inequality given in Section 1.2.1 of Henry (1981). □

Proof of Theorem 3.1. Our basic idea is to compare the numerical solution, obtained with the linearly implicit method, with the solution of a suitably chosen implicit Runge–Kutta method.

(a) First we apply a linearly implicit method to (2.1) which gives (2.4). In order to write (2.4) more compactly, we employ the following vector notation

$$\begin{aligned} &U_n = (U_{n1}, \ldots, U_{ns})^T, \quad U_n' = (U_{n1}', \ldots, U_{ns}')^T \\ &F_n = \big(f(t_n + \alpha_1 h, U_{n1}), \ldots, f(t_n + \alpha_s h, U_{ns})\big)^T. \end{aligned} \tag{3.2}$$

Together with (2.5) we get

$$U_n' + (\mathcal{I} \otimes A)U_n = F_n + (\Gamma \otimes hJ_n)U_n' + \gamma \otimes hg_n \tag{3.3a}$$

$$U_n = \mathbb{1} \otimes u_n + \big((\mathcal{Q} - \Gamma) \otimes hI\big)U_n' \tag{3.3b}$$

$$u_{n+1} = u_n + (b^T \otimes hI)U_n'. \tag{3.3c}$$

Here we have used Kronecker product notation. Thus the (k, m)-th component of $\mathcal{B} \otimes A$, where A is a linear operator and $\mathcal{B}$ an arbitrary matrix with coefficients b_{ij}, is given by $b_{km}A$. For notational simplicity, we write $\mathcal{B} \otimes hA$ instead of $\mathcal{B} \otimes (hA)$. We further distinguish between the identity matrix $\mathcal{I}$ on $\mathbb{R}^s$ and the identity operator I on X or V.

Inserting (3.3b) into (3.3a) and setting $D_n = J_n + A$, we get

$$U_n' = (\mathcal{I} \otimes I + \mathcal{O} \otimes hA)^{-1}\big(-1\!\!1 \otimes (Au_n) + F_n + (\Gamma \otimes hD_n)U_n' + \gamma \otimes hg_n\big).$$

Together with (3.3c) this yields the recursion

$$u_{n+1} = \mathcal{R}(-hA)\, u_n + (b^T \otimes hI)(\mathcal{I} \otimes I + \mathcal{O} \otimes hA)^{-1} \big(F_n + (\Gamma \otimes hD_n)U_n' + \gamma \otimes hg_n\big). \quad (3.4)$$

(b) Next we compare this numerical solution with the following implicit Runge–Kutta discretization

$$\widetilde{U}_n' + (\mathcal{I} \otimes A)\widetilde{U}_n = \widetilde{F}_n \quad (3.5a)$$

$$\widetilde{U}_n = 1\!\!1 \otimes \widetilde{u}_n + (\mathcal{O} \otimes hI)\widetilde{U}_n' \quad (3.5b)$$

$$\widetilde{u}_{n+1} = \widetilde{u}_n + (b^T \otimes hI)\widetilde{U}_n'. \quad (3.5c)$$

Here we have used the same abbreviations as in (3.2) (replacing U_n with $\widetilde{U}_n$, etc). In particular we set

$$\widetilde{F}_n = \big(f(t_n + c_1 h, \widetilde{U}_{n1}), \ldots, f(t_n + c_s h, \widetilde{U}_{ns})\big)^T \qquad \text{with} \quad c = \mathcal{O}1\!\!1.$$

A similar calculation as before yields

$$\widetilde{u}_{n+1} = \mathcal{R}(-hA)\, \widetilde{u}_n + (b^T \otimes hI)(\mathcal{I} \otimes I + \mathcal{O} \otimes hA)^{-1} \widetilde{F}_n. \quad (3.6)$$

(c) The difference between the linearly implicit solution u_n and the Runge–Kutta solution $\widetilde{u}_n$ is denoted by $e_n = u_n - \widetilde{u}_n$. In accordance with that, E_n and E_n' are defined by $E_n = U_n - \widetilde{U}_n$ and $E_n' = U_n' - \widetilde{U}_n'$, respectively. Taking the difference between (3.4) and (3.6) gives

$$e_{n+1} = \mathcal{R}(-hA)\, e_n + (b^T \otimes hI)(\mathcal{I} \otimes I + \mathcal{O} \otimes hA)^{-1} \big(F_n - \widetilde{F}_n + (\Gamma \otimes hD_n)U_n' + \gamma \otimes hg_n\big).$$

Solving this recursion yields

$$e_{n+1} = (b^T \otimes hI) \sum_{\nu=0}^{n} (\mathcal{I} \otimes I + \mathcal{O} \otimes hA)^{-1} \big(\mathcal{I} \otimes \mathcal{R}(-hA)^{n-\nu}\big) \big(F_\nu - \widetilde{F}_\nu + (\Gamma \otimes hD_\nu)U_\nu' + \gamma \otimes hg_\nu\big), \quad (3.7)$$

where we have already used the fact that both methods start with the same initial value u_0. Since f is locally Lipschitz-continuous, we have

$$|F_n - \widetilde{F}_n| \leqslant L\big(h \max_{1\leqslant i\leqslant s} |\alpha_i - c_i| + \|E_n\|\big)$$

for $\|U_n\|$, $\|\widetilde{U}_n\| \leqslant R$. We suppose for a moment that the radius R can be chosen independently of n. This will be justified at the end of the proof.

From the last equation and the uniform boundedness of D_n and g_n we get

$$\|e_{n+1}\| \leqslant Ch \sum_{\nu=0}^{n} \left\| (\mathcal{I} \otimes I + \mathcal{O} \otimes hA)^{-1} \left(\mathcal{I} \otimes \mathcal{R}(-hA)^{n-\nu}\right) \right\|_{V \leftarrow X} \Big(\|E_\nu\| + h \left\|U'_\nu\right\| + h\Big). \quad (3.8)$$

(d) We now derive several relations that are necessary to bound e_{n+1}. First we consider hU'_n. From (3.3b) and (3.5b) we get

$$(\Gamma \otimes hI)U'_n = 1\!\!1 \otimes e_n + (\mathcal{O} \otimes hI)E'_n - E_n. \quad (3.9)$$

In order to eliminate E'_n, we multiply (3.9) by $(\mathcal{I} \otimes J_n)$ and insert it into the difference of (3.3a) and (3.5a). This yields

$$\begin{aligned}(\mathcal{I} \otimes I + \mathcal{O} \otimes hA)E'_n = &-1\!\!1 \otimes (Ae_n) + 1\!\!1 \otimes (D_n e_n) + F_n - \widetilde{F}_n \\ &+ (\mathcal{O} \otimes hD_n)E'_n - (\mathcal{I} \otimes D_n)E_n + \gamma \otimes hg_n. \quad (3.10)\end{aligned}$$

We multiply this identity by $(\mathcal{I} \otimes I + \mathcal{O} \otimes hA)^{-1}$. The existence of this operator is guaranteed by Lemma 6.5. Applying (6.2) to the term involving Ae_n and (6.3) with $\rho = \alpha$ to the remaining expressions gives

$$h\|E'_n\| \leqslant C\,\|e_n\| + Ch^{1-\alpha}\|E_n\| + Ch^{2-\alpha}$$

for h sufficiently small. Together with (3.9) we get

$$h\|U'_n\| \leqslant C\,\|e_n\| + C\,\|E_n\| + Ch^{2-\alpha}. \quad (3.11)$$

It remains to express E_n in terms of e_n. Regrouping (3.9) we have

$$E_n = 1\!\!1 \otimes e_n + ((\mathcal{O} - \Gamma) \otimes hI)E'_n - (\Gamma \otimes hI)\widetilde{U}'_n. \quad (3.12)$$

In order to estimate $h\widetilde{U}'_n$, we use (3.5b) in the form

$$(\mathcal{O} \otimes hI)\widetilde{U}'_n = -1\!\!1 \otimes \widetilde{u}_n + \widetilde{U}_n. \quad (3.13)$$

Each component on the right-hand side of (3.13) can be written as

$$u(t_n) - \widetilde{u}_n + \widetilde{U}_{ni} - u(t_n + c_i h) + \int_{t_n}^{t_n + c_i h} u'(\tau)\,\mathrm{d}\tau. \quad (3.14)$$

Applying the triangular inequality as well as Lemma 6.4 and Lemma A.1 gives

$$h\|\widetilde{U}'_n\| \leqslant C\Big(t_n^{-1}h + t_n^{-\alpha}h\,|\log h|\Big) \qquad \text{for } n \geqslant 1.$$

Therefore we finally get from (3.12)

$$\|E_n\| \leqslant C\,\|e_n\| + C\Big(t_n^{-1}h + t_n^{-\alpha}h\,|\log h|\Big) \qquad \text{for } n \geqslant 1. \quad (3.15)$$

Note that $\|E_0\|$ and thus $h\|U_0'\|$ are bounded by a constant.

(e) Inserting (3.11) and (3.15) into (3.8), we obtain with (6.4) and Lemma 6.1

$$\|e_n\| \leqslant Ch \sum_{\nu=1}^{n-1} t_{n-\nu}^{-\alpha} \|e_\nu\| + C\Big(t_n^{-1}h + t_n^{-\alpha}h\,|\log h|\Big)\,.$$

Applying the discrete Gronwall Lemma 6.2, we get

$$\|u_n - \widetilde{u}_n\| = \|e_n\| \leqslant C\Big(t_n^{-1}h + t_n^{-\alpha}h\ |\log h|\Big) \tag{3.16}$$

due to linearity. The desired estimate

$$\|u_n - u(t_n)\| \leqslant \|u_n - \widetilde{u}_n\| + \|\widetilde{u}_n - u(t_n)\| \leqslant C\Big(t_n^{-1}h + t_n^{-\alpha}h\ |\log h|\Big)$$

finally follows with Lemma A.1.

(f) We still have to show that the numerical solution remains in a ball of radius R. Note that the exact solution as well as the Runge–Kutta solution are bounded on $[0, T]$. We take R sufficiently large and choose a smooth cut-off function

$$\chi : V \to [0, 1] \qquad \text{with} \quad \chi(v) = \begin{cases} 1 & \text{if } \|v\| \leqslant R, \\ 0 & \text{if } \|v\| \geqslant 2R. \end{cases}$$

Since $f(t, \chi(u)\cdot u)$ has a global Lipschitz constant, we infer from (3.16) that the numerical solution, obtained with this new f, is bounded by R for h sufficiently small. It thus coincides with the numerical solution, obtained with the original f. This concludes the proof of Theorem 3.1. □

4. Applications

As already mentioned in the introduction, Theorem 3.1 together with Corollary 3.1 can be used to study the question of whether the continuous dynamics of a parabolic equation is correctly represented in its discretization. A result in this direction is given in Lubich & Ostermann (1996) for Runge–Kutta discretizations of periodic orbits. The proof there carries over literally to linearly implicit Runge–Kutta methods. Note that the necessary bounds for smooth initial data are provided by Lubich & Ostermann (1995) and Schwitzer (1995). We do not give the details here.

Another immediate consequence of our non-smooth data error estimates are long-term error bounds. As an illustration, we show below that exponentially stable solutions of (2.1a) are uniformly approximated by linearly implicit Runge–Kutta methods over arbitrarily long time intervals. Our presentation follows an idea of Larsson (1992). Alternatively one might use directly the results of Stuart (1995). A close examination of their proofs shows that they are applicable despite the additional $|\log h|$ term in our error estimate.

We recall that a solution u of (2.1a) is *exponentially stable* if there exist positive constants τ and δ such that any solution v of (2.1a) with initial value $v(\tau_0) \in V$ and $\|u(\tau_0) - v(\tau_0)\| \leqslant \delta$ satisfies

$$\|u(t) - v(t)\| \leqslant \tfrac{1}{2}\|u(\tau_0) - v(\tau_0)\| \qquad \text{for } t \geqslant \tau_0 + \tau. \tag{4.1}$$

This condition holds, for example, in the neighbourhood of an asymptotically stable fixed-point due to its exponential attractivity.

THEOREM 4.1 In addition to the assumptions of Theorem 3.1, let the solution u be exponentially stable and globally bounded. Then, for any choice of $t^* > 0$, there are positive constants C and h_0 such that for all stepsizes $0 < h \leqslant h_0$ we have

$$\|u_n - u(t_n)\| \leqslant Ch\,|\log h| \qquad \text{for } t_n \in [t^*, \infty). \tag{4.2}$$

The constant C depends on t^* and on τ, given by (4.1). Moreover it depends on the quantities appearing in Assumptions 2.1 and 2.2, on the numerical method, and on the bound for the solution.

It is remarkable that (4.2) holds for quite crude approximations to the Jacobian. For example, the choice $J_n = A$ is possible without any assumption on the growth of the semigroup, i.e. on the sign of the constant ω appearing in (2.2).

Proof. Henceforth, the constants δ and τ have the same meaning as in the definition of exponentially stable solutions. Since the solution u is globally bounded in V, it stays in a ball of radius $R/2$, say. We may assume that $t^* \leqslant \tau$ and set $T = 2\tau + t^*$. Then Theorem 3.1 shows the existence of a constant $C^* = C^*(R, t^*, T)$ with

$$\|u_n - u(t_n)\| \leqslant C^* h\,|\log h| \qquad \text{for } t^* \leqslant t_n \leqslant T \text{ and } 0 < h \leqslant h_0. \tag{4.3a}$$

After a possible reduction of δ and h_0, we have $h_0 \leqslant \tau$ and

$$C^* h\,|\log h| \leqslant \delta/2 \leqslant R/4 \qquad \text{for } 0 < h \leqslant h_0. \tag{4.3b}$$

Assume for a moment that the estimate

$$\|u_n - u(t_n)\| \leqslant 2C^* h\,|\log h|\,, \qquad t^* + k\tau < t_n \leqslant t^* + (k+1)\tau$$

holds for some $k \geqslant 2$, and let m be such that $(m-1)h < \tau \leqslant mh$. Further denote by $v_n(t)$ the solution of (2.1a) with initial value $v_n(t_n) = u_n$. From Theorem 3.1 and the exponential stability (4.1), we get

$$\begin{aligned}\|u_{n+m} - u(t_{n+m})\| &\leqslant \|u_{n+m} - v_n(t_{n+m})\| + \|v_n(t_{n+m}) - u(t_{n+m})\| \\ &\leqslant C^* h\,|\log h| + \tfrac{1}{2}\|u_n - u(t_n)\| \leqslant 2C^* h\,|\log h| \leqslant \delta.\end{aligned}$$

The bound (4.2) with $C = 2C^*$ thus follows from (4.3) by induction. □

The above result can also be used to obtain stability bounds for splitting methods. As an example, we consider the linear problem

$$u' + Au = Bu, \qquad u(0) = u_0 \tag{4.4}$$

and its discretization by the linearly implicit Euler method

$$u_{n+1} = (I + hA)^{-1}(I + hB)u_n.$$

Since A and B are treated in a different way, this scheme can be interpreted as a splitting method.

We assume that the operator A satisfies Assumption 2.1 and that B is bounded as an operator from V to X. Thus $B - A$ is the infinitesimal generator of an analytic semigroup on V, see Corollary 1.4.5 of Henry (1981). We further suppose that this semigroup satisfies

$$\left\| e^{-t(A-B)} \right\| \leqslant C e^{-\kappa t} \qquad \text{for } t \geqslant 0 \tag{4.5}$$

with some $\kappa > 0$. We then have the following result.

COROLLARY 4.1 Under the above assumptions, for any $\widetilde{\kappa} < \kappa$, there are positive constants C and h_0 such that for $0 < h \leqslant h_0$

$$\left\| \left((I + hA)^{-1} (I + hB) \right)^n \right\| \leqslant C e^{-\tilde{\kappa} n h} \qquad \text{for } n \geqslant 1.$$

The constant C depends on κ and $\widetilde{\kappa}$, on the quantities appearing in Assumption 2.1, and on $\|B\|_{X \leftarrow V}$.

This proves the stability of the above splitting scheme for sufficiently small stepsizes. We are not aware of any other proof for this result, apart from the case $\alpha = 0$ where B has to be bounded on X.

Proof. For given κ, we choose $0 \leqslant \widetilde{\kappa} < \mu < \kappa$ and consider the equation

$$w' + Aw = \widetilde{B} w \qquad \text{with} \quad \widetilde{B} = \mu I + (1 + h\mu) B.$$

For h sufficiently small, the solutions of this problem are exponentially stable, since there exists some $\varepsilon > 0$ such that

$$\left\| e^{-t(A - \widetilde{B})} \right\| \leqslant C e^{-\varepsilon t} \qquad \text{for } t \geqslant 0.$$

This follows from Theorem 3.2.1 of Pazy (1983). Note that $\widetilde{B}$ has a Lipschitz constant that depends on μ, h_0, $\|B\|_{X \leftarrow V}$, but not on h. Due to linearity, we obtain from Theorem 3.1 and Theorem 4.1 the estimate

$$\left\| \left((I + hA)^{-1} (I + h\widetilde{B}) \right)^n \right\| \leqslant C \qquad \text{for } n \geqslant 1.$$

The desired bound finally follows from $I + h\widetilde{B} = (1 + h\mu)(I + hB)$. □

5. Refined error estimate

Theorem 3.1 essentially yields convergence of order one. In this section we show that we can raise the order of convergence under slightly stronger assumptions on the data. To be more specific, we require that f satisfies the following property.

ASSUMPTION 5.1 Let $f : [0, T] \times V \to X$ be locally Lipschitz-continuous with respect to the norms $\|A_a^{-\beta} \cdot \|$ and $|A_a^{-\beta} \cdot |$ for some $0 < \beta < 1 - \alpha$, i.e.

$$\left| A_a^{-\beta} \left(f(t_1, v_1) - f(t_2, v_2) \right) \right| \leqslant L \left(|t_1 - t_2| + \|A_a^{-\beta} (v_1 - v_2)\| \right) \tag{5.1}$$

for all $t_i \in [0, T]$ and $v_i \in V$ with $\|v_i\| \leqslant R$, $i = 1, 2$. Further suppose that the first- and second-order derivatives of f are locally Lipschitz bounded with respect to these norms also.

Let $X^\rho = \mathcal{D}(A_a^\rho)$ for $\rho > 0$, and let $X^{-\rho}$ denote the completion of X with respect to the norm $|A_a^{-\rho} \cdot |$. We require that the approximations D_n to the Jacobian $D_u f(t_n, u_n)$ are uniformly bounded as mappings from $X^{\alpha-\beta}$ to $X^{-\beta}$, i.e.

$$\|A_a^{-\beta} D_n A_a^\beta\|_{X \leftarrow V} \leqslant C \qquad \text{for } 0 \leqslant t_n < T, \tag{5.2}$$

with the same β as above.

For the convenience of the reader, we recall that the coefficients of a Rosenbrock method of order $p = 2$ satisfy

$$b^T \mathbb{1} = 1 \qquad \text{and} \qquad b^T \mathcal{O}\mathbb{1} = b^T c = \tfrac{1}{2}, \tag{5.3a}$$

whereas general linearly implicit Runge–Kutta methods of order two further fulfil the order conditions

$$b^T \alpha = \tfrac{1}{2}, \qquad b^T \gamma = 0, \qquad b^T \Gamma \mathbb{1} = 0. \tag{5.3b}$$

We are now in a position to state the refined error estimate.

THEOREM 5.1 In addition to the assumptions of Theorem 3.1, let Assumption 5.1 and (5.2) hold. Further, suppose that the method (2.4) has order $p \geqslant 2$. Then there exist constants h_0 and C such that for all stepsizes $0 < h \leqslant h_0$ the numerical solution u_n satisfies the estimate

$$\|u_n - u(t_n)\| \leqslant C\, t_n^{-1-\beta} h^{1+\beta} \qquad \text{for } 0 < t_n \leqslant T.$$

The constants h_0 and C depend on the constants appearing in Assumption 5.1 and in (5.2), as well as on the quantities specified in Theorem 3.1.

Proof. This proof is an extension of the proof of Theorem 3.1. We thus concentrate on those aspects that go beyond that proof.

(a) We have to estimate the difference $F_n - \widetilde{F}_n$ in (3.7) more carefully. Taylor series expansion gives

$$F_n - \widetilde{F}_n = (\alpha - c) \otimes h D_t f(t_n, u_n) + \big(\mathcal{I} \otimes D_u f(t_n, u_n)\big) E_n + \Delta_n. \tag{5.4}$$

We note for later use that the remainder Δ_n is bounded by

$$\left|\left(\mathcal{I} \otimes A_a^{-\beta}\right) \Delta_n\right| \leqslant C\, \|e_n\| + C\, t_n^{\beta-1} h^{1+\beta}. \tag{5.5}$$

This follows from Assumption 5.1, the preliminary bound $\|e_n\| \leqslant C$, and

$$\begin{aligned} h \left\|\left(\mathcal{I} \otimes A_a^{-\beta}\right) E_n'\right\| &\leqslant C\Big(h^\beta \|e_n\| + t_n^{\beta-1} h^{1+\beta}\Big) \\ h \left\|\left(\mathcal{I} \otimes A_a^{-\beta}\right) U_n'\right\| + \left\|\left(\mathcal{I} \otimes A_a^{-\beta}\right) E_n\right\| &\leqslant C\Big(\|e_n\| + t_n^{\beta-1} h\Big). \end{aligned} \tag{5.6}$$

The boundedness of e_n is an immediate consequence of Theorem 3.1, whereas (5.6) is obtained in a similar way to the bound for $h\|E_n'\|$. Using (5.1) and (5.2), we get from (3.10)

$$h\|(\mathcal{I} \otimes A_a^{-\beta}) E_n'\| \leqslant C\left(h^\beta \|e_n\| + h^{1-\alpha} \|(\mathcal{I} \otimes A_a^{-\beta}) E_n\| + h^{2-\alpha} |A^{-\beta} g_n|\right),$$

and since $\alpha + \beta < 1$, this implies

$$h\|(\mathcal{I} \otimes A_a^{-\beta})E_n'\| \leqslant C\left(h^\beta \|e_n\| + h^\beta \|(\mathcal{I} \otimes A_a^{-\beta})E_n\| + h^{1+\beta}\right). \tag{5.7}$$

Further, a direct estimate of (3.12) gives

$$\|(\mathcal{I} \otimes A_a^{-\beta})E_n\| \leqslant C\Big(\|e_n\| + h\|(\mathcal{I} \otimes A_a^{-\beta})E_n'\| + h\|(\mathcal{I} \otimes A_a^{-\beta})\widetilde{U}_n'\|\Big).$$

We thus have to bound $h(\mathcal{I} \otimes A_a^{-\beta})\widetilde{U}_n'$. Using Lemma 6.4 and (A.4), we obtain from (3.13)

$$h\|(\mathcal{I} \otimes A_a^{-\beta})\widetilde{U}_n'\| \leqslant C\, t_n^{\beta-1} h.$$

Reinserting this bound into the above estimates together with (3.9) finally gives (5.6).

(b) We first give the proof for Rosenbrock methods. Recall that in this case, the identities $D_n = D_u f(t_n, u_n)$ and $g_n = D_t f(t_n, u_n)$ as well as $\alpha + \gamma = c$ hold. The latter follows from (2.7) and (2.5). From (5.4) we obtain with (3.9)

$$F_n - \widetilde{F}_n + (\Gamma \otimes hD_n)U_n' + \gamma \otimes hg_n = 1\!\!1 \otimes (D_n e_n) + (\mathcal{O} \otimes hD_n)E_n' + \Delta_n. \tag{5.8}$$

We now insert this relation into (3.7) and start to estimate the recursion more carefully. For this we denote the left-hand side of (5.8) by x and the operator on the right-hand side of (3.7) by B. Using

$$|A_a^\alpha Bx| \leqslant |A_a^{\alpha+\beta}B| \cdot |A_a^{-\beta}x|$$

together with (5.5), (5.6) and Lemma 6.5, we obtain

$$\|e_n\| \leqslant Ch\sum_{\nu=1}^{n-1} t_{n-\nu}^{-\alpha-\beta}\|e_\nu\| + C\, t_n^{\beta-1}h^{1+\beta}. \tag{5.9}$$

The discrete Gronwall Lemma 6.2 and the corresponding bound (A.3) for Runge–Kutta methods finally yield the desired result.

(c) For general linearly implicit Runge–Kutta methods, the identities $D_n = D_u f(t_n, u_n)$ and $g_n = D_t f(t_n, u_n)$ are not necessarily valid. Instead, we have to use the additional order conditions (5.3), combined with an elimination process. We illustrate this with the term $(\alpha - c) \otimes hD_t f(t_n, u_n)$ from (5.4). Inserted in (3.7), it gives

$$(b^T \otimes hI)\sum_{\nu=0}^{n}(\mathcal{I} \otimes I + \mathcal{O} \otimes hA)^{-1}\left(\mathcal{I} \otimes \mathcal{R}(-hA)^{n-\nu}\right)(\alpha - c) \otimes hD_t f(t_\nu, u_\nu) \tag{5.10}$$

where a direct estimate would only give order one. We first split

$$\mathcal{R}(-hA)^n = r^n + \left(\mathcal{R}(-hA)^n - r^n\right) \qquad \text{with } r = \mathcal{R}(\infty).$$

The term with r^n can be estimated as in the proof of Theorem 3.1. Since $|r| < 1$, we get an additional factor $h^{1-\alpha}$ and hence the desired factor h^β. For the second term, we use the identity

$$(b^T \otimes I)(\mathcal{I} \otimes I + \mathcal{O} \otimes hA)^{-1} = b^T \otimes I - (b^T\mathcal{O} \otimes hA)(\mathcal{I} \otimes I + \mathcal{O} \otimes hA)^{-1}$$

together with the order conditions (5.3). This yields

$$\begin{aligned}&\big\|(b^T \otimes I)(\mathcal{I} \otimes I + \mathcal{O} \otimes hA)^{-1}\big((\alpha - c) \otimes \big(\mathcal{R}(-hA)^n - r^n\big)\big)\big\|_{V \leftarrow X}\\ &\quad \leqslant Ch^{\beta}\big|\big(b^T \mathcal{O} \otimes (hA_a)^{1-\beta}\big)(\mathcal{I} \otimes I + \mathcal{O} \otimes hA)^{-1}\big| \cdot \big|A_a^{\alpha+\beta}\big(\mathcal{R}(-hA)^n - r^n\big)\big|.\end{aligned}$$

An application of Lemma 6.5 thus shows that (5.10) gives a contribution of order $h^{1+\beta}$. The other terms in (3.7) are treated similarly and we again obtain (5.9). This concludes the proof. □

6. Lemmas for Section 3 and Section 5

In this section we collect several results that we have used in the proofs of Theorem 3.1, Corollary 3.1, and Theorem 5.1. We start with a discrete convolution of weakly singular functions.

LEMMA 6.1 For $n \in \mathbb{N}$ and $h > 0$, let $t_n = nh$. Then the following relation holds for $0 \leqslant \rho < 1$

$$h \sum_{\nu=1}^{n-1} t_{n-\nu}^{-\rho} t_{\nu}^{-\sigma} \leqslant \begin{cases} C t_n^{1-\rho-\sigma} & \text{for } 0 \leqslant \sigma < 1, \\ C t_n^{-\rho} \,|\log h| & \text{for } \sigma = 1, \\ C t_n^{1-\rho-\sigma} n^{\sigma-1} & \text{for } \sigma > 1. \end{cases}$$

Proof. We interpret the left-hand side as a Riemann-sum and estimate it by the corresponding integral. □

An integrable function $\varepsilon : [0, T] \to \mathbb{R}$ with the property

$$0 \leqslant \varepsilon(t) \leqslant a \int_0^t (t-\tau)^{-\rho} \varepsilon(\tau)\,\mathrm{d}\tau + b\,t^{-\sigma} \qquad \text{for } 0 \leqslant \rho, \sigma < 1$$

fulfils the estimate $0 \leqslant \varepsilon(t) \leqslant Ct^{-\sigma}$, see Section 1.2.1 of Henry (1981). We next formulate a discrete version of this Gronwall lemma.

LEMMA 6.2 For $h > 0$ and $T > 0$, let $0 \leqslant t_n = nh \leqslant T$. Further assume that the sequence of non-negative numbers ε_n satisfies the inequality

$$\varepsilon_n \leqslant ah \sum_{\nu=1}^{n-1} t_{n-\nu}^{-\rho} \varepsilon_{\nu} + b\,t_n^{-\sigma}$$

for $0 \leqslant \rho < 1$ and $a, b \geqslant 0$. Then the following estimate holds

$$\varepsilon_n \leqslant \begin{cases} Cb\,t_n^{-\sigma} & \text{for } 0 \leqslant \sigma < 1, \\ Cb\big(t_n^{-1} + t_n^{-\rho}\,|\log h|\big) & \text{for } \sigma = 1, \end{cases}$$

where the constant C depends on ρ, σ, a, and on T.

Proof. This can be shown by using similar arguments as in the proof of Theorem 1.5.5 in Brunner & van der Houwen (1986). We omit the details. □

For the remainder of this section, we suppose that the assumptions of Theorem 3.1 are fulfiled. In particular we have $0 \leqslant \alpha < 1$.

LEMMA 6.3 The analytic semigroup e^{-tA} satisfies the bound

$$\left|A_a^\rho \mathrm{e}^{-tA}\right| \leqslant C\mathrm{e}^{-at}t^{-\rho} \qquad \text{for } t > 0 \text{ and } \rho \geqslant 0.$$

Proof. This is Theorem 1.4.3 of Henry (1981). □

LEMMA 6.4 Let u denote the solution of (2.1) with initial value $u_0 \in V$, and let $0 \leqslant \rho \leqslant 1$. Then the derivative of u with respect to t satisfies the estimate

$$\|A_a^{-\rho} u'(t)\| \leqslant Ct^{\rho-1} \qquad \text{for } 0 < t \leqslant T.$$

Proof. For $\alpha - \rho \geqslant 0$ this bound is given in Theorem 3.5.2 of Henry (1981). In the remaining case, it follows from the identity

$$\begin{aligned} A_a^{\alpha-\rho} u'(t) = &- A_a^{1-\rho}\mathrm{e}^{-tA} \cdot A_a^\alpha u_0 - \int_0^t A_a^{1+\alpha-\rho}\mathrm{e}^{-(t-\tau)A} f(\tau, u(\tau))\,\mathrm{d}\tau \\ &+ A_a^{\alpha-\rho} f(t, u(t)) \end{aligned}$$

and Lemma 6.3. □

We close this section with some estimates for the numerical discretization.

LEMMA 6.5 Under the assumptions of Theorem 3.1, the following bounds hold for $0 \leqslant \rho < 1$ and $0 < nh \leqslant T$

$$\left|A_a^\rho\big(\mathcal{R}(-hA)^n - \mathcal{R}(\infty)^n\big)\right| \leqslant Ct_n^{-\rho}, \tag{6.1}$$

$$\left\|(\mathcal{I} \otimes hA)(\mathcal{I} \otimes I + \mathcal{O} \otimes hA)^{-1}\right\| \leqslant C, \tag{6.2}$$

$$\left|(\mathcal{I} \otimes A_a^\rho)(\mathcal{I} \otimes I + \mathcal{O} \otimes hA)^{-1}\right| \leqslant Ch^{-\rho}, \tag{6.3}$$

$$\left|(\mathcal{I} \otimes A_a^\rho)(\mathcal{I} \otimes I + \mathcal{O} \otimes hA)^{-1}\left(\mathcal{I} \otimes \mathcal{R}(-hA)^n\right)\right| \leqslant Ct_n^{-\rho}. \tag{6.4}$$

Proof. These estimates are standard. They follow from the resolvent condition (2.2) and the interpolation result (see Theorem 1.4.4 in Henry 1981)

$$\left|A_a^\rho(\lambda I + A)^{-1}\right| \leqslant C\left|A(\lambda I + A)^{-1}\right|^\rho \cdot \left|(\lambda I + A)^{-1}\right|^{1-\rho}$$

together with the Cauchy integral formula. Similar bounds are given in Lemma 2.3 of Lubich & Ostermann (1993), and in Section 3 of Nakaguchi & Yagi (1997). Note that (6.1) can also be derived from the proof of Theorem 3.5 in Lubich & Nevanlinna (1991). □

Acknowledgement

We thank the anonymous referee for his comments that helped to improve the presentation of this paper.

REFERENCES

ALOUGES, F. & DEBUSSCHE, A. 1993 On the discretization of a partial differential equation in the neighborhood of a periodic orbit. *Numer. Math.* **65**, 143–175.

BORNEMANN, F. 1990 An adaptive multilevel approach to parabolic equations. I. General theory and 1D implementation. *Impact Comput. Sci. Eng.* **2**, 279–317.

BRUNNER, H. & VAN DER HOUWEN, P. J. 1986 *The numerical solution of Volterra equations*. Amsterdam: North-Holland.

DEUFLHARD, P. & BORNEMANN, F. 1994 *Numerische Mathematik II. Integration Gewöhnlicher Differentialgleichungen*. Berlin: de Gruyter.

HAIRER, E. & WANNER, G. 1996 *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems*, 2nd revised edn. Berlin: Springer.

HENRY, D. 1981 *Geometric Theory of Semilinear Parabolic Equations* (*LNM 840*). Berlin: Springer.

LANG, J. 1995 Two-dimensional fully adaptive solutions of reaction–diffusion equations. *Appl. Numer. Math.* **18**, 223–240.

LARSSON, S. 1992 Nonsmooth data error estimates with applications to the study of the long-time behavior of finite element solutions of semilinear parabolic problems. *Report 1992–36*, Department of Mathematics, Chalmers University, Göteborg.

LE ROUX, M.-N. 1979 Semidiscretization in time for parabolic problems. *Math. Comput.* **33**, 919–931.

LUBICH, CH. & NEVANLINNA, O. 1991 On resolvent conditions and stability estimates. *BIT* **31**, 293–313.

LUBICH, CH. & OSTERMANN, A. 1993 Runge–Kutta methods for parabolic equations and convolution quadrature. *Math. Comput.* **60**, 105–131.

LUBICH, CH. & OSTERMANN, A. 1995 Linearly implicit time discretization of non-linear parabolic equations. *IMA J. Numer. Anal.* **15**, 555–583.

LUBICH, CH. & OSTERMANN, A. 1996 Runge–Kutta time discretization of reaction–diffusion and Navier–Stokes equations: nonsmooth-data error estimates and applications to long-time behaviour. *Appl. Numer. Math.* **22**, 279–292.

NAKAGUCHI, E. & YAGI, A. 1997 Error estimates of implicit Runge–Kutta methods for quasilinear abstract equations of parabolic type in Banach spaces. *Preprint*, Osaka University.

NOWAK, U. 1993 Adaptive Linienmethoden für parabolische Systeme in einer Raumdimension. *Technical Report TR 93-14*, Konrad-Zuse-Zentrum Berlin.

OSTERMANN, A. & ROCHE, M. 1993 Rosenbrock methods for partial differential equations and fractional orders of convergence. *SIAM J. Numer. Anal.* **30**, 1084–1098.

PAZY, A. 1983 *Semigroups of Linear Operators and Applications to Partial Differential Equations*. New York: Springer.

SCHWITZER, F. 1995 W-methods for semilinear parabolic equations. *Appl. Numer. Math.* **18**, 351–366.

STREHMEL, K. & WEINER, R. 1992 *Linear-implizite Runge–Kutta Methoden und ihre Anwendungen*. Stuttgart: Teubner.

STUART, A. 1995 Perturbation theory for infinite dimensional dynamical systems. *Theory and Numerics of Ordinary and Partial Differential Equations* (*Advances in Numerical Analysis IV*) (M. Ainsworth, J. Levesley, W. A. Light and M. Marletta, eds). Oxford: Clarendon.

VAN DORSSELAER, J. 1998 Inertial manifolds under multistep discretization. *Preprint*, Universität Tübingen.

VAN DORSSELAER, J. & LUBICH, CH. 1999 Inertial manifolds of parabolic differential equations under higher-order discretizations. *IMA J. Numer. Anal.* **19**, 455.

Appendix A: Error estimates for Runge–Kutta methods

The present analysis relies strongly on non-smooth data error estimates for Runge–Kutta methods. For the convenience of the reader, we recall the convergence results from Lubich & Ostermann (1996).

For a given linearly implicit Runge–Kutta method of order p, we consider the corresponding *Runge–Kutta* discretization of (2.1)

$$\begin{aligned}\widetilde{U}'_{ni} + A\widetilde{U}_{ni} &= f(t_n + c_i h, \widetilde{U}_{ni})\\ \widetilde{U}_{ni} &= \widetilde{u}_n + h\sum_{j=1}^{s} a_{ij}\widetilde{U}'_{nj}, \qquad \widetilde{u}_{n+1} = \widetilde{u}_n + h\sum_{j=1}^{s} b_j\widetilde{U}'_{nj}\end{aligned} \tag{A.1}$$

with the coefficients a_{ij}, b_i, c_i as in (2.5). This diagonally implicit Runge–Kutta method enjoys the following properties: It has order p, since the order conditions for Runge–Kutta methods form a subset of those for linearly implicit methods. Due to $c = \mathcal{O}1\!\!1$, its stage order q is at least one. Moreover it has the same stability function and thus the same linear stability properties as the underlying linearly implicit method. The existence of the Runge–Kutta solution for $A(\vartheta)$-stable methods follows from Theorem 2.1 in Lubich & Ostermann (1996).

In Section 3 we have used the subsequent convergence result.

LEMMA A.1 Under the assumptions of Theorem 3.1, the following estimate holds for $0 < h \leqslant h_0$ and $0 < t_n \leqslant T$

$$\|\widetilde{u}_n - u(t_n)\| + \|\widetilde{U}_{ni} - u(t_n + c_i h)\| \leqslant C\left(t_n^{-1}h + t_n^{-\alpha}h\,|\log h|\right).$$

For $n = 0$ the same bound holds as for $n = 1$. The constants C and h_0 depend on the quantities specified in Theorem 3.1.

Proof. This result is a sharper version of Theorem 2.1 in Lubich & Ostermann (1996). It follows from (4.15) of *loc. cit.* with $r = \min(p, q+1) \geqslant 1$. Note that the first iterate of the fixed-point iteration is not given correctly there. In the fourth line above formula (4.15) of *loc. cit.*, it should read

$$U_{ni}^{(1)} = X_{ni} + Y_{ni} + d_{ni} \qquad \text{with} \qquad d_n = h\sum_{\nu=0}^{n} W_{n-\nu}\big(F_\nu(U_\nu^{(0)}) - G_\nu\big).$$

Using the Lipschitz condition for f, the bound then follows from Lemmas 4.2 and 4.3 of *loc. cit.* □

Under the assumptions of Theorem 5.1, a refinement of Lemma A.1 is possible. For this we note that the function $g(t) = f(t, u(t))$ satisfies

$$|A_a^{-\beta} g'(t)| \leqslant K t^{\beta-1}, \qquad 0 < t \leqslant T. \tag{A.2}$$

This follows from Assumption 5.1 and Lemma 6.4. We are now in a position to state this refinement.

LEMMA A.2 Under the assumptions of Theorem 5.1, the following estimates hold for $0 < h \leqslant h_0$ and $0 < t_n \leqslant T$

$$\|\widetilde{u}_n - u(t_n)\| + \|\widetilde{U}_{ni} - u(t_n + c_i h)\| \leqslant C \left(t_n^{-2} h^2 + t_n^{-\alpha-\beta} h^{1+\beta}\right), \tag{A.3}$$

$$\left\|A_a^{-\beta}\bigl(\widetilde{u}_n - u(t_n)\bigr)\right\| + \left\|A_a^{-\beta}\bigl(\widetilde{U}_{ni} - u(t_n + c_i h)\bigr)\right\| \leqslant C\, t_n^{\beta-1} h. \tag{A.4}$$

For $n = 0$ the same bounds hold as for $n = 1$. The constants C and h_0 depend on the quantities specified in Theorem 5.1.

Proof. This lemma is a sharper version of Theorem 2.3 of Lubich & Ostermann (1996). The bound (A.3) follows essentially from Lemma 4.4 of *loc. cit.* There, a similar result is proved under an additional assumption on $g''(t)$ which enters the estimate of $E_h g_\delta(t)$. Since we use here only information on $g'(t)$, we have to estimate this term differently. We proceed as in the proof of Lemma 4.3 of *loc. cit.* and split the integral

$$\int_0^t \left\|E_h \mathbf{1}(t-\tau) g_\delta'(\tau)\right\| \mathrm{d}\tau \leqslant \int_0^{t/2} \left\|E_h \mathbf{1}(t-\tau)\right\|_{X^\alpha \leftarrow X^{-\beta}} \left|A_a^{-\beta} g_\delta'(\tau)\right| \mathrm{d}\tau + \int_{t/2}^t \left\|E_h \mathbf{1}(t-\tau)\right\|_{X^\alpha \leftarrow X} \left|g_\delta'(\tau)\right| \mathrm{d}\tau.$$

The desired result

$$\|E_h g_\delta(t_n)\| \leqslant C t_n^{-\alpha-\beta} h^{1+\beta}$$

then follows from (A.2) and the bounds

$$\|E_h \mathbf{1}(t)\|_{X^\alpha \leftarrow X} \leqslant C \min\left(t^{-1-\alpha} h^2, h^{1-\alpha}\right)$$

$$\|E_h \mathbf{1}(t)\|_{X^\alpha \leftarrow X^{-\beta}} \leqslant C \min\left(t^{-1-\alpha-\beta} h^2, h^{1-\alpha-\beta}\right)$$

for $0 \leqslant t \leqslant T$. Since $r = \min(p, q+1) \geqslant 2$, we obtain (A.3) as in Lubich & Ostermann (1996).

In order to verify (A.4), we consider the Runge–Kutta discretization of

$$x' + Ax = 0, \qquad x(0) = u_0.$$

The proof of Theorem 1.2 in Le Roux (1979) shows that

$$\left\|A_a^{-\beta}\bigl(\widetilde{x}_n - x(t_n)\bigr)\right\| + \left\|A_a^{-\beta}\bigl(\widetilde{X}_{ni} - x(t_n + c_i h)\bigr)\right\| \leqslant C\, t_n^{\beta-2} h^2,$$

where x_n denotes the Runge–Kutta approximation to $x(t_n)$ and $\widetilde{X}_{ni}$ the corresponding stage values. With this bound at hand, the desired result then follows as in Lemma A.1. □