# Exponential Taylor methods: analysis and implementation

Antti Koskela[a,*], Alexander Ostermann[a,**]

[a]*Department of Mathematics, University of Innsbruck, Technikerstraße 13, A-6020 Innsbruck, Austria*

---

**Abstract**

For the time integration of semilinear systems of differential equations, a class of multiderivative exponential integrators is considered. The methods are based on a Taylor series expansion of the semilinearity about the numerical solution, the required derivatives are computed by automatic differentiation. Inserting these derivatives into the variation-of-constants formula results in an exponential integrator which requires the action of the exponential of an augmented Jacobian only.

The convergence properties of such exponential integrators are analyzed, and potential sources of numerical instabilities are identified. In particular, it is shown that local linearization gives rise to better stability for stiff problems. A number of numerical experiments illustrate the theoretical results.

*Keywords:* Exponential integrators, Taylor series methods, Multiderivative methods, Local linearization, Stability
*2000 MSC:* 65L20, 65L05, 65M12

---

## 1. Introduction

In this paper we are concerned with the numerical analysis of multiderivative exponential integrators for stiff problems. In particular, we consider methods for the solution of semilinear systems of initial value problems

$$u'(t) = Au(t) + g(t, u(t)), \quad u(t_0) = u_0. \tag{1}$$

---

*The work of the first author was supported by the FWF doctoral program 'Computational Interdisciplinary Modelling' W1227.

**Corresponding author.

*Email addresses:* `antti.koskela@uibk.ac.at` (Antti Koskela), `alexander.ostermann@uibk.ac.at` (Alexander Ostermann)

In contrast to standard methods, exponential integrators are based on the mild solution (variation-of-constants formula)

$$u(t) = \mathrm{e}^{(t-t_0)A} u_0 + \int_{t_0}^{t} \mathrm{e}^{(t-\tau)A} g(\tau, u(\tau)) \,\mathrm{d}\tau. \tag{2}$$

The idea behind exponential Taylor methods is to replace the nonlinearity $g(\tau, u(\tau))$ in (2) by its Taylor polynomial at $\tau = t_0$ and to compute the integrals exactly. This procedure defines a numerical scheme that makes explicit use of certain derivatives of the nonlinearity.

Exponential Taylor methods were recently proposed in a paper by Al-Mohy and Higham [1]. The computational attractiveness of these methods comes from the fact that one time step can be computed by the action of a single exponential matrix of an augmented Jacobian of the right-hand side of (1). The required derivatives can be computed by standard automatic differentiation, see [2, Sect. I.8] and [3]. For non-stiff problems, such Taylor methods are well understood, see [2, 4]. A possible extension to stiff problems is described in [5]. Here, we analyze an exponential version of such methods for stiff problems.

In section 2 we derive the class of exponential Taylor methods and recall its possible numerical implementation from [1]. The stability and convergence properties of these methods are analyzed in section 3. In contrast to the linear problem, where methods of arbitrary high order exist, the situation is more involved for semilinear problems. It is shown that the exponential Taylor method of classical order two is indeed second order convergent for stiff problems; higher order methods, however, suffer from instabilities when applied to stiff problems. For methods that linearize the vector field in each step, the situation is slightly more favorable. It is shown that linearized exponential Taylor methods up to order three do not suffer from instabilities. The efficient implementation of exponential Taylor integrators is addressed in section 4.

The remaining sections are devoted to numerical experiments. Section 5 illustrates the instabilities that are inherent in high order exponential Taylor methods when applied to semilinear stiff problems. Section 6 shows how the instabilities are triggered by round-off errors near the stability border. Finally, in section 7, the numerical efficiency of exponential Taylor methods for linear problems is demonstrated by a comparison with standard MATLAB integrators.

## 2. Exponential Taylor methods

For the time integration of semilinear systems of stiff differential equations

$$u'(t) = f(t, u(t)), \quad u(t_0) = u_0, \tag{3a}$$

with

$$f(t, u(t)) = Au(t) + g(t, u(t)) \tag{3b}$$

we consider exponential integrators. Our focus will be on equations that arise from spatial semidiscretization of partial differential equations of evolutionary type. In the above problem $A$ denotes a real $d \times d$ matrix (with $d$ large), and $g$ is a nonlinear function with a moderate Lipschitz constant.

Exponential integrators propagate the linear part $A$ of (3) exactly and can be constructed with the help of the variation-of-constants formula

$$u(t) = \mathrm{e}^{(t-t_0)A} u_0 + \int_{t_0}^{t} \mathrm{e}^{(t-\tau)A} g(\tau, u(\tau)) \,\mathrm{d}\tau. \tag{4}$$

The simplest method is obtained by replacing the nonlinearity with the known value $g(t_0, u_0)$, giving the exponential Euler method

$$\begin{aligned} u_1 &= \mathrm{e}^{hA} u_0 + h\varphi_1(hA) g(t_0, u_0) \\ &= u_0 + h\varphi_1(hA) f(t_0, u_0), \end{aligned} \tag{5}$$

where $h$ denotes the step size, $u_1$ is the numerical approximation to $u(t_0 + h)$, and $\varphi_1$ is the entire function

$$\varphi_1(z) = \frac{\mathrm{e}^z - 1}{z}.$$

Higher order exponential multistep methods are obtained in a similar way by replacing the nonlinearity in (4) by an appropriate interpolation polynomial, see [6]. The construction of high order exponential Runge–Kutta methods, however, is more involved, see [7].

Here, we consider one-step methods that are based on the Taylor expansion of the nonlinear term $g(\tau, u(\tau))$. For its construction, we replace the nonlinearity in (4) by its Taylor polynomial of degree $p - 1$

$$g(\tau, u(\tau)) \approx \sum_{k=0}^{p-1} \frac{(\tau - t_0)^k}{k!} \frac{\mathrm{d}^k}{\mathrm{d}t^k} g(t, u(t))\big|_{t=t_0}. \tag{6}$$

In order to obtain a numerical scheme, we still have to approximate the derivatives of $g$ by known quantities. For linear problems where $g$ does not depend on $u$, this is a straightforward task. Otherwise, we invoke the differential equation and the chain rule to get the required approximations.

For notational simplicity, we illustrate this process for autonomous problems where $g = g(u)$. Using the chain rule, we get for $p = 5$

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} g(u) &= g'(u)u' \\
\frac{\mathrm{d}^2}{\mathrm{d}t^2} g(u) &= g''(u)(u', u') + g'(u)u'' \\
\frac{\mathrm{d}^3}{\mathrm{d}t^3} g(u) &= g'''(u)(u', u', u') + 3g''(u)(u'', u') + g'(u)u''' \\
\frac{\mathrm{d}^4}{\mathrm{d}t^4} g(u) &= g^{(4)}(u)(u', u', u', u') + 6g'''(u)(u'', u', u') \\
&\qquad + 3g''(u)(u'', u'') + 4g''(u)(u''', u') + g'(u)u^{(4)}.
\end{aligned} \tag{7}$$

Using the differential equation, we define recursively

$$u_0^{(k)} = Au_0^{(k-1)} + w_k, \quad k \geq 1$$

and

$$\begin{aligned}
w_1 &= g(u_0) \\
w_2 &= g'(u_0)u_0' \\
w_3 &= g''(u_0)(u_0', u_0') + g'(u_0)u_0'' \\
w_4 &= g'''(u_0)(u_0', u_0', u_0') + 3g''(u_0)(u_0'', u_0') + g'(u_0)u_0''' \\
w_5 &= g^{(4)}(u_0)(u_0', u_0', u_0', u_0') + 6g'''(u_0)(u_0'', u_0', u_0') + 3g''(u_0)(u_0'', u_0'') \\
&\qquad + 4g''(u_0)(u_0''', u_0') + g'(u_0)u_0^{(4)}.
\end{aligned} \tag{8}$$

Inserting the approximations

$$w_k \approx \frac{\mathrm{d}^{k-1}}{\mathrm{d}t^{k-1}} g(t, u(t))\big|_{t=t_0}$$

into (6), (4) and computing the arising integrals defines the numerical scheme

$$u_1 = \mathrm{e}^{hA}u_0 + \sum_{k=1}^{p} h^k \varphi_k(hA) w_k \tag{9a}$$

$$= u_0 + h\varphi_1(hA) f(t_0, u_0) + \sum_{k=2}^{p} h^k \varphi_k(hA) w_k, \tag{9b}$$

which henceforth will be called the *exponential Taylor method*. Recall that the entire functions

$$\varphi_k(z) = \int_0^1 \mathrm{e}^{(1-\theta)z} \frac{\theta^{k-1}}{(k-1)!}\,\mathrm{d}\theta, \quad k \geq 1 \tag{10}$$

satisfy the normalization $\varphi_k(0) = 1/k!$ and the recurrence relation

$$\varphi_{k+1}(z) = \frac{\varphi_k(z) - \varphi_k(0)}{z}, \quad \varphi_0(z) = \mathrm{e}^z. \tag{11}$$

As shown in [1, Thm. 2.1], the series (9a) can be represented as the product of the exponential of a $(d+p) \times (d+p)$ matrix and a vector. Therefore an explicit computation of the actions of the operators $\varphi_k(hA)$, $k \geq 1$, is avoided, which makes the method computationally attractive. In the following lemma we give an alternative proof for this result.

**Lemma 1.** *Let* $A \in \mathbb{R}^{d\times d}$, $W = [w_p, w_{p-1}, \ldots, w_1] \in \mathbb{R}^{d\times p}, h \in \mathbb{R}$ *and*

$$\widetilde{A} = \begin{bmatrix} A & W \\ 0 & J \end{bmatrix} \in \mathbb{R}^{(d+p)\times(d+p)}, \quad J = \begin{bmatrix} 0 & I_{p-1} \\ 0 & 0 \end{bmatrix}, \quad v_0 = \begin{bmatrix} u_0 \\ e_p \end{bmatrix} \in \mathbb{R}^{d+p} \tag{12}$$

*with* $e_p = [0,\ldots,0,1]^{\mathsf{T}}$. *Then the approximation* (9) *is obtained from*

$$u_1 = \begin{bmatrix} I_d & 0 \end{bmatrix} \mathrm{e}^{h\widetilde{A}} v_0. \tag{13}$$

*Proof.* Using the integral equation (4) and the representation (10) of the $\varphi_k$-functions, one easily sees that the approximation

$$\widehat{u}(t) = \mathrm{e}^{tA} u_0 + \sum_{k=1}^{p} t^k \varphi_k(tA) w_k$$

is the exact solution of the differential equation

$$\widehat{u}'(t) = A\widehat{u}(t) + w_1 + tw_2 + \ldots + \frac{t^{p-1}}{(p-1)!} w_p, \qquad \widehat{u}(0) = u_0.$$

The coefficients of the inhomogeneity,

$$y_i(t) = \frac{t^{i-1}}{(i-1)!}, \quad i \geq 1,$$

verify $y_1'(t) = 0$ and $y_i'(t) = y_{i-1}$ for $i \geq 2$. Thus the function

$$v(t) = \begin{bmatrix} \widehat{u}(t) \\ y_p(t) \\ \vdots \\ y_1(t) \end{bmatrix} \in \mathbb{R}^{d+p}$$

satisfies the initial value problem

$$v'(t) = \begin{bmatrix} A & W \\ 0 & J \end{bmatrix} v(t) = \widetilde{A} v(t), \qquad v_0 = \begin{bmatrix} u_0 \\ e_p \end{bmatrix}$$

with exact solution $v(t) = \mathrm{e}^{t\widetilde{A}} v_0$. From this the claim follows. $\square$

*Remark.* In the case of (9b), we get the approximation $u_1$ by replacing the vectors $w_1$ by $f(t_0, u_0)$ and $v_0$ by $e_{d+p}$, respectively. When $f$ becomes small, this approach is beneficial for the numerical approximation of the matrix exponential.

## 3. Stability and convergence

In this section we analyze the convergence properties of exponential Taylor methods for stiff problems. Let the involved matrix exponential satisfy the bound

$$\|\mathrm{e}^{tA}\| \le C_0 \mathrm{e}^{\omega t}, \qquad t \ge 0 \tag{14}$$

for some nonnegative constants $C_0$ and $\omega$. It turns out that our convergence bounds will only depend on these constants and *not* on the very form of $A$, in particular *not* on its dimension $d$. This fact is important for applications where $A$ is a discretization of a spatial differential operator.

We further note for later use that (14) implies the bound

$$\|\varphi_k(tA)\| \le \frac{C_0 \mathrm{e}^{\omega t}}{k!}, \qquad t \ge 0. \tag{15}$$

This follows at once from the representation (10).

### *3.1. The linear case*

Applied to the linear problem

$$u'(t) = Au(t) + g(t), \quad u(t_0) = u_0, \qquad t_0 \le t \le T \tag{16}$$

the exponential Taylor method assumes the form

$$u_{n+1} = \mathrm{e}^{hA} u_n + \sum_{k=1}^{p} h^k \varphi_k(hA) g^{(k-1)}(t_n), \quad n \ge 0. \tag{17}$$

Its error $e_{n+1} = u_{n+1} - u(t_{n+1})$ satisfies the recursion

$$e_{n+1} = \mathrm{e}^{hA} e_n - \delta_{n+1},$$

where

$$\delta_{n+1} = \int_0^h \mathrm{e}^{(h-\tau)A} \int_0^\tau \frac{(\tau-\xi)^{p-1}}{(p-1)!} g^{(p)}(t_n + \xi)\,\mathrm{d}\xi\,\mathrm{d}\tau.$$

This implies the following theorem whose proof is straightforward.

**Theorem 2.** *Let the inhomogeneity $g$ in* (16) *be $p$ times differentiable with $g^{(p)} \in L^1(0,T)$. Then, the exponential Taylor method* (17) *is convergent of order $p$.* $\square$

We remark that this convergence result extends to variable step sizes in an obvious way.

*3.2. The semilinear case*

Exponential Taylor methods for semilinear problems (3) possess an inherent instability for $p \geq 3$, cf. our discussion below. Therefore, we consider here only[1] the Taylor scheme with $p = 2$. It has the form

$$\begin{aligned} u_{n+1} = \mathrm{e}^{hA} u_n + h\varphi_1(hA) g(t_n, u_n) \\ + h^2 \varphi_2(hA)\big(g_t(t_n,u_n) + g_u(t_n,u_n)(Au_n + g(t_n,u_n))\big). \end{aligned} \tag{18}$$

Let $\widetilde{u}_n = u(t_n)$ denote the exact solution of (3). We assume that $\psi(t) = g(t, u(t))$ is sufficiently smooth. The exact solution satisfies the recursion

$$\begin{aligned} \widetilde{u}_{n+1} = \mathrm{e}^{hA} \widetilde{u}_n + h\varphi_1(hA) g(t_n, \widetilde{u}_n) \\ + h^2 \varphi_2(hA)\big(g_t(t_n,\widetilde{u}_n) + g_u(t_n,\widetilde{u}_n)(A\widetilde{u}_n + g(t_n,\widetilde{u}_n))\big) + \delta_{n+1}, \end{aligned} \tag{19}$$

where

$$\delta_{n+1} = \int_0^h \mathrm{e}^{(h-\tau)A} \int_0^\tau (\tau - \xi) \frac{\mathrm{d}^2 g}{\mathrm{d}t^2}\big(t, u(t)\big)\big|_{t=t_n+\xi} \,\mathrm{d}\xi\, \mathrm{d}\tau. \tag{20}$$

Let $e_n = u_n - \widetilde{u}_n$. Taking the difference of (18) and (19) gives the error recursion

$$\begin{aligned} e_{n+1} = \mathrm{e}^{hA} e_n + h\varphi_1(hA)\big(g(t_n,u_n) - g(t_n,\widetilde{u}_n)\big) \\ + h^2\varphi_2(hA)\big(g_t(t_n,u_n) - g_t(t_n,\widetilde{u}_n)\big) \\ + h^2\varphi_2(hA) g_u(t_n,u_n)\big(Ae_n + g(t_n,u_n) - g(t_n,\widetilde{u}_n)\big) \\ + h^2\varphi_2(hA)\big(g_u(t_n,u_n) - g_u(t_n,\widetilde{u}_n)\big)u'(t_n) - \delta_{n+1} \end{aligned} \tag{21}$$

which we have to solve. As $g$ is smooth, we can bound the differences of $g$, and of its derivatives, respectively, by Lipschitz conditions. In order to adequately treat the term $Ae_n$ which appears on the right-hand side of (21), we premultiply the whole recursion by $hA$ and use the recurrence relation (11) of the $\varphi$-functions. This gives a coupled system of recursions that can be

[1] The scheme with $p = 1$ is the well-known exponential Euler method. It was analyzed in [7].

easily solved with the help of a standard Gronwall lemma. One finally obtains the estimate

$$\|e_n\| + h\|Ae_n\| \le C \sum_{j=1}^{n} \Big( \|\mathrm{e}^{(n-j)hA}\delta_j\| + h\|A\mathrm{e}^{(n-j)hA}\delta_j\| \Big) \tag{22}$$

with a constant $C$ that is independent of the norm of $A$.

**Theorem 3.** *Let the inhomogeneity $\psi(t) = g(t, u(t))$ in (3) be twice differentiable and $\psi'' \in L^1(0,T)$. Then, the exponential Taylor method (18) is second order convergent.*

*Proof.* Inserting the defects (20) into the estimate (22) and integrating once by parts yields the desired result. □

The above proof also shows that the scheme (18) is stable with respect to perturbations. Let

$$\begin{aligned} \widehat{u}_{n+1} &= \mathrm{e}^{hA}\widehat{u}_n + h\varphi_1(hA)g(t_n, \widehat{u}_n) \\ &\quad + h^2\varphi_2(hA)\big(g_t(t_n,\widehat{u}_n) + g_u(t_n,\widehat{u}_n)(A\widehat{u}_n + g(t_n,\widehat{u}_n))\big) + \theta_{n+1} \end{aligned} \tag{23}$$

denote the perturbed scheme with perturbations $\theta_{n+1}$. In the same way as above we get

$$\|u_n - \widehat{u}_n\| \le C \sum_{j=1}^{n} \Big( \|\mathrm{e}^{(n-j)hA}\theta_j\| + h\|A\mathrm{e}^{(n-j)hA}\theta_j\| \Big) \tag{24}$$

which can simply be estimated by

$$\|u_n - \widehat{u}_n\| \le Cn(1 + h\|A\|) \max_{1\le j\le n} \|\theta_j\|. \tag{25}$$

Better estimates are possible if $A$ or parts of it commute with $g_u$. We omit the details.

Exponential Taylor methods with $p \ge 3$, however, are inherently unstable for stiff problems. This is already obvious from the linear problem

$$u'(t) = Au(t) + u(t)$$

for which the exponential Taylor method with $p = 3$ reads

$$u_{n+1} = \mathrm{e}^{hA}u_n + h\varphi_1(hA)u_n + h^2\varphi_2(hA)(A+I)u_n + h^3\varphi_3(hA)(A+I)^2u_n.$$

This recursion contains the term

$$h\varphi_3(hA)(hA)^2$$

which will give rise to exponential instabilities, see also section 5.

*3.3. A linearized scheme*

The analysis of the previous paragraph limits the order of stable exponential Taylor methods to two. For *linearized* schemes, however, order three is possible, as will be shown now. For notational simplicity, we restrict the presentation here to autonomous problems

$$u'(t) = Au(t) + g(u(t)), \quad u(t_0) = u_0, \tag{26}$$

where we denote again the right-hand side by $f(u(t))$ for short. Given the numerical approximation $u_n$ at time $t_n$, we linearize the differential equation at this state. This yields

$$v'(t) = J_n v(t) + g_n(v(t)), \quad v(t_n) = u_n, \tag{27}$$

where $J_n$ denotes the Fréchet derivative of $f$ and $g_n$ the remainder, i.e.

$$\begin{aligned} J_n &= f'(u_n) = A + g'(u_n), \\ g_n(u) &= f(u) - J_n u = g(u) - g'(u_n)u. \end{aligned}$$

Applying an exponential Taylor method to (27) yields a so-called *linearized* exponential Taylor method. Linearized exponential integrators were first proposed by Pope [8], see also [9, 7]. By construction, the remainder in (27) satisfies

$$g_n'(u_n) = 0.$$

Therefore, the linearized exponential Taylor scheme for $p = 3$ has the form[2]

$$\begin{aligned} u_{n+1} &= \mathrm{e}^{hJ_n} u_n + h\varphi_1(hJ_n) g_n(u_n) \\ &\quad + h^3 \varphi_3(hJ_n) g''(u_n)\big(J_n u_n + g_n(u_n), J_n u_n + g_n(u_n)\big) \\ &= u_n + h\varphi_1(hJ_n) f(u_n) + h^3\varphi_3(hJ_n) f''(u_n)\big(f(u_n), f(u_n)\big). \end{aligned} \tag{28}$$

For $p \geq 4$ and $h\|A\|$ large the schemes are unstable. In order to show that (28) is indeed a third order method we proceed as in section 3.2. The exact solution $\widetilde{u}_n = u(t_n)$ satisfies the recursion

$$\begin{aligned} \widetilde{u}_{n+1} &= \mathrm{e}^{h\widetilde{J}_n} \widetilde{u}_n + h\varphi_1(h\widetilde{J}_n) \widetilde{g}_n(\widetilde{u}_n) \\ &\quad + h^3 \varphi_3(h\widetilde{J}_n) g''(\widetilde{u}_n)\big(\widetilde{J}_n \widetilde{u}_n + \widetilde{g}_n(\widetilde{u}_n), \widetilde{J}_n \widetilde{u}_n + \widetilde{g}_n(\widetilde{u}_n)\big) + \delta_{n+1}, \end{aligned} \tag{29}$$

where

$$\widetilde{J}_n = f'(u(t_n)), \qquad \widetilde{g}_n(u) = f(u) - \widetilde{J}_n u$$

[2]For $p = 1$ and $p = 2$, the method reduces to the exponential Rosenbrock–Euler method, which is of order two (see [7]).

and

$$\delta_{n+1} = \int_0^h \mathrm{e}^{(h-\tau)A} \int_0^\tau \frac{(\tau-\xi)^2}{2} \frac{\mathrm{d}^3 \widetilde{g}_n}{\mathrm{d}t^3}\big(u(t)\big)\big|_{t=t_n+\xi} \,\mathrm{d}\xi\,\mathrm{d}\tau. \tag{30}$$

Let $e_n = u_n - u(t_n)$. In order to solve the error recursion, we need the following lemma.

**Lemma 4.** *Under the above assumptions, the following estimates hold*

$$\|\widetilde{J}_{n+1} - \widetilde{J}_n\| \le Ch, \tag{31a}$$

$$\|J_{n+1} - J_n\| \le C(h + \|e_n\| + \|e_{n+1}\|), \tag{31b}$$

$$\|\mathrm{e}^{hJ_n} - \mathrm{e}^{h\widetilde{J}_n}\| \le Ch\|e_n\|, \tag{31c}$$

$$\|\varphi_k(hJ_n) - \varphi_k(h\widetilde{J}_n)\| \le Ch\|e_n\|, \quad k \ge 1, \tag{31d}$$

$$\|g_n(u_n) - \widetilde{g}_n(\widetilde{u}_n)\| \le C\|e_n\|. \tag{31e}$$

*Proof.* The estimates (31a) and (31b) follow from the smoothness of the exact solution and the Lipschitz condition on $g$. The bound (31c) follows from the differential equation

$$v'(t) = hJ_n v(t) = h\widetilde{J}_n v(t) + h(J_n - \widetilde{J}_n)v(t), \qquad v(0) = I$$

with solution

$$v(1) = \mathrm{e}^{hJ_n} = \mathrm{e}^{h\widetilde{J}_n} + h\int_0^1 \mathrm{e}^{(1-\tau)h\widetilde{J}_n}(J_n - \widetilde{J}_n)\mathrm{e}^{\tau h J_n}\,\mathrm{d}\tau.$$

The bound (31d) follows in a similar way by taking into account that $\varphi_k(hJ)$ is the solution of the differential equation

$$v'(t) = Jv(t) + \frac{t^{k-1}}{(k-1)!}I, \qquad v(0) = 0$$

at $t = 1$. The last bound is a consequence of

$$g_n(u_n) = g(u_n) - g'(u_n)u_n, \quad \widetilde{g}_n(\widetilde{u}_n) = g(\widetilde{u}_n) - g'(\widetilde{u}_n)\widetilde{u}_n$$

and the Lipschitz conditions on $g$ and $g'$. □

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

**Theorem 5.** *Let the inhomogeneity $\psi(t) = g(u(t))$ in (26) be three times differentiable with $\psi''' \in L^1(0,T)$. Then, the exponential Taylor method (28) is convergent of order three, i.e.*

$$\|u_n - u(t_n)\| \le Ch^3$$

*with a constant $C$ that is uniform on compact intervals $0 \le nh \le T$.*

*Proof.* The proof is straightforward and carried out in a similar way as that of Theorem 3. The fact that the numerical and the exact solution involve different matrices $J_n$ and $\widetilde{J}_n$ makes the proof, however, tedious. For instance, the term

$$\mathrm{e}^{hJ_n}u_n - \mathrm{e}^{h\widetilde{J}_n}\widetilde{u}_n = \mathrm{e}^{hJ_n}e_n + \bigl(\mathrm{e}^{hJ_n} - \mathrm{e}^{h\widetilde{J}_n}\bigr)\widetilde{u}_n,$$

which is part of the error recursion, becomes after multiplication with $J_{n+1}$ and $\widetilde{J}_{n+1}$, respectively

$$J_{n+1}\mathrm{e}^{hJ_n}u_n - \widetilde{J}_{n+1}\mathrm{e}^{h\widetilde{J}_n}\widetilde{u}_n = \bigl(J_{n+1} - \widetilde{J}_{n+1}\bigr)\mathrm{e}^{hJ_n}u_n + \widetilde{J}_{n+1}\bigl(\mathrm{e}^{hJ_n}u_n - \mathrm{e}^{h\widetilde{J}_n}\widetilde{u}_n\bigr).$$

Further, we have

$$\begin{aligned}\widetilde{J}_{n+1}\bigl(\mathrm{e}^{hJ_n}u_n - \mathrm{e}^{h\widetilde{J}_n}\widetilde{u}_n\bigr) = \bigl(\widetilde{J}_{n+1} - \widetilde{J}_n\bigr)&\bigl(\mathrm{e}^{hJ_n}u_n - \mathrm{e}^{h\widetilde{J}_n}\widetilde{u}_n\bigr)\\ &+ (\widetilde{J}_n - J_n)\mathrm{e}^{hJ_n}u_n + \mathrm{e}^{hJ_n}J_nu_n - \mathrm{e}^{h\widetilde{J}_n}\widetilde{J}_n\widetilde{u}_n\end{aligned}$$

and it remains to rewrite

$$\mathrm{e}^{hJ_n}J_nu_n - \mathrm{e}^{h\widetilde{J}_n}\widetilde{J}_n\widetilde{u}_n = \mathrm{e}^{hJ_n}(J_nu_n - \widetilde{J}_n\widetilde{u}_n) + \bigl(\mathrm{e}^{hJ_n}u_n - \mathrm{e}^{h\widetilde{J}_n}\bigr)\widetilde{J}_n\widetilde{u}_n.$$

All these terms are now appropriately bounded with the help of Lemma 4. The remaining terms in the recursion are treated in a similar way and left as an exercise. The required stability bounds for the products $\mathrm{e}^{hJ_n}\cdots\mathrm{e}^{hJ_\ell}$, $\ell \le n$ are given in [9]. $\square$

## 4. Implementation

The implementation of exponential Taylor methods essentially consists of two parts: the computation of the derivatives $w_i$ (8), and the approximation of the matrix exponential times a vector (13). In this section efficient methods for both parts are provided, and a standard step size control scheme is described. For computing the vectors $w_i$ only the semilinear case is considered as the linear case (16) is trivial.

### *4.1. Computation of the derivatives for semilinear problems*

Using ideas of automatic differentiation (AD), the computation of the derivatives $w_i$ for semilinear equations is straightforward. For the implementation of Taylor methods AD was already considered in [2, 5]; an extensive discussion of the AD paradigm can be found in [3].

For the purpose of this article the following recursion proved useful. It was derived using the chain rule and is easy to implement for the equations considered. For the autonomous problem (26), the vectors $w_i$ of the exponential Taylor method (9b) are given by

$$
\begin{aligned}
w_1 &= f^{(0)} \\
w_2 &= J^{(0)} f^{(0)} \\
i \geq 2 \quad \left\{ \begin{aligned} f^{(i-1)} &= A f^{(i-2)} + w_i \\ w_{i+1} &= \sum_{k=0}^{i-1} \tbinom{i-1}{k} J^{(i-1-k)} f^{(k)}, \end{aligned} \right.
\end{aligned}
$$

where

$$
f^{(k)} = \frac{\mathrm{d}^k}{\mathrm{d}t^k} f(u(t))\big|_{u=u_n} \quad \text{and} \quad J^{(k)} = \frac{\mathrm{d}^k}{\mathrm{d}t^k} g'(u(t))\big|_{u=u_n}.
$$

A similar recursion holds for the linearized scheme (27).

*4.2. Computation of the matrix exponential times a vector*

For sparse matrices, methods that employ matrix times vector multiplications give efficient means for evaluating the action of a matrix function on a vector. For approximating the product (13) we use Krylov methods.

Krylov subspace methods are based on the idea of projecting the large problem onto a lower-dimensional Krylov subspace $\mathcal{K}_k(A, b)$, which is defined for a matrix $A \in \mathbb{R}^{d \times d}$ and a vector $b \in \mathbb{R}^d$ by

$$
\mathcal{K}_k(A, b) = \operatorname{span}\{b, Ab, A^2 b, ..., A^{k-1} b\}.
$$

The well-known Arnoldi iteration performs the Gram–Schmidt orthogonalization for this basis and gives as a result an orthonormal matrix $V_k \in \mathbb{R}^{d \times k}$ and a Hessenberg matrix $H_k = V_k^{\mathsf{T}} A V_k \in \mathbb{R}^{k \times k}$, which represents the action of $A$ in the subspace $\mathcal{K}_k(A, b)$. If $A$ is symmetric, $H_k$ is tridiagonal and we get the Lanczos iteration. The product $\mathrm{e}^A b$ is then approximated by

$$
\mathrm{e}^A b \approx V_k \mathrm{e}^{H_k} V_k^{\mathsf{T}} b = \|b\|_2 V_k \mathrm{e}^{H_k} e_1. \tag{32}
$$

The exponential of the smaller matrix $H_k$ can be computed in MATLAB by the `expm` function which uses the diagonal Padé approximant combined with scaling and squaring [10].

The convergence of the Krylov approximation (32) generally depends on the field of values of the matrix [11]. Although the field of values of the augmented matrix (12) may be considerably larger than that of $A$, numerically the convergence is found satisfying. We note that the spectrum satisfies the relation $\operatorname{spec}(\widetilde{A}) = \operatorname{spec}(A) \cup \{0\}$.

To illustrate the good convergence behavior of the Arnoldi iteration when applied to the augmented matrix, we consider a finite difference spatial discretization of the semilinear problem

$$\partial_t u = \partial_{xx} u + u(1-u), \quad x \in \left[-\tfrac{5}{2}, \tfrac{5}{2}\right],$$

subject to periodic boundary conditions, and with initial value

$$u_0(x) = \mathrm{e}^{-10x^2}.$$

The number of spatial discretization points is 500, the time step $h$ is chosen such that $\|hA\|_2 \approx 80$, where $A$ is the discretized Laplacian. The augmented matrix $\widetilde{A}$ is formed for the exponential Taylor method with $p=5$ and recursions (8) at $t=0$. The quantities $v_0$ and $W$ are as in (12). In this example it holds $\|hW\|_2 \approx 2.7 \cdot 10^5$.

In Figure 1 (left) the solid line depicts the error of the Lanczos approximation $\|\mathrm{e}^{hA}u_0 - V_k \mathrm{e}^{H_k} V_k^\mathsf{T} b\|_2$, and the dashed line the corresponding error for the product $\mathrm{e}^{h\widetilde{A}}v_0$ using the Arnoldi iteration. We note that as $A$ is symmetric, the Lanczos approximation for $\mathrm{e}^{hA}u_0$ converges fast, and thus the convergence for $\mathrm{e}^{h\widetilde{A}}v_0$ is also very satisfactory.

As noted by Al-Mohy and Higham [1], the relation (13) can be equivalently expressed as

$$\begin{bmatrix} I_d & 0 \end{bmatrix} \exp\left( h \begin{bmatrix} A & W \\ 0 & J \end{bmatrix} \right) \begin{bmatrix} u_0 \\ e_p \end{bmatrix} = \begin{bmatrix} I_d & 0 \end{bmatrix} \exp\left( h \begin{bmatrix} A & \eta W \\ 0 & J \end{bmatrix} \right) \begin{bmatrix} u_0 \\ \eta^{-1} e_p \end{bmatrix},$$

where $\eta \in \mathbb{R}$. This is easy to see from the proof of Lemma 1. In double precision arithmetic, however, the choice of $\eta$ can be seen to affect the numerical stability of the Krylov iteration. Figure 1 (right) depicts the convergence of the Krylov approximation of $\mathrm{e}^{h\widetilde{A}}v_0$ for several choices of $\eta$. Here, the exponential of $hH_k$ was simply computed by MATLAB's `expm` function. The quantities $\widetilde{A}$, $v_0$ and $h$ are chosen as above.

### *4.3. Local error and step size control*

An error control for the exponential Taylor integrator

$$u_1^{[p]} = \mathrm{e}^{hA}u_0 + \sum_{k=1}^{p} h^k \varphi_k(hA) w_k$$

can be obtained from the last term of the sum, since for small $h$ we may approximate

$$\|u(t_0+h) - u_1^{[p-1]}\|_2 \approx \|u_1^{[p]} - u_1^{[p-1]}\|_2 = \|h^p \varphi_p(hA) w_p\|_2.$$

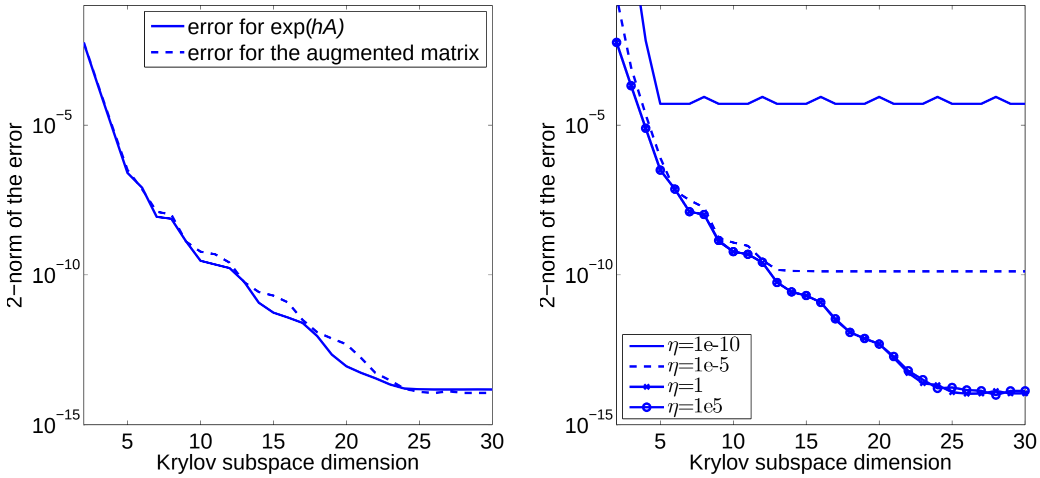


Figure 1: The errors of the Krylov approximations vs. the dimension of the Krylov subspace (left); the influence of the scaling parameter $\eta$ (right).

Setting $\widehat{W} = [w_p, 0, \dots, 0] \in \mathbb{R}^{d\times p}$ and

$$\widehat{A} = \begin{bmatrix} A & \widehat{W} \\ 0 & J \end{bmatrix} \in \mathbb{R}^{(d+p)\times(d+p)},$$

we get

$$\begin{bmatrix} I_d & 0 \end{bmatrix} e^{h\widehat{A}} e_{d+p} = h^p \varphi_p(hA) w_p.$$

So this term can be computed by one matrix exponential.

For the step size control a standard algorithm as described in [12] is implemented. The local error $v = h^p \varphi_p(hA) w_p$ is measured in the weighted and scaled norm

$$\|v\|_E = \sqrt{\frac{1}{d}\sum_{i=1}^{d}\left(\frac{v_i}{\mathrm{scal}_i}\right)^2}, \quad \mathrm{scal}_i = \mathrm{atol} + \mathrm{rtol}\cdot\max(|u_{n,i}|, |u_{n+1,i}|),$$

where the time step is accepted whenever $\|v\|_E \le 1$.

The new step size is computed both after accepted and rejected steps by

$$h_{\mathrm{new}} = h\min\{\nu_{\max}, \max\{\nu_{\min}, \nu\|v\|_E^{-1/p}\}\},$$

where the values $\nu_{\max} = 1.5$, $\nu_{\min} = 0.5$, and $\nu = 0.85$ are used.

## 5. A simple test equation and accumulation of the local errors

To illustrate the instabilities of exponential Taylor methods for $p \geq 3$ (and $p \geq 4$ for the linearized scheme) we consider a simple one-dimensional partial differential equation

$$\partial_t u = \partial_{xx} u + \gamma u(1-u), \qquad x \in \left[-\tfrac{5}{2}, \tfrac{5}{2}\right], \tag{33}$$

subject to periodic boundary conditions, and with initial value

$$u(x,0) = \mathrm{e}^{-10x^2}.$$

We choose $\gamma = 10$ and discretize (33) in space using standard finite differences, which gives for the linear part the discretized Laplacian

$$A = \frac{1}{(\Delta x)^2} \begin{bmatrix} -2 & 1 & & & 1 \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ 1 & & & 1 & -2 \end{bmatrix}, \tag{34}$$

where $\Delta x = L/d$. As we know, $A$ has its eigenvalues on the negative real axis, the smallest satisfying $\lambda_{\min} \approx -4/(\Delta x)^2$. As a result of this spatial discretization we get the semilinear problem

$$u'(t) = Au(t) + g(u(t)), \quad u(0) = u_0.$$

When applying the non-linearized scheme, we see from the expressions (8) that the term $w_p$ contains the highest powers of $A$, namely $g'(u_n)A^{p-1}u_n$. Using the rough approximation

$$|g'(u)| = \gamma|1-2u| \approx \gamma$$

we find that the stability of the exponential Taylor scheme is governed by the factor

$$\left|\gamma h^p \varphi_p(h\lambda_{\min})\lambda_{\min}^{p-1}\right|$$

which has to be power-bounded, i.e.

$$\gamma h^p \varphi_p(h\lambda_{\min})\, |\lambda_{\min}|^{p-1} \leq 1. \tag{35}$$

This (approximate) stability condition for high order exponential Taylor methods restricts the step sizes for stable computations.

| $p$ | $d = 500$ | $d = 1000$ |
|---|---|---|
| 3 | 1.4E-2 / 1.5E-2 | 8.6E-3 / 8.5E-3 |
| 4 | 3.7E-3 / 3.3E-3 | 1.9E-3 / 1.6E-3 |
| 5 | 1.8E-3 / 1.6E-3 | 7.9E-4 / 5.9E-4 |

Table 1: Step sizes determined by condition (35) (left column) vs. experimentally observed maximal step sizes for stable computations (right column) for exponential Taylor methods of (classical) order $p$ and $d = 500$ and 1000 spatial discretization points, respectively.

Table 1 gives a comparison of the step sizes computed from condition (35) with numerically observed maximal step sizes for stable computations, respectively. The results show once more that instabilities occur for $p \geq 3$. Figure 2 shows an exponential growth of the 2-norms of the terms $h^k \varphi_k(hA) w_k$ for step sizes slightly beyond the stability limit.

When applying the linearized scheme of order $p$, we have $g_n'(u_n) = 0$. Therefore, the term $(p-1)g''(u_n)(u_n^{(p-2)}, u_n')$ is expected to start growing first. Now $g''(u) = -2\gamma$, and based on numerical experiments we approximate $u_n' = f(u_n) \approx 1$. Then, similarly as above, we derive an approximate stability condition

$$2(p-1)\gamma h^p \varphi(h\lambda_{\min}) \, |\lambda_{\min}|^{p-2} \leq 1. \tag{36}$$

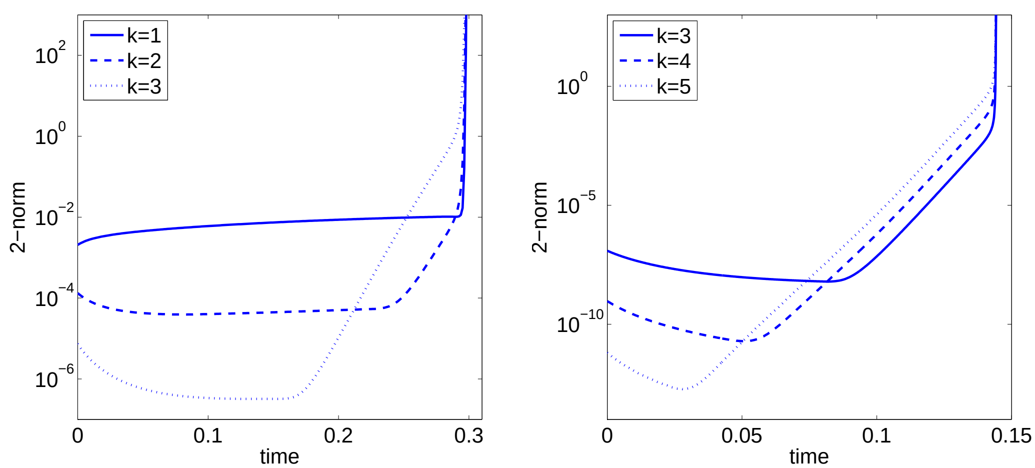


Figure 2: 2-norms of the terms $h^k \varphi_k(hA) w_k$ as a function of time for the exponential Taylor methods with $p = 3$ (left) and $p = 5$ (right), respectively. The experiment was carried out with $d = 1000$, and the step sizes were $h = 1.3 \cdot 10^{-3}$ (left) and $h = 1.8 \cdot 10^{-4}$ (right).

| $p$ | $n = 500$ | $n = 1000$ |
|---|---|---|
| 4 | 1.7E-2 / 1.6E-2 | 10.0E-3 / 8.0E-3 |
| 5 | 3.1E-3 / 3.1E-3 | 2.2E-3 / 1.6E-3 |
| 6 | 1.1E-3 / 1.3E-3 | 8.9E-4 / 6.8E-4 |

Table 2: Step sizes determined by condition (36) (left column) vs. experimentally observed maximal step sizes for stable computations (right column) for linearized exponential Taylor methods of (classical) order $p$ and $d = 500$ and 1000 spatial discretization points, respectively.

In contrast to condition (35), it contains one power of $|\lambda_{\min}|$ less. Table 2 gives again a comparison of the step sizes computed from condition (36) with numerically observed maximal step sizes for stable computations, respectively. We see that the methods are unstable for $p \geq 4$. Figure 3 shows an exponential growth of the 2-norms of the terms $h^k \varphi_k(hJ_n)w_k$ for step sizes slightly beyond the stability limit.

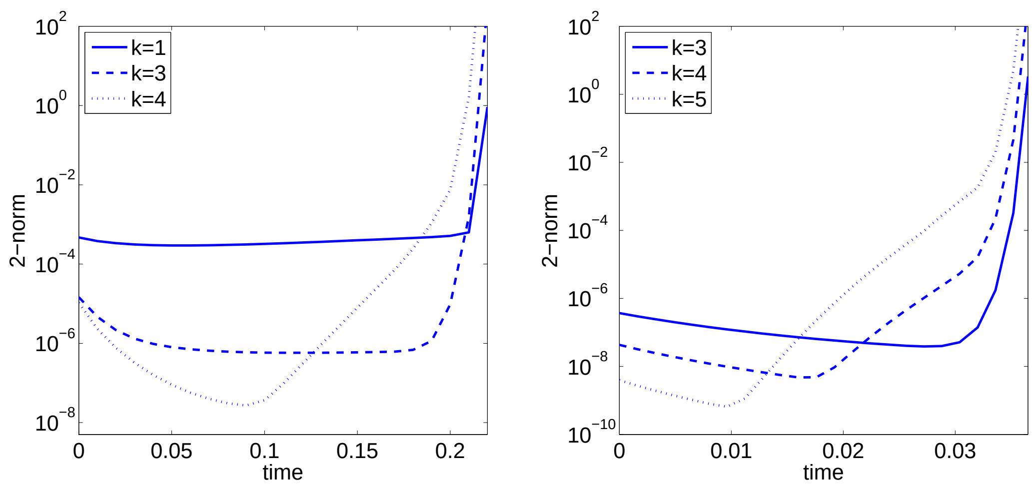


Figure 3: 2-norms of the terms $h^k \varphi_k(hJ_n)w_k$ as a function of time for the linearized exponential Taylor methods with $p = 4$ (left) and $p = 5$ (right), respectively. The experiment was carried out with $d = 500$, and the step sizes were $h = 1.7 \cdot 10^{-2}$ (left) and $h = 3.3 \cdot 10^{-3}$ (right).

## 6. Accumulation of round-off errors

When applying the non-linearized integrator (with $p = 6$) to the test equation (33) with step size $h = 10^{-4}$, we find that the higher order terms $w_i$ are strongly affected by round-off errors (see Figure 4). Due to the smoothing

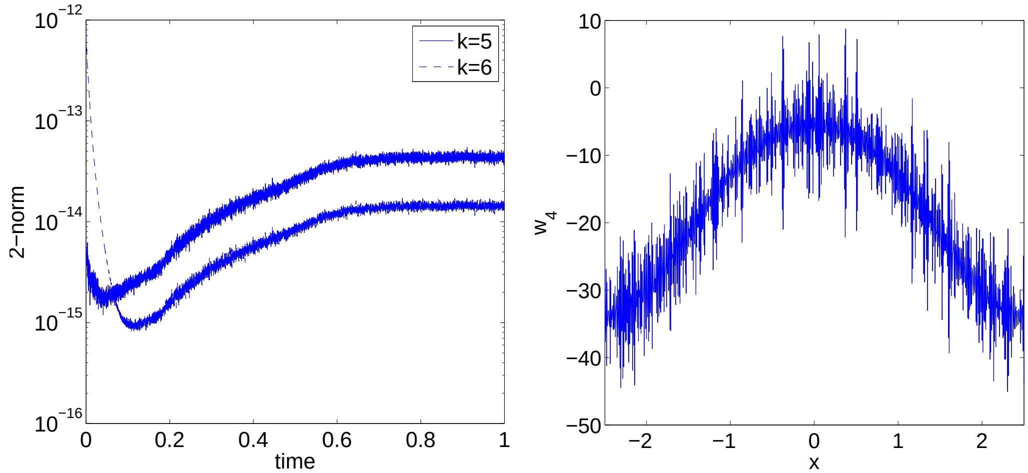


Figure 4: Left: 2-norm of the terms $h^k\varphi_k(hA)w_k$ with $h = 10^{-4}$ as a function of time for $k = 5$ and $k = 6$. Right: the term $w_4$ at time $t = 1$.

property of the equation, however, the round-off errors settle in a neighborhood of the solution and will not amplify before condition (35) is violated.

To observe the amplification of the round-off errors numerically, we apply the method to a hyperbolic equation, the nonlinear Schrödinger equation

$$\mathrm{i}\partial_t\psi = -\Delta\psi - |\psi|^2\,\psi, \qquad x \in [-\pi, \pi], \tag{37}$$

subject to periodic boundary conditions. As initial value we take

$$\psi(x, 0) = \psi_0(x) = \frac{1 + \mathrm{i}}{1 + \sin^2 x}. \tag{38}$$

The spatial discretization is performed with central differences, the discretized Laplacian $A$ is again given by (34). The discretization of the initial value gives a perturbed vector

$$\widetilde{\psi}_0 = \psi_0 + \varepsilon,$$

where $\varepsilon \in \mathbb{R}^d$ is the round-off error. It is observed numerically that these errors are approximately normally distributed, see Figure 5.

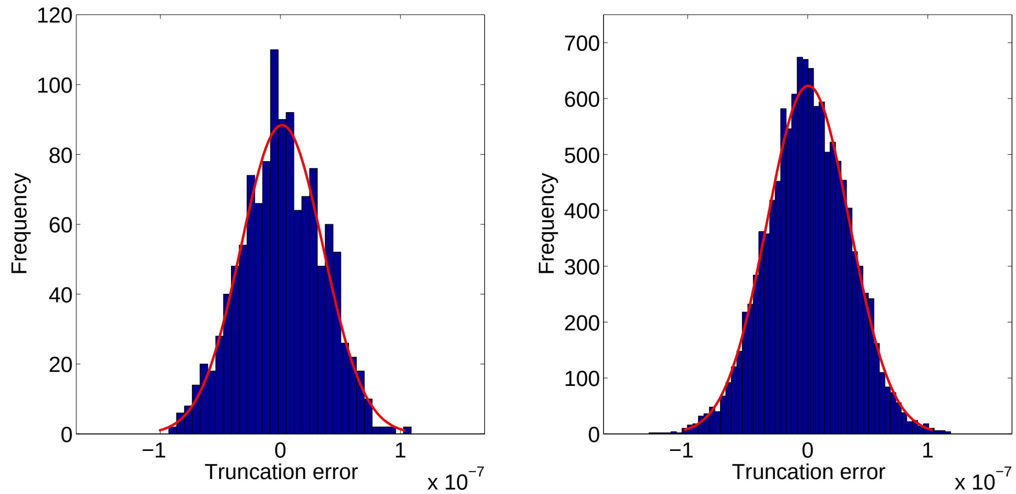


Figure 5: Histogram of the round-off errors $\varepsilon_i$, $i = 1, \ldots, d$ in single precision arithmetic for $d = 1200$ and $d = 12000$ discretization points, respectively. The corresponding normal distributions are fitted with MATLAB's `histfit` command.

This experiment justifies the assumption that the elements of $\varepsilon$ are normally distributed, i.e., for all $1 \le i \le d$

$$\varepsilon_i \sim \mathcal{N}(0, \sigma^2) \quad \text{i.i.d.}$$

The assumption that the round-off errors are statistical variables was already used, for example, in [13]. A discussion including several references can be found in [14].

If $A$ arises from the Laplacian with periodic boundary conditions, we see that

$$[A\varepsilon]_i = \frac{\varepsilon_{i-1} - 2\varepsilon_i + \varepsilon_{i+1}}{(\Delta x)^2} \sim \mathcal{N}\big(0, 6\sigma^2(\Delta x)^{-4}\big),$$

where the indices are taken modulo $d$. Since

$$(\Delta x)^4 \|A\varepsilon\|_2^2 = \sum_{i=1}^{d} \big(6\varepsilon_i^2 + \varepsilon_{i-1}\varepsilon_{i+1} - 4\varepsilon_i\varepsilon_{i+1}\big),$$

the expectation values satisfy $\mathbb{E}(\varepsilon_i^2) = \sigma^2$ and the variables $\varepsilon_i$ are assumed to be independent, we get

$$\mathbb{E}(\|A\varepsilon\|_2^2) = \frac{6d\sigma^2}{(\Delta x)^4},$$

and further

$$\frac{\mathbb{E}\left(\|A\varepsilon\|_2^2\right)}{\mathbb{E}(\|\varepsilon\|_2^2)} = \frac{6}{(\Delta x)^4}.$$

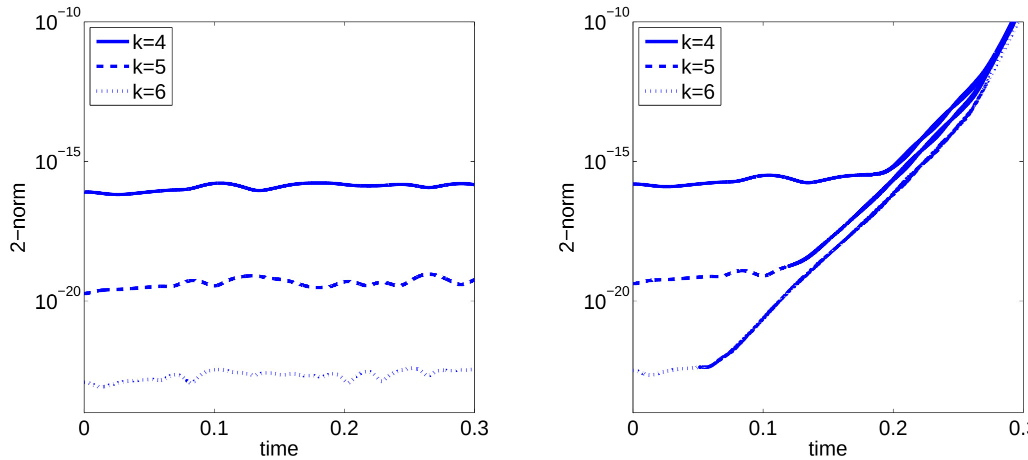


Figure 6: 2-norms of the terms $h^k\varphi_k(hA)w_k$ when using the exponential Taylor method with $p = 6$ and $d = 1200$ discretization points. Left figure: $\alpha \approx 0.8$ ($h = 3.4 \cdot 10^{-5}$); right figure: $\alpha \approx 1.0$ ($h = 4 \cdot 10^{-5}$).

Let

$$\alpha = \frac{\sqrt{6}h}{(\Delta x)^2}.$$

Figure 6 shows the evolution of the 2-norms of the terms $h^k\varphi_k(hA)w_k$ for $k = 4, 5, 6$ when using the 6th order scheme for equation (37) with $d = 1200$. The figure indicates that the growth of the round-off errors starts already near $\alpha = 1$.

## 7. A linear example with inhomogeneity

To illustrate the favorable properties of exponential Taylor methods when applied to equations with linear inhomogeneities, we consider a finite difference spatial discretization (with 500 points)

$$\begin{aligned} \partial_t u &= \partial_{xx} u + 10\,\mathrm{e}^{-10t} x(1-x), \quad x \in [0,1], \quad t \in [0, 0.1] \\ u(x,0) &= 16x^2(1-x)^2, \\ u(0,t) &= u(1,t) = 0. \end{aligned} \tag{39}$$

As expected, no instabilities occur in the linear case. Using the step size control as described in chapter 4, we find that the method with $p = 5$ is able to take larger step sizes than the standard implicit integrators `ode15s` and `ode23s` of MATLAB. Figure 7 shows the resulting step size sequences when requiring a relative error $10^{-7}$ at time $t = 0.1$. The methods take 11, 148 and 868 steps, respectively. To enhance the performance of the MATLAB integrators, a function handle is provided to evaluate the linear part.

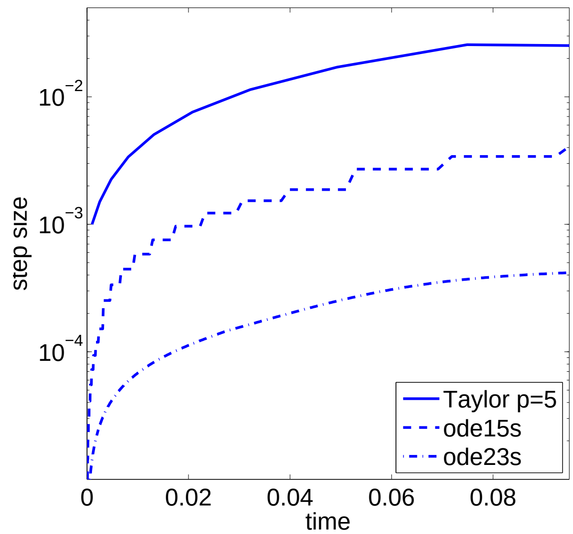


Figure 7: Step sizes taken by the exponential Taylor method with $p = 5$, ode15s and ode23s, respectively, when applied to problem (39).

## References


[1] A. H. Al-Mohy, N. J. Higham, Computing the action of the matrix exponential, with an application to exponential integrators, SIAM J. Sci. Comput. 33 (2010) 488–511.

[2] E. Hairer, S. Nørsett, G. Wanner, Solving Ordinary Differential Equations I. Nonstiff Problems, 2nd Edition, Springer, Berlin, 1993.

[3] A. Griewank, A. Walther, Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation, 2nd Edition, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008.

[4] R. Barrio, F. Blesa, M. Lara, VSVO formulation of the Taylor method for the numerical solution of ODEs, Comput. Math. Appl. 50 (2005) 93–111.

[5] G. Corliss, A. Griewank, P. Henneberger, G. Kirlinger, F. Potra, H. Stetter, High-order stiff ODE solvers via automatic differentiation and rational prediction, in: L. Vulkov, J. Wasniewski, P. Yalamov (Eds.), Nu-

merical Analysis and Its Applications, Vol. 1196 of Lecture Notes in Computer Science, Springer Berlin, 1997, pp. 114–125.

[6] M. Hochbruck, A. Ostermann, Exponential multistep methods of Adams-type, BIT 51 (2011) 889–908.

[7] M. Hochbruck, A. Ostermann, Exponential integrators, Acta Numerica 19 (2010) 209–286.

[8] D. A. Pope, An exponential method of numerical integration of ordinary differential equations, Comm. ACM 6 (1963) 491–493.

[9] M. Hochbruck, A. Ostermann, J. Schweitzer, Exponential Rosenbrock-type methods, SIAM J. Numer. Anal. 47 (2009) 786–803.

[10] N. J. Higham, The scaling and squaring method for the matrix exponential revisited, SIAM J. Matrix Anal. Appl. 26 (2005) 1179–1193.

[11] M. Hochbruck, C. Lubich, On Krylov subspace approximations to the matrix exponential operator, SIAM J. Numer. Anal. 34 (1997) 1911–1925.

[12] E. Hairer, G. Wanner, Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems, 2nd Edition, Springer, Berlin, 1996.

[13] P. Henrici, Error Propagation for Difference Methods, John Wiley, New York, 1963.

[14] N. J. Higham, Accuracy and Stability of Numerical Algorithms, 2nd Edition, Society for Industrial and Applied Mathematics, 2002.