On the convergence of Lawson methods for semilinear stiff problems

Marlis Hochbruck, Alexander Ostermann

KARLSRUHE INSTITUTE OF TECHNOLOGY

CRC 1173

Participating universities

Universität Stuttgart

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Funded by

DFG

On the convergence of Lawson methods for semilinear stiff problems

MARLIS HOCHBRUCK†

Institut für Angewandte und Numerische Mathematik, Karlsruher Institut für Technologie,
D-76149 Karlsruhe, Germany

AND

ALEXANDER OSTERMANN‡

Institut für Mathematik, Universität Innsbruck, A-6020 Innsbruck, Austria

Since their introduction in 1967, Lawson methods have achieved constant interest in the time discretization of evolution equations. The methods were originally devised for the numerical solution of stiff differential equations. Meanwhile, they constitute a well-established class of exponential integrators. The popularity of Lawson methods is in some contrast to the fact that they may have a bad convergence behaviour, since they do not satisfy any of the stiff order conditions. The aim of this paper is to explain this discrepancy. It is shown that non-stiff order conditions together with appropriate regularity assumptions imply high-order convergence of Lawson methods. Note, however, that the term regularity here includes the behaviour of the solution at the boundary. For instance, Lawson methods will behave well in the case of periodic boundary conditions, but they will show a dramatic order reduction for, e.g., Dirichlet boundary conditions. The precise regularity assumptions required for high-order convergence are worked out in this paper and related to the corresponding assumptions for splitting schemes. In contrast to previous work, the analysis is based on expansions of the exact and the numerical solution along the flow of the homogeneous problem. Numerical examples for the Schrödinger equation are included.

Keywords: exponential integrators; Lawson methods; linear and nonlinear Schrödinger equations; evolution equations; order conditions.

1. Introduction

Exponential integrators are a well-established class of methods for the numerical solution of semilinear stiff differential equations. If the stiff initial value problem stems from a spatial semi-discretization of an evolutionary partial differential equation (PDE), the very form of the domain of the spatial differential operator enters the convergence analysis. The *stiff order conditions*, which guarantee a certain order of convergence independently of the considered problem, must be independent of the domain of this operator (which, in general, involves certain boundary conditions). This is the main reason why stiff order conditions for exponential integrators are quite involved (see Hochbruck & Ostermann (2005a) and Luan & Ostermann (2013)).

For particular problems, however, less conditions are required for obtaining a certain order of convergence. (The same is true for ordinary differential equations (ODEs), where linear problems, e.g., require less order conditions for Runge–Kutta methods than nonlinear ones.) It was already observed in Hochbruck & Ostermann (2005b) that periodic boundary conditions do not give any order reduction in

---

†marlis.hochbruck@kit.edu
‡alexander.ostermann@uibk.ac.at

exponential integrators of collocation type in contrast to homogeneous Dirichlet boundary conditions, which restrict the order of convergence considerably (close to the stage order, depending on the precise situation). Full-order convergence for periodic boundary conditions was also noticed in Kassam & Trefethen (2005) and Besse *et al.* (2017).

A similar behaviour can be observed for Lawson methods which are obtained by a linear variable transformation from (explicit) Runge–Kutta methods (see Lawson (1967) and Section 2 below). These methods are very attractive, since they can be easily constructed from any known Runge–Kutta method. Unfortunately, Lawson methods exhibit a strong order reduction, in general. For particular problems, however, they show full order of convergence (see Cano & González-Pachón (2015), Balac *et al.* (2016), and Montanelli & Bootland (2016)). By construction, Lawson methods do satisfy the order conditions for non-stiff problems. Such conditions will be called *non-stiff* or *conventional order conditions* henceforth. However, Lawson methods *do not* satisfy any of the stiff order conditions, as detailed in Hochbruck & Ostermann (2005a), Hochbruck & Ostermann (2010), and Luan & Ostermann (2013). This fact can result in a dramatic order reduction, even down to order one for parabolic problems with homogeneous Dirichlet boundary conditions.

So far, the derivation of (stiff) order conditions for exponential integrators was based on standard expansions of the exact and the numerical solution. There, the main assumption on the problem is that the exact solution and its composition with the nonlinearity are both sufficiently smooth in time (see Hochbruck & Ostermann (2005a) and Luan & Ostermann (2013)). Any additional regularity in space is not of immediate benefit in this analysis. This is in contrast to splitting methods, where spatial regularity usually shows up in form of commutator bounds (see, e.g., Jahnke & Lubich (2000)).

In this paper, we study the convergence behaviour of Lawson methods for semilinear problems. One of the main contributions of this paper is a different expansion of the solution. It is still based on the variation-of-constants formula but the nonlinearity is expanded along the flow of the homogeneous problem. This expansion can be derived in a systematic way using trees as in Hairer *et al.* (1993) and Luan & Ostermann (2013). The expansion of the exact solution is carried out in terms of elementary integrals, that of the numerical solution in terms of elementary quadrature rules. We show that conventional, non-stiff order conditions together with (problem-dependent) assumptions on the exact solution give full order of convergence. This involves regularity of the solution in *space and time*. Our main result for Lawson methods is stated in Theorems 4.6 and 4.7. We prove that Lawson methods converge with order p, if the order of the underlying Runge–Kutta methods is at least p and the solution satisfies appropriate regularity assumptions. These conditions are studied in detail for methods of orders one and two, respectively, and they are related to the corresponding conditions that arise in the analysis of splitting methods. In particular, this is worked out for the nonlinear Schrödinger equation. Our error analysis also reveals a different behaviour between the first-order Lawson method and the exponential Euler method, which is visible in numerical experiments.

The outline of the paper is as follows. In Section 2, we recall the construction of Lawson methods. The expansion of the numerical and the exact solution in terms of elementary integrals is given in Section 3. There, we also introduce the analytic (finite dimensional) framework which typically occurs when discretizing a semilinear parabolic or hyperbolic PDE in space. Order conditions and convergence results are given in Section 4. The resulting regularity assumptions are discussed in Section 5. These assumptions are related to the corresponding conditions for splitting methods. Numerical examples that illustrate the required regularity assumptions and the proven convergence behaviour are also presented.

2. Lawson methods

Consider a semilinear system of stiff differential equations

$$u'(t) + Au(t) = g(t, u(t)), \qquad u(0) = u_0, \tag{2.1}$$

where the stiffness stems from the linear part of the equation, i.e., from A, which is either an unbounded linear operator or its spatial discretization, i.e., a matrix. The precise assumptions on A and g will be given in Section 3. For the numerical solution of (2.1), Lawson (1967) considered the following change of variables:

$$w(t) = e^{tA} u(t).$$

Note that when applied to evolution equations, this transformation has to be done in a formal way, since e^{tA} might not be a meaningful object in our general framework.

Inserting the new variables into (2.1) gives the transformed differential equation

$$\begin{aligned} w'(t) &= e^{tA} \big(u'(t) + Au(t) \big) \\ &= e^{tA} g \big(t, e^{-tA} w(t) \big), \qquad w(0) = u_0. \end{aligned} \tag{2.2}$$

For the solution of this problem, an s-stage explicit Runge–Kutta method with coefficients b_i, c_i, a_{ij} is considered. The method is assumed to satisfy the simplifying assumptions $c_1 = 0$ and

$$\sum_{j=1}^{i-1} a_{ij} = c_i, \qquad i = 2, \dots, s. \tag{2.3}$$

Transforming the Runge–Kutta discretization of (2.2) back to the original variables yields the corresponding Lawson method for (2.1)

$$u_{n+1} = e^{-hA} u_n + h \sum_{i=1}^{s} b_i e^{-(1-c_i)hA} G_{ni}, \tag{2.4a}$$

$$G_{ni} = g \big(t_n + c_i h, U_{ni} \big), \tag{2.4b}$$

$$U_{ni} = e^{-c_i hA} u_n + h \sum_{j=1}^{i-1} a_{ij} e^{-(c_i - c_j)hA} G_{nj}, \qquad i = 1, \dots, s. \tag{2.4c}$$

Here, u_n is the numerical approximation to the exact solution $u(t)$ at time $t = t_n = nh$, and h is the step size. Note that this method makes explicit use of the action of the matrix exponential function. Depending on the properties of A, the nodes $c_1, \dots, c_s$ have to fulfill particular assumptions, see Assumption 3.1 in the next section. Because of these actions of the matrix exponential, Lawson methods form a particular class of exponential integrators. For a review on such integrators, we refer to Hochbruck & Ostermann (2010).

For a non-stiff ordinary differential equation (2.1), it is obvious that the order of the Runge–Kutta method applied to (2.2) coincides with that of the corresponding Lawson method applied to (2.1). It is the aim of this paper to show that this is also true in the stiff situation, if appropriate regularity assumptions hold (we will explain the meaning of regularity in the context of discretized PDEs in Section 5).

3. Expansion of the exact and the numerical solution

By adding $t' = 1$ to (2.1), the differential equation is transformed to autonomous form. It is well known that Runge–Kutta methods of order at least one satisfying (2.3) are invariant under this transformation. Therefore, we restrict ourselves henceforth to the autonomous problem

$$u'(t) + Au(t) = g\big(u(t)\big), \qquad u(0) = u_0. \tag{3.1}$$

Let X be a Hilbert space or a Banach space with norm $\|\cdot\|$. Our main assumptions on A and g are as follows.

ASSUMPTION 3.1 Let A belong to a family $\mathscr{F}$ of linear operators on X such that $-A$ generates a group satisfying

$$\big\|\mathrm{e}^{-tA}\big\| \leqslant C_{\mathscr{F}}, \tag{3.2}$$

with a moderate constant $C_{\mathscr{F}}$, uniformly for all $t \in \mathbb{R}$ and all operators $A \in \mathscr{F}$. It is sufficient to require that $-A$ generates a bounded semigroup (i. e., (3.2) for $t \geqslant 0$), if the nodes c_i of the considered explicit Runge–Kutta method are ordered as $0 = c_1 \leqslant c_2 \leqslant \ldots \leqslant c_s \leqslant 1$.

The set of infinitesimal generators of non-expansive (semi)groups in X is a possible choice for the family $\mathscr{F}$. In addition, the above assumption is typically satisfied in situations where (3.1) stems from a spatial discretization of a semilinear parabolic or hyperbolic partial differential equation. The important fact here is that the constant $C_{\mathscr{F}}$ is independent of the spatial mesh width for finite difference and finite element methods, and independent of the number of ansatz functions in spectral methods. As our error bounds derived below do not depend on A itself but only on the constant $C_{\mathscr{F}}$, they also apply to spatially discretized systems.

ASSUMPTION 3.2 For a given integer $p \geqslant 0$, the nonlinearity g is p times differentiable with bounded derivatives in a neighborhood of the solution of (3.1).

We recall that the solution of (3.1) can be represented in terms of the variation-of-constants formula

$$u(\theta h) = \mathrm{e}^{-\theta hA} u_0 + h \int_0^\theta \mathrm{e}^{-(\theta-\sigma)hA} g\big(u(\sigma h)\big)\mathrm{d}\sigma.$$

Applying this formula recursively and expanding the nonlinearity along the flow of the homogeneous problem yields the following expansion of the exact solution

$$
\begin{aligned}
u(h) &= \mathrm{e}^{-hA} u_0 + h \int_0^1 \mathrm{e}^{-(1-\sigma)hA} g\Big(\mathrm{e}^{-\sigma hA} u_0 + h \int_0^\sigma \mathrm{e}^{-(\sigma-\eta)hA} g\big(u(\eta h)\big)\mathrm{d}\eta\Big)\mathrm{d}\sigma \\
&= \mathrm{e}^{-hA} u_0 + h \int_0^1 \mathrm{e}^{-(1-\sigma)hA} g_\sigma \mathrm{d}\sigma \\
&\quad + h^2 \int_0^1 \mathrm{e}^{-(1-\sigma)hA} g'_\sigma \int_0^\sigma \mathrm{e}^{-(\sigma-\eta)hA} g_\eta \mathrm{d}\eta \mathrm{d}\sigma \\
&\quad + h^3 \int_0^1 \mathrm{e}^{-(1-\sigma)hA} g'_\sigma \int_0^\sigma \mathrm{e}^{-(\sigma-\eta)hA} g'_\eta \int_0^\eta \mathrm{e}^{-(\eta-\xi)hA} g_\xi \mathrm{d}\xi \mathrm{d}\eta \mathrm{d}\sigma \\
&\quad + \tfrac{1}{2}h^3 \int_0^1 \mathrm{e}^{-(1-\sigma)hA} g''_\sigma \Big(\int_0^\sigma \mathrm{e}^{-(\sigma-\eta)hA} g_\eta \mathrm{d}\eta, \int_0^\sigma \mathrm{e}^{-(\sigma-\xi)hA} g_\xi \mathrm{d}\xi\Big)\mathrm{d}\sigma + \mathcal{O}(h^4),
\end{aligned}
\tag{3.3}
$$

where we have used the shorthand notation

$$g_\eta = g_\eta(u_0) = g\big(\mathrm{e}^{-\eta hA} u_0\big), \qquad g_\eta^{(k)} = g_\eta^{(k)}(u_0) = g^{(k)}\big(\mathrm{e}^{-\eta hA} u_0\big), \qquad k \geqslant 1. \tag{3.4}$$

Note that here and throughout the whole section, the constant in the Landau symbol $\mathcal{O}$ only depends on $C_{\mathscr{F}}$ and the derivatives of g, but not explicitly on A itself, i.e., not on the stiffness. Also note that this expansion differs considerably from the previous work (see, e.g., Hochbruck & Ostermann (2005a); Luan & Ostermann (2013)) where the nonlinearity $g(u(t))$ was expanded with respect to t.

Next we perform a similar expansion of the numerical solution (2.4), which yields (again in the autonomous case)

$$
\begin{aligned}
u_1 = {} & \mathrm{e}^{-hA}u_0 + h\sum_{i=1}^{s} b_i \mathrm{e}^{-(1-c_i)hA} g_{c_i} \\
& + h^2 \sum_{i=1}^{s} b_i \mathrm{e}^{-(1-c_i)hA} g'_{c_i} \sum_{j=1}^{i-1} a_{ij} \mathrm{e}^{-(c_i-c_j)hA} g_{c_j} \\
& + h^3 \sum_{i=1}^{s} b_i \mathrm{e}^{-(1-c_i)hA} g'_{c_i} \sum_{j=1}^{i-1} a_{ij} \mathrm{e}^{-(c_i-c_j)hA} g'_{c_j} \sum_{k=1}^{j-1} a_{jk} \mathrm{e}^{-(c_j-c_k)hA} g_{c_k} \\
& + \tfrac{1}{2} h^3 \sum_{i=1}^{s} b_i \mathrm{e}^{-(1-c_i)hA} g''_{c_i} \left(\sum_{j=1}^{i-1} a_{ij} \mathrm{e}^{-(c_i-c_j)hA} g_{c_j}, \sum_{k=1}^{i-1} a_{ik} \mathrm{e}^{-(c_i-c_k)hA} g_{c_k} \right) + \mathcal{O}(h^4).
\end{aligned}
\tag{3.5}
$$

As we have used the variation-of-constants formula and its discrete counterpart, respectively, the expansions of the exact and the numerical solution reflect the well-known tree structure of (explicit) Runge–Kutta methods. In the following we use the classic trees which are well-established for studying the non-stiff order conditions for Runge–Kutta methods, see (Hairer *et al.*, 1993, Section II.2), (Hairer *et al.*, 2006, Section III.1), and references given there.

By $\mathscr{T}$ we denote the set of unlabeled rooted trees. We recall that these trees are defined recursively by

(a) $\bullet \in \mathscr{T}$,

(b) if $\tau_1, \ldots, \tau_k \in \mathscr{T}$, then $[\tau_1, \ldots, \tau_k] \in \mathscr{T}$.

Here, $[\tau_1, \ldots, \tau_k]$ (a k tuple without ordering) denotes the tree which is obtained by concatenating the roots of the trees $\tau_1, \ldots, \tau_k$ via k branches with a new node. This node becomes the root of the tree $[\tau_1, \ldots, \tau_k]$.

For $\tau \in \mathscr{T}$ the elementary differential $D(\tau)$ of a smooth function g is defined recursively in the following way. For $\tau = \bullet$ we have $D(\bullet)(w) = g(w)$, and for $\tau = [\tau_1, \ldots, \tau_k]$ we have

$$
D(\tau)(w) = g^{(k)}(w)\big(D(\tau_1)(w), \ldots, D(\tau_k)(w)\big).
$$

By $\rho(\tau)$ we denote the order of the tree τ which is defined as the number of nodes of $\tau \in \mathscr{T}$. The trees of order less or equal then p are denoted by

$$
\mathscr{T}_p = \{\tau \in \mathscr{T} \mid \rho(\tau) \leqslant p\}.
$$

Motivated by the expansion (3.3) of the exact solution we define elementary integrals.

DEFINITION 3.3 For $\tau \in \mathscr{T}$ and $0 \leqslant \zeta \leqslant 1$ we define the *elementary integral* $G_\zeta(\tau)$, its *integrand* $\Psi_\zeta(\tau)$ and the *multivariate integration operator* $I_\zeta(\tau)$ recursively in the following way.

(a) For $\tau = \bullet$ and a univariate function f, we set

$$\Psi_\zeta(\bullet)(\sigma, w) = e^{-(\zeta-\sigma)hA} g_\sigma(w),$$

$$I_\zeta(\bullet)f = \int_0^\zeta f(\sigma)d\sigma.$$

(b) For $\tau = [\tau_1, \ldots, \tau_k]$ and a multivariate function f in $\rho(\tau)$ variables, we set

$$\Psi_\zeta(\tau)(\sigma, \cdot_1, \ldots, \cdot_k, w) = e^{-(\zeta-\sigma)hA} g_\sigma^{(k)}(w)\left(\Psi_\sigma(\tau_1)(\cdot_1, w), \ldots, \Psi_\sigma(\tau_k)(\cdot_k, w)\right),$$

$$I_\zeta(\tau)f = \int_0^\zeta I_\sigma(\tau_1)\cdots I_\sigma(\tau_k)f(\sigma, \cdot_1, \ldots, \cdot_k)d\sigma.$$

Here, $\cdot_j$ refers to the variables corresponding to the jth subtree τ_j.

Finally, we define for all $\tau \in \mathscr{T}$ the *elementary integrals* as

$$G_\zeta(\tau)(w) = I_\zeta(\tau)\Psi_\zeta(\tau)(\cdot, w)$$

and set $\Psi(\tau) = \Psi_1(\tau)$, $I(\tau) = I_1(\tau)$, and $G(\tau) = G_1(\tau)$.

For example, we have $\mathbf{\cdot\!\!\cdot} = [\bullet]$ and

$$\Psi(\mathbf{\cdot\!\!\cdot})(\sigma_1, \sigma_2, w) = e^{-(1-\sigma_1)hA} g_{\sigma_1}'(w) e^{-(\sigma_1-\sigma_2)hA} g_{\sigma_2}(w).$$

It is straightforward to verify that the elementary integrals satisfy the recurrence relation

$$G_\zeta(\bullet)(w) = \int_0^\zeta e^{-(\zeta-\sigma)hA} g_\sigma(w)d\sigma,$$

$$G_\zeta(\tau)(w) = \int_0^\zeta e^{-(\zeta-\sigma)hA} g_\sigma^{(k)}(w)\left(G_\sigma(\tau_1)(w), \ldots, G_\sigma(\tau_k)(w)\right)d\sigma$$

for $\tau = [\tau_1, \ldots, \tau_k]$.

Our assumptions on g and A ensure that the integrand $\Psi_\zeta(\tau)(\cdot, w)$ is bounded if $\tau \in \mathscr{T}_{p+1}$ for w in a neighborhood of the exact solution of (3.1) and h sufficiently small.

REMARK 3.4 In the nonstiff situation, where $A \equiv 0$, all evaluations of g or its derivatives are at the fixed value w. Thus $G_\zeta(\tau)(w)$ reduces to a multivariate integral over the constant integrand $\Psi_\zeta(\tau)(\cdot, w) \equiv D(\tau)(w)$.

The following theorem shows how the expansion (3.3) can be expressed as a (truncated) B-series. Here we use the notation from (Hairer *et al.*, 2006, Section III.1).

THEOREM 3.5 The exact solution of (2.1) satisfies

$$u(\zeta h) = e^{-\zeta hA}u_0 + B_p(u_0)(\zeta) + \mathscr{O}(h^{p+1}), \qquad \zeta \in [0, 1], \tag{3.6}$$

where we define the B-series for $w \in X$ and $\zeta \in [0, 1]$ as

$$B_p(w)(\zeta) = \sum_{\tau \in \mathscr{T}_p} \frac{h^{\rho(\tau)}}{\sigma(\tau)} G_\zeta(\tau)(w),$$

with the symmetry coefficients $\sigma(\bullet) = 1$ and $\sigma([\tau_1, \ldots, \tau_k]) = \sigma(\tau_1)\cdots\sigma(\tau_k)\mu_1!\mu_2!\cdots$. The integers $\mu_1, \mu_2, \ldots$ specify the number of equal trees among $\tau_1, \ldots, \tau_k$.

Proof. The proof is done by induction. For $p = 0$, the claim follows from the variation-of-constants formula and $B_0(u_0)(\zeta) = 0$, since

$$u(\zeta h) = \mathrm{e}^{-\zeta hA} u_0 + h \int_0^\zeta \mathrm{e}^{-(\zeta-\sigma)hA} g\big(u(\sigma h)\big) \mathrm{d}\sigma = \mathrm{e}^{-\zeta hA} u_0 + \mathcal{O}(h). \tag{3.7}$$

The induction step follows the lines of the proof of (Hairer *et al.*, 2006, Lemma III.1.9) with the following modifications: we use the variation-of-constants formula and truncate the series in such a way that only the first p derivatives of g enter the expansion. We omit the details. $\square$

Now we proceed analogously for the numerical solution starting with the definition of elementary quadrature rules.

DEFINITION 3.6 For $\tau \in \mathscr{T}$ we define the *multivariate quadrature operators* $\widehat{I}(\tau)$ and $\widehat{I}_i(\tau)$, $i = 1, \ldots, s$, recursively in the following way.

(a) For $\tau = \bullet$ and a univariate function f, we set

$$\widehat{I}(\bullet)f = \sum_{j=1}^s b_j f(c_j), \qquad \widehat{I}_i(\bullet)f = \sum_{j=1}^{i-1} a_{ij} f(c_j), \quad 1 \leqslant i \leqslant s.$$

(b) For $\tau = [\tau_1, \ldots, \tau_k]$ and a multivariate function f in $\rho(\tau)$ variables, we set

$$\widehat{I}(\tau)f = \sum_{j=1}^s b_j \widehat{I}_j(\tau_1) \cdots \widehat{I}_j(\tau_k) f(c_j, \cdot_1, \ldots, \cdot_k),$$

$$\widehat{I}_i(\tau)f = \sum_{j=1}^{i-1} a_{ij} \widehat{I}_j(\tau_1) \cdots \widehat{I}_j(\tau_k) f(c_j, \cdot_1, \ldots, \cdot_k), \quad 1 \leqslant i \leqslant s.$$

Finally, we define the *elementary quadrature rules* in the following way

$$\widehat{G}(\tau)(w) = \widehat{I}(\tau)\Psi(\tau)(\cdot, w), \qquad \widehat{G}_i(\tau)(w) = \widehat{I}_i(\tau)\Psi_{c_i}(\tau)(\cdot, w), \quad 1 \leqslant i \leqslant s.$$

It is easy to see from the recursive definitions that the elementary quadrature rules satisfy

$$\widehat{G}(\bullet)(w) = \sum_{j=1}^s b_j \mathrm{e}^{-(1-c_j)hA} g_{c_j}(w),$$

$$\widehat{G}_i(\bullet)(w) = \sum_{j=1}^{i-1} a_{ij} \mathrm{e}^{-(c_i-c_j)hA} g_{c_j}(w), \quad 1 \leqslant i \leqslant s$$

and

$$\widehat{G}(\tau)(w) = \sum_{j=1}^s b_j \mathrm{e}^{-(1-c_j)hA} g_{c_j}^{(k)}(w) \big(\widehat{G}_j(\tau_1)(w), \ldots, \widehat{G}_j(\tau_k)(w)\big),$$

$$\widehat{G}_i(\tau)(w) = \sum_{j=1}^{i-1} a_{ij} \mathrm{e}^{-(c_i-c_j)hA} g_{c_j}^{(k)}(w) \big(\widehat{G}_j(\tau_1)(w), \ldots, \widehat{G}_j(\tau_k)(w)\big), \quad 1 \leqslant i \leqslant s$$

for $\tau = [\tau_1, \ldots, \tau_k]$. This allows us to express the expansion (3.5) of the numerical solution in terms of elementary quadrature rules.

THEOREM 3.7 The numerical solution of (3.1) satisfies

$$u_1 = \mathrm{e}^{-hA} u_0 + \widehat{B}_p(u_0) + \mathscr{O}(h^{p+1}), \tag{3.8a}$$

$$U_{0i} = \mathrm{e}^{-c_i hA} u_0 + \widehat{B}_p^{(i)}(u_0) + \mathscr{O}(h^{p+1}), \qquad i = 1, \ldots, s, \tag{3.8b}$$

where we define the numerical B-series for $w \in X$ as

$$\widehat{B}_p(w) = \sum_{\tau \in \mathscr{T}_p} \frac{h^{\rho(\tau)}}{\sigma(\tau)} \widehat{G}(\tau)(w),$$

$$\widehat{B}_p^{(i)}(w) = \sum_{\tau \in \mathscr{T}_p} \frac{h^{\rho(\tau)}}{\sigma(\tau)} \widehat{G}_i(\tau)(w), \quad i = 1, \ldots, s.$$

Proof. The proof is done analogously to the proof of Theorem 3.7. □

4. Order conditions and convergence

In this section we present a systematic way of deriving general stiff convergence results for Lawson methods based on trees.

The expansions of the exact and the numerical solution in terms of elementary integrals and elementary quadrature rules derived in the previous section allow us to study the local error in the same way as for classical Runge–Kutta methods. In fact we show that the orders of these quadrature rules determine the local error of the Lawson method. A similar strategy was used in the analysis of splitting methods by Jahnke & Lubich (2000). General stiff order conditions for exponential Runge–Kutta methods have been derived in Luan & Ostermann (2013) and for splitting methods in Hansen & Ostermann (2016).

As usual, we say that the Lawson method is of (stiff) order p if the local error satisfies

$$\|u(h) - u_1\| \leqslant C h^{p+1}$$

uniformly for smooth nonlinearities and operators A satisfying Assumption 3.1, meaning that the constant C depends on the constant $C_{\mathscr{F}}$ defined in (3.2) but not on A itself.

THEOREM 4.1 The Lawson method (2.4) is of order p if

$$\widehat{G}(\tau)(u_0) - G(\tau)(u_0) = \mathscr{O}(h^{p+1-\rho(\tau)}), \qquad \text{for all } \tau \in \mathscr{T}_p.$$

Proof. From Theorems 3.5 and 3.7 we have

$$u(h) - u_1 = \sum_{\tau \in \mathscr{T}_p} \frac{h^{\rho(\tau)}}{\sigma(\tau)} \big(G(\tau)(u_0) - \widehat{G}(\tau)(u_0) \big) + \mathscr{O}(h^{p+1}). \tag{4.1}$$

This proves the statement. □

REMARK 4.2 The above derivation can be easily generalized to exponential integrators with a fixed linearization

$$U_i = \mathrm{e}^{-c_i hA} u_0 + h \sum_{j=1}^{i-1} a_{ij}(-hA) g(U_j), \qquad i = 1, \ldots, s,$$

$$u_1 = \mathrm{e}^{-hA} u_0 + h \sum_{i=1}^{s} b_i(-hA) g(U_i),$$

cf. Hochbruck & Ostermann (2010). If one replaces $b_i \mathrm{e}^{-(1-c_i)hA}$ by $b_i(-hA)$ and $a_{ij}\mathrm{e}^{-(c_i-c_j)hA}$ by $a_{ij}(-hA)$ in Definition 3.6, Theorems 3.7 and 4.1 also hold for general exponential Runge–Kutta methods. If the stiff order conditions derived in Hochbruck & Ostermann (2005a) and Luan & Ostermann (2013) are satisfied up to order p, then $\widehat{G}(\tau)(u_0) - G(\tau)(u_0) = \mathcal{O}(h^{p+1-\rho(\tau)})$ for all $\tau \in \mathscr{T}_p$.

EXAMPLE 4.3 For the exponential Euler method, where $s = 1$, $c_1 = 0$, and $b_1(z) = \varphi_1(z)$, we have

$$\widehat{G}(\bullet)(u_0) = b_1(-hA)g(u_0) = \varphi_1(-hA)g(u_0) = \int_0^1 \mathrm{e}^{-(1-\sigma)hA} g(u_0)\mathrm{d}\sigma$$

and thus

$$\widehat{G}(\bullet)(u_0) - G(\bullet)(u_0) = \int_0^1 \mathrm{e}^{-(1-\sigma)hA} \big(g(u_0) - g_\sigma(u_0)\big)\mathrm{d}\sigma.$$

The condition for order one requires that $\|g(u_0) - g_\sigma(u_0)\| \leqslant Ch$. In the linear case, where $g(u) = Bu$, this can be written as

$$h^{-1}\big(g(u_0) - g_\sigma(u_0)\big) = h^{-1}B\big(I - \mathrm{e}^{-\sigma hA}\big)u_0 = B\varphi_1(-\sigma hA)\sigma A u_0. \tag{4.2}$$

Hence the condition is fulfilled if Au_0 is uniformly bounded, i. e., $u_0 \in \mathscr{D}(A)$. For the convergence, we thus need $u(t) \in \mathscr{D}(A)$ for $t \in [0,T]$.

It might be interesting to compare (4.2) to the condition given in (Hochbruck & Ostermann, 2010, Lemma 2.13) which was proved by a Taylor series expansion of $g\big(u(t)\big)$. For linear problems, it reads

$$\|B(Au(t) + Bu(t))\| \leqslant C. \tag{4.3}$$

Hence both results require the same regularity, namely that $Au(t)$ is uniformly bounded. Note, however, that (4.3) does not involve the φ_1 function. The latter decays like $1/z$ as $z \to \infty$ in the closed left half-plane, hence components corresponding to eigenvalues with large negative real part are damped.

COROLLARY 4.4 If the underlying Runge–Kutta method is of (conventional) order p then

$$I(\tau)1 = \widehat{I}(\tau)1 \qquad \text{for all } \tau \in \mathscr{T}_p, \tag{4.4}$$

where $1 : [0,1]^{\rho(\tau)} \to \mathbb{R} : \sigma \mapsto 1$ denotes the multivariate constant function with value one.

Proof. First note that for problems with $A \equiv 0$, we have

$$G(\tau)(w) = I(\tau)D(\tau)(w), \qquad \widehat{G}(\tau)(w) = \widehat{I}(\tau)D(\tau)(w).$$

On the one hand, classical Runge–Kutta theory implies that the local error (4.1) behaves as $\mathcal{O}(h^{p+1})$ for any sufficiently smooth g. On the other hand, the elementary differentials $D(\tau)$ are known to be independent. Hence, we obtain that $G(\tau)(w) = \widehat{G}(\tau)(w)$. The statement follows because the integrand $\Psi(\tau)(\cdot, w) \equiv D(\tau)(w)$ is a constant. This yields (4.4). $\qquad\square$

Since the convergence analysis of Lawson methods also employs Taylor expansion, we next study quadrature of monomials. For $\sigma = (\sigma_1, \ldots, \sigma_q) \in [0,1]^q$ and a given vector $\kappa = (\kappa_1, \ldots, \kappa_q) \in \mathbb{N}_0^q$ of non-negative integers, we define as usual

$$\sigma^\kappa = \sigma_1^{\kappa_1} \cdot \ldots \cdot \sigma_q^{\kappa_q}$$

and set $\kappa! = \kappa_1! \cdots \kappa_q!$ and $|\kappa| = \kappa_1 + \ldots + \kappa_q$. Moreover, we denote the q-variate monomial function of degree $|\kappa|$ by

$$\mathbf{z}^\kappa : [0,1]^q \to \mathbb{R} : \sigma \mapsto \sigma^\kappa.$$

We note that $\mathbf{z}^\kappa = \mathbf{z}_1^{\kappa_1} \ldots \mathbf{z}_q^{\kappa_q}$ and $\mathbf{z}^0 = 1$.

It turns out that a multivariate integration (or quadrature) w.r.t. τ of such monomials corresponds to the integration (or quadrature) of the constant one function w.r.t. a particular higher order tree stemming from τ.

LEMMA 4.5 Let $\tau \in \mathscr{T}$ and $\kappa \in \mathbb{N}_0^{\rho(\tau)}$. Labeling the nodes of τ with the numbers $\{1, \ldots, \rho(\tau)\}$ (in an arbitrary order) we denote by $\tau^{(\kappa)} \in \mathscr{T}$ the tree stemming from τ where κ_j leafs are added to its jth node. Then $\rho(\tau^{(\kappa)}) = \rho(\tau) + |\kappa|$ and

$$I_\zeta(\tau)\mathbf{z}^\kappa = I_\zeta(\tau^{(\kappa)})1, \qquad \widehat{I}(\tau)\mathbf{z}^\kappa = \widehat{I}(\tau^{(\kappa)})1, \qquad \widehat{I}_i(\tau)\mathbf{z}^\kappa = \widehat{I}_i(\tau^{(\kappa)})1, \qquad i = 1, \ldots, s.$$

Proof. We prove the lemma by induction on $\rho(\tau)$. The tree $\tau = \bullet$ is the unique tree with $\rho(\tau) = 1$. For $k \in \mathbb{N}_0$, we have $\bullet^{(k)} = [\bullet]^k$, where $[\bullet]^k = [\bullet, \ldots, \bullet]$ denotes the bush with k leafs. Using $\sigma = I_\sigma(\bullet)1$ and the recursive definition of $I_\zeta(\tau)$ we obtain

$$I_\zeta(\bullet)\mathbf{z}^k = \int_0^\zeta \sigma^k \mathrm{d}\sigma = \int_0^\zeta \big(I_\sigma(\bullet)1\big)^k \mathrm{d}\sigma = I_\zeta([\bullet]^k)1 = I_\zeta(\bullet^{(k)})1.$$

Analogously, for $i = 1, \ldots, s$, the simplifying assumptions yield $c_i = \widehat{I}_i(\bullet)1$ and this gives

$$\widehat{I}_i(\bullet)\mathbf{z}^k = \sum_{j=1}^{i-1} a_{ij} c_j^k = \sum_{j=1}^{i-1} a_{ij} \big(\widehat{I}_j(\bullet)1\big)^k = \widehat{I}_i([\bullet]^k)1 = \widehat{I}_i(\bullet^{(k)})1$$

and $\widehat{I}(\bullet)\mathbf{z}^k = \widehat{I}(\bullet^{(k)})1$.

For the induction step, consider the tree $\tau = [\tau_1, \ldots, \tau_m]$. We regroup σ and κ according to the tree structure of τ as $\sigma = (\mu, \sigma_1, \ldots, \sigma_m) \in \mathbb{R}^{\rho(\tau)}$, $\kappa = (k, \kappa_1, \ldots, \kappa_m) \in \mathbb{N}_0^{\rho(\tau)}$ with $\mu \in \mathbb{R}$, $k \in \mathbb{N}_0$, $\sigma_j \in \mathbb{R}^{\rho(\tau_j)}$, $\kappa_j \in \mathbb{N}_0^{\rho(\tau_j)}$ for $j = 1, \ldots, m$. Then it holds $\sigma^\kappa = \mu^k \cdot \sigma_1^{\kappa_1} \cdot \ldots \cdot \sigma_m^{\kappa_m}$, $\mathbf{z}^\kappa = \mathbf{z}_0^k \mathbf{z}_1^{\kappa_1} \ldots \mathbf{z}_m^{\kappa_m}$, and $\mathbf{z}^0 = 1$. The recursive definition of $\widehat{I}_i(\tau)$ and the induction hypothesis imply

$$\begin{aligned}
\widehat{I}_i(\tau)\mathbf{z}^\kappa &= \sum_{j=1}^{i-1} a_{ij} \widehat{I}_j(\tau_1) \cdots \widehat{I}_j(\tau_m) c_j^k \mathbf{z}_1^{\kappa_1} \cdots \mathbf{z}_m^{\kappa_m} \\
&= \sum_{j=1}^{i-1} a_{ij} c_j^k \big(\widehat{I}_j(\tau_1)\mathbf{z}_1^{\kappa_1}\big) \cdots \big(\widehat{I}_j(\tau_m)\mathbf{z}_m^{\kappa_m}\big) \\
&= \sum_{j=1}^{i-1} a_{ij} \big(\widehat{I}_j(\bullet)1\big)^k \big(\widehat{I}_j(\tau_1^{(\kappa_1)})1\big) \cdots \big(\widehat{I}_j(\tau_m^{(\kappa_m)})1\big) \\
&= \widehat{I}_i(\tau^{(\kappa)})1
\end{aligned}$$

for $i = 1, \ldots, s$, where we used that the sprouted tree can be cast recursively as

$$\tau^{(\kappa)} = [\bullet, \ldots, \bullet, \tau_1^{(\kappa_1)}, \ldots, \tau_m^{(\kappa_m)}].$$

The assertion for $I_\zeta(\tau)$ and $\widehat{I}(\tau)$ can be shown analogously. $\qquad\square$

The following theorem provides a sufficient condition for Lawson methods being of (stiff) order p. Here, $C^{k,1}(\mathbb{R}^d, X)$ denotes the space of k time continuously differentiable functions which have a Lipschitz continuous kth derivative.

THEOREM 4.6 Let the integrand $\Psi(\tau)$ of $G(\tau)$ satisfy

$$\Psi(\tau)\big(\cdot, u(t)\big) \in C^{p-\rho(\tau),1}\big([0,1]^{\rho(\tau)}, X\big) \qquad \text{for all } \tau \in \mathscr{T}_p, \tag{4.5}$$

where $u(t) \in X$ is the solution of (3.1), $0 \leqslant t \leqslant T$. If the underlying Runge–Kutta method is of (conventional) order p, then the Lawson method (2.4) is of (stiff) order p.

Proof. Let $\tau \in \mathscr{T}$ such that $\rho(\tau) \leqslant p$. We approximate $\Psi(\tau)(\cdot, u_0) \colon [0,1]^{\rho(\tau)} \to X$ by a multivariate Taylor polynomial of degree $p - \rho(\tau)$. By assumption on $\Psi(\tau)$, the coefficients and the remainder of this Taylor polynomial are bounded. Using the linearity of the multivariate integrals and quadrature rules, we have by Lemma 4.5

$$\begin{aligned}
G(\tau)(u_0) - \widehat{G}(\tau)(u_0) &= I(\tau)\Psi(\tau)(\cdot, u_0) - \widehat{I}(\tau)\Psi(\tau)(\cdot, u_0) \\
&= \sum_{|\kappa| \leqslant p-\rho(\tau)} \frac{1}{\kappa!} D^\kappa(\Psi(\tau))(0, u_0)\left(I(\tau) - \widehat{I}(\tau)\right)\mathbf{z}^\kappa + \mathscr{O}(h^{p-\rho(\tau)+1}) \\
&= \sum_{|\kappa| \leqslant p-\rho(\tau)} \frac{1}{\kappa!} D^\kappa(\Psi(\tau))(0, u_0)\left(I(\tau^{(\kappa)})1 - \widehat{I}(\tau^{(\kappa)})1\right) + \mathscr{O}(h^{p-\rho(\tau)+1}).
\end{aligned}$$

Here we used $\|D^\kappa\Psi(\tau)\| = \mathscr{O}(h^{|\kappa|})$ to bound the remainder term. Since the Runge–Kutta method is of order p, the claim now follows from $\rho(\tau^{(\kappa)}) = \rho(\tau) + |\kappa| \leqslant p$ and Corollary 4.4 which implies $I(\tau^{(\kappa)})1 = \widehat{I}(\tau^{(\kappa)})1$. $\qquad\square$

This result now allows us to prove an error bound for Lawson methods which is uniform for all problems (3.1) with A satisfying Assumption 3.1.

THEOREM 4.7 Let u be the solution of (3.1) and let the assumptions of Theorem 4.6 be satisfied. If the underlying Runge–Kutta method is of (conventional) order p, then there exists $h_0 > 0$ such that for all $0 < h \leqslant h_0$ sufficiently small,

$$\|u(t_n) - u_n\| \leqslant Ch^p, \qquad t_n = nh \leqslant T,$$

where C and h_0 are independent of n, h, and A.

Proof. We define a norm by

$$\|v\|_\star = \sup_{t \in \mathbb{R}} \left\|\mathrm{e}^{-tA}v\right\|.$$

This norm is equivalent to $\|\cdot\|$ and we have in the corresponding operator norm

$$\left\|\mathrm{e}^{-tA}\right\|_\star \leqslant 1, \qquad \text{for all } t \in \mathbb{R}. \tag{4.6}$$

If $-A$ only generates a bounded semigroup, then taking the supremum only over $t \geqslant 0$ shows that $-A$ generates a contraction semigroup.

By assumption, g is locally Lipschitz continuous. Then (4.6) and Theorem 3.7 show that the Lawson method is locally Lipschitz with respect to the initial value with a Lipschitz constant of size $1 + \mathscr{O}(h)$. This implies the required stability.

The error bound follows in a standard way using Lady Windermere's fan. $\qquad\square$

5. Regularity conditions and applications

It remains to discuss the regularity conditions (4.5) and to give some applications. We first examine the conditions for orders one and two, respectively. The extension to higher orders is a tedious but straightforward exercise. It turns out that these regularity conditions can all be expressed in terms of commutators, very much like in the case of splitting methods.

In order to obtain simple sufficient conditions, we replace the space $C^{k,1}(\Omega, X)$ in condition (4.5) by the subspace of $k+1$ times partially differentiable functions with uniformly bounded partial derivatives on Ω in the following discussion. This is also justified by the fact that Lipschitz continuous functions are almost everywhere differentiable (Rademacher's theorem).

5.1 *Condition for order one*

Since $p = \rho(\tau) = 1$, we only have to consider the tree $\tau = \bullet$ in (4.5). Differentiating

$$\Psi(\bullet)(\sigma, w) = e^{-(1-\sigma)hA} g_\sigma(w) = e^{-(1-\sigma)hA} g\big(e^{-\sigma hA} w\big) \tag{5.1}$$

with respect to σ yields

$$\begin{aligned} \partial_\sigma \Psi(\bullet)(\sigma, w) &= h e^{-(1-\sigma)hA} \Big(A g_\sigma(w) - g'_\sigma(w) A e^{-\sigma hA} w\Big) \\ &= h e^{-(1-\sigma)hA} [F_A, g]\big(e^{-\sigma hA} w\big), \end{aligned} \tag{5.2}$$

where $[F_A, g]$ denotes the Lie commutator of g and $F_A(w) = Aw$, defined as

$$[F_A, g](w) = F'_A(w) g(w) - g'(w) F_A(w) = A g(w) - g'(w) Aw. \tag{5.3}$$

From this calculation, we conclude the following result. If the bound

$$\sup_{0 \leqslant \sigma \leqslant 1} \sup_{0 \leqslant t \leqslant T} \left\| e^{-(1-\sigma)hA} [F_A, g]\big(e^{-\sigma hA} u(t)\big) \right\| \leqslant C \tag{5.4}$$

holds with a constant C that is allowed to depend on $C_{\mathscr{F}}$, then a Lawson method of non-stiff order *one* has also *stiff* order *one*.

5.2 *Conditions for order two*

Stiff order two is achieved if we require the following two regularity conditions

$$\Psi(\bullet)(\cdot, u(t)) \in C^2([0,1], X) \qquad \text{and} \qquad \Psi(\overset{\bullet}{\bullet})(\cdot, u(t)) \in C^1([0,1]^2, X).$$

We commence with the first condition and exploit the fact that $\partial_\sigma \Psi(\bullet)(\sigma, w)$ is of exactly the same form as (5.1) with g replaced by the vector field $[F_A, g]$. Hence from (5.2) we have

$$\partial_\sigma^2 \Psi(\bullet)(\sigma, w) = h^2 e^{-(1-\sigma)hA} \big[F_A, [F_A, g]\big]\big(e^{-\sigma hA} w\big).$$

Therefore, the bound

$$\sup_{0 \leqslant \sigma \leqslant 1} \sup_{0 \leqslant t \leqslant T} \left\| e^{-(1-\sigma)hA} \big[F_A, [F_A, g]\big]\big(e^{-\sigma hA} u(t)\big) \right\| \leqslant C \tag{5.5}$$

should hold with a constant C that is independent of $\|A\|$.

Next, we move to the second condition. Differentiating

$$\Psi(\checkmark)(\sigma_1,\sigma_2,w)=\mathrm{e}^{-(1-\sigma_1)hA}g'_{\sigma_1}(w)\mathrm{e}^{-(\sigma_1-\sigma_2)hA}g_{\sigma_2}(w)$$

with respect to σ_1 and σ_2 yields

$$\begin{aligned}\partial_{\sigma_1}\Psi(\checkmark)(\sigma_1,\sigma_2,w)&=h\mathrm{e}^{-(1-\sigma_1)hA}\Big(Ag'_{\sigma_1}(w)\mathrm{e}^{-(\sigma_1-\sigma_2)hA}g_{\sigma_2}(w)\\&\quad-g''_{\sigma_1}(w)\big(\mathrm{e}^{-\sigma_1hA}Aw,\mathrm{e}^{-(\sigma_1-\sigma_2)hA}g_{\sigma_2}(w)\big)-g'_{\sigma_1}(w)A\mathrm{e}^{-(\sigma_1-\sigma_2)hA}g_{\sigma_2}(w)\Big)\\&=h\mathrm{e}^{-(1-\sigma_1)hA}[F_A,g]'\big(\mathrm{e}^{-\sigma_1hA}w\big)\mathrm{e}^{-(\sigma_1-\sigma_2)hA}g_{\sigma_2}(w),\end{aligned}$$

since by definition (5.3) the derivative of the commutator satisfies

$$[F_A,g]'(w)v=\frac{d}{dw}\big([F_A,g]\big)(w)v=Ag'(w)v-g''(w)(Aw,v)-g'(w)Av.\tag{5.6}$$

Moreover, we have

$$\partial_{\sigma_2}\Psi(\checkmark)(\sigma_1,\sigma_2,w)=h\mathrm{e}^{-(1-\sigma_1)hA}g'_{\sigma_1}(w)\mathrm{e}^{-(\sigma_1-\sigma_2)hA}[F_A,g]\big(\mathrm{e}^{-\sigma_2hA}w\big),$$

respectively. From these two relations, we infer that the bounds

$$\sup_{0\leqslant\sigma_1,\sigma_2\leqslant1}\ \sup_{0\leqslant t\leqslant T}\left\|\mathrm{e}^{-(1-\sigma_1)hA}[F_A,g]'\big(\mathrm{e}^{-\sigma_1hA}u(t)\big)\mathrm{e}^{-(\sigma_1-\sigma_2)hA}g_{\sigma_2}(u(t))\right\|\leqslant C,\tag{5.7a}$$

$$\sup_{0\leqslant\sigma_1,\sigma_2\leqslant1}\ \sup_{0\leqslant t\leqslant T}\left\|\mathrm{e}^{-(1-\sigma_1)hA}g'_{\sigma_1}(u(t))\mathrm{e}^{-(\sigma_1-\sigma_2)hA}[F_A,g]\big(\mathrm{e}^{-\sigma_2hA}u(t)\big)\right\|\leqslant C\tag{5.7b}$$

should hold with a constant C that is independent of $\|A\|$.

From the above calculations, we conclude the following result. If the conditions (5.4), (5.5), and (5.7) hold with a constant C that does not depend on $\|A\|$, then a Lawson method of non-stiff order *two* has also *stiff* order *two*.

5.3 *Conditions for higher order*

The following lemma provides the formulas to derive the order conditions for order larger than two in a systematic way.

LEMMA 5.1 Let $m\geqslant1$.

(a) For $\tau=\bullet$ we have
$$\partial_\sigma^m\Psi_\zeta(\bullet)(\sigma,w)=h^m\mathrm{e}^{-(\zeta-\sigma)hA}[F_A,g]_m\big(\mathrm{e}^{-\sigma hA}w\big),$$

where $[F_A,g]_{m+1}=\big[F_A,[F_A,g]_m\big]$ with $[F_A,g]_1=[F_A,g]$ denotes the $(m+1)$-fold commutator.

(b) For $\tau=[\tau_1,\ldots,\tau_k]$ we have

$$\partial_\sigma^m\Psi_\zeta(\tau)(\sigma,\sigma_1,\ldots,\sigma_k,w)=h\mathrm{e}^{-(\zeta-\sigma)hA}[F_A,g]_m^{(k)}\big(\mathrm{e}^{-\sigma hA}w\big)\big(\Psi_\sigma(\tau_1)(\sigma_1,w),\ldots,\Psi_\sigma(\tau_k)(\sigma_k,w)\big)$$

for $\sigma\in\mathbb{R}$ and $\sigma_j\in\mathbb{R}^{\rho(\tau_j)}$, $j=1,\ldots,k$.

Proof. Both parts are proved by induction on *m*.

(a) For $m = 1$ the statement was proved in (5.2). The induction step is proved by the same arguments as were used for $m = 2$ above.

(b) To prove the statement for $m = 1$, we first note that for $\tau = [\tau_1, \ldots, \tau_k]$ the integrand of $G_\zeta(\tau)$ is given recursively as

$$\Psi_\zeta(\tau)(\sigma, \sigma_1, \ldots, \sigma_k, w) = e^{-(\zeta-\sigma)hA} g_\sigma^{(k)}(w) \big(\Psi_\sigma(\tau_1)(\sigma_1, w), \ldots, \Psi_\sigma(\tau_k)(\sigma_k, w)\big).$$

Since $\partial_\eta \Psi_\eta(\tau) = -hA\Psi_\eta(\tau)$ for any tree τ, we obtain

$$\begin{aligned}
\partial_\sigma \Psi_\zeta(\tau)(\sigma, \sigma_1, \ldots, \sigma_k, w) = \ & he^{-(\zeta-\sigma)hA} \Big(Ag_\sigma^{(k)}(w)\big(\Psi_\sigma(\tau_1)(\sigma_1, w), \ldots, \Psi_\sigma(\tau_k)(\sigma_k, w)\big) \\
& - g_\sigma^{(k+1)}(w)\big(Ae^{-\sigma hA}w, \Psi_\sigma(\tau_1)(\sigma_1, w), \ldots, \Psi_\sigma(\tau_k)(\sigma_k, w)\big) \\
& - g_\sigma^{(k)}(w)\big(A\Psi_\sigma(\tau_1)(\sigma_1, w), \Psi_\sigma(\tau_2)(\sigma_2, w) \ldots, \Psi_\sigma(\tau_k)(\sigma_k, w)\big) \\
& - \ldots \\
& - g_\sigma^{(k)}(w)\big(\Psi_\sigma(\tau_1)(\sigma_1, w), \ldots, \Psi_\sigma(\tau_{k-1})(\sigma_{k-1}, w), A\Psi_\sigma(\tau_k)(\sigma_k, w)\big)\Big).
\end{aligned}$$

On the other hand, by definition (5.3), we have

$$[F_A, g]^{(k)}(w)(v_1, \ldots, v_k) = Ag^{(k)}(w)(v_1, \ldots, v_k) - \frac{d^k}{dw^k}\big(g'(w)Aw\big)(v_1, \ldots, v_k). \tag{5.8}$$

Using induction on *k* it is easy to see that

$$\begin{aligned}
\frac{d^k}{dw^k}\big(g'(w)Aw\big)(v_1, \ldots, v_k) = \ & g^{(k+1)}(w)(Aw, v_1, \ldots, v_k) \\
& + g^{(k)}(w)(Av_1, v_2, \ldots, v_k) + \ldots + g^{(k)}(w)(v_1, \ldots, Av_k).
\end{aligned} \tag{5.9}$$

This proves the claim for $m = 1$. If it holds for some $m \geqslant 1$ then it does also for $m+1$, since the same calculation can be done with $[F_A, g]^{(k)}$ in the role of $g^{(k)}$. $\qquad\square$

The lemma thus shows that all derivatives arising in the order conditions can be obtained recursively from the tree structure. Moreover, only commutators, iterated commutators and their derivatives appear.

5.4 *Specialisation to linear problems*

For the linear evolution equation

$$u' + Au = Bu, \qquad u(0) = u_0$$

with bounded operator *B* on *X*, the above conditions (5.4), (5.5), and (5.7) simplify a bit. Having $g(u) = Bu$, the Lie commutator coincides with the operator commutator of *A* and *B*

$$[F_A, g](w) = ABw - BAw = [A, B]w.$$

A first-order Lawson method is of *stiff* order one if

$$\sup_{0\leqslant\sigma\leqslant1} \sup_{0\leqslant t\leqslant T} \big\|e^{-(1-\sigma)hA}[A, B]e^{-\sigma hA}u(t)\big\| \leqslant C. \tag{5.10a}$$

For second order, the conditions read

$$\sup_{0\leqslant\sigma\leqslant1}\ \sup_{0\leqslant t\leqslant T}\left\|e^{-(1-\sigma)hA}[A,[A,B]]e^{-\sigma hA}u(t)\right\|\leqslant C,\tag{5.10b}$$

$$\sup_{0\leqslant\sigma_2\leqslant\sigma_1\leqslant1}\ \sup_{0\leqslant t\leqslant T}\left\|e^{-(1-\sigma_1)hA}[A,B]e^{-(\sigma_1-\sigma_2)hA}Be^{-\sigma_2hA}u(t)\right\|\leqslant C,\tag{5.10c}$$

$$\sup_{0\leqslant\sigma_2\leqslant\sigma_1\leqslant1}\ \sup_{0\leqslant t\leqslant T}\left\|e^{-(1-\sigma_1)hA}Be^{-(\sigma_1-\sigma_2)hA}[A,B]e^{-\sigma_2hA}u(t)\right\|\leqslant C.\tag{5.10d}$$

We recall that such conditions also arise in the analysis of splitting methods, cf. Jahnke & Lubich (2000).

Using Lemma 5.1, the above analysis can easily be generalized to higher order, since for linear problems, only long trees have to be considered. For all other trees, which have at least one node with two branches, the integrand Ψ vanishes.

5.5 *Nonlinear Schrödinger equations*

For the time discretization of nonlinear Schrödinger equations

$$u' = \mathrm{i}\big(\Delta u + f(|u|^2)u\big),\tag{5.11}$$

split-step methods are commonly viewed as the method of choice. In recent years, however, exponential integrators have been considered as a viable alternative for the solution of (5.11). For instance, Besse *et al.* (2017) studied exponential integrators in the context of Bose–Einstein condensates; Cano & González-Pachón (2015) and Balac *et al.* (2016) reported favorable results for Lawson integrators of the form as discussed in this paper. Rigorous convergence results, however, are still missing for these methods.

As an application of our analysis, we will use the above regularity conditions (5.4), (5.5), and (5.7) to verify second-order convergence of Lawson methods. We refrain from any particular space discretization and argue in an abstract Hilbert space framework. Note, however, that our reasoning carries over to spatial discretizations (by spectral methods, e.g.) without any difficulty.

For this purpose, we consider (5.11) with periodic boundary conditions on the d dimensional torus and smooth potential. Then it is well known (see, e.g., (Kato, 1995, Thm. 4.1)) that the problem is well posed in H^m for $m > d/2$. The regularity of an initial value $u_0 \in H^m$ is thus preserved along the solution. Henceforth we choose $m > d/2$.

Second-order Strang splitting for (5.11) with $f(u) = \pm u$ was rigorously analysed in Lubich (2008). There it was shown that commutator relations similar to our conditions (5.4), (5.5), and (5.7) play a crucial role in the convergence proof for Strang splitting. The analysis given here shows that Lawson methods converge under the same regularity assumptions as splitting schemes. This will be worked out now in detail for first and second-order methods.

Let $A = -\mathrm{i}\Delta$ and $g(u) = \mathrm{i}\beta|u|^2u$, $\beta \in \mathbb{R}$, i.e. $f = \beta I$. By

$$g(u+w) = \mathrm{i}\beta(u+w)^2(\overline{u+w}) = g(u) + \mathrm{i}\beta(u^2\overline{w} + 2u\overline{u}w) + \mathcal{O}(|w|^2),$$

the Fréchet derivative of g is given by

$$g'(u)w = \mathrm{i}\beta(u^2\overline{w} + 2|u|^2w).$$

The first commutator $[F_A, g]$ then takes the form

$$
\begin{aligned}
[F_A, g](u) &= Ag(u) - g'(u)Au \\
&= -\mathrm{i}\Delta g(u) + g'(u)(\mathrm{i}\Delta u) \\
&= \beta\nabla\cdot\nabla(u^2\overline{u}) + \mathrm{i}\beta\left(u^2\overline{(\mathrm{i}\Delta u)} + 2|u|^2\,\mathrm{i}\Delta u\right) \\
&= \beta\nabla\cdot\left(2u\overline{u}\nabla u + u^2\nabla\overline{u}\right) + \beta\left(u^2\Delta\overline{u} - 2|u|^2\,\Delta u\right) \\
&= \beta\left(2\overline{u}\nabla u\cdot\nabla u + 2u\nabla u\cdot\nabla\overline{u} + 2|u|^2\,\Delta u + 2u\nabla u\cdot\nabla\overline{u} + 2u^2\Delta\overline{u} - 2|u|^2\,\Delta u\right) \\
&= 2\beta\left(\overline{u}\nabla u\cdot\nabla u + 2u\nabla u\cdot\nabla\overline{u} + u^2\Delta\overline{u}\right). \tag{5.12}
\end{aligned}
$$

We next show that the commutator can be bounded in H^m if the solution is in H^{m+2} for $m \geqslant 0$.

LEMMA 5.2 Let $\Omega \subset \mathbb{R}^d$, $d \leqslant 3$, be a bounded Lipschitz domain. Then there exists a constant C which only depends on Ω and d such that

$$
\|[F_A, g]u\|_m \leqslant C\|u\|_{m+2}^3. \tag{5.13}
$$

Proof. Note that by the Sobolev embedding theorem we have the following bounds

$$
\begin{aligned}
\|uvw\|_0 &\leqslant C\|u\|_1\|v\|_1\|w\|_1, & \text{(5.14a)} \\
\|uvw\|_0 &\leqslant C\|u\|_0\|v\|_2\|w\|_2, & \text{(5.14b)} \\
\|uvw\|_1 &\leqslant C\|u\|_1\|v\|_2\|w\|_2, & \text{(5.14c)} \\
\|uvw\|_m &\leqslant C\|u\|_m\|v\|_m\|w\|_m, \quad m \geqslant 2, & \text{(5.14d)}
\end{aligned}
$$

cf. (Lubich, 2008, Section 8).

For $m = 0$, the bound (5.13) follows from using (5.14a) for the first two terms and (5.14b) for the last one in the explicit expression (5.12) of $[F_A, g]$. For $m = 1$ we apply (5.14c) to all terms and for $m \geqslant 2$ the bound follows from (5.14d). $\qquad\square$

For Lawson methods, a first-order convergence bound in H^m thus requires H^{m+2} regularity of the exact solution, which is the same regularity as required for the first-order Lie splitting.

For second-order methods, one has to estimate the double commutator $[F_A, [F_A, g]]$. A simple calculation shows that a bound in H^m requires H^{m+4} regularity of the exact solution. This situation is exactly the same as for second-order Strang splitting (see Lubich (2008)). Using (5.12) we conclude that the derivative of the commutator $[F_A, g]$ can be expressed as

$$
\begin{aligned}
[F_A, g]'(u)w = 2\beta\big(&\overline{w}\nabla u\cdot\nabla u + 2\overline{u}\nabla u\cdot\nabla w + 2w\nabla u\cdot\nabla\overline{u} + 2u\nabla w\cdot\nabla\overline{u} \\
&+ 2u\nabla u\cdot\nabla\overline{w} + u^2\Delta\overline{w} + 2uw\Delta\overline{u}\big).
\end{aligned}
$$

This commutator can again be bounded in H^m for $u, w \in H^{m+2}$. We thus conclude that Lawson methods require the same regularity for second-order convergence as Strang splitting.

5.6 *Numerical examples*

Lawson methods exhibit a strong order reduction, in general. For particular problems, however, they show full order of convergence (see Kassam & Trefethen (2005), Besse *et al.* (2017), Cano & González-Pachón (2015), Balac *et al.* (2016), and Montanelli & Bootland (2016)). Most of the problems considered in these papers result from space discretizations of partial differential equations posed with periodic boundary conditions.

After space discretization (by finite differences, finite elements, or spectral methods) the evolution equation (3.1) becomes an ordinary differential equation

$$\mathbf{u}'(t) + A_N \mathbf{u}(t) = g_N\big(\mathbf{u}(t)\big), \qquad \mathbf{u}(0) = \mathbf{u}_0. \tag{5.15}$$

with a matrix $A_N \in \mathbb{C}^{N \times N}$ and a discretization $g_N : \mathbb{C}^N \to \mathbb{C}^N$ of g, where N denotes the employed degrees of freedom. In order to satisfy Assumption 3.1 the space discretization is required to provide matrices A_N such that

$$\big\| \mathrm{e}^{-tA_N} \big\| \leqslant C_{\mathscr{F}} \tag{5.16}$$

holds with a constant $C_{\mathscr{F}}$ being uniform in N and $t \in \mathbb{R}$.

In the previous sections we showed that full order of convergence is only guaranteed if certain regularity conditions are satisfied. The aim of the following numerical examples is to show that order reduction can also be verified numerically, if some of these regularity assumptions are violated. In fact, such order reductions can even be observed for linear problems. Hence we resign from presenting numerical examples for semilinear problems here. Numerous such examples can be found in the literature mentioned above. We also restrict ourselves to the first order schemes covered by our analysis, the exponential Euler and the Lawson Euler method, since they already show interesting (and different) convergence behavior.

We consider the linear Schrödinger equation

$$u_t = iu_{xx} + if(x)u, \qquad x \in [-\pi, \pi], \quad u(0, \cdot) = u_0, \tag{5.17}$$

with periodic boundary conditions and discretize it using a Fourier spectral method on an equidistant grid. Let N be even and denote by $\mathscr{F}_N$ the discrete Fourier matrix. Then matrix A_N is given as

$$A_N = i\mathscr{F}_N^{-1} D_N^2 \mathscr{F}_N, \qquad \text{where} \qquad D_N = \mathrm{diag}\big(-\tfrac{N}{2} + 1, -\tfrac{N}{2} + 2, \ldots, \tfrac{N}{2}\big),$$

and

$$g_N(\mathbf{u}) = B_N \mathbf{u}, \qquad B_N = i\,\mathrm{diag}\big(f(x_{-N/2+1}), \ldots, f(x_{N/2})\big), \qquad x_m = m\tfrac{2\pi}{N}.$$

With this notation, the exact solution of (5.15) is given by

$$\mathbf{u}(t) = \mathrm{e}^{t(-A_N + B_N)} \mathbf{u}_0. \tag{5.18}$$

EXAMPLE 5.3 The aim of the first example is to explain that the concept of regularity is relevant even in the ODE context. In order to show what regularity means here, for each $N = 2^7, \ldots, 2^{12}$ we choose a a regularity parameter $\alpha \geqslant 0$ and a vector $\mathbf{r} = (r_m)_{m=-N/2+1}^{N/2} \in \mathbb{C}^N$ of Fourier coefficients whose entries contain random numbers uniformly distributed in the unit disc. Then we define an initial function as the trigonometric polynomial

$$\widetilde{u}_{0;N}(x) = \sum_{m=-N/2+1}^{N/2} v_m e^{imx}, \qquad v_m = \frac{r_m}{(1+m^2)^{\frac{1}{2}\left(\frac{1}{2}+\alpha+\varepsilon\right)}}, \qquad \varepsilon = 10^{-6}. \tag{5.19}$$

In the limit $N \to \infty$, this sequence of trigonometric polynomials converges to a function in the Sobolev space $H_{per}^\alpha = H_{per}^\alpha((-\pi, \pi))$ equipped with the norm

$$\|u\|_\alpha^2 = 2\pi \sum_{m \in \mathbb{Z}} (1+m^2)^\alpha |v_m|^2 \qquad \text{for} \qquad u = \sum_{m \in \mathbb{Z}} v_m e^{imx}.$$
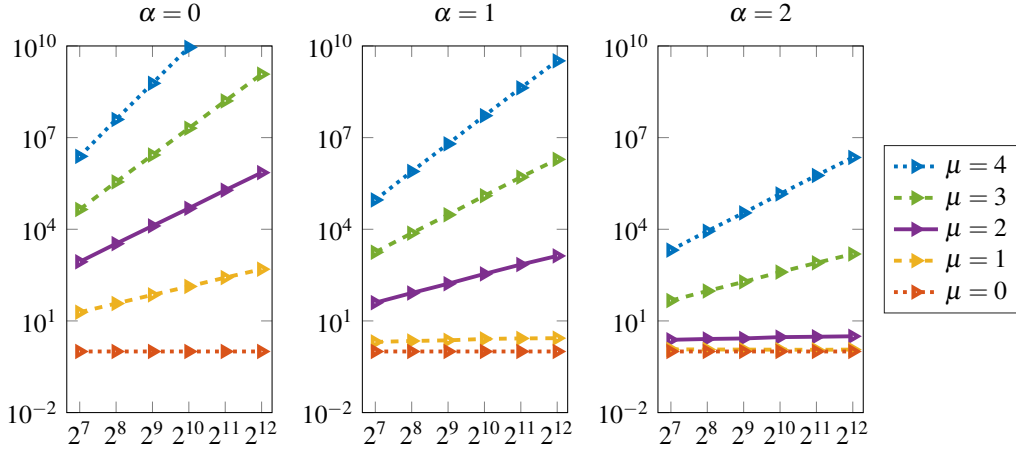
FIG. 1. Illustration of discrete regularity: the discrete H_{per}^{μ}-Sobolev seminorm $\|\mathbf{u}_0\|_{\mu,N}$ is plotted against the number N of Fourier modes, where $\mathbf{u}_0$ is chosen as in (5.20) (and thus corresponds to a function in $H_{\mathrm{per}}^{\alpha}$).

For $\alpha = 0$ we have the standard L^2 norm

$$\|u\|_0^2 = \int_{-\pi}^{\pi} |u(x)|^2 \, \mathrm{d}x.$$

Then we define an initial vector $\mathbf{u}_0 \in \mathbb{C}^N$ for (5.15) corresponding to a function $u_0 \in H_{\mathrm{per}}^{\alpha}$ by setting the jth component as

$$(\mathbf{u}_0)_j = u_{0;N}(x_j), \qquad j = -\tfrac{N}{2} + 1, \ldots, \tfrac{N}{2}, \tag{5.20}$$

where $u_{0;N} = \widetilde{u}_{0;N} / \|\widetilde{u}_{0;N}\|_0$ has unit L^2 norm. The discrete Sobolev norms in $\mathbb{C}^N$ corresponding to $\|\cdot\|_{\alpha}$ can be computed via

$$\|\mathbf{u}\|_{\alpha,N}^2 = 2\pi \left\| (I + D_N^2)^{\alpha/2} \mathscr{F}_N \mathbf{u} \right\|_{\mathbb{C}^N}^2,$$

where $\|\cdot\|_{\mathbb{C}^N}$ denotes the Euclidean norm in $\mathbb{C}^N$. This yields $\|\mathbf{u}_0\|_{\alpha,N} = \|u_{0;N}\|_{\alpha}$.

In Figure 1 we plot $\|\mathbf{u}_0\|_{\mu,N}$ for different values of μ over the number of Fourier modes N. The three graphs clearly show that $\|\mathbf{u}_0\|_{\mu,N}$ is bounded independently of the number N of Fourier modes only for $\mu \leqslant \alpha$. This corresponds to the continuous case, where obviously, the Sobolev norm $\|u\|_{\mu}$ is bounded for all functions $u \in H_{\mathrm{per}}^{\alpha}$ for $\mu \leqslant \alpha$.

The example clearly shows that regularity of the corresponding continuous function is crucial to obtain error bounds which do not deteriorate in the limit $N \to \infty$.

After these introductory explanations, we now fix the spatial discretization and set $N = 2048$. We consider (5.17) for two different functions f:

$$f(x) = \sin x, \tag{5.21a}$$

$$f(x) = (x/\pi)^2. \tag{5.21b}$$

EXAMPLE 5.4 In Figure 2 we show the numerically observed orders of the exponential Euler and the Lawson Euler method for the smooth, periodic potential (5.21a) for different values of α such that the
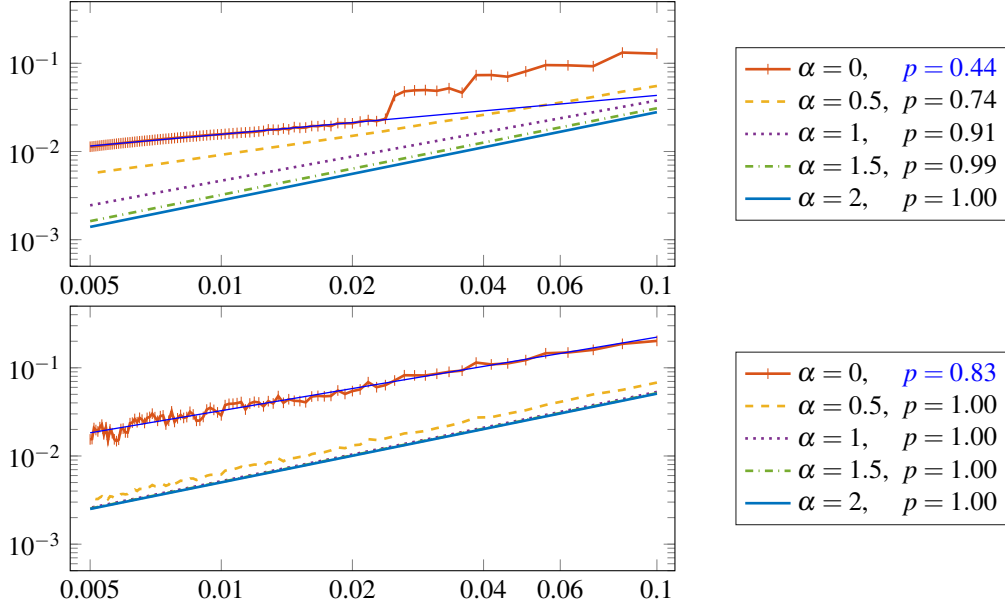
FIG. 2. Discrete $L^\infty((0,1), L^2(\Omega))$ error of the numerical solution of (5.17) with periodic potential (5.21a) for the exponential Euler method (top) and the Lawson Euler method (bottom) for starting values in H_{per}^α. The values of p in the legend show the numerically observed orders of the schemes.

corresponding initial function is contained in H_{per}^α. The leading error terms of the new analysis for the exponential Euler and the Lawson-Euler method are given in (4.2) and (5.10a), respectively. For comparison, we also added the leading error term (4.3) from our previous work.

Since $B : H_{\mathrm{per}}^\alpha \to H_{\mathrm{per}}^\alpha$ is a bounded perturbation of A, the exact solution of the continuous problem is guaranteed to stay in H_{per}^α for initial values in H_{per}^α for $\alpha \geqslant 0$. For the discrete problem, $\mathrm{e}^{-\sigma h A_N}$ and $\mathrm{e}^{\sigma h(-A_n + B_N)}$ are unitary matrices, which means that they leave all discrete Sobolev norms $\|\cdot\|_{\alpha,N}$ invariant. Thus the expression in (5.10a) can be bounded by

$$\left\| \mathrm{e}^{-(1-\sigma)h A_N}[A_N, B_N]\mathrm{e}^{-\sigma h A_N}\mathbf{u}(t) \right\|_{0,N} \leqslant c_1 \left\| \mathrm{e}^{-\sigma h A_N}\mathbf{u}(t) \right\|_{1,N} = c_1 \left\| \mathbf{u}(t) \right\|_{1,N} = c_1 \left\| \mathbf{u}_0 \right\|_{1,N}.$$

Here, the first inequality was proved in (Jahnke & Lubich, 2000, Lemma 3.1) with a constant c_1 independent of N and A_N.

Hence, the (sufficient but not necessary) order condition (5.10a) for the Lawson Euler method yields order one convergence for initial values bounded in $\|\cdot\|_\alpha$ for $\alpha \geqslant 1$. Numerically, we observe an order reduction for $\alpha = 0$ for the Lawson Euler method, while the exponential Euler method, which requires initial values in $H_{\mathrm{per}}^2 = D(A)$, cf. (4.2) or (4.3), shows order reduction for $\alpha \leqslant 1$. For $\alpha = 0$ the error of the exponential Euler method has an irregular behaviour for larger step sizes. To better visualise the order, we added thin lines (blue in the colored version) to all curves related to $\alpha = 0$. The slopes p of these lines are also given in the legends (blue in the colored version).
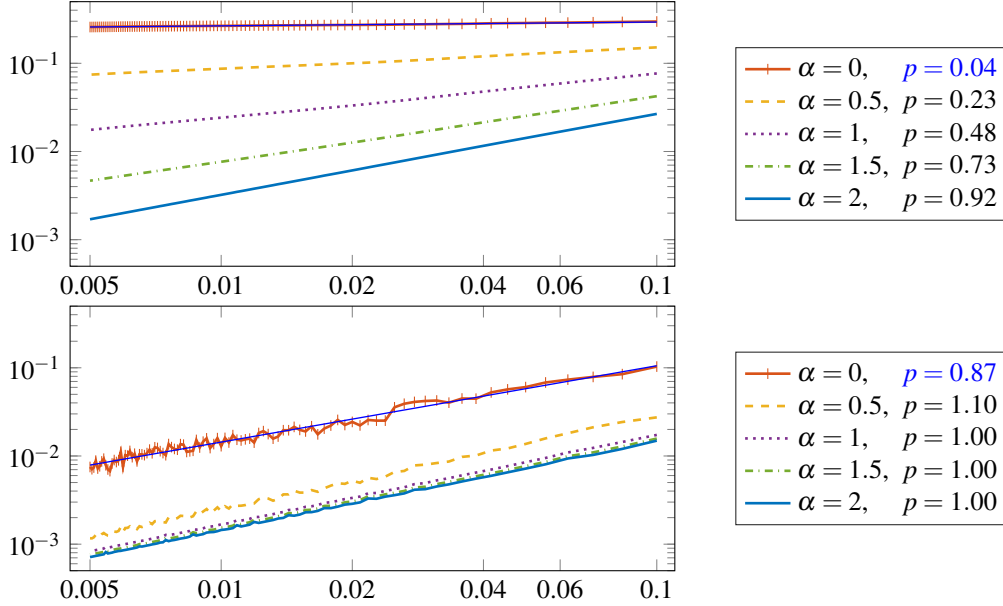
FIG. 3. Discrete $L^\infty((0,1),L^2(\Omega))$ error of the numerical solution of (5.17) with quadratic potential (5.21b) for the exponential Euler method (top) and the Lawson Euler method (bottom) for starting values in H_{per}^α. The values of p in the legend show the numerically observed orders of the schemes.

EXAMPLE 5.5 In Figure 3 we present the same experiment for the quadratic potential (5.21b). Here, the commutator bound of (Jahnke & Lubich, 2000, Lemma 3.1) does not apply, since it requires a C^5 smooth and periodic potential f. The situations differs considerably for the exponential Euler method which suffers from order reduction for all $\alpha \leqslant 2$ due to the nonsmooth potential f. In contrast, the Lawson Euler method still converges with order one for $\alpha \geqslant 0.5$.

Note that for these examples, the convergence behavior is slightly better than predicted by our theory. This is not a contradiction, because the order conditions are only sufficient but not necessary. To be more precise, our analysis contains a worst case estimation of the error propagation from the local to the global error by using Lady Windermere's fan in the proof of Theorem 4.7. Nevertheless, the examples clearly show the different behavior of the exponential Euler method and the Lawson-Euler method. Which of the two methods yields better results depends on the given problem, as reflected by our error analysis.

Acknowledgements

<div align="center">REFERENCES</div>

BALAC, S., FERNANDEZ, A., MAHÉ, F., MÉHATS, F. & TEXIER-PICARD, R. (2016) The interaction picture method for solving the generalized nonlinear Schrödinger equation in optics. *ESAIM: M2AN*, **50**, 945–964.

BESSE, C., DUJARDIN, G. & LACROIX-VIOLET, I. (2017) High order exponential integrators for nonlinear Schrödinger equations with application to rotating Bose–Einstein condensates. *SIAM J. Numer. Anal.*, **55**, 1387–1411.

CANO, B. & GONZÁLEZ-PACHÓN, A. (2015) Projected explicit Lawson methods for the integration of Schrödinger equation. *Numer. Methods Partial Differential Equations*, **31**, 78–104.

HAIRER, E., NØRSETT, S. P. & WANNER, G. (1993) *Solving ordinary differential equations I: Nonstiff problems.* Springer Series in Computational Mathematics, vol. 8, second edn. Springer-Verlag, Berlin.

HAIRER, E., LUBICH, C. & WANNER, G. (2006) *Geometric numerical integration: Structure-preserving algorithms for ordinary differential equations.* Springer Series in Computational Mathematics, vol. 31, second edn. Springer-Verlag, Berlin.

HANSEN, E. & OSTERMANN, A. (2016) High-order splitting schemes for semilinear evolution equations. *BIT*, **56**, 1303–1316.

HOCHBRUCK, M. & OSTERMANN, A. (2005a) Explicit exponential Runge–Kutta methods for semilinear parabolic problems. *SIAM J. Numer. Anal.*, **43**, 1069–1090.

HOCHBRUCK, M. & OSTERMANN, A. (2005b) Exponential Runge–Kutta methods for parabolic problems. *Appl. Numer. Math.*, **53**, 323–339.

HOCHBRUCK, M. & OSTERMANN, A. (2010) Exponential integrators. *Acta Numer.*, **19**, 209–286.

JAHNKE, T. & LUBICH, C. (2000) Error bounds for exponential operator splittings. *BIT*, **40**, 735–744.

KASSAM, A.-K. & TREFETHEN, L. N. (2005) Fourth-order time-stepping for stiff PDEs. *SIAM J. Sci. Comput.*, **26**, 1214–1233 (electronic).

KATO, T. (1995) On nonlinear Schrödinger equations. II. H^s-solutions and unconditional well-posedness. *J. Anal. Math.*, **67**, 281–306.

LAWSON, J. D. (1967) Generalized Runge–Kutta processes for stable systems with large Lipschitz constants. *SIAM J. Numer. Anal.*, **4**, 372–380.

LUAN, V. T. & OSTERMANN, A. (2013) Exponential B-series: the stiff case. *SIAM J. Numer. Anal.*, **51**, 3431–3445.

LUBICH, C. (2008) On splitting methods for Schrödinger-Poisson and cubic nonlinear Schrödinger equations. *Math. Comp.*, **77**, 2141–2153.

MONTANELLI, H. & BOOTLAND, N. (2016) Solving stiff PDEs in 1D, 2D and 3D with exponential integrators. *Preprint* arXiv:1604.08900.