Stiff convergence of force-gradient operator splitting methods

Emil Kieri*

Division of Scientific Computing, Department of Information Technology, Uppsala University, Box 337, 751 05 Uppsala, Sweden.

Abstract

We consider force-gradient, also called modified potential, operator splitting methods for problems with unbounded operators. We prove that force-gradient operator splitting schemes retain their classical orders of accuracy for linear time-dependent partial differential equations of parabolic and Schrödinger types, provided that the solution is sufficiently regular.

Keywords: exponential operator splitting, Schrödinger equation, convergence
2010 MSC: 65L04, 65L05, 65M12

1. Introduction

We consider the numerical time-integration of differential equations on the form

$$u'(t) = Au(t) + Bu(t), \quad u(0) = u_0, \tag{1}$$

evolving on a Banach space X, with the operators $A : D(A) \subset X \to X$, $B : D(B) \subset X \to X$. Splitting methods are advantageous if the subproblems with only one of the operators can be solved easily. An example of such a problem, where splitting methods often are used, is the time-dependent Schrödinger equation,

$$\mathrm{i}u_t(x,t) = -\Delta u(x,t) + V(x)u(x,t), \quad u(x,0) = u_0(x) \in L^2(\mathbb{R}^d).$$

If one of the terms is neglected we get the subproblems

$$\begin{aligned} \mathrm{i}v_t(x,t) &= -\Delta v(x,t), \quad v(x,0) = v_0(x), && \Rightarrow \quad v(x,t) = \mathrm{e}^{\mathrm{i}t\Delta}v_0(x), \quad \text{and} \\ \mathrm{i}w_t(x,t) &= V(x)w(x,t), \quad w(x,0) = w_0(x), && \Rightarrow \quad w(x,t) = \mathrm{e}^{-\mathrm{i}tV(x)}w_0(x). \end{aligned}$$

The operator exponentials are readily computed since the potential V and the Laplacian Δ are diagonal in coordinate and frequency space, respectively. The combined operator

*Corresponding author. Tel.: +46 18 471 2978
Email address: emil.kieri@it.uu.se (Emil Kieri)

$(-\Delta+V(x))$, however, is not as easily exponentiated. Splitting methods consist in taking small time steps, alternatingly applying the action of the two operators.

An s-stage standard splitting method for (1) can be written as

$$u_{n+1} = Su_n, \quad S = \prod_{j=1}^{s} \mathrm{e}^{b_j hB} \mathrm{e}^{a_j hA}, \tag{2}$$

where h is the time step and the coefficients a_j, b_j define the splitting scheme. Throughout this paper products are defined in falling order, i.e.,

$$\prod_{j=1}^{s} A_j = A_s A_{s-1} \dots A_1,$$

and $\prod_{j=k}^{l} a_j = 1$ if $k > l$. We let C denote a generic constant which may take different values at different occurrences.

We say that a splitting method is of classical order p if the local error estimate

$$\|(S - \mathrm{e}^{h(A+B)})u\| \le Ch^{p+1}$$

holds for bounded operators A and B, and sufficiently small time steps h. Such estimates are obtained using Taylor expansion or the Baker–Campbell–Hausdorff formula [12, Ch. III.5]. They are adequate, for instance, for ordinary differential equations, with $u \in \mathbb{R}^N$, and $A, B \in \mathbb{R}^{N\times N}$.

In infinite-dimensional settings, classical error analysis is no longer rigorous since the possibly unbounded operators A and B appear in the remainder. Still, splitting methods often retain their classical order also for such problems. Error estimates which hold independently of the norms of the operators are called stiff estimates. In an early work, Jahnke and Lubich proved that the Strang splitting scheme keeps its second order accuracy for Schrödinger equations [16]. Thalhammer [28], and Hansen and Ostermann [14], proved that splitting methods keep their classical order for certain classes of problems with unbounded operators, given that the solution is sufficiently regular. See also [5, 15], which enables splitting methods of order higher than two for parabolic problems using the theory of analytic semigroups. Using the calculus of Lie derivatives, the theory was extended to non-linear problems in [17, 29].

If the problem has a particular structure, this can often be used to construct more efficient methods. One example is Runge–Kutta–Nyström methods [4, 12], which apply when commutators of the operators satisfy $[B,[B,[A,B]]] = 0$. This reduces the number of order conditions, and gives more freedom to the choice of method coefficients. The condition holds when A is a differential operator of at most second order, possibly with variable coefficients, and B is a multiplication operator. The Schrödinger equation, in particular, is of this type. However, systems of Schrödinger equations will not satisfy the condition, since matrices generally do not commute. In this paper we will study force-gradient methods, which were proposed by Koseleff and Suzuki [18, 25, 26]. They rely on the same commutator assumption as Runge–Kutta–Nyström methods, and in addition that $[B,[A,B]]$ can be conveniently evaluated. A general force-gradient method for the problem (1) reads

$$u_{n+1} = Su_n, \quad S = \prod_{j=1}^{s} \mathrm{e}^{b_j hB + c_j h^3 [B,[A,B]]} \mathrm{e}^{a_j hA}. \tag{3}$$

The perhaps most used force-gradient scheme is the fourth order method of Chin [6, 8, 9, 18], with $s = 3$. An attractive property of the Chin scheme is that all its coefficients are real and non-negative. The method can therefore be used both for parabolic and Schrödinger-type problems. The parabolic correspondence of a Schrödinger problem, $u_t(x,t) = \Delta u(x,t) - V(x)u(x,t)$, is of interest since it gives the ground state of the quantum system when solved to steady-state with proper renormalisation. This way of computing the ground state is widely used by chemists, who call it imaginary time propagation. Force-gradient methods of order higher than four have been derived, e.g., by Omelyan et al. [21]. These methods do however have some negative coefficients. This makes them unsuitable for parabolic problems, which cannot be solved backwards in time. Chin has proven that a 6th order force-gradient method necessarily has negative coefficients, unless higher order commutators are included in the scheme [7]. As for standard splitting methods, this order barrier can be circumvented if the coefficients are allowed to be complex-valued. In [2], Bader et al. derive several high order force-gradient schemes with complex coefficients with positive real parts, suitable for parabolic problems.

In this paper we derive stiff order conditions for force-gradient methods. That is, conditions under which the methods are of a certain order of accuracy also for unbounded operators. We then prove that the stiff order conditions are equivalent to the classical ones, as long as the exact solution is sufficiently regular. The result holds for parabolic and Schrödinger-type partial differential equations. Problems where force-gradient methods are applicable can be found in those classes, since $[V,[\Delta,V]] = -2|\nabla V|^2$.

The remainder of this paper is organised as follows. In Section 2 we specify the assumptions on the operators and the regularity requirements, and state local and global error estimates. The global estimate follows directly from the local estimate and the assumptions. The local error estimate, which is the main result of this paper, is proven in Section 3. In Section 4 we use the calculus of Lie derivatives to show how force-gradient methods can be used for non-autonomous problems. We conclude the paper with numerical experiments in Section 5, which corroborate the theoretical results.

2. Statement of the error estimate

In this section we list the assumptions we rely on and prove the global error estimate, Theorem 2. The global estimate relies on a local error estimate, Theorem 1, which we will prove in Section 3. The first assumption specifies the classes of problems under consideration. In particular, linear autonomous problems of parabolic ($A = \Delta$, $B = -V : \mathbb{R}^d \to \mathbb{R}$) and Schrödinger ($A = \mathrm{i}\Delta$, $B = -\mathrm{i}V$) type are included, with $X = L^2$, as long as the potential V is bounded from below. The assumption will however not hold in general in the L^p-spaces with $p \neq 2$. In those cases a more intricate stability analysis is necessary, see [22].

Assumption 1. *The problem* (1) *is assumed to be well-posed, and the commutators of* A *and* B *are assumed to satisfy* $[B,[B,[A,B]]] = 0$. *We further assume either of the following two conditions.*

1. *(Schrödinger-type problem) The operators* $(A + B)$, A, B *and* $[B,[A,B]]$ *are the infinitesimal generators of strongly continuous groups of operators* $\{\mathrm{e}^{tF}\}_{t\in\mathbb{R}}$. *Fur-*

thermore, there exists a constant $\omega \geq 0$ *such that*

$$\|\mathrm{e}^{tF}\|_{X\leftarrow X} \leq \mathrm{e}^{\omega|t|} \quad \forall\ t \in \mathbb{R},$$
$$F = (A+B),\ A,\ B,\ \text{and } [B,[A,B]].$$

2. *(Parabolic problem) The operators* $(A+B)$, A, B, *and* $[B,[A,B]]$ *are the infinitesimal generators of analytic semigroups* $\{\mathrm{e}^{tF}\}_{t\in\Sigma_\theta}$, *defined in the sectors* $\Sigma_\theta = \{t \in \mathbb{C} : -\theta \leq \arg t \leq \theta\}$ *for some* $0 < \theta \leq \pi/2$. *Furthermore, there exists a constant* $\omega \geq 0$ *such that*

$$\|\mathrm{e}^{tF}\|_{X\leftarrow X} \leq \mathrm{e}^{\omega|t|} \quad \forall\ t \in \Sigma_\theta,$$
$$F = (A+B),\ A,\ B,\ \text{and } [B,[A,B]].$$

Next, we assume that the solution and the operators A and B are sufficiently regular. This assumption was justified for Schrödinger equations with sufficiently regular, bounded potential in [28, Sec. 2]. The arguments transfer directly to parabolic problems on the form $u_t(x,t) = \Delta u(x,t) - V(x)u(x,t)$. We here use the notation $|\boldsymbol{\mu}| = |\mu_1| + \cdots + |\mu_k|$ for the modulus of a multi-index, and iterated commutators are denoted

$$\mathrm{ad}_A^{p+1}(B) = [A, \mathrm{ad}_A^p(B)], \quad \mathrm{ad}_A^0(B) = B.$$

Assumption 2. *For a force-gradient splitting method of classical order* p, *assume that*

$$\Big\| \prod_{j=1}^{k} \Big(\mathrm{ad}_A^{\mu_j}(\mathcal{B}_j)\mathrm{e}^{\zeta_j A} \Big) \Big\|_{X\leftarrow Y_p} \leq C, \quad \boldsymbol{\mu} \in \mathbb{N}^k,\ |\boldsymbol{\mu}| = p+1-k, \tag{4}$$

$$\mathcal{B}_j = B \text{ and } [B,[A,B]], \quad j = 1,\ldots,k, \tag{5}$$

holds for $k = 1,\ldots,p+1$, *with* $\zeta_j \in \mathbb{R}$ *(Schrödinger) or* $\zeta_j \in \Sigma_\theta$ *(parabolic),* $|\zeta_j| \leq C$, *and some suitably chosen normed space* $Y_p \subset X$. *Assume further that*

$$\|\mathrm{e}^{\zeta F}\|_{Y_p\leftarrow Y_p} \leq C, \quad F = (A+B),\ A,\ B,\ \text{and } [B,[A,B]], \tag{6}$$

and that the exact solution to (1) *resides in the space* Y_p.

For a problem with smooth, bounded potential we get the standard Sobolev spaces $Y_p = H^p$. If the potential is unbounded, the choice of proper spaces Y_p becomes more delicate [20].

Under these assumptions we can prove the following local and global error estimates.

Theorem 1. *Assume that Assumptions 1 and 2 hold, and that the coefficients* a_j, b_j, c_j *define a force-gradient operator splitting scheme of classical order* p. *Assume further that the coefficients reside in* $\mathbb{R}$ *if* (1) *is of Schrödinger type, or in* Σ_θ *if parabolic. Then, the scheme satisfies the local error estimate*

$$\|Su(t) - u(t+h)\|_X \leq Ch^{p+1},$$

also for unbounded operators A *and* B.

Proof. We will prove this result in Section 3. □

Theorem 2. *Assume the conditions of Theorem 1. Then, the force-gradient operator splitting scheme* (3) *satisfies the global error estimate*

$$\|S^n u_0 - u(nh)\|_X \le Ch^p, \quad 0 \le nh \le T.$$

Proof. By Assumption 1, $\|S\|_{X\leftarrow X} \le \mathrm{e}^{ch}$ for some c. Using this in connection with the local error estimate, the result follows through a standard Lady Windermere's fan argument [13, Ch. II.3]. See, e.g., [14, Thm. 2.3] for a proof of a much similar result. □

3. Proof of the local error estimate

This section is devoted to the proof of Theorem 1. The line of attack is as follows. First, we summarise the proof of a similar result [28], a local error estimate for standard splitting methods on the form (2). This introduces notation and techniques which we will use when we extend the proof to force-gradient methods. Second, we take the commutators $[B,[A,B]]$ into account, and manipulate the expressions so that they fit in the framework of the stiff order conditions for standard methods. Finally, we observe that the stiff order conditions hold if and only if the classical order conditions do.

Since regularity is assumed uniformly in time, we can without restriction consider the error in the first time step, $e_0 = (S - \mathrm{e}^{h(A+B)})u_0$. As a first step, we express the exact and approximate solutions in forms that facilitate comparison. Repeating calculations from [28], we express the exact solution after one time step as

$$u(h) = \mathrm{e}^{hA}u_0 + \sum_{k=1}^{p} I_k u_0 + R_{p+1}, \qquad I_k = \int_{\Delta_k} g_k(\boldsymbol{\tau})\,\mathrm{d}\boldsymbol{\tau}, \tag{7}$$

with $\boldsymbol{\tau} = (\tau_1, \dots, \tau_k)$, $\tau_0 = h$, and

$$f_k(\boldsymbol{\tau}) = \prod_{l=1}^{k} \mathrm{e}^{(\tau_{k-l} - \tau_{k-l+1})A} B, \qquad g_k(\boldsymbol{\tau}) = f_k(\boldsymbol{\tau})\mathrm{e}^{\tau_k A}.$$

The integrals are evaluated over the domains

$$\Delta_k = \{\boldsymbol{\tau} \in \mathbb{R}^k : 0 \le \tau_k \le \tau_{k-1} \le \cdots \le \tau_1 \le h\},$$

and the remainder term reads

$$R_{p+1} = \int_{\Delta_{p+1}} f_{p+1}(\boldsymbol{\tau}) u(\tau_{p+1})\,\mathrm{d}\boldsymbol{\tau}.$$

Throughout this section, we let $\mathcal{R}'_{p+1}$ and $\mathcal{R}_{p+1}$ denote arbitrary linear combinations of terms on the form

$$\mathcal{R}'_{p+1} \sim h^{p+1}\Big(\prod_{j=1}^{k} \mathrm{ad}_A^{\mu_j}(\mathcal{B}_j)\mathrm{e}^{\zeta_j A}\Big)\chi, \qquad \mathcal{R}_{p+1} = \mathcal{R}'_{p+1}u, \quad |\boldsymbol{\mu}| \le p-k+1,$$

with the linear operator χ bounded $Y_p \to Y_p$. By Assumption 2, this means that $\|\mathcal{R}_{p+1}\|_X \le Ch^{p+1}$. Note that $R_{p+1} = \mathcal{R}_{p+1}$. Now, if we define

$$A_j = a_j hA, \qquad B_j = b_j hB, \qquad C_j = c_j h^3[B,[A,B]],$$

the approximate solution can be expressed similarly,

$$S = \prod_{j=1}^{s} \mathrm{e}^{A_j} + \sum_{k=1}^{p} Q_k + r_{p+1}.$$

Here,

$$Q_k = \sum_{\boldsymbol{\lambda}\in\Lambda_k} \kappa_{\boldsymbol{\lambda}} G(\boldsymbol{\lambda}), \quad \Lambda_k = \{\boldsymbol{\lambda} \in \mathbb{N}^k : 1 \le \lambda_k \le \lambda_{k-1} \le \cdots \le \lambda_1 \le s\},$$

$$F(\boldsymbol{\lambda}) = \prod_{l=1}^{k} \prod_{j=\lambda_{k-l+1}+1}^{\lambda_{k-l}} \mathrm{e}^{A_j}(B_{\lambda_{k-l+1}} + C_{\lambda_{k-l+1}}), \qquad G(\boldsymbol{\lambda}) = F(\boldsymbol{\lambda}) \prod_{j=1}^{\lambda_k} \mathrm{e}^{A_j},$$

with $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_k) \in \mathbb{N}^k$ and $\lambda_0 = s$. The remainder term reads

$$r_{p+1} = \sum_{j=1}^{p+1} \sum_{\boldsymbol{\lambda}\in\Lambda_{p+1}} \kappa'_{j,\boldsymbol{\lambda}} F(\boldsymbol{\lambda}) \varphi_j(B_{\lambda_k} + C_{\lambda_k}) \mathrm{e}^{A_{\lambda_k}} \prod_{j=1}^{\lambda_k - 1} \mathrm{e}^{B_j + C_j} \mathrm{e}^{A_j}.$$

The functions $\varphi_j(z)$, which appear in the remainder, are defined recursively as

$$\varphi_j(z) = \frac{1}{j!} + z\varphi_{j+1}(z), \quad j \ge 0, \qquad \varphi_0(z) = \mathrm{e}^z, \tag{8}$$

and $\kappa_{\boldsymbol{\lambda}}$ and $\kappa'_{j,\boldsymbol{\lambda}}$ are well-behaved expansion coefficients. The expression

$$\varphi_j(B_{\lambda_k} + C_{\lambda_k}) \mathrm{e}^{A_{\lambda_k}} \prod_{j=1}^{\lambda_k - 1} \mathrm{e}^{B_j + C_j} \mathrm{e}^{A_j}$$

is a bounded operator $Y_p \to Y_p$ due to (6), and therefore $r_{p+1} = \mathcal{R}'_{p+1}$. The values of $\kappa'_{\boldsymbol{\lambda}}$ are of lesser importance, but $\kappa_{\boldsymbol{\lambda}}$ appears in the order conditions. If

$$\sigma_0 = 0, \qquad \lambda_{\sigma_0+\cdots+\sigma_{l-1}+1} = \cdots = \lambda_{\sigma_0+\cdots+\sigma_l}, \quad l = 1, \ldots, m,$$

and $\sigma_0 + \cdots + \sigma_m = k$, then

$$\kappa_{\boldsymbol{\lambda}} = \frac{1}{\sigma_0! \cdots \sigma_m!}.$$

The local error can now be expressed as

$$e_0 = \Big(\mathrm{e}^{hA} - \prod_{j=1}^{s} \mathrm{e}^{A_j}\Big) u_0 + \sum_{k=1}^{p} (I_k - Q_k) u_0 + \mathcal{R}_{p+1}. \tag{9}$$

In [28] it was shown how each of the terms in (9) could be bounded separately, yielding the desired result for a standard operator splitting scheme. For force-gradient methods this is not possible since some of the order conditions for standard splitting schemes are violated. Error cancellation between the terms will instead recover the desired order of accuracy. We start off ignoring this, and re-derive the order conditions for a standard method. After that, we investigate how corrections from the $[B,[A,B]]$-terms kick in.

In what is coming, we need some notation for multi-indices. For $\boldsymbol{\mu} \in \mathbb{N}^k$ and $\mathbf{x} \in \mathbb{C}^k$, define $\boldsymbol{\mu}! = \mu_1!\mu_2! \cdots \mu_k!$ and $\mathbf{x}^{\boldsymbol{\mu}} = x_1^{\mu_1} x_2^{\mu_2} \cdots x_k^{\mu_k}$. Furthermore, if $1 \le \mu_j \le l$ and $\mathbf{x} \in \mathbb{C}^l$, let $x_{\boldsymbol{\mu}} = (x_{\mu_1}, \dots, x_{\mu_k})$.

The first error term,

$$e_0^{(0)} = \Big(\mathrm{e}^{hA} - \prod_{j=1}^{s} \mathrm{e}^{a_j h A}\Big) u_0,$$

is zero as long as the consistency condition $\sum_{j=1}^{s} a_j = 1$ is fulfilled. Taylor expansion of the integrands in (7) gives

$$I_k = \sum_{|\boldsymbol{\mu}| \le p-k} \frac{1}{\boldsymbol{\mu}!} \int_{\Delta_k} \boldsymbol{\tau}^{\boldsymbol{\mu}} \,\mathrm{d}\boldsymbol{\tau}\, \partial_{\boldsymbol{\tau}}^{\boldsymbol{\mu}} g_k(\mathbf{0}) + \int_{\Delta_k} \varrho_{k,p-k+1}(\boldsymbol{\tau}) \,\mathrm{d}\boldsymbol{\tau}, \tag{10}$$

with the remainder

$$\varrho_{k,n+1} = \sum_{\substack{\boldsymbol{\nu} \in \mathbb{N}^k \\ |\boldsymbol{\nu}| = n+1}} \frac{n+1}{\boldsymbol{\nu}!} \int_0^1 (1-s)^n \partial_{\boldsymbol{\tau}}^{\boldsymbol{\nu}} g_k(s\boldsymbol{\tau}) \boldsymbol{\tau}^{\boldsymbol{\nu}} \,\mathrm{d}s,$$

and $\varrho_{k,n+1} = \mathcal{R}'_{n+1}$. The expansion coefficients in (10) can be evaluated as

$$\int_{\Delta_k} \boldsymbol{\tau}^{\boldsymbol{\mu}} \,\mathrm{d}\boldsymbol{\tau} = h^{k+|\boldsymbol{\mu}|} \prod_{l=1}^{k} \frac{1}{\mu_l + \cdots + \mu_k + k - l + 1},$$

and the partial derivatives of g_k are

$$\partial_{\boldsymbol{\tau}}^{\boldsymbol{\mu}} g_k(\boldsymbol{\tau}) = (-1)^{|\boldsymbol{\mu}|} \prod_{l=1}^{k} \Big(\mathrm{e}^{(\tau_{k-l} - \tau_{k-l+1})A} \mathrm{ad}_A^{\mu_{k-l+1}}(B)\Big) \mathrm{e}^{\tau_k A}. \tag{11}$$

Next, we let

$$Q_k^{(0)} = Q_k|_{c_j \equiv 0},$$

and Taylor expand that. Defining $\alpha_j = \sum_{i=1}^{j} a_i$, we get

$$\begin{aligned} Q_k^{(0)} &= h^k \sum_{\boldsymbol{\lambda} \in \Lambda_k} \kappa_{\boldsymbol{\lambda}} \Big(\prod_{l=1}^{k} b_{\lambda_l}\Big) g_k(\alpha_{\boldsymbol{\lambda}} h) \\ &= \sum_{|\boldsymbol{\mu}| \le p-k} h^{k+|\boldsymbol{\mu}|} \frac{1}{\boldsymbol{\mu}!} \sum_{\boldsymbol{\lambda} \in \Lambda_k} \kappa_{\boldsymbol{\lambda}} \Big(\prod_{l=1}^{k} b_{\lambda_l}\Big) \alpha_{\boldsymbol{\lambda}}^{\boldsymbol{\mu}} \partial_{\boldsymbol{\tau}}^{\boldsymbol{\mu}} g_k(\mathbf{0}) + \\ &\quad + h^k \sum_{\boldsymbol{\lambda} \in \Lambda_k} \kappa_{\boldsymbol{\lambda}} \Big(\prod_{l=1}^{k} b_{\lambda_l}\Big) \varrho_{k,p-k+1}(\alpha_{\boldsymbol{\lambda}} h). \end{aligned}$$

Defining

$$\Pi_{\boldsymbol{\mu}} = \frac{1}{\boldsymbol{\mu}!} \prod_{l=1}^{k} \frac{1}{\mu_l + \cdots + \mu_k + k - l + 1}, \qquad \Sigma_{\boldsymbol{\mu}}^{(0)} = \frac{1}{\boldsymbol{\mu}!} \sum_{\boldsymbol{\lambda} \in \Lambda_k} \kappa_{\boldsymbol{\lambda}} \Big(\prod_{l=1}^{k} b_{\lambda_l}\Big) \alpha_{\boldsymbol{\lambda}}^{\boldsymbol{\mu}},$$

the kth order error term reads

$$I_k - Q_k^{(0)} = \sum_{|\boldsymbol{\mu}|\le p-k} h^{k+|\boldsymbol{\mu}|}\Big(\Pi_{\boldsymbol{\mu}} - \Sigma_{\boldsymbol{\mu}}^{(0)}\Big)\partial_{\boldsymbol{\tau}}^{\boldsymbol{\mu}} g_k(\mathbf{0}) + \mathcal{R}'_{p+1}.$$

We thereby achieve pth order of accuracy if $\alpha_s = 1$ and

$$\sum_{\substack{\boldsymbol{\mu}\in\mathbb{N}^k\\ |\boldsymbol{\mu}|=\sigma}} \big(\Pi_{\boldsymbol{\mu}} - \Sigma_{\boldsymbol{\mu}}^{(0)}\big)\partial_{\boldsymbol{\tau}}^{\boldsymbol{\mu}} g_k(\mathbf{0}) = 0, \quad k = 1,\dots,p, \quad \sigma = 0,\dots,p-k. \tag{12}$$

These are the stiff order conditions for standard operator splitting methods derived in [28]. For force-gradient methods these conditions are typically violated, since the contributions from C_j are not yet taken into account. We will next extend the proof to address that. The key observation is that changing a B for $[B,[A,B]]$ in the expansion of $Q_k^{(0)}$ transforms g_k into two first order partial derivatives of g_{k+1}. Take, for example,

$$g_2(\alpha_{(2,1)}h) = \prod_{j=3}^{3} \mathrm{e}^{a_j hA} B \prod_{j=2}^{2} \mathrm{e}^{a_j hA} B \prod_{j=1}^{1} \mathrm{e}^{a_j hA},$$

and compare it to

$$\begin{aligned}
&\prod_{j=3}^{3} \mathrm{e}^{a_j hA} B \prod_{j=2}^{2} \mathrm{e}^{a_j hA} [B,[A,B]] \prod_{j=1}^{1} \mathrm{e}^{a_j hA} \\
&= \prod_{j=3}^{3} \mathrm{e}^{a_j hA} B \prod_{j=2}^{2} \mathrm{e}^{a_j hA} B[A,B] \prod_{j=1}^{1} \mathrm{e}^{a_j hA} + \\
&\quad - \prod_{j=3}^{3} \mathrm{e}^{a_j hA} B \prod_{j=2}^{2} \mathrm{e}^{a_j hA} [A,B] B \prod_{j=1}^{1} \mathrm{e}^{a_j hA} \\
&= -\partial_{\boldsymbol{\tau}}^{(0,0,1)} g_3(\alpha_{(2,1,1)}h) + \partial_{\boldsymbol{\tau}}^{(0,1,0)} g_3(\alpha_{(2,1,1)}h),
\end{aligned}$$

cf. (11). We introduce the operators $\mathcal{D}_l^{\pm}$ as

$$\mathcal{D}_l^{\pm} \partial_{\boldsymbol{\tau}}^{\boldsymbol{\mu}} g_k(\boldsymbol{\tau}) = \partial_{\boldsymbol{\tau}}^{\boldsymbol{\mu}_{\pm,l}} g_k(\boldsymbol{\tau}'),$$

with $\boldsymbol{\mu} \in \mathbb{N}^k$, $\boldsymbol{\tau} \in \mathbb{R}^k$, and

$$\begin{aligned}
\boldsymbol{\mu}_{+,l} &= (\mu_1,\dots,\mu_{l-1},\mu_l,\mu_l+1,\mu_{l+1},\dots,\mu_k) \in \mathbb{N}^{k+1},\\
\boldsymbol{\mu}_{-,l} &= (\mu_1,\dots,\mu_{l-1},\mu_l+1,\mu_l,\mu_{l+1},\dots,\mu_k) \in \mathbb{N}^{k+1},\\
\boldsymbol{\tau}' &= (\tau_1,\dots,\tau_{l-1},\tau_l,\tau_l,\tau_{l+1},\dots,\tau_k) \in \mathbb{R}^{k+1}.
\end{aligned}$$

With this notation we can write

$$Q_k = h^k \sum_{\boldsymbol{\lambda}\in\Lambda_k} \kappa_{\boldsymbol{\lambda}} \prod_{l=1}^{k} \Big(b_{\lambda_{k-l+1}} - h^2 c_{\lambda_{k-l+1}} \mathcal{D}_{k-l+1}^{+} + h^2 c_{\lambda_{k-l+1}} \mathcal{D}_{k-l+1}^{-}\Big) g_k(\alpha_{\boldsymbol{\lambda}} h). \tag{13}$$

We expand the products in (13), and order the terms with respect to the number of $\mathcal{D}$-operators in them. We let

$$Q_k = \sum_{m=0}^{k} Q_k^{(m)},$$

with $Q_k^{(0)}$ as before and

$$Q_k^{(m)} = h^{k+2m} \sum_{\boldsymbol{\lambda}\in\Lambda_k} \kappa_{\boldsymbol{\lambda}} \sum_{l_1=m}^{k} \cdots \sum_{l_m=1}^{l_{m-1}-1} \Big(\prod_{\substack{l=1\\ l\neq l_j}}^{k} b_{\lambda_{k-l+1}} \Big) \times$$
$$\times \Big(\prod_{j=1}^{m} c_{\lambda_{k-l_j+1}} \big(\mathcal{D}^-_{k-l_j+1} - \mathcal{D}^+_{k-l_j+1} \big) \Big) g_k(\alpha_{\boldsymbol{\lambda}} h), \quad m = 1, \dots k.$$

The point with this notation is that the error now can be expressed as

$$e_0 = \Big(\mathrm{e}^{hA} - \prod_{j=1}^{s} \mathrm{e}^{A_j} \Big) u_0 + \sum_{k=1}^{p} \Big(I_k - \sum_{m=0}^{\lfloor \frac{k}{2} \rfloor} Q_{k-m}^{(m)} \Big) u_0 + \mathcal{R}_{p+1}.$$

The disregarded terms

$$\sum_{k=0}^{p} \sum_{m=p-k+1}^{k} Q_k^{(m)} u_0$$

are of the form $\mathcal{R}_q$, with $q \geq p+1$, and can thereby be controlled. With this formulation, we require no cross-cancellation between the terms

$$\Big(I_k - \sum_{m=0}^{\lfloor \frac{k}{2} \rfloor} Q_{k-m}^{(m)} \Big).$$

In order to match terms with the same derivative of g_k we Taylor expand $Q_{k-m}^{(m)}$. For illustration, we first expand $Q_{k-1}^{(1)}$.

$$\begin{aligned} Q_{k-1}^{(1)} &= h^{(k-1)+2} \sum_{\boldsymbol{\lambda}\in\Lambda_{k-1}} \kappa_{\boldsymbol{\lambda}} \sum_{l_1=1}^{k-1} \Big(\prod_{\substack{l=1\\ l\neq l_1}}^{k-1} b_{\lambda_{(k-1)-l+1}} \Big) c_{\lambda_{(k-1)-l_1+1}} \times \\ &\quad \times \big(\mathcal{D}^-_{(k-1)-l_1+1} - \mathcal{D}^+_{(k-1)-l_1+1} \big) g_{k-1}(\alpha_{\boldsymbol{\lambda}} h) \\ &= h^{k+1} \sum_{\boldsymbol{\lambda}\in\Lambda_{k-1}} \kappa_{\boldsymbol{\lambda}} \sum_{l_1=1}^{k-1} \Big(\prod_{\substack{l=1\\ l\neq l_1}}^{k-1} b_{\lambda_{k-l}} \Big) c_{\lambda_{k-l_1}} \times \\ &\quad \times \big(\partial_{\boldsymbol{\tau}}^{\mathbf{e}_{k-l_1}} - \partial_{\boldsymbol{\tau}}^{\mathbf{e}_{k-l_1+1}} \big) g_k(\alpha_{\boldsymbol{\lambda}^{(l)}} h) \\ &= \sum_{|\boldsymbol{\mu}|\leq p-(k+1)} h^{k+1+|\boldsymbol{\mu}|} \frac{1}{\boldsymbol{\mu}!} \sum_{\boldsymbol{\lambda}\in\Lambda_{k-1}} \kappa_{\boldsymbol{\lambda}} \sum_{l_1=1}^{k-1} \Big(\prod_{\substack{l=1\\ l\neq l_1}}^{k-1} b_{\lambda_{k-l}} \Big) c_{\lambda_{k-l_1}} \times \\ &\quad \times \alpha_{\boldsymbol{\lambda}^{(l)}}^{\boldsymbol{\mu}} \big(\partial_{\boldsymbol{\tau}}^{\boldsymbol{\mu}+\mathbf{e}_{k-l_1}} - \partial_{\boldsymbol{\tau}}^{\boldsymbol{\mu}+\mathbf{e}_{k-l_1+1}} \big) g_k(\mathbf{0}) + \mathcal{R}'_{p+1}, \end{aligned}$$

with $\boldsymbol{\lambda}^{(l)} = (\lambda_1, \dots, \lambda_{l-1}, \lambda_l, \lambda_l, \lambda_{l+1}, \dots, \lambda_{k-1}) \in \mathbb{N}^k$. We rewrite this as

$$
\begin{aligned}
Q_{k-1}^{(1)} &= \sum_{1 \le |\boldsymbol{\mu}| \le p-k} h^{k+|\boldsymbol{\mu}|} \Sigma_{\boldsymbol{\mu}}^{(1)} \partial_{\boldsymbol{\tau}}^{\boldsymbol{\mu}} g_k(\mathbf{0}) + \mathcal{R}'_{p+1}, \quad \text{with} \\
\Sigma_{\boldsymbol{\mu}}^{(1)} &= \sum_{\nu=1}^{k-1} \frac{1}{(\boldsymbol{\mu} - \mathbf{e}_\nu)!} \sum_{\boldsymbol{\lambda} \in \Lambda_{k-1}} \kappa_{\boldsymbol{\lambda}} \Big(\prod_{\substack{l=1 \\ l \neq k-\nu}}^{k-1} b_{\lambda_{k-l}} \Big) c_{\lambda_\nu} \alpha_{\boldsymbol{\lambda}^{(l)}}^{\boldsymbol{\mu} - \mathbf{e}_\nu} + \\
&\quad - \sum_{\nu=2}^{k} \frac{1}{(\boldsymbol{\mu} - \mathbf{e}_\nu)!} \sum_{\boldsymbol{\lambda} \in \Lambda_{k-1}} \kappa_{\boldsymbol{\lambda}} \Big(\prod_{\substack{l=1 \\ l \neq k-\nu+1}}^{k-1} b_{\lambda_{k-l}} \Big) c_{\lambda_{\nu-1}} \alpha_{\boldsymbol{\lambda}^{(l)}}^{\boldsymbol{\mu} - \mathbf{e}_\nu}.
\end{aligned}
$$

Terms with $\boldsymbol{\mu} - \mathbf{e}_\nu \notin \mathbb{N}^k$ are excluded from the summation. Similarly in the general case,

$$
\begin{aligned}
Q_{k-m}^{(m)} &= h^{(k-m)+2m} \sum_{\boldsymbol{\lambda} \in \Lambda_{k-m}} \kappa_{\boldsymbol{\lambda}} \sum_{l_1=m}^{k-m} \cdots \sum_{l_m=1}^{l_{m-1}-1} \Big(\prod_{\substack{l=1 \\ l \neq l_j}}^{k-m} b_{\lambda_{(k-m)-l+1}} \Big) \times \\
&\quad \times \Big(\prod_{j=1}^{m} c_{\lambda_{(k-m)-l_j+1}} \big(\mathcal{D}^-_{(k-m)-l_j+1} - \mathcal{D}^+_{(k-m)-l_j+1} \big) \Big) g_{k-m}(\alpha_{\boldsymbol{\lambda}} h) \\
&= h^{k+m} \sum_{\boldsymbol{\lambda} \in \Lambda_{k-m}} \kappa_{\boldsymbol{\lambda}} \sum_{l_1=m}^{k-m} \cdots \sum_{l_m=1}^{l_{m-1}-1} \Big(\prod_{\substack{l=1 \\ l \neq l_j}}^{k-m} b_{\lambda_{(k-m)-l+1}} \Big) \times \\
&\quad \times \Big(\prod_{j=1}^{m} c_{\lambda_{(k-m)-l_j+1}} \big(\partial_{\boldsymbol{\tau}}^{\mathbf{e}_{(k-m)-l_j+j}} - \partial_{\boldsymbol{\tau}}^{\mathbf{e}_{(k-m)-l_j+j+1}} \big) \Big) g_k(\alpha_{\boldsymbol{\lambda}^{(1)}} h) \\
&= \sum_{|\boldsymbol{\mu}| \le p-(k+m)} h^{k+m+|\boldsymbol{\mu}|} \frac{1}{\boldsymbol{\mu}!} \sum_{\boldsymbol{\lambda} \in \Lambda_{k-m}} \kappa_{\boldsymbol{\lambda}} \sum_{l_1=m}^{k-m} \cdots \sum_{l_m=1}^{l_{m-1}-1} \Big(\prod_{\substack{l=1 \\ l \neq l_j}}^{k-m} b_{\lambda_{(k-m)-l+1}} \Big) \times \\
&\quad \times \Big(\prod_{j=1}^{m} c_{\lambda_{(k-m)-l_j+1}} \big(\partial_{\boldsymbol{\tau}}^{\mathbf{e}_{(k-m)-l_j+j}} - \partial_{\boldsymbol{\tau}}^{\mathbf{e}_{(k-m)-l_j+j+1}} \big) \Big) \alpha_{\boldsymbol{\lambda}^{(1)}}^{\boldsymbol{\mu}} \partial_{\boldsymbol{\tau}}^{\boldsymbol{\mu}} g_k(\mathbf{0}) + \\
&\quad + \mathcal{R}'_{p+1}.
\end{aligned}
$$

$\boldsymbol{\lambda}^{(\mathbf{l})}$ is defined analogously to $\boldsymbol{\lambda}^{(l)}$, with all the elements λ_{l_j}, $j = 1, \dots, m$, doubled. In order to turn this expression inside out we need to introduce some notation. Let $\mathbb{B} = \{0, 1\}$, and define the function $\boldsymbol{\nu} : \mathbb{N}^m \times \mathbb{B}^m \to \mathbb{N}^m$ as

$$
\nu_j(\mathbf{l}, \mathbf{v}) = k - m - l_j + j + v_j, \quad j = 1, \dots, m, \quad \mathbf{l} = (l_1, \dots, l_m).
$$

Then the product of binomials above can be expanded as

$$
\prod_{j=1}^{m} \big(\partial_{\boldsymbol{\tau}}^{\mathbf{e}_{(k-m)-l_j+j}} - \partial_{\boldsymbol{\tau}}^{\mathbf{e}_{(k-m)-l_j+j+1}} \big) = \sum_{\mathbf{v} \in \mathbb{B}^m} (-1)^{|\mathbf{v}|} \partial_{\boldsymbol{\tau}}^{\boldsymbol{\nu}(\mathbf{l}, \mathbf{v})}.
$$

We can now rewrite the expression for $Q_{k-m}^{(m)}$ as

$$Q_{k-m}^{(m)} = \sum_{\substack{\boldsymbol{\mu}\in\mathbb{N}^k \\ m\le|\boldsymbol{\mu}|\le p-k}} h^{k+|\boldsymbol{\mu}|}\Sigma_{\boldsymbol{\mu}}^{(m)}\partial_{\boldsymbol{\tau}}^{\boldsymbol{\mu}} g_k(\mathbf{0}) + \mathcal{R}'_{p+1}, \quad \text{with}$$

$$\Sigma_{\boldsymbol{\mu}}^{(m)} = \sum_{l_1=m}^{k-m}\cdots\sum_{l_m=1}^{l_{m-1}-1}\sum_{\mathbf{v}\in\mathbb{B}^m}\frac{(-1)^{|\mathbf{v}|}}{(\boldsymbol{\mu}-\boldsymbol{\nu}(\mathbf{l},\mathbf{v}))!}\sum_{\boldsymbol{\lambda}\in\Lambda_{k-m}}\kappa_{\boldsymbol{\lambda}}\times$$
$$\times\Big(\prod_{\substack{l=1\\ l\neq l_j}}^{k-m} b_{\lambda_{(k-m)-l+1}}\Big)\Big(\prod_{j=1}^{m} c_{\lambda_{(k-m)-l_j+1}}\Big)\alpha_{\boldsymbol{\lambda}^{(1)}}^{\boldsymbol{\mu}-\boldsymbol{\nu}(\mathbf{l},\mathbf{v})}.$$

As before, terms with $\boldsymbol{\mu}-\boldsymbol{\nu}(\mathbf{l},\mathbf{v})\notin\mathbb{N}^k$ are omitted from the summation. Assuming the consistency condition $\alpha_s = 1$, we can write the local error as

$$e_0 = \sum_{k=1}^{p}\sum_{\substack{\boldsymbol{\mu}\in\mathbb{N}^k\\ |\boldsymbol{\mu}|\le p-k}} h^{k+|\boldsymbol{\mu}|}\Big(\Pi_{\boldsymbol{\mu}} - \sum_{m=0}^{\lfloor\frac{k}{2}\rfloor}\Sigma_{\boldsymbol{\mu}}^{(m)}\Big)\partial_{\boldsymbol{\tau}}^{\boldsymbol{\mu}} g_k(\mathbf{0})u_0 + \mathcal{R}_{p+1}. \tag{14}$$

Noting that

$$\partial_{\boldsymbol{\tau}}^{\boldsymbol{\mu}} g_k(\mathbf{0}) = (-1)^{|\boldsymbol{\mu}|}\mathrm{e}^{hA}\prod_{l=1}^{k}\mathrm{ad}_A^{\mu_{k-l+1}}(B),$$

we can write the stiff order conditions for force-gradient methods as[1]

$$\sum_{\substack{\boldsymbol{\mu}\in\mathbb{N}^k\\ |\boldsymbol{\mu}|=\sigma}} E_{\boldsymbol{\mu}}\prod_{l=1}^{k}\mathrm{ad}_A^{\mu_{k-l+1}}(B) = 0, \qquad E_{\boldsymbol{\mu}} = \Pi_{\boldsymbol{\mu}} - \sum_{m=0}^{\lfloor\frac{k}{2}\rfloor}\Sigma_{\boldsymbol{\mu}}^{(m)}, \tag{15}$$

for all $k = 1,\dots,p,\ \sigma = 1,\dots,p-k$. Note the similarity with the order conditions (12) for standard splitting methods. Note also that terms in (15) can be matched using the properties of commutators, in particular the Jacobi identity and the assumed relation $[B,[B,[A,B]]] = 0$ of Runge–Kutta–Nyström methods.

Finally, by repeating the arguments from [28], one can show that the conditions (15) are equivalent to the classical order conditions. We assume that A and B are bounded operators, e.g., square matrices of fixed size, and claim that (15) are sufficient and necessary conditions for pth order convergence. They are sufficient conditions for bounded operators, since they are for unbounded. By Taylor expanding e^{hA} in the partial derivatives of $g_k(\mathbf{0})$ in (14) and collecting terms of similar order in h, we see that the conditions are necessary also for non-stiff problems. □

4. Generalisation to non-autonomous problems

Using the calculus of Lie derivatives [12, Ch. III.5], the results presented in this paper can be generalised to non-linear problems. That is, the proofs of the theorems

[1] A Matlab script for evaluating the expressions $E_{\boldsymbol{\mu}}$ is available upon request.

are readily repeated, but computing the flows of the subproblems might not always be practical. The requirement that the Lie commutator $[B,[A,B]]$ should be easy to handle is also restrictive. In [1], Bader derived the Lie commutator for the Gross–Pitaevskii equation, and conducted numerical experiments with force-gradient methods. In this work, however, we stick to linear problems, and will only use the non-linear framework to generalise the results to non-autonomous linear problems.

Given two non-linear operators $A, B : D \subset X \to X$, the Lie derivative D_A of A and its exponential are defined by

$$D_A B u = B'(u)A(u), \qquad \mathrm{e}^{(t_2-t_1)D_A} B u = B(\Psi_A^{t_1,t_2} u).$$

$\Psi_A^{t_1,t_2}$ is here the exact flow of the differential equation $u'(t) = A(u(t))$ from $t = t_1$ to $t = t_2$, and $B'(u)$ denotes the Fréchet derivative of $B(u)$. Taking B as the identity mapping yields

$$D_A u = A(u), \qquad \mathrm{e}^{(t_2-t_1)D_A} u = \Psi_A^{t_1,t_2} u.$$

The Lie commutator $[D_A, D_B]$ will be an essential tool in the following. It is defined through

$$[D_A, D_B]u = D_A D_B u - D_B D_A u = B'(u)A(u) - A'(u)B(u).$$

To generalise Theorems 1 and 2 to non-linear problems on the form

$$u'(t) = A(u(t)) + B(u(t)), \quad u(0) = u_0,$$

we only need to assert that the problem satisfies a modified Assumption 2. In the modified assumption, (4)–(5) are replaced by

$$\Big\| \prod_{j=1}^{k} \Big(\mathrm{ad}_{D_A}^{\mu_j}(D_{\mathcal{B}_j}) \mathrm{e}^{\zeta_j D_A} \Big) \Big\|_{X \leftarrow Y_p} \leq C, \quad \boldsymbol{\mu} \in \mathbb{N}^k, \ |\boldsymbol{\mu}| = p + 1 - k,$$

$$D_{\mathcal{B}_j} = D_B \text{ or } [D_B, [D_A, D_B]], \quad j = 1, \ldots, k,$$

for $k = 1, \ldots, p+1$. Provided that this holds, the proofs can be carried out as previously, only with the order of the exponentials reversed.

We will now show that this assumption holds for non-autonomous linear problems

$$u'(t) = A(t)u(t) + B(t)u(t), \quad u(0) = u_0,$$

satisfying $[A(s), A(t)] = [B(s), B(t)] = 0$ for all s, t. These commutator conditions hold for Schrödinger problems with time-dependent potentials, arising when a molecule interacts with a time-dependent electric field. More generally, they hold for problems on the form

$$\mathrm{i}u_t(x,t) = -\nabla \cdot K(t) \nabla u(x,t) + V(x,t)u(x,t).$$

A restriction compared to the autonomous case is that the differential operator $\nabla \cdot K(t)\nabla$ may no longer have x-dependent coefficients. Equations on this form arise when the Gaussian wave packet transform [23] is applied. Naturally, the conditions also hold for the corresponding parabolic problems. We will furthermore show that the Lie commutator

$[D_B, [D_A, D_B]]$ has a simple form and can be used as before. We introduce time as an auxiliary variable and write the problem as

$$U'(t) = F(U(t)) + G(U(t)), \quad \text{with} \quad U(t) = \begin{pmatrix} u(t) \\ s(t) \end{pmatrix},$$

and

$$F(U) = \begin{pmatrix} A(s)u \\ 1 \end{pmatrix}, \quad G(U) = \begin{pmatrix} B(s)u \\ 0 \end{pmatrix}. \tag{16}$$

The Fréchet derivatives of F and G are then

$$F'(U) = \begin{pmatrix} A(s) & A'(s)u \\ 0 & 0 \end{pmatrix}, \quad G'(U) = \begin{pmatrix} B(s) & B'(s)u \\ 0 & 0 \end{pmatrix},$$

and the iterated Lie commutators can be determined as

$$\mathrm{ad}^p_{D_F}(D_G)U = \begin{pmatrix} \widetilde{\mathrm{ad}}^p_{A(s)}(B(s))u \\ 0 \end{pmatrix},$$

with

$$\begin{aligned} \widetilde{\mathrm{ad}}^{p+1}_{A(s)}(B(s)) &= [\widetilde{\mathrm{ad}}^p_{A(s)}(B(s)), A(s)] + \frac{\mathrm{d}}{\mathrm{d}s}(\widetilde{\mathrm{ad}}^p_{A(s)}(B(s))), \\ \widetilde{\mathrm{ad}}^0_{A(s)}(B(s)) &= B(s). \end{aligned}$$

We can conclude that in order to fulfil the modified Assumption 2 we, in addition to the conditions from the autonomous case, must require the operators A and B to be temporally smooth. This is a reasonable requirement. The force-gradient correction similarly evaluates to

$$\begin{aligned} [D_G, [D_F, D_G]]U &= \begin{pmatrix} -[B(s), [B(s), A(s)]]u + \frac{\mathrm{d}B(s)}{\mathrm{d}s}B(s)u - B(s)\frac{\mathrm{d}B(s)}{\mathrm{d}s}u \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} [B(s), [A(s), B(s)]]u \\ 0 \end{pmatrix}. \end{aligned} \tag{17}$$

Since $B(s)$ commutes with itself evaluated at any other time, it also commutes with its time derivative. The commutator in (17) is the same commutator as in the autonomous problem, evaluated at a single point in time. Furthermore, since the operator A evaluated at different times commutes, its exponentiation is given by the first term of its Magnus series,

$$\Psi_A^{t_1,t_2} = \exp\Big(\int_{t_1}^{t_2} A(\sigma)\,\mathrm{d}\sigma\Big).$$

In a computational context, it is sufficient to approximate the integral in the exponent with a pth order quadrature rule. Using the recurrence relation (8) of the functions φ_j the quadrature remainder can be isolated from the scheme. As an example, consider Strang splitting with the integral of $A(t)$ approximated with the midpoint rule. Writing

the quadrature rule as the exact integral and a truncation error, we get

$$\begin{aligned}
&\mathrm{e}^{hB/2}\exp\Big(\int_{t_n}^{t_n+h} A(\sigma)\,\mathrm{d}\sigma + Ch^3A''(\xi)\Big)\mathrm{e}^{hB/2} = \\
&\qquad \mathrm{e}^{hB/2}\exp\Big(\int_{t_n}^{t_n+h} A(\sigma)\,\mathrm{d}\sigma\Big)\mathrm{e}^{hB/2} + \\
&\qquad + Ch^3\mathrm{e}^{hB/2}\exp\Big(\int_{t_n}^{t_n+h} A(\sigma)\,\mathrm{d}\sigma\Big)A''(\xi)\varphi_1\big(Ch^3A''(\xi)\big)\mathrm{e}^{hB/2},
\end{aligned}$$

with $\xi \in [t_n, t_n + h]$.

Note that it is possible to introduce time as an auxiliary variable in several ways [3]. One may even prefer to use more than one auxiliary variable. In connection with force-gradient methods, however, the arrangement (16) seems the most practical. The reason for this is that we want time to "stand still" while $G(U)$ is integrated, otherwise the commutator (17) will become complicated.

When using schemes with complex coefficients for non-autonomous problems, one may have to evaluate the operators at complex times. This can lead to poor conditioning. In order to circumvent this difficulty, Seydaoğlu and Blanes suggest methods with $a_j \in \mathbb{R}^+, b_j \in \Sigma_\theta$ [24]. No force-gradient method with this property (and $c_j \in \Sigma_\theta$) of order higher than four has to our knowledge been proposed, but could in principle be derived.

5. Numerical experiments

Preservation of the classical order for force-gradient schemes applied to partial differential equations have been observed previously. The Chin scheme, with $s = 3$ and

$$a_1 = 0, \quad a_2 = a_3 = \frac{1}{2}, \quad b_1 = b_3 = \frac{1}{6}, \quad b_2 = \frac{2}{3}, \quad c_1 = c_3 = 0, \quad c_2 = \frac{1}{72}, \tag{18}$$

was studied for the Schrödinger equation in [8, 9]. In [23], fourth order accuracy was observed for a non-autonomous Schrödinger-like problem in four dimensions. For completeness, we present convergence studies for the Chin scheme applied to the two-dimensional Schrödinger-type and parabolic problems

$$\begin{aligned}
\mathrm{i}u_t(\mathbf{x},t) &= -\frac{1}{2}\nabla\cdot K(t)\nabla u(\mathbf{x},t) + V(\mathbf{x},t)u(\mathbf{x},t), \\
u_t(\mathbf{x},t) &= \nabla\cdot K(t)\nabla u(\mathbf{x},t) - V(\mathbf{x},t)u(\mathbf{x},t),
\end{aligned}$$

on $\mathbf{x} \in [-\pi, \pi]^2$. We use periodic boundary conditions and the Gaussian initial condition

$$u(\mathbf{x}, 0) = \mathrm{e}^{-5\,\mathbf{x}^T\mathbf{x}}.$$

The matrix K and the potential V are given by

$$K(t) = \begin{pmatrix} \mathrm{e}^{t} & 0 \\ 0 & \mathrm{e}^{-t} \end{pmatrix}, \qquad V(\mathbf{x},t) = 2 - \cos(x_1) - \sin(t)\cos(x_2).$$

We discretise in space with the Fourier pseudospectral method [11, 19] using 256 grid points per dimension. We compare to a reference solution, which is computed with

Table 1: Verification of the convergence rate for the Chin scheme (18). The method is confirmed to be 4th order accurate in L^2-norm for non-autonomous Schrödinger-type and parabolic problems.

	Schrödinger		Parabolic	
Δt	error	rate	error	rate
2^{-1}	1.364e-2	-	7.745e-5	-
2^{-2}	7.855e-4	4.12	8.197e-6	3.24
2^{-3}	3.287e-5	4.58	6.269e-7	3.71
2^{-4}	1.853e-6	4.15	4.168e-8	3.91
2^{-5}	1.130e-7	4.03	2.649e-9	3.98
2^{-6}	7.022e-9	4.01	1.663e-10	3.99
2^{-7}	4.381e-10	4.00	1.040e-11	4.00
2^{-8}	2.728e-11	4.01	6.479e-13	4.00

$\Delta t = 2^{-10}$. The results, which confirm the 4th order convergence, are displayed in Table 1.

In Figure 1, we compare the efficiency of the Chin scheme to two other splitting methods: the Yoshida scheme [10, 30] and the optimised fourth order Runge–Kutta–Nyström scheme (RKN4) of Blanes and Moan [4]. We also compare with a fourth order commutator-free scheme [27], using the Arnoldi process to approximate the matrix exponentials. We solve the same non-autonomous Schrödinger problem as previously. In Figure 1 (a) we plot the error in L^2-norm against the time step. In Figure 1 (b) we plot the error against work, in terms of the number of fast Fourier transform (FFT) pairs computed. The Chin, Yoshida and RKN4 schemes use 2, 3 and 6 FFT pairs per time step, respectively. The commutator-free scheme does not fully conform to this kind of study, since the number of Arnoldi steps depends strongly on the problem and the method parameters. A line for the commutator-free scheme would appear on top of the Chin scheme line if each exponential was computed in two steps. That is however unlikely, most often more Arnoldi steps would be necessary. We can conclude that, at least for this particular problem, the Chin and RKN4 schemes are of comparable efficiency, beating the Yoshida and commutator-free schemes.

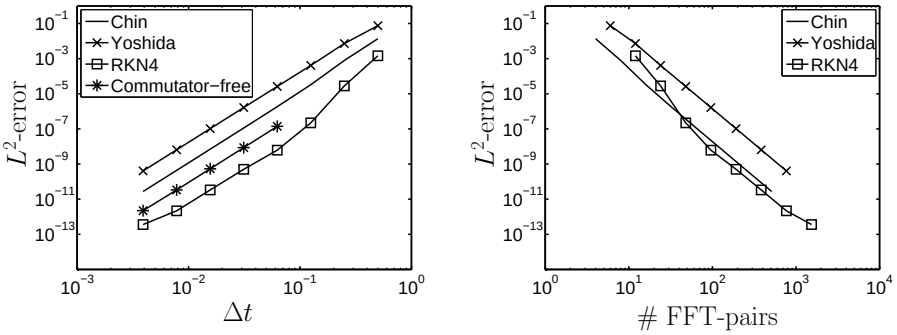
We emphasise the independence of the norms of the operators with another numerical study. We consider the harmonic oscillator in two dimensions,

$$\mathrm{i}u_t(\mathbf{x},t) = -\frac{1}{2}\Delta u(\mathbf{x},t) + \frac{1}{2}(x_1^2 + x_2^2)u(\mathbf{x},t), \tag{19}$$

and truncate the spatial domain to $[-3\pi, 3\pi]$. We start in the first eigenstate,

$$u(\mathbf{x},0) = \pi^{-1/2}\mathrm{e}^{-\frac{1}{2}\mathbf{x}^T\mathbf{x}},$$

and solve over one period, up to the time $t = 2\pi$, using 50 time steps. We use the Fourier pseudospectral method in space, and time step with the Chin scheme. The behaviour of the L^2-error, compared to the analytical solution, as the spatial step length $\Delta x_1 = \Delta x_2 = 6\pi/N$ is decreased is shown in Table 2. When the solution is well resolved and the temporal error dominates, the error stays constant as the spatial grid is refined further. The quadratic growth with N of the discrete Laplace operator does not affect the error.



(a) (b)

Figure 1: Error decay as the time step is refined for different fourth order schemes. In (a) the L^2-error is plotted against the time step. In (b) the same error is related to the computational work, measured in number of FFT pairs.

Table 2: Behaviour of the error in L^2 for the harmonic oscillator (19) when the spatial grid is refined. We use $\Delta x_1 = \Delta x_2 = 6\pi/N$, and the constant time step $\Delta t = \pi/25$. When the temporal error is dominant the error stagnates, the growing norm of the spatial operator does not influence the error.

N	error
8	2.481
16	1.261e-1
32	1.872e-5
64	3.630e-7
128	3.630e-7
256	3.630e-7
512	3.630e-7
1024	3.630e-7

Acknowledgements

Discussions with Vasile Gradinaru and Sverker Holmgren are gratefully acknowledged.

References

[1] P. Bader, Geometric Integrators for Schrödinger Equations, Ph.D. thesis, Universitat Politècnica de València, 2014.

[2] P. Bader, S. Blanes, F. Casas, Solving the Schrödinger eigenvalue problem by the imaginary time propagation technique using splitting methods with complex coefficients, J. Chem. Phys. 139 (2013) 124117.

[3] S. Blanes, F. Diele, C. Marangi, S. Ragni, Splitting and composition methods for explicit time dependence in separable dynamical systems, J. Comput. Appl. Math. 235 (2010) 646–659.

[4] S. Blanes, P.C. Moan, Practical symplectic partitioned Runge–Kutta and Runge–Kutta–Nyström methods, J. Comput. Appl. Math. 142 (2002) 313–330.

[5] F. Castella, P. Chartier, S. Descombes, G. Vilmart, Splitting methods with complex times for parabolic equations, BIT 49 (2009) 487–508.

[6] S.A. Chin, Symplectic integrators from composite operator factorizations, Phys. Lett. A 226 (1997) 344–348.

[7] S.A. Chin, Structure of positive decompositions of exponential operators, Phys. Rev. E 71 (2005) 016703.
[8] S.A. Chin, C.R. Chen, Fourth order gradient symplectic integrator methods for solving the time-dependent Schrödinger equation, J. Chem. Phys. 114 (2001) 7338–7341.
[9] S.A. Chin, C.R. Chen, Gradient symplectic algorithms for solving the Schrödinger equation with time-dependent potentials, J. Chem. Phys. 117 (2002) 1409–1415.
[10] E. Forest, R.D. Ruth, Fourth-order symplectic integration, Phys. D 43 (1990) 105–117.
[11] B. Fornberg, A Practical Guide to Pseudospectral Methods, Cambridge University Press, Cambridge, 1996.
[12] E. Hairer, C. Lubich, G. Wanner, Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations, Springer, Berlin, 2002.
[13] E. Hairer, S.P. Nørsett, G. Wanner, Solving Ordinary Differential Equations I. Nonstiff Problems, Springer, Berlin, 2nd edition, 1993.
[14] E. Hansen, A. Ostermann, Exponential splitting for unbounded operators, Math. Comput. 78 (2009) 1485–1496.
[15] E. Hansen, A. Ostermann, High order splitting methods for analytic semigroups exist, BIT 49 (2009) 527–542.
[16] T. Jahnke, C. Lubich, Error bounds for exponential operator splittings, BIT 40 (2000) 735–744.
[17] O. Koch, C. Neuhauser, M. Thalhammer, Error analysis of high-order splitting methods for nonlinear evolutionary Schrödinger equations and application to the MCTDHF equations in electron dynamics, M2AN Math. Model. Numer. Anal. 47 (2013) 1265–1286.
[18] P.V. Koseleff, Relations among Lie formal series and construction of symplectic integrators, in: G. Cohen, T. Mora, O. Moreno (Eds.), Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, volume 673 of *Lecture Notes in Computer Science*, Springer, Berlin, 1993, pp. 213–230.
[19] H.-O. Kreiss, J. Oliger, Comparison of accurate methods for the integration of hyperbolic equations, Tellus 24 (1972) 199–215.
[20] C. Neuhauser, M. Thalhammer, On the convergence of splitting methods for linear evolutionary Schrödinger equations involving an unbounded potential, BIT 49 (2009) 199–215.
[21] I.P. Omelyan, I.M. Mryglod, R. Folk, Construction of high-order force-gradient algorithms for integration of motion in classical and quantum systems, Phys. Rev. E 66 (2002) 026701.
[22] A. Ostermann, K. Schratz, Stability of exponential operator splitting methods for noncontractive semigroups, SIAM J. Numer. Anal. 51 (2013) 191–203.
[23] G. Russo, P. Smereka, The Gaussian wave packet transform: Efficient computation of the semi-classical limit of the Schrödinger equation. Part 2 – multidimensional case, J. Comput. Phys. 257 (2014) 1022–1038.
[24] M. Seydaoğlu, S. Blanes, High-order splitting methods for separable non-autonomous parabolic equations, Appl. Numer. Math. 84 (2014) 22–32.
[25] M. Suzuki, Hybrid exponential product formulas for unbounded operators with possible applications to Monte Carlo simulations, Phys. Lett. A 201 (1995) 425–428.
[26] M. Suzuki, New scheme of hybrid exponential product formulas with applications to quantum Monte-Carlo simulations, in: D.P. Landau, K.K. Mon, H.B. Schüttler (Eds.), Computer Simulation Studies in Condensed-Matter Physics VIII, volume 80 of *Springer Proceedings in Physics*, Springer, Berlin, 1995, pp. 169–174.
[27] M. Thalhammer, A fourth-order commutator-free exponential integrator for nonautonomous differential equations, SIAM J. Numer. Anal. 44 (2006) 851–864.
[28] M. Thalhammer, High-order exponential operator splitting methods for time-dependent Schrödinger equations, SIAM J. Numer. Anal. 46 (2008) 2022–2038.
[29] M. Thalhammer, Convergence analysis of high-order time-splitting pseudospectral methods for nonlinear Schrödinger equations, SIAM J. Numer. Anal. 50 (2012) 3231–3258.
[30] H. Yoshida, Construction of higher order symplectic integrators, Phys. Lett. A 150 (1990) 262–268.