IMA Journal of Numerical Analysis (2018) **38**, 1294–1323
doi: 10.1093/imanum/drx047
Advance Access publication on August 19, 2017

Avoiding order reduction when integrating linear initial boundary value problems with exponential splitting methods

I. Alonso-Mallo and B. Cano*
IMUVA, Departamento de Matemática Aplicada, Facultad de Ciencias,
Universidad de Valladolid, Paseo de Belén 7, 47011 Valladolid, Spain
*Corresponding author: bego@mac.uva.es isaias@mac.uva.es

and

N. Reguera
IMUVA, Departamento de Matemáticas y Computación, Escuela Politécnica Superior,
Universidad de Burgos, Avda. Cantabria, 09006 Burgos, Spain
nreguera@ubu.es

[Received on 6 September 2016; revised on 7 July 2017]

It is well known the order reduction phenomenon which arises when exponential methods are used to integrate time-dependent initial boundary value problems, so that the classical order of these methods is reduced. In particular, this subject has been recently studied for Lie–Trotter and Strang exponential splitting methods, and the order observed in practice has been exactly calculated. In this article, a technique is suggested to avoid that order reduction. We deal directly with nonhomogeneous time-dependent boundary conditions, without having to reduce the problem to the homogeneous ones. We give a thorough error analysis of the full discretization and justify why the computational cost of the technique is negligible in comparison with the rest of the calculations of the method. Some numerical results for dimension splittings are shown, which corroborate that much more accuracy is achieved.

Keywords: exponential Lie-Trotter; exponential Strang; avoiding order reduction; initial boundary value problems.

1. Introduction

Splitting methods are well known to be of interest for differential problems in which the numerical integration of separated parts of the equation is much easier or cheaper than the numerical integration as a whole (Yanenko, 1971; Hundsdorfer & Verwer, 2003). Moreover, if the stiff part of those separated problems is linear, it can be solved in an explicit way using exponential-type methods without showing stability problems. This makes exponential splitting methods very much suitable for the numerical integration of partial differential equations and, in particular, for multidimensional problems in simple domains, where considering alternatively each direction of the differential operator leads to simpler integrators.

In Faou *et al*. (2015), a thorough analysis is given for the classical first-order Lie–Trotter and classical second-order Strang exponential methods when integrating linear initial boundary value parabolic problems under homogeneous Dirichlet boundary conditions. The general conclusion there is that order reduction to 1 appears for the local error with Lie–Trotter method, although there is no order reduction for the global one. With Strang method, order reduction to 1 for the global error is shown. When the boundary condition is not homogeneous, but it is the restriction of a known smooth function on the total domain, then the problem can be reduced to one in which the boundary condition is homogeneous, but in which the source term contains derivatives of that smooth function, which must be calculated. In any case, the previous mentioned order reduction would turn up.

Our aim in this article is to approximate regular solutions of linear differential problems with generalizations of Lie–Trotter and Strang methods that avoid completely order reduction. Moreover, we will deal directly with nonhomogeneous and time-dependent boundary conditions. We will give a technique to do it, which requires a computational cost that is negligible compared with the total cost of the method, because it just adds calculations with grid values on the boundaries and not with the number of grid values on the total domain. In this sense, this technique is as cheap as that suggested in Alonso-Mallo *et al.* (2016) for exponential Lawson methods and, among others, in Alonso-Mallo (2002) and Alonso-Mallo *et al.* (2004) for other standard Runge–Kutta type methods. The idea, in a similar way as in Connors *et al*. (2014), LeVeque (1986) and LeVeque & Oliger (1983), is to consider suitable intermediate boundary conditions for the split evolutionary problems. The main difference with LeVeque (1986) and LeVeque & Oliger (1983) is that they just consider the one-dimensional first order in time hyperbolic problem, where one of the splitting parts is assumed to be smooth (or vary slowly) and the suggestion of the intermediate boundary conditions is very much based on a particular space discretization. As distinct, in the present article, both the problem and the space discretization are much more general. As for Connors *et al*. (2014), although the problem is more general than in LeVeque (1986) and LeVeque & Oliger (1983), numerical differentiation is required to approximate the boundary conditions of the intermediate evolutionary problems, while here they are given directly in terms of data. Besides, as we consider exponential methods and use exact boundary values, no stability requirement is needed and, as final differences, not only the class of linear problems is more general here, but also the way to measure the error in the analysis is more standard.

Moreover, in contrast to other examples of analysis on order reduction in the literature (Einkemmer & Ostermann, 2015, 2016a; Faou *et al*., 2015), we consider, not only the time discretization, but also the space discretization for each part of the differential operator splitting. This is important, not only because in practice a space discretization is necessary and therefore the errors that come from space must be controlled, but also because the complete description of the suggested method must be given to those who are just interested in applying the method and not on the analysis. More particularly, we consider spatial schemes satisfying quite general hypotheses, which include, for example, simple finite differences or collocation spectral methods. Moreover, the exact formulas which must be implemented after full discretization to avoid order reduction in the local and global error are given in (6.3)–(6.5) for Lie–Trotter and in (7.5)–(7.9) for Strang method. In Section 8.1, we justify that, for dimension splittings, the terms corresponding to the boundary in those formulas can always be calculated in terms of the data of the problem for Lie–Trotter method and, when the splitting terms of the differential operator commute, also for Strang method. Nevertheless, for the latter integrator, when the splitting operators do not commute, we offer the alternative (7.9)–(7.13), which boundaries can always be calculated in terms of data. In such a way, just order 2 instead of 3 is obtained for the local error but, in any case, no order reduction is shown for the global error in practice.

Although it is not an aim of this article, there are already results on applying a similar technique to nonlinear problems (Alonso-Mallo *et al*., 2017; Cano & Reguera, 2017) and in Einkemmer & Ostermann, 2016b they try to compare with the technique in Einkemmer & Ostermann (2015, 2016a) for reaction-diffusion problems.

The article is structured as follows. Section 2 gives some preliminaries on the abstract formulation of the problem and the definition of the time integrators. Section 3 describes the technique to avoid order reduction after time discretization with Lie–Trotter method as well as the analysis on the local error.

Section 4 does the same for Strang method. In Section 5, the hypotheses on the spatial discretization are stated. Sections 6 and 7 describe the formulas for the implementation after full discretization for Lie–Trotter and Strang methods, respectively, and the local and global errors are then analysed. Finally, in Section 8, it is justified that the dimension splitting problem fits into the abstract framework and the information which is needed on the boundary can be calculated from data. Besides, some numerical results are given, which corroborate the results of previous sections.

2. Preliminaries

Let X and Y be Banach spaces and let $L : D(L) \to X$ and $\partial : D(L) \to Y$ be linear operators. Our goal is to study full discretizations of the linear abstract inhomogeneous initial boundary value problem

$$\begin{aligned} u'(t) &= Lu(t) + f(t), \quad 0 \le t \le T, \\ u(0) &= u_0 \in X, \\ \partial u(t) &= g(t) \in Y, \quad 0 \le t \le T. \end{aligned} \tag{2.1}$$

The abstract setting (2.1) permits to cover a wide range of evolutionary problems governed by linear partial differential equations. To assure that (2.1) is well posed, we assume that (cf. Palencia & Alonso-Mallo, 1994):

(i) The boundary operator $\partial : D(L) \subset X \to Y$ is onto.

(ii) $\mathrm{Ker}(\partial)$ is dense in X and $L_0 : D(L_0) = \ker(\partial) \subset X \to X$, the restriction of L to $\mathrm{Ker}(\partial)$, is the infinitesimal generator of a C_0-semigroup $\{e^{tL_0}\}_{t\ge 0}$ in X, which type ω is assumed to be negative.

(iii) If $\lambda \in \mathbb{C}$ satisfies $Re(\lambda) > \omega$ and $v \in Y$, then the steady-state problem

$$\begin{aligned} Lx &= \lambda x, \\ \partial x &= v, \end{aligned} \tag{2.2}$$

possesses a unique solution denoted by $x = K(\lambda)v$. Moreover, the linear operator $K(\lambda) : Y \to D(L)$ satisfies

$$\|K(\lambda)v\|_X \le M\|v\|_Y,$$

where the constant M holds for any λ such that $Re(\lambda) \ge \omega_0 > \omega$.

The main goal of this work is to propose a suitable generalization, for initial boundary value problems, of two popular exponential splitting time integrators, the Lie–Trotter and the Strang methods. Therefore, we suppose that

$$L = A + B, \tag{2.3}$$

where $A : D(A) \to X$ and $B : D(B) \to X$ are linear operators that are assumed to be simpler than L in some sense, and $D(L) \subseteq D(A) \cap D(B)$. We also suppose that, for some Banach spaces Y_A and Y_B, the linear operators $\partial_A : D(A) \to Y_A$ and $\partial_B : D(B) \to Y_B$ satisfy the following assumptions:

(A1) $\mathrm{Ker}(\partial) = \mathrm{Ker}(\partial_A) \cap \mathrm{Ker}(\partial_B)$.

(A2) $A_0 : \mathrm{Ker}(\partial_A) \subset X \to X$ and $B_0 : D(B_0) = \mathrm{Ker}(\partial_B) \subset X \to X$, the restrictions of A (resp. B) to $\mathrm{Ker}(\partial_A)$ (resp. $\mathrm{Ker}(\partial_B)$) are the infinitesimal generators of C_0- semigroups in X: $\{e^{tA_0}\}_{t\geq 0}$, with type ω_A, and $\{e^{tB_0}\}_{t\geq 0}$, with type ω_B. We assume that $\max(\omega_A, \omega_B) < 0$.

(A3) If $\lambda \in \mathbb{C}$ satisfies $Re(\lambda) > \max(\omega_A, \omega_B)$ and $v_A \in Y_A$, $v_B \in Y_B$, then the steady-state problems

$$\begin{aligned} Ax &= \lambda x, & By &= \lambda y, \\ \partial_A x &= v_A, & \partial_B x &= v_B, \end{aligned} \tag{2.4}$$

possess unique solutions denoted by $x = K_A(\lambda)v_A$, $y = K_B(\lambda)v_B$. Moreover, these operators $K_A(\lambda) : Y_A \to D(A)$, $K_B(\lambda) : Y_B \to D(B)$, satisfy

$$\|K_A(\lambda)v_A\|_X \leq L_A\|v_A\|_{Y_A}, \quad \|K_B(\lambda)v_A\|_X \leq L_B\|v_B\|_{Y_B}, \tag{2.5}$$

where the constants L_A, L_B hold for any λ such that $Re(\lambda) \geq \omega_1 > \max(\omega_A, \omega_B)$.

To define the time integrators that are used in this article, we will consider initial boundary value problems which can be written as

$$\begin{aligned} u'(s) &= Au(s), \\ u(0) &= u_0, \\ \partial_A u(s) &= v_0 + v_1 s + v_2 s^2, \end{aligned} \tag{2.6}$$

where $u_0 \in X$ and $v_0, v_1, v_2 \in Y_A$. (Similar problems with B instead of A are also used.) The study of the well-posedness of these problems is not the objective of this article, but when the initial value is regular and compatible with the boundary datum at $s = 0$, we can explicitly obtain the solution using the hypotheses (A2) and (A3).

LEMMA 2.1 (Pazy, 1983; Alonso-Mallo *et al.*, 2017) If $f \in C^1([0, T], X)$, then $\int_0^t e^{sA_0} f(t-s)\,\mathrm{d}s \in D(A_0)$ and

$$A_0 \int_0^t e^{sA_0} f(t-s)\,\mathrm{d}s = e^{tA_0} f(0) - f(t) + \int_0^t e^{sA_0} f'(t-s)\,\mathrm{d}s,$$

for $t \geq 0$.

PROPOSITION 2.2 Assume that $u_0 \in D(A)$ and $\partial_A u_0 = v_0$, then the solution of (2.6) is given by

$$\begin{aligned} u(t) = {} & e^{tA_0}\left(u_0 - K_A(0)v_0\right) + K_A(0)(v_0 + v_1 t + v_2 t^2) \\ & - \int_0^t e^{sA_0} K_A(0)(v_1 + 2v_2(t-s))\,\mathrm{d}s. \end{aligned} \tag{2.7}$$

Proof. Since $\partial_A u_0 = v_0$, $u_0 - K_A(0)v_0 \in D(A_0)$ and therefore

$$\partial_A\left(e^{tA_0}(u_0 - K_A(0)v_0)\right) = 0.$$

On the other hand, from Lemma 2.1,

$$\partial_A \left(\int_0^t e^{sA_0} K_A(0)(v_1 + 2v_2(t-s))\,\mathrm{d}s \right) = 0,$$

and we deduce that $\partial_A u(t) = v_0 + v_1 t + v_2 t^2$. Moreover,

$$u'(t) = e^{tA_0} A_0 \left(u_0 - K_A(0)v_0 \right) + K_A(0)(v_1 + 2v_2 t) - e^{tA_0} K_A(0) v_1 - \int_0^t e^{sA_0} K_A(0) 2v_2\,\mathrm{d}s.$$

On the other hand, using Lemma 2.1 and the definition of $K_A(0)$,

$$\begin{aligned} Au(t) &= e^{tA_0} A_0 \left(u_0 - K_A(0)v_0 \right) - A_0 \int_0^t e^{sA_0} K_A(0)(v_1 + 2v_2(t-s))\,\mathrm{d}s \\ &= e^{tA_0} A_0 \left(u_0 - K_A(0)v_0 \right) - e^{tA_0} K_A(0) v_1 + K_A(0)(v_1 + 2v_2 t) - \int_0^t e^{sA_0} K_A(0) 2v_2\,\mathrm{d}s. \end{aligned}$$

Finally, it is straightforward that $u(0) = u_0$. □

REMARK 2.3 Notice that (2.7) is well defined for any $u_0 \in X$ and $v_0, v_1, v_2 \in Y_A$; therefore, it may be considered as a generalized solution of (2.6), which can be used even when u_0 is not regular or not compatible with the boundary values. We will use this fact to establish the time integrator method in the following section.

Because of hypothesis (A2), $\{\varphi_j(tA_0)\}_{j=1}^3$ and $\{\varphi_j(tB_0)\}_{j=1}^3$ are bounded operators for $t > 0$, where $\{\varphi_j\}$ are the standard functions used in exponential methods (Hochbruck & Ostermann, 2010), which are defined by

$$\varphi_j(tA_0) = \frac{1}{t^j} \int_0^t e^{(t-\tau)A_0} \frac{\tau^{j-1}}{(j-1)!}\,\mathrm{d}\tau, \quad j \geq 1, \tag{2.8}$$

and can be calculated in a recursive way through the formulas

$$\varphi_{j+1}(z) = \frac{\varphi_j(z) - 1/j!}{z}, \quad z \neq 0, \qquad \varphi_{j+1}(0) = \frac{1}{(j+1)!}, \qquad \varphi_0(z) = e^z. \tag{2.9}$$

These functions are well known to be bounded on the complex plane when $\mathrm{Re}(z) \leq 0$.

For the time integration, we will center on Lie–Trotter and Strang methods, which when applied to a finite-dimensional linear problem, such as

$$U'(t) = M_1 U(t) + M_2 U(t) + F(t), \tag{2.10}$$

where M_1 and M_2 are matrices, is described by the following formulas at each step

$$U_{n+1} = e^{kM_1} e^{kM_2} \big(U_n + kF(t_n) \big), \tag{2.11}$$

$$U_{n+1} = e^{\frac{k}{2}M_1} e^{\frac{k}{2}M_2} \left(e^{\frac{k}{2}M_2} e^{\frac{k}{2}M_1} U_n + kF\left(t_n + \frac{k}{2}\right)\right). \tag{2.12}$$

where $k > 0$ is time step size and $t_n = nk$ for $n \geq 0$.

For the study of the Lie–Trotter method, we also assume that the solution of (2.1) satisfies the following:

(A4) For every $t \in [0,T]$, and for any natural numbers l_1, l_2, j such that $l_1 + l_2 + j \le 2$, $u^{(j)}(t) \in D(A^{l_1}B^{l_2})$ and $A^{l_1}B^{l_2}u^{(j)}(t) \in C([0,T],X)$.

REMARK 2.4 From the hypothesis (A4) and the formulas

$$u' = Au + Bu + f, \quad Au' = A^2u + ABu + Af, \quad Bu' = BAu + B^2u + Bf,$$

we deduce that $f(t) \in D(A) \cap D(B)$ for $t \in [0,T]$ and $f, Af, Bf \in C([0,T],X)$.

For the finer results on Strang method, we assume that

(A4′) For every $t \in [0,T]$, and for any natural numbers l_1, l_2, l_3, l_4, j such that $l_1 + l_2 + l_3 + l_4 + j \le 3$, $u^{(j)}(t) \in D(A^{l_1}B^{l_2}A^{l_3}B^{l_4})$ and $A^{l_1}B^{l_2}A^{l_3}B^{l_4}u^{(j)}(t) \in C([0,T],X)$.

REMARK 2.5 From the hypothesis (A4′) we deduce that $f(t) \in D(A) \cap D(B) \cap D(A^2) \cap D(B^2) \cap D(AB)$ for $t \in [0,T]$ and $f, Af, Bf, A^2f, B^2f, ABf \in C([0,T],X)$.

REMARK 2.6 Although the assumptions (A4) and (A4′) seem complicated, we emphasize that, in the context of partial differential problems, they only imply that the solution $u(t)$ is regular enough in time and space.

We would like to clarify that, in the literature, A and B usually denote what we now call the operators A_0 and B_0 (which are the restriction of the present A and B to the kernels of the boundary operators ∂_A and ∂_B). In such a case, hypotheses similar to (A4) and (A4′) are artificial since belonging to the domain of a power of A_0 (or B_0) implies vanishing conditions in the boundary, which are not necessarily satisfied by the solution of (2.1). Because of that order reduction turns up in the general case and that is what we manage to avoid in the present article.

3. Time semidiscretization: exponential Lie–Trotter splitting

In this section, we give the technique to generalize the Lie–Trotter exponential method for the solution of initial boundary value problems with nonvanishing boundary conditions. Besides, we prove the full order of the local error of the time semidiscretization, that is, the time order reduction is completely avoided.

3.1 *Description of the technique*

The technique that we suggest is based on the following:

When L_0 is the infinitesimal generator of a C_0-semigroup e^{tL_0}, $t \ge 0$, and $u_0 \in D(L_0)$, the solution of the problem

$$\begin{aligned} u'(t) &= L_0 u(t), \\ u(0) &= u_0, \end{aligned} \tag{3.1}$$

is given by $u(t) = e^{tL_0}u_0$. In this way, we are able to use exponential methods when we want to approximate the solution of an ordinary differential system and, in the case of a partial differential problem, we can approximate the solution of a pure initial value problem or an initial boundary value problem with homogeneous or periodic boundary values.

For unbounded operators that are not associated with vanishing boundary conditions, as those in (2.1), we mimic (3.1) by solving analogous differential problems where some boundaries must be specified. As we aim to generalize C_0-semigroups, for the boundaries, we consider Taylor expansions until the order of accuracy we want to achieve. More precisely, considering, for $\chi = A$ or $\chi = B$, the notation $\phi_{\chi,\eta_0,\hat{\eta}}(s)$ for the solution of

$$\begin{aligned} \eta'(s) &= \chi\eta(s), \\ \eta(0) &= \eta_0, \\ \partial_\chi\eta(s) &= \partial_\chi\hat{\eta}(s), \end{aligned} \tag{3.2}$$

we first consider

$$v_n(s) = \phi_{B,u_n+kf(t_n),\hat{v}_n}(s), \tag{3.3}$$

where

$$\hat{v}_n(s) = u(t_n) + kf(t_n) + sBu(t_n). \tag{3.4}$$

Then, we take

$$w_n(s) = \phi_{A,v_n(k),\hat{w}_n}(s), \tag{3.5}$$

where

$$\hat{w}_n(s) = u(t_n) + kf(t_n) + kBu(t_n) + sAu(t_n). \tag{3.6}$$

(Notice that, although $v_n(s)$, $\hat{v}_n(s)$, $w_n(s)$ and $\hat{w}_n(s)$ do in fact depend on k, we do not include the parameter k as a subindex to simplify the notation.)

In such a way, the numerical method is given by

$$u_{n+1} = w_n(k). \tag{3.7}$$

In practice, we need to calculate the boundary values $\partial_B\hat{v}_n(s)$ and $\partial_A\hat{w}_n(s)$ in terms of data. In Section 8, we study how to calculate these boundary values taking hypothesis (A4) into account when the splitting is dimensional.

3.2 *Local error of the time semidiscretization*

To study the local error, we consider the value obtained in (3.7) starting from $u(t_n)$ instead of u_n. More precisely, we consider

$$\overline{u}_{n+1} = \overline{w}_n(k),$$

where $\overline{w}_n(s) = \phi_{A,\overline{v}_n(k),\hat{w}_n}(s)$ with $\hat{w}_n(s)$ as in (3.6) and $\overline{v}_n(s) = \phi_{B,u(t_n)+kf(t_n),\hat{v}_n}(s)$ where $\hat{v}_n(s)$ as in (3.4).

Before bounding the local error $\rho_{n+1} = \bar{u}_{n+1} - u(t_{n+1})$, let us first study more thoroughly $\overline{w}_n(s)$ and $\overline{v}_n(s)$.

LEMMA 3.1 Under hypotheses (A1)–(A4),

$$\overline{v}_n(s) = u(t_n) + kf(t_n) + sBu(t_n) + ks\varphi_1(kB_0)Bf(t_n) + s^2\varphi_2(sB_0)B^2u(t_n),$$
$$\overline{w}_n(s) = u(t_n) + k(Bu(t_n) + f(t_n)) + sAu(t_n) + k^2e^{sA_0}[\varphi_1(kB_0)Bf(t_n) + \varphi_2(kB_0)B^2u(t_n)] + ks\varphi_1(sA_0)A(Bu(t_n) + f(t_n)) + s^2\varphi_2(sA_0)A^2u(t_n),$$

where $\varphi_1(z)$ and $\varphi_2(z)$ are defined in (2.8).

Proof. Notice that

$$\hat{v}_n'(s) = Bu(t_n) = B\hat{v}_n(s) - sB^2u(t_n) - kBf(t_n).$$

Therefore,

$$\begin{aligned}\overline{v}_n'(s) - \hat{v}_n'(s) &= B(\overline{v}_n(s) - \hat{v}_n(s)) + sB^2u(t_n) + kBf(t_n),\\ \overline{v}_n(0) - \hat{v}_n(0) &= 0,\\ \partial_B(\overline{v}_n(s) - \hat{v}_n(s)) &= 0.\end{aligned}$$

Then,

$$\begin{aligned}\overline{v}_n(s) &= \hat{v}_n(s) + \int_0^s e^{(s-\tau)B_0}[\tau B^2u(t_n) + kBf(t_n)]\,\mathrm{d}\tau\\ &= u(t_n) + kf(t_n) + sBu(t_n) + s^2\varphi_2(sB_0)B^2u(t_n) + ks\varphi_1(sB_0)Bf(t_n).\end{aligned}$$

On the other hand,

$$\begin{aligned}\hat{w}_n'(s) &= Au(t_n) = A\hat{w}_n(s) - kAf(t_n) - kABu(t_n) - sA^2u(t_n),\\ \hat{w}_n(0) &= u(t_n) + kBu(t_n) + kf(t_n).\end{aligned}$$

Therefore,

$$\begin{aligned}\overline{w}_n'(s) - \hat{w}_n'(s) &= A(\overline{w}_n(s) - \hat{w}_n(s)) + kAf(t_n) + kABu(t_n) + sA^2u(t_n),\\ \overline{w}_n(0) - \hat{w}_n(0) &= k^2\varphi_2(kB_0)B^2u(t_n) + k^2\varphi_1(kB_0)Bf(t_n),\\ \partial_A(\overline{w}_n(s) - \hat{w}_n(s)) &= 0,\end{aligned}$$

from what

$$\begin{aligned}\overline{w}_n(s) - \hat{w}_n(s) &= k^2e^{sA_0}[\varphi_2(kB_0)B^2u(t_n) + \varphi_1(kB_0)Bf(t_n)]\\ &\quad + \int_0^s e^{(s-\tau)A_0}[kA(Bu(t_n) + f(t_n)) + \tau A^2u(t_n)]\,\mathrm{d}\tau,\end{aligned}$$

which proves the lemma taking the definition (2.8) of φ_1 and φ_2 into account. □

From this, we deduce the full order of consistency.

THEOREM 3.2 Under assumptions (A1)–(A4), when integrating (2.1) along with (2.3) with Lie–Trotter method using (3.7), the local error satisfies

$$\rho_{n+1} \equiv \overline{u}_{n+1} - u(t_{n+1}) = O(k^2).$$

Proof. By considering $s = k$ in Lemma 3.1 and using (2.1), it is clear that

$$\begin{aligned}\bar{u}_{n+1} &= u(t_n) + k[Bu(t_n) + Au(t_n) + f(t_n)] \\ &\quad + k^2 e^{kA_0}\varphi_2(kB_0)B^2u(t_n) + k^2\varphi_1(kA_0)ABu(t_n) + k^2\varphi_2(kA_0)A^2u(t_n) \\ &\quad + k^2 e^{kA_0}\varphi_1(kB_0)Bf(t_n) + k^2\varphi_1(kA_0)Af(t_n) \\ &= u(t_n) + ku'(t_n) + O(k^2) = u(t_{n+1}) + O(k^2). \qquad \square\end{aligned}$$

4. Time semidiscretization: exponential Strang splitting

With the same idea as in Section 3, we describe now how to generalize Strang exponential method to initial boundary value problems with nonvanishing boundary values in such a way that time order reduction is completely avoided.

4.1 *Description of the technique*

For the time integration of (2.1) along with (2.3), we first consider

$$v_n(s) = \phi_{A,u_n,\widehat{v}_n}(s), \tag{4.1}$$

where

$$\widehat{v}_n(s) = u(t_n) + sAu(t_n) + \frac{s^2}{2}A^2u(t_n); \tag{4.2}$$

secondly,

$$w_n(s) = \phi_{B,v_n(\frac{k}{2}),\widehat{w}_n}(s), \tag{4.3}$$

where

$$\widehat{w}_n(s) = u(t_n) + \frac{k}{2}Au(t_n) + \frac{k^2}{8}A^2u(t_n) + sBu(t_n) + s\frac{k}{2}BAu(t_n) + \frac{s^2}{2}B^2u(t_n); \tag{4.4}$$

thirdly,

$$r_n(s) = \phi_{B,w_n(\frac{k}{2})+kf(t_n+\frac{k}{2}),\widehat{r}_n}(s), \tag{4.5}$$

where

$$\begin{aligned}\widehat{r}_n(s) &= u(t_n) + \frac{k}{2}Au(t_n) + \frac{k}{2}Bu(t_n) + kf\left(t_n + \frac{k}{2}\right) \\ &\quad + \frac{k^2}{8}A^2u(t_n) + \frac{k^2}{4}BAu(t_n) + \frac{k^2}{8}B^2u(t_n) \\ &\quad + sBu(t_n) + s\frac{k}{2}BAu(t_n) + s\frac{k}{2}B^2u(t_n) + skBf\left(t_n + \frac{k}{2}\right) + \frac{s^2}{2}B^2u(t_n);\end{aligned} \tag{4.6}$$

and, finally,

$$z_n(s) = \phi_{A, r_n(\frac{k}{2}), \widehat{z}_n}(s), \tag{4.7}$$

where

$$\begin{aligned}\widehat{z}_n(s) &= u(t_n) + \frac{k}{2}Au(t_n) + kBu(t_n) + kf\left(t_n + \frac{k}{2}\right) \\ &\quad + \frac{k^2}{8}A^2u(t_n) + \frac{k^2}{2}BAu(t_n) + \frac{k^2}{2}B^2u(t_n) + \frac{k^2}{2}Bf\left(t_n + \frac{k}{2}\right) \\ &\quad + sAu(t_n) + s\frac{k}{2}A^2u(t_n) + skABu(t_n) + skAf\left(t_n + \frac{k}{2}\right) + \frac{s^2}{2}A^2u(t_n).\end{aligned} \tag{4.8}$$

Then, we advance a step by taking

$$u_{n+1} = z_n\left(\frac{k}{2}\right). \tag{4.9}$$

In practice, we need to calculate the boundary values in (4.1), (4.3), (4.5) and (4.7) in terms of data f and g. In Section 8, we study how to calculate these boundary values taking hypothesis (A4′) into account when alternating directions are used.

4.2 *Local error of the time semidiscretization*

To study the local error, we consider the functions $\overline{v}_n$, $\overline{w}_n$, $\overline{r}_n$ and $\overline{z}_n$, obtained in (4.1), (4.3), (4.5), (4.7), starting from $u_n = u(t_n)$ in (4.1). Following a similar argument as that of Lemma 3.1, this result follows:

LEMMA 4.1 Under hypotheses (A1)–(A3) and (A4′),

$$\begin{aligned}\overline{v}_n(s) &= \hat{v}_n(s) + s^3\varphi_3(sA_0)A^3u(t_n), \\ \overline{w}_n(s) &= \hat{w}_n(s) + \frac{k^3}{8}e^{sB_0}\varphi_3\left(\frac{k}{2}A_0\right)A^3u(t_n) + \frac{k^2}{8}s\varphi_1(sA_0)BA^2u(t_n) + \frac{k}{2}s^2\varphi_2(sA_0)B^2Au(t_n) \\ &\quad + s^3\varphi_3(sA_0)B^3u(t_n), \\ \overline{r}_n(s) &= \hat{r}_n(s) - e^{sB_0}\left[\frac{k^3}{8}e^{\frac{k}{2}B_0}\varphi_3\left(\frac{k}{2}A_0\right)A^3u(t_n) + \frac{k^3}{16}\varphi_1\left(\frac{k}{2}A_0\right)BA^2u(t_n)\right. \\ &\quad \left. + \frac{k^3}{8}\varphi_2\left(\frac{k}{2}A_0\right)B^2Au(t_n) + \frac{k^3}{8}\varphi_3\left(\frac{k}{2}A_0\right)B^3u(t_n)\right] \\ &\quad + s\varphi_1(sB_0)\left[\frac{k^2}{4}B^2Au(t_n) + \frac{k^2}{8}B^3u(t_n)\right] \\ &\quad + s^2\varphi_2(sB_0)\left[\frac{k}{2}B^2Au(t_n) + \frac{k}{2}B^3u(t_n) + kB^2f\left(t_n + \frac{k}{2}\right)\right] + s^3\varphi_3(sB_0)B^3u(t_n),\end{aligned}$$

$$
\begin{aligned}
\overline{z}_n(s) = \hat{z}_n(s) - e^{sA_0} e^{\frac{k}{2}B_0} \Bigg[& \frac{k^3}{8} e^{\frac{k}{2}B_0} \varphi_3\left(\frac{k}{2}A_0\right) A^3 u(t_n) + \frac{k^3}{16}\varphi_1\left(\frac{k}{2}A_0\right) BA^2 u(t_n) \\
& + \frac{k^3}{8}\varphi_2\left(\frac{k}{2}A_0\right) B^2 A u(t_n) + \frac{k^3}{8}\varphi_3\left(\frac{k}{2}A_0\right) B^3 u(t_n) \Bigg] \\
& + sk^2\varphi_1(sA_0)\left[\frac{1}{8}A^3 u(t_n) + \frac{1}{2}ABAu(t_n) + \frac{1}{2}AB^2 u(t_n) + \frac{1}{2}ABf\left(t_n + \frac{k}{2}\right)\right] \\
& + s^2 k\varphi_2(sA_0)\left[\frac{1}{2}A^3 u(t_n) + A^2 Bu(t_n) + A^2 f\left(t_n + \frac{k}{2}\right)\right] + s^3\varphi_3(sA_0)A^3 u(t_n).
\end{aligned}
$$

From Lemma 4.1, it is clear that

$$
\begin{aligned}
\overline{z}_n\left(\frac{k}{2}\right) &= \widehat{z}_n\left(\frac{k}{2}\right) + O(k^3) \\
&= u(t_n) + k(A+B)u(t_n) + kf(t_n) \\
&\quad + \frac{k^2}{2}\big((A+B)^2 u(t_n) + (A+B)f(t_n) + f'(t_n)\big) + O(k^3) \\
&= u(t_n) + ku'(t_n) + \frac{k^2}{2}u''(t_n) + O(k^3) = u(t_n + k) + O(k^3).
\end{aligned}
$$

Then, if we define

$$
\overline{u}_{n+1} = \overline{z}_n\left(\frac{k}{2}\right),
$$

we have proved the following result:

THEOREM 4.2 Under assumptions (A1)–(A3), (A4′), when integrating (2.1) along with (2.3) with Strang method using the procedures (4.1)–(4.9), the local error satisfies

$$
\rho_{n+1} = u(t_{n+1}) - \overline{u}_{n+1} = O(k^3).
$$

We will show in Section 8 that, when the splitting is dimensional, the terms of second order in s and k in the functions (4.2)–(4.4)–(4.6)–(4.8) can be calculated in terms of the data f and g only when the operators A and B commute. However, the alternative boundary values

$$
\begin{aligned}
\widehat{v}_n(s) &= u(t_n) + sAu(t_n), \\
\widehat{w}_n(s) &= u(t_n) + \tfrac{k}{2}Au(t_n) + sBu(t_n), \\
\widehat{r}_n(s) &= u(t_n) + \tfrac{k}{2}Au(t_n) + \tfrac{k}{2}Bu(t_n) + kf(t_n + \tfrac{k}{2}) + sBu(t_n), \\
\widehat{z}_n(s) &= u(t_n) + \tfrac{k}{2}Au(t_n) + kBu(t_n) + kf(t_n + \tfrac{k}{2}) + sAu(t_n),
\end{aligned}
\tag{4.10}
$$

can always be calculated and we obtain

THEOREM 4.3 Under assumptions (A1)–(A4), when integrating (2.1) along with (2.3) with Strang method using the technique (4.1), (4.3), (4.5), (4.7), (4.9), with the alternative boundary values (4.10), the local error satisfies

$$\rho_{n+1} = u(t_{n+1}) - \overline{u}_{n+1} = O(k^2).$$

5. Spatial discretization

In this section, we describe a quite general procedure to discretize in space the corresponding evolutionary problems.

Although the previous analysis is valid for other types of boundary conditions, we consider here, for the sake of simplicity, an abstract spatial discretization which is suitable for Dirichlet boundary conditions. (Look at Alonso-Mallo *et al.*, 2017, for a complete analysis of a similar technique with Neumann or Robin boundary conditions for nonlinear problems, where the nonlinear part is a smooth operator. With linear problems, although both operators are unbounded, the analysis there would be extended in a simpler way, because the boundary conditions can always be exactly calculated in terms of data instead of just approximately, as it happens in Alonso-Mallo *et al.*, 2017).

Without loss of generality, we will assume that we have the same parameter of space discretization for A and B. Let us denote it by $h \in (0, h_0]$. Let X_h be a family of finite dimensional spaces, approximating X. The norm in X_h is denoted by $\|\cdot\|_h$. We suppose that

$$X_h = X_{h,0} \oplus X_{h,b}$$

in such a way that the internal approximation is collected in $X_{h,0}$ and $X_{h,b}$ accounts for the boundary values.

The elements in $D(A_0) \cap D(B_0)$, which are regular in space and have vanishing boundary conditions, can be approximated in $X_{h,0}$. However, it is possible to consider elements $u \in X$ that are regular in space, but with nonvanishing boundary conditions, i.e. $u \in D(A) \cap D(B)$. Then, it is necessary to use the whole discrete space X_h.

Since the solution is known at the boundary, our goal is to obtain a value in $X_{h,0}$, which is a good approximation inside the domain. Let us take a projection operator

$$P_h : X \to X_{h,0}.$$

When $x \in D(A_0) \cap D(B_0)$, $P_h x$ will be its *best* approximation in $X_{h,0}$. We also assume that there exist interpolation operators

$$Q_{h,A} : Y_A \to X_{h,b}, \quad Q_{h,B} : Y_B \to X_{h,b},$$

which permit to discretize spatially the boundary values.

On the other hand, the operators A and B are approximated by means of the operators

$$A_h : X_h \to X_{h,0}, \quad B_h : X_h \to X_{h,0},$$

in such a way that $A_{h,0}$ and $B_{h,0}$, the restrictions of A_h and B_h to the subspaces $X_{h,0}$, are approximations of A_0 and B_0. Therefore, when $x_h = x_{h,0} + x_{h,b} \in X_{h,0} \oplus X_{h,b} = X_h$, we have

$$A_h x_h = A_{h,0} x_{h,0} + A_h x_{h,b}, \quad B_h x_h = B_{h,0} x_{h,0} + B_h x_{h,b}.$$

Using this, the following semidiscrete problem arises after discretizing (2.1) along with (2.3) in space,

$$\begin{aligned} U_h'(t) &= A_{h,0}U_h(t) + B_{h,0}U_h(t) \\ &\quad + A_hQ_{h,A}\partial_A g(t) + B_hQ_{h,B}\partial_B g(t) + P_hf(t), \\ U_h(0) &= P_hu(0). \end{aligned} \tag{5.1}$$

The subsequent analysis is carried out under the following hypotheses, which are related to those in Alonso-Mallo *et al.* (2016) (see also Brenner *et al.*, 1982).

(H1) The operators $A_{h,0}$ and $B_{h,0}$ are invertible and generate uniformly bounded C_0-semigroups $e^{tA_{h,0}}$, $e^{tB_{h,0}}$, on $X_{h,0}$ satisfying

$$||e^{tA_{h,0}}||_h, ||e^{tB_{h,0}}||_h \leq M, \tag{5.2}$$

where $M \geq 1$ is a constant.

(H2) For each $u \in X$, $v_A \in Y_A$ and $v_B \in Y_B$, there exist constants C, C_A' and C_B' such that

$$\|P_hu\|_h \leq C\|u\|_X, \ \left\|Q_{h,A}v_A\right\|_h \leq C_A'\|v_A\|_{Y_A}, \ \left\|Q_{h,B}v_B\right\|_h \leq C_B'\|v_B\|_{Y_B}. \tag{5.3}$$

(H3) We define the elliptic projections $R_{h,A} : D(A) \to X_{h,0}$ and $R_{h,B} : D(B) \to X_{h,0}$ as the solutions of

$$A_h(R_{h,A}u + Q_{h,A}\partial_A u) = P_hAu, \quad B_h(R_{h,B}u + Q_{h,B}\partial_B u) = P_hBu. \tag{5.4}$$

We assume that there exists a subspace Z of X, such that, for $u \in Z$,

(a) $A_0^{-1}u, B_0^{-1}u \in Z$ and $e^{tA_0}u, e^{tB_0}u \in Z$, for $t \geq 0$,

(b) for some $\varepsilon_{h,A}$ and $\varepsilon_{h,B}$ which are small with h,

$$\left\|A_{h,0}(P_hu - R_{h,A}u)\right\|_h \leq \varepsilon_{h,A}\,\|u\|_Z\,, \ \left\|B_{h,0}(P_hu - R_{h,B}u)\right\|_h \leq \varepsilon_{h,B}\,\|u\|_Z.$$

6. Full discretization: exponential Lie–Trotter splitting

Instead of integrating firstly in space (5.1) and then in time, which is the standard method of lines for the integration of (2.1), in this section, we apply the space discretization to the intermediate evolutionary problems that were described in Section 3 when integrating firstly in time. In such a way, the following final formulas are obtained.

6.1 *Final formula for the implementation*

We apply the space discretization described above to the operators A and B, which appear in the evolutionary problems (3.2) corresponding to (3.3) and (3.5), and we obtain $V_{h,n}(s), W_{h,n}(s) \in X_{h,0}$ as the solutions of

$$\begin{aligned} V_{h,n}'(s) &= B_h(V_{h,n}(s) + Q_{h,B}\partial_B\hat{v}_n(s)), \\ V_{h,n}(0) &= U_{h,n} + kP_hf(t_n), \end{aligned} \tag{6.1}$$

where $\hat{v}_n(s)$ is that in (3.4), $U_{h,n} \in X_{h,0}$ is the numerical solution in the interior of the domain after full discretization at n steps, and

$$\begin{aligned} W'_{h,n}(s) &= A_h(W_{h,n}(s) + Q_{h,A}\partial_A \hat{w}_n(s)),\\ W_{h,n}(0) &= V_{h,n}(k), \end{aligned} \tag{6.2}$$

where $\hat{w}_n(s)$ is that in (3.6). In such a way, using the variations of constants formula,

$$V_{h,n}(k) = e^{kB_{h,0}}[U_{h,n} + kP_h f(t_n)] + \int_0^k e^{(k-s)B_{h,0}} B_h Q_{h,B}\partial_B[u(t_n) + kf(t_n) + sBu(t_n)]\,\mathrm{d}s,$$

$$W_{h,n}(k) = e^{kA_{h,0}} V_{h,n}(k) + \int_0^k e^{(k-s)A_{h,0}} A_h Q_{h,A}\partial_A[u(t_n) + kBu(t_n) + kf(t_n) + sAu(t_n)]\,\mathrm{d}s,$$

and, using then the definition of the functions φ_1 and φ_2 in (2.8),

$$\begin{aligned} V_{h,n}(k) &= e^{kB_{h,0}}[U_{h,n} + kP_h f(t_n)] \\ &\quad + k\Big[\varphi_1(kB_{h,0})B_h Q_{h,B}\partial_B[u(t_n) + kf(t_n)] + k\varphi_2(kB_{h,0})B_h Q_{h,B}\partial_B Bu(t_n)\Big], \end{aligned} \tag{6.3}$$

$$\begin{aligned} W_{h,n}(k) &= e^{kA_{h,0}} V_{h,n}(k) \\ &\quad + k\Big[\varphi_1(kA_{h,0})A_h Q_{h,A}\partial_A[u(t_n) + k(Bu(t_n) + f(t_n))] + k\varphi_2(kA_{h,0})A_h Q_{h,A}\partial_A Au(t_n)\Big], \end{aligned} \tag{6.4}$$

and the numerical solution at step $n+1$ is therefore given by

$$U_{h,n+1} = W_{h,n}(k). \tag{6.5}$$

Moreover, we will take, as initial condition,

$$U_{h,0} = P_h u(0). \tag{6.6}$$

REMARK 6.1 Notice that, when

$$\partial u(t_n) = \partial Au(t_n) = \partial Bu(t_n) = 0, \tag{6.7}$$

it is also deduced from (2.1) along with (2.3) that $\partial f(t_n) = 0$. Therefore, formulas (6.3)–(6.4) just reduce to the standard time integration with Lie–Trotter method of the corresponding differential system

$$U'_h(t) = A_{h,0}U_h(t) + B_{h,0}(t)U_h(t) + P_h f(t).$$

Although the order for the local error under these assumptions is not explicitly stated in Faou *et al.* (2015), when the exact solution of (2.1) satisfies (6.7), we are implicitly proving that there is no order reduction in the local error with the standard Lie–Trotter method.

REMARK 6.2 The calculation of the terms in (6.3) and (6.4) which contain the exponential-type functions can be performed with Krylov techniques in general Gockler & Grimm (2013) and with discrete sine transforms for some particular cases. In any case, we would like to notice that, for many space discretizations, for $v_A \in Y_A$ and $v_B \in Y_B$, $A_h Q_{h,A} v_A$ and $B_h Q_{h,B} v_B$ are local in the sense that they vanish on the interior grid nodes (or great part of them) (see Section 8). Because of this, for the terms that contain the functions φ_1 and φ_2, another possibility when the step size k is fixed during all integration is to calculate once and for all at the very beginning just some columns of the matrices, which represent $\varphi_1(kA_{h,0})$, $\varphi_2(kA_{h,0})$, $\varphi_1(kB_{h,0})$ and $\varphi_2(kB_{h,0})$. After that, at each step, just a linear combination of those columns would be necessary. In practice, $A_{h,0}$ and $B_{h,0}$ can be represented by block-diagonal matrices (where the blocks in the diagonal can even be the same in some cases, as when $A+B$ is the Laplacian) and therefore, φ_1 or φ_2 over k times those matrices is also block diagonal. Due to that, the number of nonvanishing elements of each necessary column of $\varphi_i(kA_{h,0})$ and $\varphi_i(kB_{h,0})$ would just be $O(J)$ if J is the number of nodes in each direction.

6.2 *Local errors*

To define the local error, we consider

$$\overline{U}_{h,n+1} = \overline{W}_{h,n}(k), \tag{6.8}$$

where $\overline{W}_{h,n}(s)$ is the solution of

$$\begin{aligned} \overline{W}'_{h,n}(s) &= A_h(\overline{W}_{h,n}(s) + Q_{h,A}\partial_A \hat{w}_n(s)), \\ \overline{W}_{h,n}(0) &= \overline{V}_{h,n}(k), \end{aligned} \tag{6.9}$$

with $\hat{w}_n(s)$ that in (3.6) and $\overline{V}_{h,n}(s)$ the solution of

$$\begin{aligned} \overline{V}'_{h,n}(s) &= B_h(\overline{V}_{h,n}(s) + Q_{h,B}\partial_B \hat{v}_n(s)), \\ \overline{V}_{h,n}(0) &= P_h[u(t_n) + kf(t_n)], \end{aligned} \tag{6.10}$$

with $\hat{v}_n(s)$ in (3.4). We now define the local error at $t = t_n$ as

$$\rho_{h,n} = P_h u(t_n) - \overline{U}_{h,n},$$

and study its behaviour in the following theorem.

THEOREM 6.3 Under assumptions (A1)–(A4) and (H1)–(H3), when integrating (2.1) with Lie–Trotter method as described in (6.3), (6.4), (6.5), (6.6), whenever the functions in (A4) belong to the space Z which is introduced in (H3), the local error after full discretization satisfies

$$\rho_{h,n+1} = O(k\varepsilon_{h,A} + k\varepsilon_{h,B} + k^2),$$

where $\varepsilon_{h,A}$ and $\varepsilon_{h,B}$ are those in (H3b).

Proof. From the definition of $\rho_{h,n}$,

$$\begin{aligned} \rho_{h,n+1} &= (P_h u(t_{n+1}) - P_h \overline{u}_{n+1}) + (P_h \overline{u}_{n+1} - \overline{U}_{h,n+1}) \\ &= P_h \rho_{n+1} + (P_h \overline{w}_n(k) - \overline{W}_{h,n}(k)). \end{aligned}$$

Using Theorem 3.2 and (5.3), the first term in parenthesis is $O(k^2)$. To bound the second term, we apply the operator P_h to (3.2) (corresponding to $\overline{w}_n(s)$) and use (5.4),

$$\begin{aligned} P_h\overline{w}'_n(s) &= P_hA\overline{w}_n(s) \\ &= A_h(R_{h,A}\overline{w}_n(s) + Q_{h,A}\partial_A\hat{w}_n(s)) \\ &= A_{h,0}P_h\overline{w}_n(s) + A_{h,0}(R_{h,A} - P_h)\overline{w}_n(s) + A_hQ_{h,A}\partial_A\hat{w}_n(s) \\ P_h\overline{w}_n(0) &= P_h\overline{v}_n(k). \end{aligned}$$

Then, subtracting (6.9),

$$\begin{aligned} P_h\overline{w}'_n(s) - \overline{W}'_{h,n}(s) &= A_{h,0}(P_h\overline{w}_n(s) - \overline{W}_{h,n}(s)) + A_{h,0}(R_{h,A} - P_h)\overline{w}_n(s), \\ P_h\overline{w}_n(0) - \overline{W}_{h,n}(0) &= P_h\overline{v}_n(k) - \overline{V}_{h,n}(k). \end{aligned}$$

Solving this problem exactly,

$$\begin{aligned} P_h\overline{w}_n(k) - \overline{W}_{h,n}(k) = e^{kA_{h,0}}(P_h\overline{v}_n(k) - \overline{V}_{h,n}(k)) \\ + \int_0^k e^{(k-s)A_{h,0}}A_{h,0}(R_{h,A} - P_h)\overline{w}_n(s)\,\mathrm{d}s. \end{aligned} \tag{6.11}$$

Making the difference now between (3.2) multiplied by P_h (and corresponding to $\overline{v}_n(s)$) and (6.10),

$$\begin{aligned} &P_h\overline{v}'_n(s) - \overline{V}'_{h,n}(s) = B_{h,0}(P_h\overline{v}_n(s) - \overline{V}_{h,n}(s)) + B_{h,0}(R_{h,B} - P_h)\overline{v}_n(s), \\ &P_h\overline{v}_n(0) - \overline{V}_{h,n}(0) = 0, \end{aligned}$$

which implies that

$$P_h\overline{v}_n(k) - \overline{V}_{h,n}(k) = \int_0^k e^{(k-s)B_{h,0}}B_{h,0}(R_{h,B} - P_h)\overline{v}_n(s)\,\mathrm{d}s = O(k\varepsilon_{h,B}),$$

due to (5.2) and (H3b) considering that $\overline{v}_n(s) \in Z$ because of Lemma 3.1, the hypotheses on u and f, (H3a) and the recursive definition of $\{\varphi_j\}$. Using this in (6.11) together with (5.2), (H3), and Lemma 3.1 again with $\overline{w}_n(s) \in Z$ now, it follows that $P_h\overline{w}_n(k) - \overline{W}_{h,n}(k) = O(k\varepsilon_{h,A} + k\varepsilon_{h,B})$, which proves the result. □

6.3 *Global errors*

We now study the global errors at $t = t_n$,

$$e_{h,n} = P_hu(t_n) - U_{h,n}.$$

THEOREM 6.4 Under the same assumptions of Theorem 6.3 and assuming also that there exists a constant C such that, whenever $nk \in [0, T]$,

$$\|(e^{kA_{h,0}}e^{kB_{h,0}})^n\|_h \le C, \tag{6.12}$$

the global error satisfies

$$e_{h,n} = O(k + \varepsilon_{h,A} + \varepsilon_{h,B}),$$

where $\varepsilon_{h,A}$ and $\varepsilon_{h,B}$ are those in (H3b).

Proof. It suffices to notice that

$$e_{h,n+1} = [P_h u(t_{n+1}) - \overline{U}_{h,n+1}] + [\overline{U}_{h,n+1} - U_{h,n+1}] = \rho_{h,n+1} + \overline{W}_{h,n}(k) - W_{h,n}(k),$$

where the definition of $\rho_{h,n+1}$, (6.5) and (6.8) have been used. Then, considering (6.1), (6.2), (6.9) and (6.10),

$$\overline{W}_{h,n}(k) - W_{h,n}(k) = e^{kA_{h,0}}(\overline{V}_{h,n}(k) - V_{h,n}(k)) = e^{kA_{h,0}}e^{kB_{h,0}}(P_h u(t_n) - U_{h,n}),$$

and we obtain the recursive formula

$$e_{h,n+1} = \rho_{h,n+1} + e^{kA_{h,0}}\, e^{kB_{h,0}} e_{h,n}.$$

Since $e_{h,0} = 0$ because of (6.6), this implies that

$$e_{h,n} = \sum_{l=1}^{n} (e^{kA_{h,0}}\, e^{kB_{h,0}})^{n-l} \rho_{h,l},$$

which, together with Theorem 6.3 and (6.12), proves the result. □

REMARK 6.5 Condition (6.12) is directly deduced from (5.2) whenever $A_{h,0}$ and $B_{h,0}$ commute. The more general nonconmutative case has been studied in Ostermann & Schratz (2013) in an abstract setting (that is, without considering the spatial discretizations) with other assumptions which imply that stability bound for exponential splitting methods. In particular, the authors assume that the operators A_0, B_0 and L_0 generate analytic semigroups on X. In this way, they are able to prove the stability for dimensional splitting for second order strongly elliptic operator and its splitting in L^p.

7. Full discretization: exponential Strang splitting

7.1 *Final formula for the implementation*

First, we consider the spatial discretization of the problems (4.1), (4.3), (4.5) and (4.7), which is given by

$$\begin{aligned} V'_{h,n}(s) &= A_h(V_{h,n}(s) + Q_{h,A}\partial_A \widehat{v}_n(s)), \\ V_{h,n}(0) &= U_{h,n}, \end{aligned} \tag{7.1}$$

$$\begin{aligned} W'_{h,n}(s) &= B_h(W_{h,n}(s) + Q_{h,B}\partial_B \widehat{w}_n(s)), \\ W_{h,n}(0) &= V_{h,n}\left(\tfrac{k}{2}\right), \end{aligned} \tag{7.2}$$

$$\begin{aligned} R'_{h,n}(s) &= B_h(R_{h,n}(s) + Q_{h,B}\partial_B \widehat{r}_n(s)), \\ R_{h,n}(0) &= W_{h,n}\left(\tfrac{k}{2}\right) + kP_h f\left(t_n + \tfrac{k}{2}\right), \end{aligned} \tag{7.3}$$

$$\begin{aligned} Z'_{h,n}(s) &= A_h(Z_{h,n}(s) + Q_{h,A}\partial_A \widehat{z}_n(s)), \\ Z_{h,n}(0) &= R_{h,n}\left(\tfrac{k}{2}\right). \end{aligned} \tag{7.4}$$

Then, we obtain recursively the exact solution of these full discrete problems at $s = \frac{k}{2}$:

$$\begin{aligned} V_{h,n}\left(\frac{k}{2}\right) &= e^{\frac{k}{2}A_{h,0}} U_{h,n} + \int_0^{\frac{k}{2}} e^{\left(\frac{k}{2}-\tau\right)A_{h,0}} A_h Q_{h,A}\partial_A \widehat{v}_n(\tau)\, \mathrm{d}\tau, \\ W_{h,n}\left(\frac{k}{2}\right) &= e^{\frac{k}{2}B_{h,0}} V_{h,n}\left(\frac{k}{2}\right) + \int_0^{\frac{k}{2}} e^{\left(\frac{k}{2}-\tau\right)B_{h,0}} B_h Q_{h,B}\partial_B \widehat{w}_n(\tau)\, \mathrm{d}\tau, \\ R_{h,n}\left(\frac{k}{2}\right) &= e^{\frac{k}{2}B_{h,0}} \left(W_{h,n}\left(\frac{k}{2}\right) + kP_h f\left(t_n + \frac{k}{2}\right)\right) + \int_0^{\frac{k}{2}} e^{\left(\frac{k}{2}-\tau\right)B_{h,0}} B_h Q_{h,B}\partial_B \widehat{r}_n(\tau)\, \mathrm{d}\tau, \\ Z_{h,n}\left(\frac{k}{2}\right) &= e^{\frac{k}{2}A_{h,0}} R_{h,n}\left(\frac{k}{2}\right) + \int_0^{\frac{k}{2}} e^{\left(\frac{k}{2}-\tau\right)A_{h,0}} A_h Q_{h,A}\partial_A \widehat{z}_n(\tau)\, \mathrm{d}\tau. \end{aligned}$$

If we use the values (4.2), (4.4), (4.6) and (4.8) to reach *local order 3*, in terms of φ_1 and φ_2, this can be written as

$$\begin{aligned} V_{h,n}\left(\frac{k}{2}\right) &= e^{\frac{k}{2}A_{h,0}} U_{h,n} + \frac{k}{2}\varphi_1\left(\frac{k}{2}A_{h,0}\right) A_h Q_{h,A}\partial_A u(t_n) \\ &\quad + \frac{k^2}{4}\varphi_2\left(\frac{k}{2}A_{h,0}\right) A_h Q_{h,A}\partial_A A u(t_n) + \frac{k^3}{8}\varphi_3\left(\frac{k}{2}A_{h,0}\right) A_h Q_{h,A}\partial_A A^2 u(t_n), \end{aligned} \tag{7.5}$$

$$\begin{aligned} W_{h,n}\left(\frac{k}{2}\right) &= e^{\frac{k}{2}B_{h,0}} V_{h,n}\left(\frac{k}{2}\right) \\ &\quad + \frac{k}{2}\varphi_1\left(\frac{k}{2}B_{h,0}\right) B_h Q_{h,B}\partial_B \left(u(t_n) + \frac{k}{2}Au(t_n) + \frac{k^2}{8}A^2 u(t_n)\right) \\ &\quad + \frac{k^2}{4}\varphi_2\left(\frac{k}{2}B_{h,0}\right) B_h Q_{h,B}\partial_B \left(Bu(t_n) + \frac{k}{2}BAu(t_n)\right) \\ &\quad + \frac{k^3}{8}\varphi_3\left(\frac{k}{2}B_{h,0}\right) B_h Q_{h,B}\partial_B B^2 u(t_n), \end{aligned} \tag{7.6}$$

$$\begin{aligned} R_{h,n}\left(\frac{k}{2}\right) &= e^{\frac{k}{2}B_{h,0}} \left(W_{h,n}\left(\frac{k}{2}\right) + kP_h f\left(t_n + \frac{k}{2}\right)\right) \\ &\quad + \frac{k}{2}\varphi_1\left(\frac{k}{2}B_{h,0}\right) B_h Q_{h,B}\partial_B \left(u(t_n) + \frac{k}{2}(A+B)u(t_n)\right. \\ &\quad \left. + \frac{k^2}{8}(A^2 + 2AB + B^2)u(t_n) + kf\left(t_n + \frac{k}{2}\right)\right) \\ &\quad + \frac{k^2}{4}\varphi_2\left(\frac{k}{2}B_{h,0}\right) B_h Q_{h,B}\partial_B \left(Bu(t_n) + \frac{k}{2}(BA + B^2)u(t_n) + kBf\left(t_n + \frac{k}{2}\right)\right) \\ &\quad + \frac{k^3}{8}\varphi_3\left(\frac{k}{2}B_{h,0}\right) B_h Q_{h,B}\partial_B B^2 u(t_n), \end{aligned} \tag{7.7}$$

$$
\begin{aligned}
Z_{h,n}\left(\frac{k}{2}\right) &= e^{\frac{k}{2}A_{h,0}}R_{h,n}\left(\frac{k}{2}\right) \\
&+ \frac{k}{2}\varphi_1\left(\frac{k}{2}A_{h,0}\right)A_hQ_{h,A}\partial_A\Big(u(t_n) + k\left(\left(\frac{1}{2}A + B\right)u(t_n) + f\left(t_n + \frac{k}{2}\right)\right) \\
&+ \frac{k^2}{2}\left(\left(\frac{1}{4}A^2 + BA + B^2\right)u(t_n) + Bf\left(t_n + \frac{k}{2}\right)\right)\Big) \\
&+ \frac{k^2}{4}\varphi_2\left(\frac{k}{2}A_{h,0}\right)A_hQ_{h,A}\partial_A\Big(Au(t_n) + k\left(\frac{1}{2}A^2 + AB\right)u(t_n) + kAf\left(t_n + \frac{k}{2}\right)\Big) \\
&+ \frac{k^3}{8}\varphi_3\left(\frac{k}{2}A_{h,0}\right)A_hQ_{h,A}\partial_A A^2u(t_n).
\end{aligned} \tag{7.8}
$$

Then, we take

$$
U_{h,n+1} = Z_{h,n}\left(\frac{k}{2}\right). \tag{7.9}
$$

Alternatively, if we use the values (4.10) to reach *local order 2*, we obtain with the same procedure

$$
\begin{aligned}
V_{h,n}\left(\frac{k}{2}\right) &= e^{\frac{k}{2}A_{h,0}}U_{h,n} + \frac{k}{2}\varphi_1\left(\frac{k}{2}A_{h,0}\right)A_hQ_{h,A}\partial_A u(t_n) \\
&+ \frac{k^2}{4}\varphi_2\left(\frac{k}{2}A_{h,0}\right)A_hQ_{h,A}\partial_A Au(t_n),
\end{aligned} \tag{7.10}
$$

$$
\begin{aligned}
W_{h,n}\left(\frac{k}{2}\right) &= e^{\frac{k}{2}B_{h,0}}V_{h,n}\left(\frac{k}{2}\right) + \frac{k}{2}\varphi_1\left(\frac{k}{2}B_{h,0}\right)B_hQ_{h,B}\partial_B\left(u(t_n) + \frac{k}{2}Au(t_n)\right) \\
&+ \frac{k^2}{4}\varphi_2\left(\frac{k}{2}B_{h,0}\right)B_hQ_{h,B}\partial_B Bu(t_n),
\end{aligned} \tag{7.11}
$$

$$
\begin{aligned}
R_{h,n}\left(\frac{k}{2}\right) &= e^{\frac{k}{2}B_{h,0}}\left(W_{h,n}\left(\frac{k}{2}\right) + kP_hf\left(t_n + \frac{k}{2}\right)\right) \\
&+ \frac{k}{2}\varphi_1\left(\frac{k}{2}B_{h,0}\right)B_hQ_{h,B}\partial_B\left(u(t_n) + \frac{k}{2}Au(t_n) + \frac{k}{2}Bu(t_n) + kf\left(t_n + \frac{k}{2}\right)\right) \\
&+ \frac{k^2}{4}\varphi_2\left(\frac{k}{2}B_{h,0}\right)B_hQ_{h,B}\partial_B\left(Bu(t_n)\right),
\end{aligned} \tag{7.12}
$$

$$
\begin{aligned}
Z_{h,n}\left(\frac{k}{2}\right) &= e^{\frac{k}{2}A_{h,0}}R_{h,n}\left(\frac{k}{2}\right) \\
&+ \frac{k}{2}\varphi_1\left(\frac{k}{2}A_{h,0}\right)A_hQ_{h,A}\partial_A\left(u(t_n) + \frac{k}{2}Au(t_n) + kBu(t_n) + kf\left(t_n + \frac{k}{2}\right)\right) \\
&+ \frac{k^2}{4}\varphi_2\left(\frac{k}{2}A_{h,0}\right)A_hQ_{h,A}\partial_A Au(t_n).
\end{aligned} \tag{7.13}
$$

7.2 *Local errors*

To define the local error, we consider $\overline{V}_{h,n}$, $\overline{W}_{h,n}$, $\overline{R}_{h,n}$ and $\overline{Z}_{h,n}$ the solutions of (7.1)–(7.4) starting from $U_{h,n} = P_h u(t_n)$. Then, $\overline{U}_{h,n+1} = \overline{Z}_{h,n}\left(\frac{k}{2}\right)$ and the behaviour of the local error

$$\rho_{h,n} = P_h u(t_n) - \overline{U}_{h,n},$$

is given in the following theorem.

THEOREM 7.1 Under assumptions (A1)–(A3), (A4′), and (H1)–(H3), when integrating (2.1) along with (2.3) with Strang method as described in (7.5)–(7.9), whenever the functions in (A4′) belong to the space Z which is introduced in (H3),

$$\rho_{h,n+1} = O(k\varepsilon_{h,A} + k\varepsilon_{h,B} + k^3), \tag{7.14}$$

where $\varepsilon_{h,A}$ and $\varepsilon_{h,B}$ are those in (H3b).

Proof. Making the same decomposition as in the proof of Theorem 6.3,

$$\rho_{h,n+1} = P_h\rho_{n+1} + \left(P_h\overline{u}_{n+1} - \overline{U}_{h,n+1}\right).$$

As distinct, using now Theorem 4.2, the first term in parenthesis is $O(k^3)$. To bound the second term, we now have

$$P_h\overline{u}_{n+1} - \overline{U}_{h,n+1} = P_h\overline{z}_n\left(\frac{k}{2}\right) - \overline{Z}_{h,n}\left(\frac{k}{2}\right). \tag{7.15}$$

Following a similar argument as that of the proof of Theorem 6.3,

$$P_h\overline{z}_n\left(\frac{k}{2}\right) - \overline{Z}_{h,n}\left(\frac{k}{2}\right) = e^{\frac{k}{2}A_{h,0}}\left(P_h\overline{r}_n\left(\frac{k}{2}\right) - \overline{R}_{h,n}\left(\frac{k}{2}\right)\right) + \int_0^{\frac{k}{2}} e^{\left(\frac{k}{2}-s\right)A_{h,0}}A_{h,0}(R_{h,A} - P_h)\overline{z}_n(s)\,\mathrm{d}s. \tag{7.16}$$

$$P_h\overline{r}_n\left(\frac{k}{2}\right) - \overline{R}_{h,n}\left(\frac{k}{2}\right) = e^{\frac{k}{2}B_{h,0}}\left(P_h\overline{w}_n\left(\frac{k}{2}\right) - \overline{W}_{h,n}\left(\frac{k}{2}\right)\right) + \int_0^{\frac{k}{2}} e^{\left(\frac{k}{2}-s\right)B_{h,0}}B_{h,0}(R_{h,B} - P_h)\overline{r}_n(s)\,\mathrm{d}s. \tag{7.17}$$

$$P_h\overline{w}_n\left(\frac{k}{2}\right) - \overline{W}_{h,n}\left(\frac{k}{2}\right) = e^{\frac{k}{2}B_{h,0}}\left(P_h\overline{v}_n\left(\frac{k}{2}\right) - \overline{V}_{h,n}\left(\frac{k}{2}\right)\right) + \int_0^{\frac{k}{2}} e^{\left(\frac{k}{2}-s\right)B_{h,0}}B_{h,0}(R_{h,B} - P_h)\overline{w}_n(s)\,\mathrm{d}s. \tag{7.18}$$

$$P_h\overline{v}_n\left(\frac{k}{2}\right) - \overline{V}_{h,n}\left(\frac{k}{2}\right) = \int_0^{\frac{k}{2}} e^{\left(\frac{k}{2}-s\right)A_{h,0}}A_{h,0}(R_{h,A} - P_h)\overline{v}_n(s)\,\mathrm{d}s.$$

Now, as $\overline{v}_n \in Z$ because of Lemma 4.1, again with the same argument as in the proof of Theorem 6.3, the last formula is $O(k\varepsilon_{h,A})$. Inserting this in (7.18) and using also that $\overline{w}_n \in Z$, that formula is $O(k\varepsilon_{h,A} + k\varepsilon_{h,B})$. Doing the same with (7.17) and (7.16) and taking also into account that $\overline{r}_n(s), \overline{z}_n(s) \in Z$, the result follows. □

In a similar way,

THEOREM 7.2 Under assumptions (A1)–(A4) and (H1)–(H3), when integrating (2.1) with Strang method as described in (7.10)–(7.9), whenever the functions in (A4) belong to the space Z which is introduced in (H3),

$$\rho_{h,n+1} = O(k\varepsilon_{h,A} + k\varepsilon_{h,B} + k^2), \tag{7.19}$$

where $\varepsilon_{h,A}$ and $\varepsilon_{h,B}$ are those in (H3b).

7.3 *Global errors*

For the global errors $e_{h,n} = P_h u(t_n) - U_{h,n}$, we now have the following result.

THEOREM 7.3 Under the same assumptions of Theorem 7.1 and assuming also that there exists a constant C such that, whenever $nk \in [0, T]$,

$$\|(e^{\frac{k}{2}A_{h,0}}\, e^{kB_{h,0}}\, e^{\frac{k}{2}A_{h,0}})^n\|_h \le C, \tag{7.20}$$

$$e_{h,n} = O(k^2 + \varepsilon_{h,A} + \varepsilon_{h,B}),$$

where $\varepsilon_{h,A}$ and $\varepsilon_{h,B}$ are those in (H3b).

Proof. The only difference with the proof of Theorem 6.4 is that now

$$e_{h,n+1} = \rho_{h,n+1} + \overline{Z}_{h,n}\left(\frac{k}{2}\right) - Z_{h,n}\left(\frac{k}{2}\right),$$

where $Z_{h,n}(\frac{k}{2})$ is that in (7.4). Considering also now (7.1)–(7.3),

$$\overline{Z}_{h,n}(k) - Z_{h,n}(k) = e^{\frac{k}{2}A_{h,0}}\left(\overline{R}_{h,n}\left(\frac{k}{2}\right) - R_{h,n}\left(\frac{k}{2}\right)\right) = e^{\frac{k}{2}A_{h,0}}\, e^{\frac{k}{2}B_{h,0}}\left(\overline{W}_{h,n}\left(\frac{k}{2}\right) - W_{h,n}\left(\frac{k}{2}\right)\right)$$
$$= e^{\frac{k}{2}A_{h,0}}\, e^{kB_{h,0}}\left(\overline{V}_{h,n}\left(\frac{k}{2}\right) - V_{h,n}\left(\frac{k}{2}\right)\right) = e^{\frac{k}{2}A_{h,0}}\, e^{kB_{h,0}}\, e^{\frac{k}{2}A_{h,0}}\left(P_h u(t_n) - U_{h,n}\right).$$

Then, the recursive formula for the error is

$$e_{h,n+1} = \rho_{h,n+1} + e^{\frac{k}{2}A_{h,0}}\, e^{kB_{h,0}}\, e^{\frac{k}{2}A_{h,0}}\, e_{h,n},$$

which implies that

$$e_{h,n} = \sum_{l=1}^{n} (e^{\frac{k}{2} kA_{h,0}} e^{kB_{h,0}} e^{\frac{k}{2} kA_{h,0}})^{n-l} \rho_{h,l},$$

and, together with (7.14) and (7.20), this proves the result. □

REMARK 7.4 As for condition (6.12), (7.20) is directly deduced from (5.2) whenever $A_{h,0}$ and $B_{h,0}$ commute and other assumptions, which imply that bound, appear in Ostermann & Schratz (2013) in the abstract setting of exponential operator splitting methods.

On the other hand, with the same proof, if just the assumptions of Theorem 7.2 can be done:

THEOREM 7.5 Under the same assumptions of Theorem 7.2 and assuming also (7.20),

$$e_{h,n} = O(k + \varepsilon_{h,A} + \varepsilon_{h,B}),$$

where $\varepsilon_{h,A}$ and $\varepsilon_{h,B}$ are those in (H3b).

REMARK 7.6 In spite of the previous result, the numerical experiments in Section 8.2 show that the optimal global order 2 is reached when the values (4.10) are used. It seems that this improvement is caused by a summation by parts argument similar to the one used in Faou *et al.* (2015).

8. Examples and numerical results

In this section, we corroborate the results of previous sections by integrating parabolic problems with homogeneous and nonhomogeneous Dirichlet boundary conditions with a dimension splitting.

8.1 *Dimension splitting*

We assume that a and b are sufficiently smooth positive coefficients that are bounded away from zero, and we consider the parabolic problem which is defined, for the sake of simplicity, on $0 \leq x, y \leq 1$, $0 \leq t \leq T$, as

$$\begin{aligned} u_t(t,x,y) &= (a(x,y)u_x(t,x,y))_x + (b(x,y)u_y(t,x,y))_y + f(t,x,y), \quad (8.1)\\ u(0,x,y) &= u_0(x,y),\\ u(t,0,y) &= g_{1,0}(t,y),\\ u(t,1,y) &= g_{1,1}(t,y),\\ u(t,x,0) &= g_{2,0}(t,x),\\ u(t,x,1) &= g_{2,1}(t,x). \end{aligned}$$

To adjust this problem to the abstract formulation (2.1) and to use the theory given in Palencia & Alonso-Mallo (1994), we take $\Omega = (0,1) \times (0,1)$, $X = L^2(\Omega)$, $Y = H^{3/2}(\partial\Omega)$, and L the strongly elliptic operator defined by $L = D_x(aD_x) + D_y(bD_y)$, with domain $D(L) = H^2(\Omega)$. The boundary operator $\partial : D(L) \to Y$ is the trace operator on $\partial\Omega$. In such a way, $\mathrm{Ker}(\partial) = D(L_0) = H^2(\Omega) \cap H_0^1(\Omega)$ and $L_0 = L_{|\mathrm{Ker}(\partial)}$ is a sectorial operator, which generates an analytical semigroup on X.

With the idea of using an alternating directions scheme, we consider the splitting

$$A = D_x(a(x,y)D_x), \quad B = D_y(b(x,y)D_y),$$

with

$$\begin{aligned} D(A) &= \{u \in L^2(\Omega) : D_x u, D_{xx} u \in L^2(\Omega)\}, \\ D(B) &= \{u \in L^2(\Omega) : D_y u, D_{yy} u \in L^2(\Omega)\}, \end{aligned} \tag{8.2}$$

and then

$$\begin{aligned} D(A_0) &= \{u \in L^2(\Omega) : D_x u, D_{xx} u \in L^2(\Omega),\ u(0,y) = u(1,y) = 0 \text{ for a.e. } y \in (0,1)\}, \\ D(B_0) &= \{u \in L^2(\Omega) : D_y u, D_{yy} u \in L^2(\Omega),\ u(x,0) = u(x,1) = 0 \text{ for a.e. } x \in (0,1)\}. \end{aligned}$$

Notice that $\partial\Omega = \partial_A\Omega \cup \partial_B\Omega$, with $\partial_A\Omega = \{0,1\} \times [0,1]$ and $\partial_B\Omega = [0,1] \times \{0,1\}$. Besides, $\partial_A : D(A) \to Y_A$ and $\partial_B : D(B) \to Y_B$ are the restrictions to $\partial_A\Omega$ and $\partial_B\Omega$, respectively. In such a case, hypothesis (A1) is satisfied. Regarding the assumption (A2), we note that a direct consequence of Theorem 6.6 in Ostermann & Schratz (2013) is that $A_0 : D(A_0) = \mathrm{Ker}(\partial_A) \subset X \to X$ and $B_0 : D(B_0) = \mathrm{Ker}(\partial_B) \subset X \to X$, are the infinitesimal generators of analytic semigroups on X such that $0 \in \rho(A_0),\ \rho(B_0)$. Therefore, (A2) is satisfied. We also deduce that the resolvents $(A_0 - \lambda I)^{-1} : X \to D(A_0)$ and $(B_0 - \lambda I)^{-1} : X \to D(B_0)$ are well defined for each $\lambda > 0$.

We now prove that (A3) is also satisfied. We consider the operator A (the case of B is similar). For each $y \in (0,1)$, we take $X_y = \{u(\cdot,y) \in L^2(0,1)\}$, $D(A_y) = \{u(\cdot,y) \in W^2(0,1)\}$, and we define $A_y : D(A_y) \to X_y$ by means of $A_y u(\cdot,y) = D_x(a(\cdot,y)D_x u(\cdot,y))$. Note that, with $D(A_{0,y}) = \{u(\cdot,y) \in W^2(0,1) \cap W_0^1(0,1)\}$, the operator $A_{0,y} : D(A_{0,y}) \to X_y$, given by the restriction of A_y to the subspace $D(A_{0,y})$, generates an analytic semigroup on X_y with negative type. Therefore, the resolvents $(A_{0,y} - \lambda I)^{-1}$ are well defined for $\lambda > 0$.

We take $v \in Y_A$ and we write $v_0(y) = v(0,y), v_1(y) = v(1,y)$. Then, we define $K_{A,y}(\lambda)(v)(\cdot,y) \in D(A_y)$, $\lambda > 0$, as the solution of the one-dimensional elliptic problem,

$$A_y u = \lambda u, \quad u(0,y) = v_0(y), \quad u(1,y) = v_1(y).$$

We now consider, for $(x,y) \in \overline{\Omega}$, $w_A(x,y) = v_0(y) + x(v_1(y) - v_0(y))$, which satisfies

$$A_y w_A(\cdot,y) = a_x(\cdot,y)(v_1(y) - v_0(y)), \quad w_A(0,y) = v_0(y), \quad w_A(1,y) = v_1(y),$$

and we deduce that

$$\begin{aligned} (A_y - \lambda I)(K_{A,y}(\lambda)(v)(\cdot,y) - w_A(\cdot,y)) &= \lambda(w_A(\cdot,y)) - a_x(\cdot,y)(v_1(y) - v_0(y)), \\ K_{A,y}(\lambda)(v)(0,y) - w_A(0,y) &= 0, \\ K_{A,y}(\lambda)(v)(1,y) - w_A(1,y) &= 0. \end{aligned}$$

Therefore,

$$K_{A,y}(\lambda)(v)(\cdot,y) = (A_{0,y} - \lambda I)^{-1}\Big(\lambda(w_A(\cdot,y)) - a_x(\cdot,y)(v_1(y) - v_0(y))\Big) + w_A(\cdot,y).$$

In the proof of Lemma 7.2 in Ostermann & Schratz (2013), it is proved that the resolvents $(A_{0,y} - \lambda I)^{-1}$ depend continuously on $y \in [0,1]$. Then, we can bound uniformly these resolvents and, from the regularity of a, we deduce that

$$\|K_{A,y}(\lambda)(v)(\cdot,y)\| \leq C\|(v_0(y), v_1(y))\|,$$

where C is a constant which is independent of $y \in [0,1]$. Since v_0, v_1 are continuous, (A3) is deduced.

We now study when the boundary values of the evolutionary problems that define the method can be calculated in terms of the data. More particularly, the boundaries in (6.3)–(6.4) for Lie–Trotter and those of (7.10)–(7.13) for Strang method can be obtained since

(i) $\partial_B Au$ and $\partial_A Bu$ can be calculated directly from g.

(ii) $\partial_B Bu$ can be calculated from (2.1) along with (2.3) just by considering that the rest of terms of the equation can already be calculated ($\partial_B u_t = \partial_B g_t$). The same applies for $\partial_A Au$.

In the particular case that $a(x,y) = b(x,y) = 1$, A and B commute and the boundaries in (7.5)–(7.8) can be obtained since:

(i) $\partial_B A^2 u = \partial_B u_{xxxx}$ can be calculated directly from g and $\partial_B ABu$ can also be calculated from equation

$$Au_t = A^2 u + ABu + Af, \tag{8.3}$$

which results from (2.1) by applying the operator A. (Notice that $\partial_B Au_t = \partial_B u_{txx}$ can also be calculated from g.) In a similar way, $\partial_A B^2 u$ and $\partial_A BAu$ can also be calculated from the data.

(ii) $\partial_A A^2 u = \partial_A u_{xxxx}$ can be calculated from (8.3) just by considering that the rest of terms of the equation can already be calculated because $ABu = BAu$ and because differentiating (2.1) with respect to time

$$\partial_A Au_t = \partial_A (u_{tt} - Bu_t - f_t).$$

In a symmetric way, $\partial_B B^2 u$ can also be calculated from the data.

Similar to space discretization for the first derivative, we have considered the standard symmetric second-order difference scheme. In such a way, in Section 5, we can interpret that we are considering as space X_h any which is determined by some nodal values (x_m, y_l) in a uniform grid in the square with $(N-1)\times(N-1)$ nodes and $h = 1/N$, P_h is just the projection onto the interior nodal values and Q_h the projection onto the nodal values of the boundary. We can consider as $\|\cdot\|_h$ the discrete L^2-norm and then, (H2) is immediately satisfied. Moreover, $A_{h,0}$ can be represented by a block-diagonal matrix whose base matrix for each $l \in \{1,\ldots,N-1\}$ is given by

$$\frac{1}{h^2}\mathrm{tridiag}(a(x_{m-\frac{1}{2}}, y_l), -(a(x_{m-\frac{1}{2}}, y_l) + a(x_{m+\frac{1}{2}}, y_l)), a(x_{m+\frac{1}{2}}, y_l)),$$

where $x_{m-\frac{1}{2}} = (x_{m-1} + x_m)/2$, and something similar happens for $B_{h,0}$. Since this matrix is irreducibly diagonally dominant, it is invertible and, using Gerschgorin Theorem and its symmetry, its eigenvalues are negative and therefore (H1) is satisfied. Besides, $A_h Q_{h,A} \partial_A u$ and $B_h Q_{h,B} \partial_B u$ can be represented by vectors with many zero values, except for the ones that come from the values of the boundary after applying the second-order difference scheme. More precisely, $A_h Q_{h,A} \partial_A u$ is a block vector, where each l-block has the form

$$\frac{1}{h^2}\begin{bmatrix} a(x_{\frac{1}{2}}, y_l)u(x_0, y_l) \\ 0 \\ \vdots \\ 0 \\ a(x_{N-\frac{1}{2}}, y_l)u(x_N, y_l) \end{bmatrix}.$$

Moreover, hypothesis (H3) is also satisfied with $Z = H^4(\Omega)$ and $\varepsilon_{h,A}$ and $\varepsilon_{h,B}$ of order $O(h^2)$ (Strikwerda, 1989). Notice that the technique to avoid order reduction in (6.3)–(6.4), (7.5)–(7.8) and (7.10)–(7.13) is cheap, because the additional cost of the method just requires the first and last column of each block of $\varphi_j(kA_{h,0})$ and $\varphi_j(kB_{h,0})$ $(j = 1, 2)$ and when k is fixed, those can be calculated once and for all at the very beginning.

8.2 *Numerical results*

Let us first use Lie–Trotter method to time integrate problem (8.1) with

$$a(x, y) = 1 + x + y, \quad b(x, y) = 1 + 2x + 3y. \tag{8.4}$$

Notice that, in this case, operators A and B do not commute.

For the first experiment we will consider

$$u_0(x, y) = (x^2 - 1/4)(y^2 - 1/4), \tag{8.5}$$

$$f(x, y, t) = \frac{e^{-t}}{16}(15 + 56y - 28y^2 - 32y^3 + x(32 - 64y^2) - 4x^2(7 + 48y + 4y^2) - 64x^3),$$

in which case, the exact solution is $u(x, y, t) = e^{-t}(x^2 - 1/4)(y^2 - 1/4)$. According to (8.2), hypotheses (A4) and (A4′) are satisfied. In Fig. 1, we can see the results after integrating (5.1) till time $T = 1$ directly with (2.11) considering the last three terms of (5.1) as a source term. More precisely, we have considered in (2.11)

$$\begin{aligned} F &= A_h Q_h \partial_A g(t) + B_h Q_h \partial_B g(t) + P_h f(t), \\ M_1 &= A_{h,0}, \\ M_2 &= B_{h,0}, \end{aligned}$$

with $h = 10^{-2}$. Moreover, we can also observe the results after applying formulas (6.3)–(6.5) to avoid order reduction. In the first case, we can observe that the results are very poor, whereas orders 2 and 1 are observed for the local and global errors, respectively, when applying the technique that is suggested

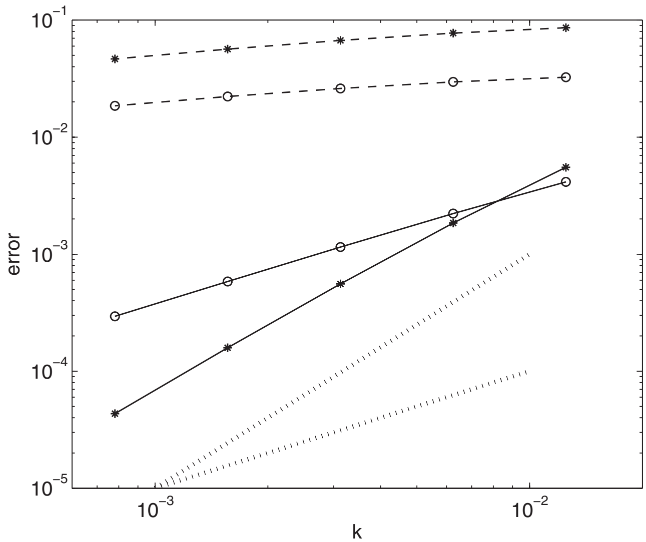

FIG. 1. Local error (*) and global error (o) without avoiding (discontinuous) and avoiding (continuous) order reduction when integrating problem (8.1) with Lie–Trotter method with noncommutable operators A and B given through (8.4) and data (8.5). Dotted lines represent the slope for orders 1 and 2.

in this article. This corroborates Theorems 6.3 and 6.4 when h^2 is negligible against k. Moreover, we see that, not only the order increases, but also the size of the errors considerably diminishes. We also notice that the same results are obtained with the suggested technique when h diminishes, so that no Courant-Friedrichs-Lewy (CFL) condition is required.

Let us now consider, as a second experiment,

$$u_0(x, y) = x(1 - x)y(1 - y) \tag{8.6}$$

$$f(x, y, t) = e^{-t}(-4x^3 + y + y^2 - 2y^3 + x(-1 + 15y - 3y^2) - x^2(-5 + 11y + y^2)).$$

For such a problem, the exact solution is $u(x, y, t) = e^{-t}x(1 - x)y(1 - y)$, which has homogeneous boundary values. Therefore, the only correction needed is due to the inhomogeneity f which is not zero on the boundary. In any case, (A4) and (A4′) are again satisfied.

When integrating directly (5.1) with Lie–Trotter, orders around 1.25 and 1 are observed for local and global errors, respectively, as stated in Faou *et al.* (2015) for vanishing boundary conditions of the exact solution. When applying (6.3)–(6.5), the results fit quite well with the orders 1 and 2, which are assured by Theorems 6.3 and 6.4, respectively. In this case, although there is not a big gain in the order of the global error, the size of the errors is about three times smaller for a same value of k, as it can be observed in Fig. 2. As the computational cost of the technique to avoid order reduction is negligible against the rest of the calculations of the method, the strategy that is suggested here is clearly better.

Let us now use Strang method for the same problem with data (8.5). As the operators A and B do not commute, we have to use formulas (7.10)–(7.13). The order 2 of the local error which is given by Theorem 7.2 is clearly seen in Fig. 3, and it seems that a summation-by-parts argument similar to that

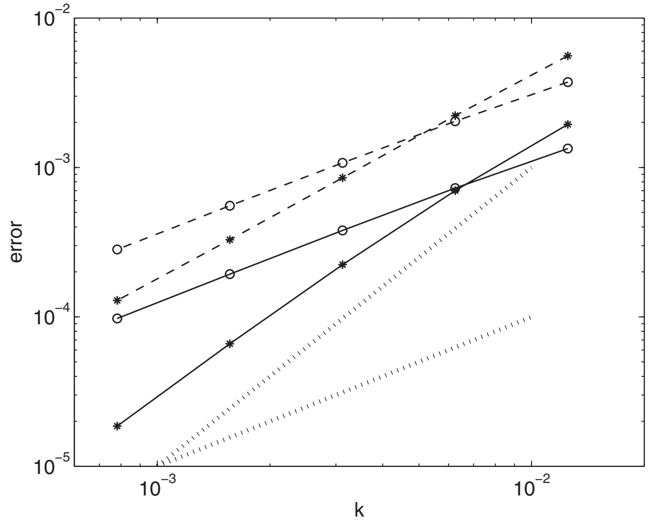

FIG. 2. Local error (*) and global error (o) without avoiding (discontinuos) and avoiding (continuous) order reduction when integrating problem (8.1) with Lie–Trotter method with noncommutable operators A and B given through (8.4) and data (8.6). Dotted lines represent the slope for orders 1 and 2.

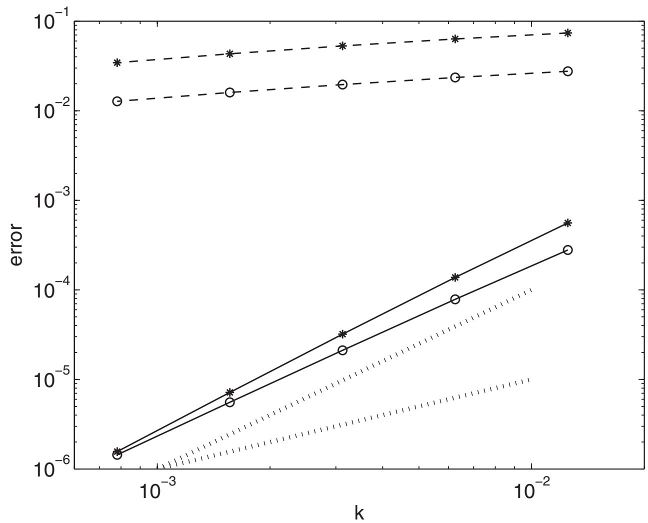

FIG. 3. Local error (*) and global error (o) without avoiding (discontinuous) and avoiding (continuous) order reduction when integrating problem (8.1) with Strang method with noncommutable operators A and B given through (8.4) and data (8.5). Dotted lines represent the slope for orders 1 and 2.

shown in Faou *et al.* (2015) for vanishing boundary problems is also working here since the global error in fact behaves as $O(k^2)$ instead of $O(k)$, as assured by Theorem 7.5. Moreover, the difference between avoiding and not avoiding order reduction is seen to be even higher than with Lie–Trotter for the same problem.

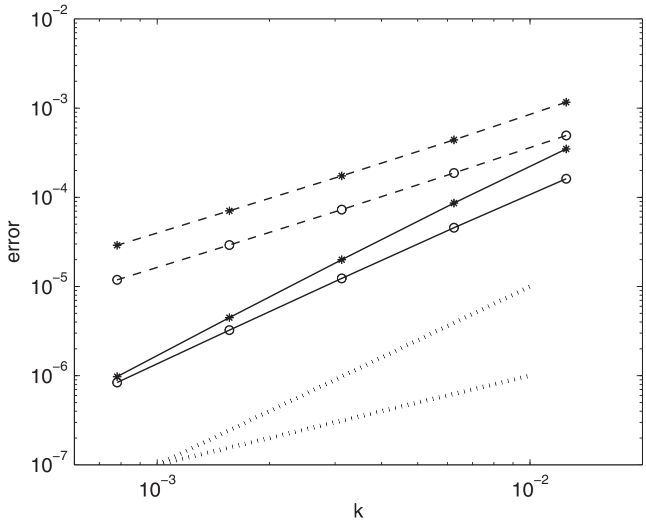

FIG. 4. Local error (*) and global error (o) without avoiding (discontinuous) and avoiding (continuous) order reduction when integrating problem (8.1) with Strang method with noncommutable operators A and B given through (8.4) and data (8.6). Dotted lines represent the slope for orders 1 and 2.

Let us now use Strang method for solving the same problem, but with data (8.6). When not avoiding order reduction, local and global orders are around 1.25, as stated in Faou *et al.* (2015). With the technique suggested here, we achieve order near 2 for both the local and global error, as in the example before. Figure 4 also shows that the size of the errors is much smaller with our technique at a very low additional cost. We would also like to remark that, using more complicated functions $a(x, y)$, $b(x, y)$ in (8.1), we have numerically checked that order reduction is also avoided.

Finally, let us consider problem (8.1) with the Laplacian operator. That is,

$$a(x, y) = b(x, y) = 1.$$

Clearly in this case the operators A and B commute, and therefore technique (7.5)–(7.8) can be applied when integrating in time with Strang method. For this experiment, we will use

$$\begin{aligned} u_0(x, y) &= (x^2 - 1/4)(y^2 - 1/4), \\ f(x, y, t) &= \frac{e^{-t}}{16}(15 - 28y^2 - 4x^2(7 + 4y^2)). \end{aligned} \tag{8.7}$$

Now the exact solution of the problem is $u(x, y) = (x^2 - 1/4)(y^2 - 1/4)\, e^{-t}$, which again satisfies regularity hypotheses (A4) and (A4′), although not vanishing at the boundary. With our technique, the local order is clearly 3 and the global one is 2, as stated by Theorems 7.1 and 7.3. However, as the solution does not vanish on the boundary, the results without avoiding order reduction are very poor. Figure 5 confirms that even in terms of the size of the errors.

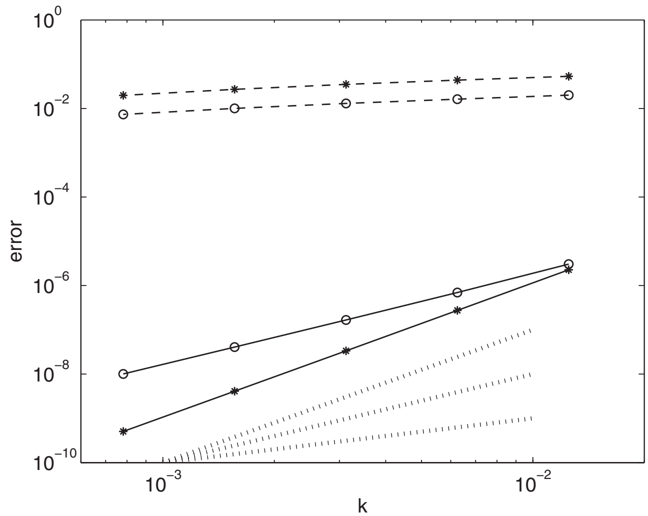

FIG. 5. Local error (*) and global error (o) without avoiding (discontinuous) and avoiding (continuous) order reduction when integrating problem (8.1) with Strang method with commutable operators *A* and *B* corresponding to the Laplacian and data (8.7). Dotted lines represent the slope for orders 1, 2 and 3.

Funding

Ministerio de Ciencia e Innovación project [MTM2015-66837-P].

REFERENCES

ALONSO-MALLO, I. (2002) Runge–Kutta methods without order reduction for linear initial boundary value problems. *Numer. Math.*, **91**, 577–603.

ALONSO-MALLO, I., CANO, B. & JORGE, J. C. (2004) Spectral-fractional step Runge–Kutta discretizations for initial boundary value problems with time-dependent boundary conditions. *Math. Comp.*, **73**, 1801–1825.

ALONSO-MALLO, I., CANO, B. & REGUERA, N. (2016) Avoiding order reduction when integrating linear initial boundary value problems with Lawson methods. *IMA J. Numer. Anal.*, 1–29, doi:10.1093/imanum/drw052.

ALONSO-MALLO, I., CANO, B. & REGUERA, N. (2017) Analysis of order reduction when integrating linear initial boundary value problems with Lawson methods. *Appl. Numer. Math.*, **118**, 64–74.

ALONSO-MALLO, I., CANO, B. & REGUERA, N. (2017) Avoiding order reduction when integrating reaction-diffusion boundary value problems with exponential splitting methods. arXiv:1705.01857.

BRENNER, P., CROUZEIX, M. & THOMÉE, V. (1982) Single step methods for non-homogeneous linear differential equations in Banach spaces. *RAIRO Anal. Numer*., **16**, 5–26.

CANO, B. & REGUERA, N. (2017) Avoiding order reduction when integrating nonlinear Schrödinger equation with Strang method. *J. Comput. Appl. Math.*, **316**, 86–99.

CONNORS, J., BANKS, J. W., HITTINGER, J. A. & WOODWARD, C. S. (2014) Quantification of errors for operator-split advection-diffusion calculations. *Comput. Methods Appl. Mech. Engrg.*, **272**, 181–197.

EINKEMMER, L. & OSTERMANN, A. (2015) Overcoming order reduction in diffusion-reaction splitting. Part 1: Dirichlet boundary conditions. *SIAM J. Sci. Comput.*, **37**, A1577–A1592.

EINKEMMER, L. & OSTERMANN, A. (2016a) Overcoming order reduction in diffusion-reaction splitting. Part 2: Oblique boundary conditions. *SIAM J. Sci. Comput.*, **38**, A3741–A3757.

Einkemmer, L. & Ostermannm, A. (2016b) A comparison of boundary correction methods for Strang splitting. arXiv:1609.05505.
Faou, E., Ostermann, A. & Schratz, K. (2015) Analysis of exponential splitting methods for inhomogeneous parabolic equations. *IMA J. Numer. Anal.*, **35**, 161–178.
Gockler, T. & Grimm, V. (2013) Convergence analysis of an extended Krylov subspace method for the approximation of operator functions in exponential integrators. *SIAM J. Numer. Anal.*, **51**, 2189–2213.
Hochbruck, M. & Ostermann, A. (2010) Exponential integrators. *Acta Numer.*, **19**, 209–286.
Hundsdorfer, W. & Verwer, J. (2003) *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*. Springer Series in Computational Mathematics, vol. 33. Berlin: Springer.
LeVeque, R. J. (1986) Intermediate boundary conditions for time-split methods applied to hyperbolic partial differential equations. *Math. Comp.*, **47**, 37–54.
LeVeque, R. J. & Oliger, J. (1983) Numerical methods based on additive splittings for hyperbolic partial differential equations. *Math. Comp.*, **40**, 469–497.
Ostermann, A. & Schratz, K. (2013) *Stability of exponential operator splitting methods for non-contractive semigroups*. *SIAM J. Numer. Anal.*, **51**, 191–203.
Palencia, C. & Alonso-Mallo, I. (1994) *Abstract initial-boundary value problems*. *Proc. Roy. Soc. Edinburgh Sect. A.*, **124**, 879–908.
Pazy, A. (1983) *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Applied Mathematical Sciences, vol. 44. New York: Springer.
Strikwerda, J. C. (1989) *Finite Difference Schemes and Partial Differential Equations*. USA: Wadsworth & Brooks.
Yanenko, N. (1971) *The Method of Fractional Steps.* Heidelberg, Berlin, New York: Springer.