Order reduction in operator splitting methods*

Roman Kozlov[†] Brynjulf Owren[‡]

July 13, 1999

Abstract

Operator splitting is used in many applications. By decomposing the operator into a sum, each term can be integrated separately, either exactly or by an efficient numerical method, and an approximation to the original problem is obtained by composition. When the splitting represents the stiff and nonstiff parts of ODE, it is known that one may observe order reduction similar to the B-convergence theory for Runge-Kutta methods. This phenomenon can not be explained by the classical error analysis. Based on the Taylor series expansions of the exact solution and the splitting method solution the classical error analysis fails for large time steps.

In this paper we propose a framework to analyze the order reduction of the splitting methods. Several types of splitting methods are examined for linear ODE systems and linear PDEs (which can be viewed as ODEs formulated in abstract spaces). The results are supported by numerical experiments.

1 Introduction

Operator splitting methods are well known in the field of numerical solutions of ordinary and partial differential equations. For PDEs there are generally two reasons to use the operator splitting: The first one arises from multidimensional PDEs where it may be useful to treat each dimension separately. This technique is called dimensional splitting. The second motivation is to split the differential operator into several parts according to different physical phenomena, such as for instance convection and diffusion. For ODEs the splitting techniques can be used, for example, to separate highly oscillatory or stiff components of the solution.

For both ordinary and partial differential equations, a numerical method is obtained by composing approximations of each of the split problems. Because the parts of the original operator are treated independently, a splitting technique may give rise to very efficient numerical methods. However, owing to the noncommutativity of the operators in the splitting, an error caused by the uncoupling is introduced. The classical error analysis uses expansions of the exponential operators. Such results can be found in [1], [2] for the linear case and in [3], [4] for the nonlinear case, where it is necessary to consider Lie derivatives.

Generally, the order of the splitting method is bounded by the lowest order of a solver used in the splitting. Thus it is of great interest to increase the order of a splitting. Strang [1] proposed symmetric splitting schemes of the second order which are based on first order solvers. Further development was done by Yoshida [5]. He presented an approach to raise the order from 2 to 4 by composing three Strang splittings with different time steps.

It is known that in case of stiff problems the splitting methods may suffer an order reduction for sufficiently large time steps Δt. Let us illustrate this phenomenon on the example equation

$$y' = (A+B)y, \qquad t>0, \qquad y(0) = y_0. \tag{1.1}$$

*This work was in part sponsored by The Norwegian Research Council under contract no. 111038/410, through the SYNODE project. WWW: `http://www.math.ntnu.no/num/synode/`

[†]Email: `Roman.Kozlov@math.ntnu.no`, WWW: `http://www.math.ntnu.no/~kozlov/`

[‡]Email: `Brynjulf.Owren@math.ntnu.no`, WWW: `http://www.math.ntnu.no/~bryn/`

with

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \qquad B = \frac{1}{\varepsilon} \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}, \qquad \varepsilon = 10^{-5} \tag{1.2}$$

The value ε can be viewed as a stiffness parameter. The solution of this problem

$$y(t) = \exp(t(A+B))y_0$$

can be approximated by the splitting method

$$y(\Delta t) = \exp(\Delta t B/2)\exp(\Delta t A)\exp(\Delta t B/2)y_0$$

We can take the Frobenius norm

$$|| \exp(\Delta t(A+B)) - \exp(\Delta t B/2)\exp(\Delta t A)\exp(\Delta t B/2)||_F$$

as a measure of the error due to uncoupling of the operator. The results of the computation are shown in Figure 1.

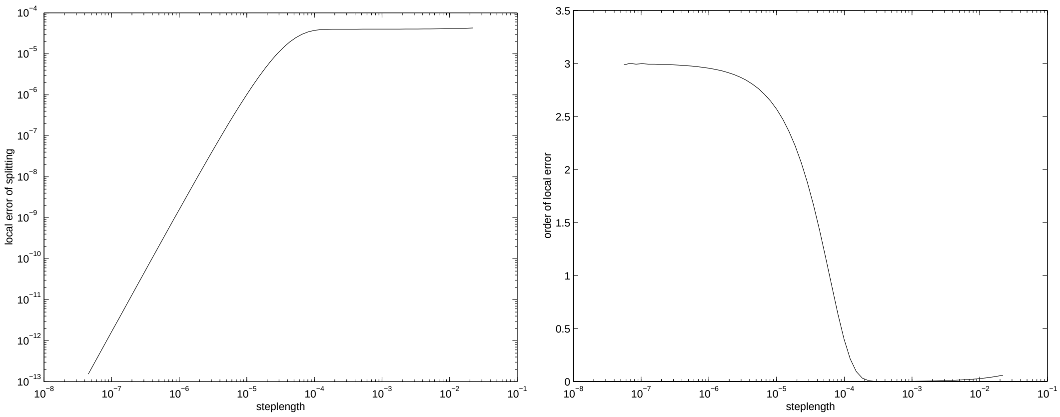

Figure 1: The local error (left plot) and the order of the local error (right plot) of the problem (1.1), (1.2).

We see that the order drops from 3 to 0 at $\Delta t \approx 10^{-4}$. The classical error analysis fails to explain this because the expansions of the exact solution and the splitting method solution it uses are not valid for sufficiently large time steps.

A new approach for studying the order using the decoupling on slow and fast components of the solutions was recently developed in [6], [7] and [8]. The core of this approach is a reduction of stiff problems to differential algebraic systems. The error estimates are provided by the reduced differential algebraic systems, which already contain approximation errors (see [6] for details and references).

In this paper we intend to study splitting errors by comparing the exact solution and the splitting methods solutions. Thus we avoid introducing any approximation error in our approach. From the entire error expression we extract the leading error term and analyze the order reduction of some splitting methods. The technique we use is based on spectral properties of the operators. In Section 2 we consider splitting methods for ODEs whereas Section 3 is devoted to splitting methods for PDEs.

The paper consists of two main sections devoted to applications of splitting methods to ODE and PDE problems respectively.

2 The ODE case

In this section we consider linear ODE systems in $\mathbb{R}^n$ of the form

$$y' = (A+B)y, \qquad t>0, \qquad y(0)=y_0. \tag{2.1}$$

In case of constant matrices A and B this problem has the solution

$$y(t) = \exp(t(A+B))y_0. \tag{2.2}$$

Thus $\exp(t(A+B))$ can be viewed as the solution operator for the equation (2.1). In what follows we will define splitting methods in terms of their solution operators, and analyze errors in terms of the norm of the corresponding error operators defined below.

2.1 The splitting analysis method

We will consider first order splitting methods (in the classical sense)

$$\mathbf{AB}: \quad \exp(\Delta tA)\exp(\Delta tB), \qquad \mathbf{BA}: \quad \exp(\Delta tB)\exp(\Delta tA), \tag{2.3}$$

and the second order splitting methods (introduced by Strang [1])

$$\begin{aligned} \mathbf{BAB} \quad &: \quad \exp(\Delta tB/2)\exp(\Delta tA)\exp(\Delta tB/2), \\ \mathbf{ABA} \quad &: \quad \exp(\Delta tA/2)\exp(\Delta tB)\exp(\Delta tA/2), \\ \mathbf{AB}+\mathbf{BA} \quad &: \quad \tfrac{1}{2}(\exp(\Delta tA)\exp(\Delta tB)+\exp(\Delta tB)\exp(\Delta tA)), \end{aligned} \tag{2.4}$$

which approximate $\exp(\Delta t(A+B))$. The idea behind these expressions for the splittings is to apply successively different parts of the original operator. For example, the splitting **AB** gives the solution of the system:

$$\begin{cases} \dfrac{dy^*}{dt} = By^*, \qquad y^*(0)=y_0 \quad \text{on} \quad [0,\Delta t], \\ \dfrac{dy^{**}}{dt} = Ay^{**}, \qquad y^{**}(0)=y^*(\Delta t) \quad \text{on} \quad [0,\Delta t]. \end{cases}$$

If A and B commute, all five types of splitting give solutions equal to the exact solution (2.2). For noncommuting A and B the splitting methods have errors due to uncoupling of the operator. Because the splittings are of interest for numerical computation of the solution of the problem (2.1) we will investigate their errors.

The order analysis of the various splitting techniques above are quite similar to each other. For this reason, we have chosen to give a detailed treatment of the particular case **BAB**, and merely list the main results for the other cases.

Thus we investigate the error

$$E = \exp(\Delta t(A+B)) - \exp(\Delta tB/2)\exp(\Delta tA)\exp(\Delta tB/2). \tag{2.5}$$

If we assume that $\|\Delta tA\| \ll 1$ and $\|\Delta tB\| \ll 1$, then we can expand exponentials by Taylor series and we find that $E = O(\Delta t^3)$, the classical local error of the splitting. This means that the splitting has second order accuracy.

We now consider constant coefficient linear ODE systems (2.1) with stiff part B and nonstiff part A. Stiffness implies in our setting that for the range of considered step sizes Δt the matrix ΔtB has "large" norm. Thus we have

$$\|\Delta tB\| \gg 1 \qquad \text{and} \qquad \|\Delta tA\| \ll 1. \tag{2.6}$$

For such step sizes Δt we observe order reductions for the splitting methods. For these values of Δt we can express $\exp(\Delta tA)$ by its Taylor series: $\exp(\Delta tA) \approx I + \Delta tA + \Delta t^2 A^2/2! + O(\Delta t^3)$. The same assumption is not valid for $\exp(\Delta tB)$ and we have to employ more delicate comparison of the exact solution operator and the approximate solution given by the splitting. For the further analysis we will make the following assumptions

Assumption 2.1

1. The matrix B can be written in the form

$$B = B_0 + \frac{B_1}{\varepsilon}, \quad \|B_0\| \le C_0, \quad \|B_1\| \le C_1, \ \varepsilon > 0, \tag{2.7}$$

   by means of a *stiffness parameter*. Typically $\varepsilon \ll 1$.

2. The matrix B_1 has eigenvalues with nonpositive real parts (This implies that for $\Delta t \in [0, T]$, $\exp(\Delta tB)$ is uniformly bounded with respect to $\varepsilon > 0$).

3. B is diagonalizable with a well–conditioned eigensystem

$$B = S\Lambda S^{-1}, \qquad \Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n), \tag{2.8}$$

   $\mathrm{Re}\,\lambda_i \le a, \ i = 1, \dots, n, \ \|S\| \le C$ and $\|S^{-1}\| \le C$ for some constant C. In the general case B has both stiff eigenvalues $\lambda_{st} = O\left(\frac{1}{\varepsilon}\right)$ and nonstiff eigenvalues $\lambda_{nst} = O(1)$.

4. We consider the order reduction in the range of time steps $|\lambda_{nst}|\Delta t \ll 1$ and $|\lambda_{st}|\Delta t \gg 1$, where $\exp(\lambda_{st}\Delta t)$ is negligible in size compared to $\exp(\lambda_{nst}\Delta t)$ and can thus be ignored in the analysis.

◊

Note that highly oscillatory problems also fit this framework with the exception of the assumption 4.

For the exact solution operator, we now compute [1]

$$\exp\left(\Delta t(A+B)\right) = \sum_{k\ge 0} \frac{\Delta t^k}{k!}(A+B)^k = \sum_{k\ge 0} \frac{1}{k!}(\Delta tB)^k$$
$$+\Delta t \sum_{n,m\ge 0} \frac{(\Delta tB)^n A (\Delta tB)^m}{(n+m+1)!} + \Delta t^2 \sum_{n,m,k\ge 0} \frac{(\Delta tB)^n A (\Delta tB)^m A (\Delta tB)^k}{(n+m+k+2)!} + O(\Delta t^3). \tag{2.9}$$

Our analysis is based on the comparison of the expansion of the exact solution (2.9) with the expansion of the splitting method. For the **BAB** case we have the following expansion

$$\exp\left(\frac{\Delta t}{2}B\right)\exp\left(\Delta tA\right)\exp\left(\frac{\Delta t}{2}B\right)$$
$$= \exp\left(\frac{\Delta t}{2}B\right)\left(I + \Delta tA + \frac{\Delta t^2}{2}A^2 + O(\Delta t^3)\right)\exp\left(\frac{\Delta t}{2}B\right) \tag{2.10}$$
$$= \exp\left(\Delta tB\right) + \Delta t \exp\left(\frac{\Delta t}{2}B\right) A \exp\left(\frac{\Delta t}{2}B\right) + \frac{\Delta t^2}{2!}\exp\left(\frac{\Delta t}{2}B\right) A^2 \exp\left(\frac{\Delta t}{2}B\right) + O(\Delta t^3).$$

[1] Although denoted as $O(\Delta t^3)$ the higher order terms are subjected to order reduction in the case of stiff problems. We let ourselves to use the classical order because these terms are negligible with respect to the previous terms.

By subtracting (2.10) from (2.9) we obtain

$$E = E_1 + E_2 + \cdots, \tag{2.11}$$

where

$$E_1 = \Delta t \left(\sum_{n,m\geq 0} \frac{(\Delta t B)^n A (\Delta t B)^m}{(n+m+1)!} - \exp\left(\frac{\Delta t}{2}B\right) A \exp\left(\frac{\Delta t}{2}B\right) \right), \tag{2.12}$$

and

$$E_2 = \Delta t^2 \left(\sum_{n,m,k\geq 0} \frac{(\Delta t B)^n A (\Delta t B)^m A (\Delta t B)^k}{(n+m+k+2)!} - \exp\left(\frac{\Delta t}{2}B\right) \frac{A^2}{2!} \exp\left(\frac{\Delta t}{2}B\right) \right). \tag{2.13}$$

By analyzing the errors E_1 and E_2, we will show that E_1 represents the dominant error term in the general case, i.e. the effect of E_2 can be ignored. This will allow us to estimate the local error of the problem (2.1) $e(t)$ given as

$$e(t) = Ey_0 \approx e^1 = E_1 y_0.$$

It may of course happen that $y_0 \in \mathrm{Ker}(E_1)$ such that e^1 is vanishing. If so, the error of the splitting can be estimated as the next term $e^2 = E_2 y_0$.

Before we proceed to analyze the leading error term E_1 we simplify the equation (2.1). Using the matrix S whose columns are the eigenvectors of B, and the corresponding diagonal eigenvalue matrix Λ, we can transform the equation (2.1) to an equivalent but simpler form, introducing a new dependent variable $x = S^{-1}y$:

$$x'(t) = (\hat{A} + \Lambda)x, \qquad t > 0, \qquad x(0) = x_0 = S^{-1}y_0, \tag{2.14}$$

where $\hat{A} = S^{-1}AS$. The error E and the error terms E_1 and E_2 of a splitting method, applied to the original problem (2.1), and the errors $\hat{E}$, $\hat{E}_1$ and $\hat{E}_2$ of the same splitting method, applied to the transformed problem (2.14), are connected as

$$E = S\hat{E}S^{-1}, \qquad E_1 = S\hat{E}_1 S^{-1} \qquad E_2 = S\hat{E}_2 S^{-1}. \tag{2.15}$$

Below we will consider the error terms for the transformed problem (2.14). Relations (2.15) will be used to transform the results back to the original problem (2.1). The errors are connected by the "weight" matrix S. In the classical case, when $E = O(\Delta t^3)$ and $\hat{E} = O(\Delta t^3)$, this transformation does not affect the order because all terms of E and $\hat{E}$ have the same order $O(\Delta t^3)$. It is not the same in the stiff case when different elements of E and $\hat{E}$ can have different order behaviors.

2.2 The first error term

Let us examine the leading error term E_1. For the transformed problem (2.14) the error $\hat{E}_1$ has a simpler form due to the diagonal form of the matrix Λ:

$$\hat{E}_1 = \Delta t \left(\sum_{n,m\geq 0} \frac{(\Delta t \Lambda)^n \hat{A} (\Delta t \Lambda)^m}{(n+m+1)!} - \exp\left(\frac{\Delta t}{2}\Lambda\right) \hat{A} \exp\left(\frac{\Delta t}{2}\Lambda\right) \right).$$

We can find the elements of the matrix $\hat{E}_1$ as follows

$$(\hat{E}_1)_{k,l} = \Delta t \left(\sum_{n,m\geq 0} \frac{(\Delta t \Lambda)^n \hat{A} (\Delta t \Lambda)^m}{(n+m+1)!} - \exp\left(\frac{\Delta t}{2}\Lambda\right) \hat{A} \exp\left(\frac{\Delta t}{2}\Lambda\right) \right)_{k,l}$$

$$= \Delta t \left(\sum_{n,m\geq 0} \frac{(\Delta t\lambda_k)^n \hat{A}_{k,l} (\Delta t\lambda_l)^m}{(n+m+1)!} - \exp\left(\frac{\Delta t}{2}\lambda_k\right) \hat{A}_{k,l} \exp\left(\frac{\Delta t}{2}\lambda_l\right) \right)$$

$$= \Delta t \hat{A}_{k,l} \left(\sum_{n,m\geq 0} \frac{(\Delta t\lambda_k)^n (\Delta t\lambda_l)^m}{(n+m+1)!} - \exp\left(\Delta t \frac{\lambda_k + \lambda_l}{2}\right) \right),$$

and obtain the error expression

$$(\hat{E}_1)_{k,l} = \Delta t \beta_{k,l} \hat{A}_{k,l}, \tag{2.16}$$

with

$$\beta_{k,l} = \sum_{n,m\geq 0} \frac{(\Delta t\lambda_k)^n (\Delta t\lambda_l)^m}{(n+m+1)!} - \exp\left(\Delta t \frac{\lambda_k + \lambda_l}{2}\right). \tag{2.17}$$

The following lemma is needed to estimate the different values of $\beta_{k,l}$.

Lemma 2.1 *For $a,\ b \in \mathbb{C}$ it is true that*

$$\sum_{n,m\geq 0} \frac{a^n b^m}{(n+m+1)!} = \begin{cases} \dfrac{\exp(b) - \exp(a)}{b-a}, & \text{if} \quad a \neq b, \\ \exp(b), \quad \text{if} \quad a = b. \end{cases}$$

Proof: We first examine the case $a \neq b$:

$$(b-a) \sum_{n,m\geq 0} \frac{a^n b^m}{(n+m+1)!} = \sum_{n,m\geq 0} \frac{a^n b^{m+1}}{(n+m+1)!} - \sum_{n,m\geq 0} \frac{a^{n+1} b^m}{(n+m+1)!}$$

$$= \left(\sum_{n,m\geq 0} \frac{a^n b^m}{(n+m)!} - \sum_{n\geq 0} \frac{a^n}{n!} \right) - \left(\sum_{n,m\geq 0} \frac{a^n b^m}{(n+m)!} - \sum_{m\geq 0} \frac{b^m}{m!} \right)$$

$$= -\sum_{n\geq 0} \frac{a^n}{n!} + \sum_{m\geq 0} \frac{b^m}{n!} = \exp(b) - \exp(a),$$

from which we obtain the first part of the lemma.
If $a = b$, then

$$\sum_{n,m\geq 0} \frac{a^n b^m}{(n+m+1)!} = \sum_{n,m\geq 0} \frac{b^{n+m}}{(n+m+1)!} = \sum_{k\geq 0} (k+1) \frac{b^k}{(k+1)!} = \exp(b).$$

□

Using Lemma 2.1, we find expressions for $\beta_{k,l}$:

$$\beta_{k,l} = \begin{cases} \dfrac{\exp(\Delta t\lambda_k) - \exp(\Delta t\lambda_l)}{\Delta t\lambda_k - \Delta t\lambda_l} - \exp\left(\Delta t \dfrac{\lambda_k + \lambda_l}{2}\right) & \text{if} \quad \lambda_k \neq \lambda_l, \\ 0, \quad \text{if} \quad \lambda_k = \lambda_l. \end{cases} \tag{2.18}$$

Now we are in a position to consider some example values of $\beta_{k,l}$. We will list several cases of eigenvalues pairs λ_k and λ_l and give the estimation of $\beta_{k,l}$:

I. For $\lambda_k = \lambda_l$ we get $\beta_{k,l} = 0$.

II. Assume $\lambda_k \neq \lambda_l$:

(a) If λ_k and λ_l are nonstiff, then $\beta_{k,l} = O(\Delta t^2)$ that corresponds to the classical order of the considered splitting.

(b) If λ_k is stiff and λ_l is nonstiff, then

$$\beta_{k,l} \approx -\frac{\exp(\Delta t \lambda_l)}{\Delta t \lambda_k} \approx -\frac{1}{\Delta t \lambda_k} = O\left(\frac{\varepsilon}{\Delta t}\right).$$

(c) For two stiff eigenvalues λ_k and λ_l $\beta_{k,l} \approx 0$ is negligible.

(d) If λ_k and λ_l are imaginary and complex conjugate, then

$$\beta_{k,l} = \frac{\sin(\Delta t \lambda_k)}{\Delta t \lambda_k} - 1 \approx -1 \quad \text{for} \quad \Delta t |\lambda_k| \gg 1.$$

Thus we see that we can expect different types of the order reduction corresponding to different eigenvalue pairs. The lowest order which $\beta_{k,l}$ can have is (-1); it is given by the pair of one stiff and one nonstiff eigenvalues (case II.b). This $\beta_{k,l}$ contributes to the splitting error as a zero order term: $\Delta t \cdot O(\frac{\varepsilon}{\Delta t}) = O(\varepsilon)$.

Remark. The results we have found so far are valid for the case $y(0) \notin \mathrm{Ker}(E_1)$. The error of the solution of the initial value problem (2.1) is of higher order if $y(0) \in \mathrm{Ker}(E_1)$. It is also possible that the splitting has no error for some initial data. It is easy to see that if A and B share a common eigenvector η, then $\eta \in \mathrm{Ker}(E_1)$. Moreover, any linear combination of common eigenvectors belongs to $\mathrm{Ker}(E_1)$. We note that for such initial data each splitting (2.3)–(2.4) is exact, i.e. the uncoupling does not introduce any error whatsoever.

2.3 The second error term

We proceed to consider the next order error term E_2. To justify our approach we should show that E_2 is much smaller than E_1. As in the previous point we transform the error operator E_2 to a simpler form $\hat{E}_2$, which corresponds to the second error term of the problem (2.14):

$$\hat{E}_2 = \Delta t^2 \left(\sum_{n,m,k \geq 0} \frac{(\Delta t \Lambda)^n \, \hat{A} \, (\Delta t \Lambda)^m \, \hat{A} \, (\Delta t \Lambda)^k}{(n+m+k+2)!} - \exp\left(\frac{\Delta t}{2}\Lambda\right) \frac{\hat{A}^2}{2!} \exp\left(\frac{\Delta t}{2}\Lambda\right) \right). \tag{2.19}$$

We compute the elements of the matrix $\hat{E}_2$. Using

$$\begin{aligned}
&\left(\sum_{n,m,k \geq 0} \frac{(\Delta t \Lambda)^n \, \hat{A} \, (\Delta t \Lambda)^m \, \hat{A} \, (\Delta t \Lambda)^k}{(n+m+k+2)!} \right)_{i,j} \\
&= \sum_{n,m,k \geq 0} \frac{1}{(n+m+k+2)!} \sum_{1 \leq p,q,c,d \leq n} ((\Delta t \Lambda)^n)_{i,p} \hat{A}_{p,q} ((\Delta t \Lambda)^m)_{q,r} \hat{A}_{r,t} ((\Delta t \Lambda)^k)_{t,j} \\
&= \sum_{n,m,k \geq 0} \frac{1}{(n+m+k+2)!} \sum_{1 \leq q \leq n} ((\Delta t \Lambda)^n)_{i,i} \hat{A}_{i,q} ((\Delta t \Lambda)^m)_{q,q} \hat{A}_{q,j} ((\Delta t \Lambda)^k)_{j,j} \\
&= \sum_{1 \leq q \leq n} \hat{A}_{i,q} \hat{A}_{q,j} \sum_{n,m,k \geq 0} \frac{(\Delta t \lambda_i)^n (\Delta t \lambda_q)^m (\Delta t \lambda_j)^k}{(n+m+k+2)!},
\end{aligned}$$

and

$$\left(\exp\left(\frac{\Delta t}{2}\Lambda\right) \hat{A}^2 \, \exp\left(\frac{\Delta t}{2}\Lambda\right) \right)_{i,j} = \exp\left(\frac{\Delta t}{2}\lambda_i\right) (\hat{A}^2)_{i,j} \exp\left(\frac{\Delta t}{2}\lambda_j\right),$$

we obtain the error term $\hat{E}_2$ in the form

$$(\hat{E}_2)_{i,j} = \Delta t^2 \sum_{1\le q\le n} \gamma_{i,q,j} \hat{A}_{i,q} \hat{A}_{q,j}, \tag{2.20}$$

where

$$\gamma_{i,q,j} = \sum_{n,m,k\ge 0} \frac{(\Delta t\lambda_i)^n (\Delta t\lambda_q)^m (\Delta t\lambda_j)^k}{(n+m+k+2)!} - \frac{1}{2}\exp\left(\Delta t \frac{\lambda_i+\lambda_j}{2}\right). \tag{2.21}$$

To estimate the different γ–values we need the following result whose proof is similar to that of Lemma 2.1.

Lemma 2.2 *For $a,\ b,\ c \in \mathbb{C}$ it is true that*

$$\sum_{n,m,k\ge 0} \frac{a^n b^m c^k}{(n+m+k+2)!}$$

$$= \begin{cases} \dfrac{(c-b)\exp(a) + (a-c)\exp(b) + (b-a)\exp(c)}{(a-b)(b-c)(c-a)}, & \text{if } a\neq b,\ a\neq c,\ b\neq c, \\[2ex] \dfrac{\exp(c)-\exp(a)}{(a-c)^2} + \dfrac{\exp(a)}{(a-c)}, \quad \text{if} \quad a = b \neq c, & \\[2ex] \exp(a)/2, \quad \text{if} \quad a=b=c. & \end{cases}$$

Now the behavior of $\gamma_{i,q,j}$ and thereby E_2 in terms of $\lambda_i, \lambda_q, \lambda_j$ can be examined by means of the Lemma 2.2 and the formula (2.21). We summarize as follows. It is easy to see that if all eigenvalues are equal $\lambda_i = \lambda_q = \lambda_j$, then $\gamma_{i,q,j} = 0$. For three nonstiff eigenvalues $\gamma_{i,q,j} = O(\Delta t)$ that gives the classical error $O(\Delta t^3)$. If all eigenvalues are stiff, then $\gamma_{i,q,j}$ has a negligible value. Thus we are left with the combinations $\{\lambda_i, \lambda_q, \lambda_j\}$, including both stiff and nonstiff eigenvalues. Considering different combinations, we obtain the possible values of $\gamma_{i,q,j}$. They are composed by terms $O(1)$, $O\left(\frac{\varepsilon}{\Delta t}\right)$ and $O\left(\frac{\varepsilon^2}{\Delta t^2}\right)$, contributing to the error E_2 as $O(\Delta t^2)$, $O(\varepsilon\Delta t)$ and $O(\varepsilon^2)$ terms respectively.

It follows that for the stiff problems (2.1) the splitting **BAB** has the following error terms in the order reduction case ($\Delta t \gg \varepsilon$):

$$E_1 = O(\Delta t, \varepsilon) \qquad \text{and} \qquad E_2 = O(\Delta t^2, \Delta t\varepsilon, \varepsilon^2). \tag{2.22}$$

This result assures that $E_2 = o(E_1)$ and that the order reduction is defined by the leading error term E_1.

2.4 The other splitting cases

In this section we study the rest of the splitting types mentioned in (2.3) and (2.4). The approach is similar to the analysis of the splitting **BAB**, it differs only in the expressions for the values β and γ and we shall therefore only present the final results. We first note that for all the cases, $\beta_{k,l} = 0$ if $\lambda_k = \lambda_l$.

1. For the splitting **AB** we get the expansion

$$\begin{aligned} \exp(\Delta tA)\exp(\Delta tB) &= \left(I + \Delta tA + \frac{1}{2}\Delta t^2 A^2 + O(\Delta t^3)\right)\exp(\Delta tB) \\ &= \exp(\Delta tB) + \Delta tA\exp(\Delta tB) + \frac{1}{2}\Delta t^2 A^2 \exp(\Delta tB) + O(\Delta t^3). \end{aligned} \tag{2.23}$$

The leading error term is

$$E_1 = \Delta t \left(\sum_{n,m \geq 0} \frac{(\Delta t B)^n A (\Delta t B)^m}{(n+m+1)!} - A \exp{(\Delta t B)} \right), \tag{2.24}$$

and the second error term is

$$E_2 = \Delta t^2 \left(\sum_{n,m,k \geq 0} \frac{(\Delta t B)^n \, A \, (\Delta t B)^m \, A \, (\Delta t B)^k}{(n+m+k+2)!} - \frac{A^2}{2} \exp(\Delta t B) \right). \tag{2.25}$$

We obtain

$$\beta_{k,l} = \frac{\exp{(\Delta t \lambda_k)} - \exp{(\Delta t \lambda_l)}}{\Delta t \lambda_k - \Delta t \lambda_l} - \exp{(\Delta t \lambda_l)}, \qquad \lambda_k \neq \lambda_l, \tag{2.26}$$

and

$$\gamma_{i,q,j} = \sum_{n,m,k \geq 0} \frac{(\Delta t \lambda_i)^n (\Delta t \lambda_q)^m (\Delta t \lambda_j)^k}{(n+m+k+2)!} - \frac{1}{2} \exp{(\Delta t \lambda_j)}. \tag{2.27}$$

2. The splitting **BA** has the expansion

$$\begin{aligned} \exp(\Delta t B) \exp(\Delta t A) &= \exp(\Delta t B) \left(I + \Delta t A + \frac{1}{2} \Delta t^2 A^2 + O(\Delta t^3) \right) \\ &= \exp(\Delta t B) + \Delta t \exp(\Delta t B) A + \frac{1}{2} \Delta t^2 \exp(\Delta t B) A^2 + O(\Delta t^3), \end{aligned} \tag{2.28}$$

giving the first error terms

$$E_1 = \Delta t \left(\sum_{n,m \geq 0} \frac{(\Delta t B)^n A (\Delta t B)^m}{(n+m+1)!} - \exp{(\Delta t B)} \, A \right), \tag{2.29}$$

and

$$E_2 = \Delta t^2 \left(\sum_{n,m,k \geq 0} \frac{(\Delta t B)^n \, A \, (\Delta t B)^m \, A \, (\Delta t B)^k}{(n+m+k+2)!} - \exp(\Delta t B) \frac{A^2}{2} \right), \tag{2.30}$$

such that

$$\beta_{k,l} = \frac{\exp{(\Delta t \lambda_k)} - \exp{(\Delta t \lambda_l)}}{\Delta t \lambda_k - \Delta t \lambda_l} - \exp{(\Delta t \lambda_k)}, \qquad \lambda_k \neq \lambda_l, \tag{2.31}$$

and

$$\gamma_{i,q,j} = \sum_{n,m,k \geq 0} \frac{(\Delta t \lambda_i)^n (\Delta t \lambda_q)^m (\Delta t \lambda_j)^k}{(n+m+k+2)!} - \frac{1}{2} \exp{(\Delta t \lambda_i)}. \tag{2.32}$$

3. The cases **ABA** and **AB+BA** have the following expansions.
ABA:

$$\exp\left(\frac{\Delta t}{2} A \right) \exp{(\Delta t B)} \exp\left(\frac{\Delta t}{2} A \right) \tag{2.33}$$

$$= \left(I + \frac{\Delta t}{2}A + \frac{1}{2}\left(\frac{\Delta t}{2}\right)^2 A^2 + O(\Delta t^3)\right) \exp\left(\Delta t B\right) \left(I + \frac{\Delta t}{2}A + \frac{1}{2}\left(\frac{\Delta t}{2}\right)^2 A^2 + O(\Delta t^3)\right)$$

$$= \exp\left(\Delta t B\right) + \Delta t \frac{\exp\left(\Delta t B\right) A + A \exp\left(\Delta t B\right)}{2}$$

$$+ \Delta t^2 \frac{A^2 \exp\left(\Delta t B\right) + 2A \exp\left(\Delta t B\right) A + \exp\left(\Delta t B\right) A^2}{8} + O(\Delta t^3).$$

AB+BA:

$$\frac{1}{2}\left(\exp(\Delta t A)\exp(\Delta t B) + \exp(\Delta t B)\exp(\Delta t A)\right) \tag{2.34}$$

$$= \frac{1}{2}\left(\left(I + \Delta t A + \frac{1}{2}\Delta t^2 A^2 + O(\Delta t^3)\right) \exp\left(\Delta t B\right) + \exp\left(\Delta t B\right) \left(I + \Delta t A + \frac{1}{2}\Delta t^2 A^2 + O(\Delta t^3)\right)\right)$$

$$= \exp\left(\Delta t B\right) + \Delta t \frac{\exp\left(\Delta t B\right) A + A \exp\left(\Delta t B\right)}{2} + h^2 \frac{A^2 \exp\left(\Delta t B\right) + \exp\left(\Delta t B\right) A^2}{4} + O(\Delta t^3).$$

In both cases there are leading errors of the form

$$E_1 = \Delta t \left(\sum_{n,m \geq 0} \frac{(\Delta t B)^n A (\Delta t B)^m}{(n+m+1)!} - \frac{\exp\left(\Delta t B\right) A + A \exp\left(\Delta t B\right)}{2} \right), \tag{2.35}$$

and consequently, they have the same $\beta_{k,l}$ values

$$\beta_{k,l} = \frac{\exp\left(\Delta t \lambda_k\right) - \exp\left(\Delta t \lambda_l\right)}{\Delta t \lambda_k - \Delta t \lambda_l} - \frac{\exp\left(\Delta t \lambda_k\right) + \exp\left(\Delta t \lambda_l\right)}{2}, \qquad \lambda_k \neq \lambda_l. \tag{2.36}$$

Thus, we expect a similar order reduction for **ABA** and **AB+BA** cases. The second error terms are different for these cases. For **ABA** we get

$$E_2 = \Delta t^2 \left(\sum_{n,m,k \geq 0} \frac{(\Delta t B)^n A (\Delta t B)^m A (\Delta t B)^k}{(n+m+k+2)!} \right.$$
$$\left. - \frac{A^2 \exp\left(\Delta t B\right) + 2A \exp\left(\Delta t B\right) A + \exp\left(\Delta t B\right) A^2}{8} \right), \tag{2.37}$$

and for **AB+BA**

$$E_2 = \Delta t^2 \left(\sum_{n,m,k \geq 0} \frac{(\Delta t B)^n A (\Delta t B)^m A (\Delta t B)^k}{(n+m+k+2)!} - \frac{A^2 \exp\left(\Delta t B\right) + \exp\left(\Delta t B\right) A^2}{4} \right). \tag{2.38}$$

Thus we obtain

$$\gamma_{i,q,j} = \sum_{n,m,k \geq 0} \frac{(\Delta t \lambda_i)^n (\Delta t \lambda_q)^m (\Delta t \lambda_j)^k}{(n+m+k+2)!} - \frac{\exp\left(\Delta t \lambda_i\right) + 2\exp\left(\Delta t \lambda_q\right) + \exp\left(\Delta t \lambda_j\right)}{8}, \tag{2.39}$$

for **ABA** and

$$\gamma_{i,q,j} = \sum_{n,m,k \geq 0} \frac{(\Delta t \lambda_i)^n (\Delta t \lambda_q)^m (\Delta t \lambda_j)^k}{(n+m+k+2)!} - \frac{\exp\left(\Delta t \lambda_i\right) + \exp\left(\Delta t \lambda_j\right)}{4}, \tag{2.40}$$

for **AB+BA**.

It should be pointed out that for the cases **ABA** and **AB+BA** the elements $\beta_{k,l}$ are symmetric: $\beta_{k,l} = \beta_{l,k}$. This is not true for the cases **AB** and **BA**.

We now list the various cases of eigenvalues pairs λ_k and λ_l and give the estimate for $\beta_{k,l}$ for all four splitting methods discussed above:

I. For $\lambda_k = \lambda_l$ we get $\beta_{k,l} = 0$;

II. Assume $\lambda_k \neq \lambda_l$:

(a) If λ_k and λ_l are nonstiff, then $\beta_{k,l} = O(\Delta t)$ for the cases **AB** and **BA** and $\beta_{k,l} = O(\Delta t^2)$ for the cases **ABA** and **AB+BA** that corresponds to the classical order of the considered splitting methods;

(b) If λ_k is stiff and λ_l is nonstiff, then

$$\beta_{k,l} = \begin{cases} O(1), & \text{for } \mathbf{AB}, \mathbf{BA} \text{ and } \mathbf{AB+BA} \\ O\left(\dfrac{1}{\Delta t \lambda_k}\right) = O\left(\dfrac{\varepsilon}{\Delta t}\right), & \text{for } \mathbf{BA} \end{cases}$$

If we take a nonstiff λ_k and a stiff λ_l, the results for the cases **AB** and **BA** will interchange.

(c) For two stiff eigenvalues λ_k and λ_l the value $\beta_{k,l} \approx 0$ is negligible;

(d) If λ_k and λ_l are imaginary and complex conjugate, then

$$|\beta_{k,l}| = \begin{cases} O(1), & \text{for } \mathbf{AB}, \mathbf{BA}, \mathbf{ABA} \\ |\cos(\Delta t |\lambda_k|)|, & \text{for } \mathbf{AB+BA} \end{cases}$$

It follows that we can expect different types of order reduction corresponding to different eigenvalue pairs. The lowest order which $\beta_{k,l}$ can have is (-1); it is given by the pair of one stiff and one nonstiff eigenvalues for the cases **AB** and **BA**. This $\beta_{k,l}$ contributes to the splitting error as a zero order term: $\Delta t \cdot O\left(\frac{\varepsilon}{\Delta t}\right) = O(\varepsilon)$.

For the γ values we obtain the same results as we got for the splitting **BAB** for the other splitting types. The examination of possible $\beta_{k,l}$ and $\gamma_{i,q,j}$ values provides the main result of this section.

Theorem 2.3 *For sufficiently small steplengths Δt such that the problem (2.1), (2.7), (2.8) is stiff: $|\lambda_{nst}|\Delta t \ll 1$ and $|\lambda_{st}|\Delta t \gg 1$, the order of the splitting methods (2.3), (2.4) is given by the leading error term $E_1 = O(\Delta t, \varepsilon)$. The next error term $E_2 = O(\Delta t^2, \Delta t\varepsilon, \varepsilon^2)$ is much smaller, i.e. $E_2 = o(E_1)$.*

2.5 Numerical experiments

The simplest example in which it is possible to observe the order reduction phenomenon is a special case of the problem (2.1):

$$y' = (A + B)\, y, \qquad y \in \mathbb{R}^2, \qquad A,\ B \in \mathbb{R}^{2\times 2} \tag{2.41}$$

with constant matrices A and B. In this case there are only two nonzero β–values: $\beta_{1,2}$ and $\beta_{2,1}$. As it was mentioned before, for the symmetric cases **BAB**, **ABA** and **AB+BA** we have $\beta_{1,2} = \beta_{2,1}$. Therefore for these splitting types we should observe the order

$$E \approx E_1 = O(\Delta t \beta_{1,2}).$$

For the cases **AB** and **BA** both $\beta_{1,2}$ and $\beta_{2,1}$ contribute to the error. Below we consider three different examples of the problem (2.41), which demonstrate different types of the order reduction. We will apply all five splittings methods (2.3), (2.4) and investigate the order reduction analytically as well as numerically.

Example 2.1 A stiff problem.

We choose matrices A and B which represent the nonstiff and stiff parts of the equation (2.41) as follows

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \qquad B = \frac{1}{\varepsilon}\begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}, \qquad \varepsilon = 10^{-5}$$

The matrix B has two eigenvalues $\lambda_1 = 0$ and $\lambda_2 = -1/\varepsilon$. We give the nonzero β–values for all types of splittings for sufficiently large steplength Δt.

$$\begin{array}{ll} \mathbf{AB} & |\beta_{1,2}| = \dfrac{\varepsilon}{\Delta t}, \qquad |\beta_{2,1}| \approx 1 \\[2ex] \mathbf{BA} & |\beta_{1,2}| = 1, \qquad |\beta_{2,1}| \approx \dfrac{\varepsilon}{\Delta t} \\[2ex] \mathbf{BAB} & |\beta_{1,2}| = |\beta_{2,1}| = \dfrac{\varepsilon}{\Delta t} \\[2ex] \mathbf{ABA},\ \mathbf{AB}+\mathbf{BA} & |\beta_{1,2}| = |\beta_{2,1}| = \dfrac{1}{2} \end{array} \tag{2.42}$$

The particular form of the matrices A and B chosen for this example, gives $\hat{A}_{2,1} = (S^{-1}AS)_{2,1} = 0$ for any choice of the matrix S giving the presentation $B = S\Lambda S^{-1}$ with $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2)$. It implies that for this example

$$\hat{E}_1 = \begin{bmatrix} 0 & O(\Delta t\beta_{1,2}) \\ 0 & 0 \end{bmatrix} = O(\Delta t\beta_{1,2}) \quad \text{and} \quad E_1 = O(\Delta t\beta_{1,2}).$$

Hence the cases **AB** and **BA** are expected to have different types of order reduction. The errors of the splitting are plotted in Figure 2 and their orders in Figure 3.

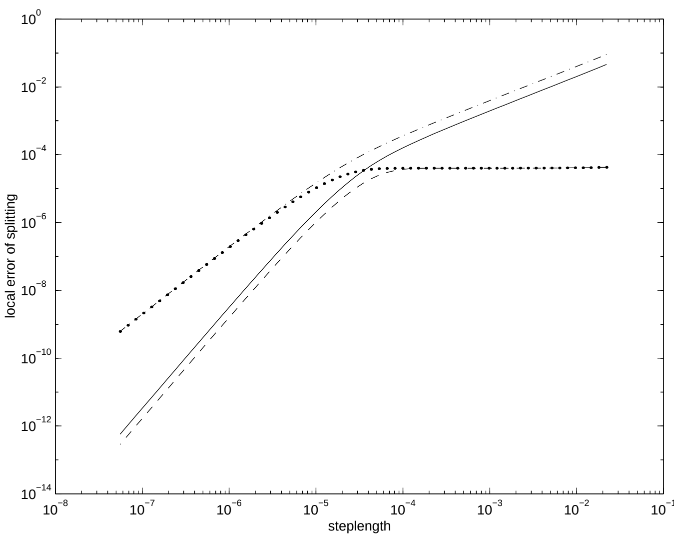

Figure 2: Frobenius norm of the error matrix E for the stiff problem. The solid line stands for **ABA** (**AB+BA**) splitting, the dashed line for **BAB**, the dot–dot line for **AB** and the dash–dot line for **BA**.

The cases **AB** and **BA** have their local errors reduced from $O(\Delta t^2)$ to $O(\Delta t)$ and $const = O(\varepsilon)$, respectively. The cases **ABA** and **AB+BA** show the same performance. They are each represented by a line in the figures. These splitting types have the local error $O(\Delta t^3)$ in the classical range of steplengths and $O(\Delta t)$ for the reduction range of Δt. The **BAB** splitting has

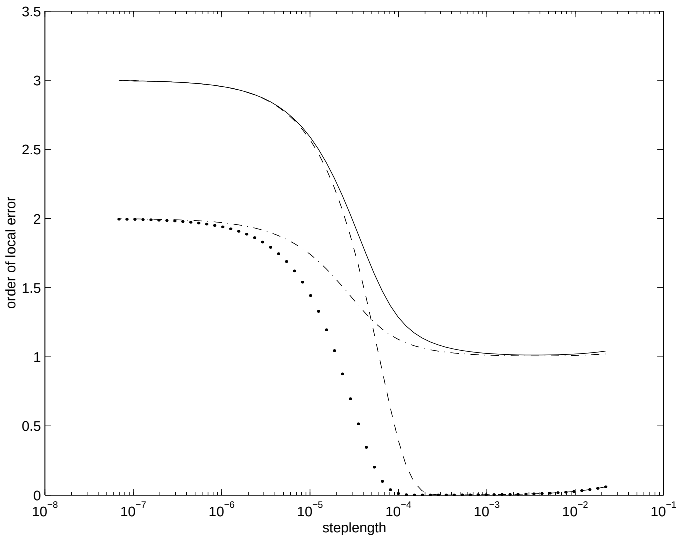

Figure 3: The order of the local error for the stiff problem. The notations are the same as in Figure 2.

reduction from $O(\Delta t^3)$ to $const = O(\varepsilon)$. The results of the computations are completely consistent with the analytical estimations of the order reduction given by (2.42).

$\Diamond$

Example 2.2 Another extreme case of the problem (2.41) is the oscillatory one. Now the matrix A correspond to *slow* components of the solution and the matrix B contains both *slow* components and *fast*, highly oscillatory ones. For instance, we can take

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \qquad B = \frac{1}{\varepsilon}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \qquad \text{with} \quad \varepsilon = 10^{-5}.$$

B has two complex conjugate eigenvalues $\lambda = \pm i/\varepsilon$ and we obtain the following nonzero β–values:

$$\begin{array}{cl} \mathbf{AB},\ \mathbf{BA} & |\beta_{1,2}| \approx |\beta_{2,1}| \approx 1 \\ \mathbf{BAB} & |\beta_{1,2}| = |\beta_{2,1}| \approx 1 \\ \mathbf{ABA},\ \mathbf{AB+BA} & |\beta_{1,2}| = |\beta_{2,1}| = |\cos(\Delta t/\varepsilon)| \end{array} \tag{2.43}$$

The results of the computations are presented in Figure 4. The cases **AB** and **BA** are indistinguishable in the plot. Both suffer order reduction in the local error from $O(\Delta t^2)$ to $O(\Delta t)$ with some oscillations near the order breaking points. These oscillations are coursed by a substantial value $\frac{\sin(\Delta t/\varepsilon)}{\Delta t/\varepsilon}$, which is present in all β–values (2.43). As $\frac{\sin(\Delta t/\varepsilon)}{\Delta t/\varepsilon}$ becomes negligible as Δt increases, the oscillations disappear. The same oscillation are shown by **BAB** while the order is getting reduced from $O(\Delta t^3)$ to $O(\Delta t)$. For the splitting **ABA** (**AB**+**BA**) the error is oscillating in terms of Δt for sufficiently large time steps when we observe the order reduction. The upper envelope of the error curve is $O(\Delta t)$ that can be seen both from the Figure 4 and from the value $|\beta_{1,2}|$ in (2.43).

$\Diamond$

Example 2.3 The third example is chosen to demonstrate the order reduction in the case when the first error term $e^1 = E_1 y_0$ is zero, i.e. $y_0 \in \mathrm{Ker}(E_1)$, and the order is defined by the next term

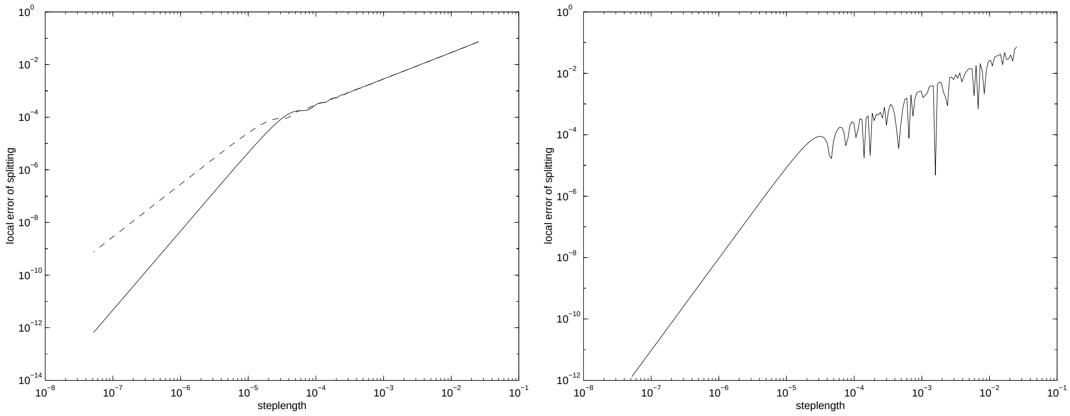

Figure 4: Frobenius norm of the error matrix E for the oscillatory problem. In the left figure the solid line stands for **BAB** splitting, the dashed line for **AB** (**BA**). The right figure presents the splitting **ABA** (**AB**+**BA**).

$e^2 = E_2 y_0$. We choose

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \\ 1 & -1 & 2 & 1 \\ 1 & -1 & 1 & 2 \end{bmatrix}, \qquad B = \frac{1}{\varepsilon}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \qquad y_0 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \tag{2.44}$$

where $\varepsilon = 10^{-5}$. The matrix B is already diagonal. Because we have only two eigenvalues of double multiplicity each: $\lambda_1 = \lambda_2 = 0$, $\lambda_3 = \lambda_4 = -\frac{1}{\varepsilon}$, there are only two nonzero β–values and the first error term matrix E_1 has the form

$$E_1 = \Delta t \begin{bmatrix} 0 & 0 & 3\beta_{1,3} & 4\beta_{1,3} \\ 0 & 0 & 2\beta_{1,3} & \beta_{1,3} \\ \beta_{3,1} & -\beta_{3,1} & 0 & 0 \\ \beta_{3,1} & -\beta_{3,1} & 0 & 0 \end{bmatrix}$$

We see that that $e^1 = E_1 y_0$ is the null vector. We also easily see that for initial data $y_0 \notin \mathrm{Ker}(E_1)$ the error is composed by terms $O(\Delta t \beta_{1,3})$ and $O(\Delta t \beta_{3,1})$, which are equal for the symmetrical cases **BAB**, **ABA** and **AB**+**BA** providing the same order behavior as we observed in Example 2.1, but are different for the splitting **AB** and **BA**. Thus for the two last cases we can get different order behavior ($e^1 = O(h)$ or $e^1 = const$) of the error for different initial data.

We thus need to examine the second error term which gives the order behavior of the problem (2.1),(2.44):

$$e^2 = E_2 y_0 = \Delta t^2 \begin{bmatrix} \sum_{q=1}^4 (\gamma_{1,q,1} A_{1,q} A_{q,1} + \gamma_{1,q,2} A_{1,q} A_{q,2}) \\ \sum_{q=1}^4 (\gamma_{2,q,1} A_{2,q} A_{q,1} + \gamma_{2,q,2} A_{2,q} A_{q,2}) \\ \sum_{q=1}^4 (\gamma_{3,q,1} A_{3,q} A_{q,1} + \gamma_{3,q,2} A_{3,q} A_{q,2}) \\ \sum_{q=1}^4 (\gamma_{4,q,1} A_{4,q} A_{q,1} + \gamma_{4,q,2} A_{4,q} A_{q,2}) \end{bmatrix}$$

There are many equal values among the γs. In particular, we have $\gamma_{i,q,1} = \gamma_{i,q,2}$ for any i and q and $\gamma_{i,q,j} = 0$ if the indices i, q and j correspond to equal eigenvalues. Using these properties and that $A_{3,1} + A_{3,2} = 0$ and $A_{4,1} + A_{4,2} = 0$, we get the order behavior of the error:

$$e^2 = \begin{bmatrix} 0 \\ 0 \\ O(\Delta t^2 \gamma_{3,1,1}) \\ O(\Delta t^2 \gamma_{3,1,1}) \end{bmatrix}$$

Now we can consider the γ expressions for different splitting types. They give the reduced order

$$e^2 = \begin{cases} O(\Delta t), & \text{for} \quad \mathbf{BA},\ \mathbf{BAB} \\ \\ O(\Delta t^2), & \text{for} \quad \mathbf{AB},\ \mathbf{ABA},\ \mathbf{AB}+\mathbf{BA} \end{cases} \tag{2.45}$$

We should note that the different order behavior of the cases **AB** and **BA** is due to the particular form of this example problem. The computations of the splitting orders are shown in Figures 5 and 6 (Because of the jump we omit the order plot for **AB+BA**). We see that the reduced orders obtained from the numerical computations are as predicted theoretically in (2.45).

It is worth mentioning that for this example all five cases have the second classical order. The cases **AB** and **BA** get higher classical order because $(AB - BA)y_0$ is the null vector.

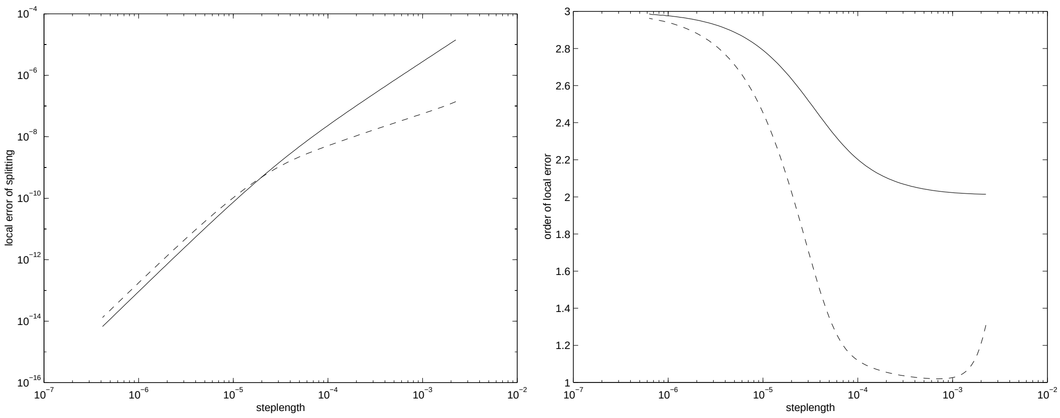

Figure 5: The local error (left plot) and the order of the local error (right plot) of the problem (2.44). The solid line stands for **AB** splitting and the dashed line for **BA**.

◊

We have considered three examples to demonstrate the order reduction. The approach which was used in our analysis can be applied to more complicated systems of equations (2.1). In the general case, the order behavior can be complicated because a pair of eigenvalues of the matrix B contributes to the order reduction in its own, special way. We believe that results can be obtained for eigensystems which possess some structure. A well known source of eigenvalues with structure are approximations of differential operators. This introduces us a new problem: order reduction of splitting methods applied to partial differential equations and their semi–discretizations. Semi–discretized PDEs of the from (2.1) represent a class of equations whose order reduction we already know how to investigate. In the next section we will extend our order analysis method to a class of PDEs whose solutions can be approximated by the splitting methods (2.3), (2.4).

3 The PDE case

Partial differential equations can be considered as ODEs in appropriate function spaces. This approach was used, for example, to apply Runge–Kutta methods to certain classes of PDEs (see [9], [10], [11]). In this section we will apply splitting methods to a class of PDEs, which can be viewed as a continuous generalization of the problem (2.1).

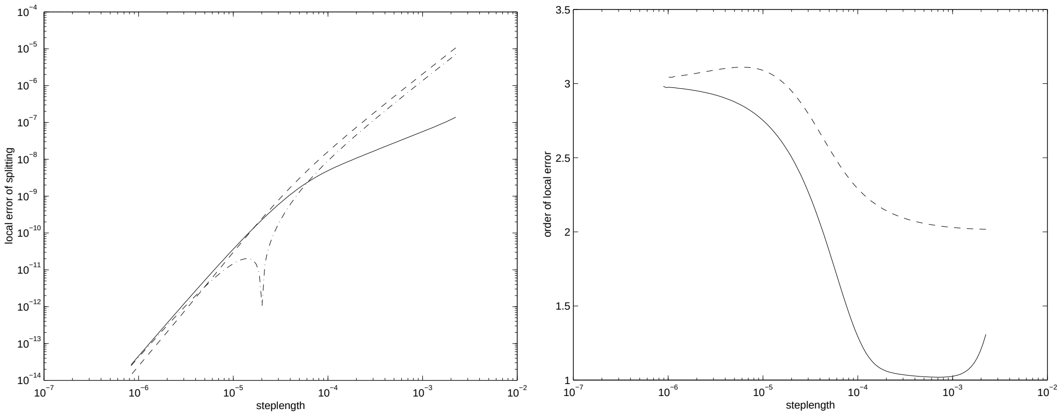

Figure 6: The local error (left plot) and the order of the local error (right plot) of the problem (2.44). The solid line stands for **ABA** splitting, the dashed line for **BAB** and the dash–dot line for **AB+BA**.

3.1 The problem formulation and the order analysis

Let us consider the following linear partial differential equation

$$
\begin{aligned}
&u_t(x,t) = A(x,\partial)u(x,t) + B(x,\partial)u(x,t), \qquad x \in \Omega, \qquad 0 \le t \le T, \\
&u(x,0) = u_0(x), \qquad x \in \Omega;
\end{aligned} \tag{3.1}
$$

with homogeneous boundary conditions. The domain Ω is an open and bounded subset of $\mathbb{R}^d$ with sufficiently smooth boundary $\partial\Omega$. The differential operators $A(x,\partial)$ and $B(x,\partial)$ are densely defined in $L^2(\Omega)$ and they do not generally commute $[A,B] \neq 0$. We will make some assumptions on the problem (3.1) partially taken from [9]:

Assumption 3.1

1. The linear operator $-B(x,\partial)$ is sectorial [13], i.e. it is a closed densely defined operator such that for some $\varphi \in (0,\pi/2)$ and some $M_1 \ge 1$ and real a_1, the sector

   $$S_{a_1,\varphi} = \{\lambda \in \mathbb{C} \mid \varphi \le |\arg(\lambda - a_1)| \le \pi, \lambda \neq a_1\}$$

   is in the resolvent $\rho(-B)$ and for some norm $||\cdot||$

   $$||(\lambda I + B)^{-1}|| \le \frac{M_1}{|\lambda - a_1|} \qquad \text{for all} \quad \lambda \in S_{a_1,\varphi}.$$

   Under this hypothesis, the operator B is the infinitesimal generator of an analytic semigroup $\{\exp(tB)\}_{t\ge 0}$ with the exponential of the unbounded operator B defined by

   $$\exp(tB) = \frac{1}{2\pi i}\int_\Gamma (\lambda I - B)^{-1}\exp(\lambda t)d\lambda, \tag{3.2}$$

   where Γ is a contour in the resolvent of B with $\arg(\lambda) \to \pm\theta$ as $|\lambda| \to \infty$ for some $\theta \in (\pi/2,\pi)$.

2. The operator B has a pure point spectrum $\{\lambda_1, \lambda_2, \lambda_3, \dots\}$:

   $$B\phi_k = \lambda_k\phi_k, \qquad k = 1,\dots,\infty, \tag{3.3}$$

   and $\text{Re}\lambda_k \to -\infty$ as $k \to \infty$. The eigenfunctions ϕ_k are assumed to satisfy the following properties:

(a) They form a basis of $L^2(\Omega)$, i.e. any $u \in L^2(\Omega)$ can be expressed by the series

$$u = \sum_{k=1}^{\infty} a_k \phi_k \quad \text{in} \quad L^2(\Omega); \tag{3.4}$$

(b) The mapping

$$\begin{cases} L^2(\Omega) \to l^2 \\ u = \sum a_k \phi_k \to \{a_k\} \end{cases} \quad \text{is a homeomorphism.} \tag{3.5}$$

We denote by l^2 the Hilbert space of sequences $\{a_k\}_{k=1,2,\ldots}$ which satisfy $\sum |a_k| < \infty$.

3. The operator A is bounded and its action on each eigenfunction of the operator B can be expressed in terms of the eigenfunctions

$$A\phi_k = \sum_{i=1}^{\infty} \alpha_k^i \phi_i. \tag{3.6}$$

$\Diamond$

Under our assumptions, the operator $-(A+B)$ is also sectorial with sector constants a_2 and φ, and for some $M_2 \geq 1$ [13]

$$\|(\lambda I + A + B)^{-1}\| \leq \frac{M_2}{|\lambda - a_2|} \qquad \text{for all} \quad \lambda \in S_{a_2,\varphi}.$$

The solution of the problem (3.1) can be written in the from

$$u(x,t) = \exp(t(A(x,\partial) + B(x,\partial)))u_0(x), \tag{3.7}$$

where the exponential is defined as

$$\exp(t(A+B)) = \frac{1}{2\pi i}\int_{\Gamma} (\lambda I - A - B))^{-1} \exp(\lambda t) d\lambda, \qquad \Gamma \in \rho(A+B). \tag{3.8}$$

We will study the order behavior of the splitting methods (2.3)–(2.4) applied to (3.1), i.e. to approximate the solution operator $\exp(t(A+B))$. Comparing (3.1) with the ODE problem (2.1), we can say that the operator B is the "stiff" part and the operator A is the nonstiff one. The difference between the ODE problem (2.1) and the PDE problem (3.1) is that the operator B is "stiff" for any choice of the time step Δt, not only for large step sizes as in the ODE case.

As in the ODE case we will choose one splitting, say **BAB**, for the detailed analysis. Therefore we need to examine the error operator

$$E^c = \exp(\Delta t(A(x,\partial) + B(x,\partial))) - \exp\left(\frac{\Delta t}{2}B(x,\partial)\right)\exp(\Delta t A(x,\partial))\exp\left(\frac{\Delta t}{2}B(x,\partial)\right), \tag{3.9}$$

applied to the initial data

$$u_0(x) = \sum_{k=1}^{\infty} a_k^0 \phi_k. \tag{3.10}$$

Using the identity

$$(\lambda I - A - B)^{-1} = (\lambda I - B)^{-1} + (\lambda I - B)^{-1} A (\lambda I - B)^{-1}$$

$$+(\lambda I - A - B)^{-1}A(\lambda I - B)^{-1}A(\lambda I - B)^{-1},$$

we obtain the following expression for the solution operator

$$\exp(\Delta t(A+B)) = \exp(\Delta t B) + \frac{1}{2\pi i}\int_\Gamma (\lambda I - B)^{-1}A(\lambda I - B)^{-1}\exp(\lambda \Delta t)d\lambda$$
$$+\frac{1}{2\pi i}\int_\Gamma (\lambda I - A - B)^{-1}A(\lambda I - B)^{-1}A(\lambda I - B)^{-1}\exp(\lambda\Delta t)d\lambda, \qquad (3.11)$$

where the contour $\Gamma \in \rho(B)\cap\rho(A+B)$. For the **BAB** splitting we get the same expansion as in the ODE case (2.10). The leading error term of the error E^c is

$$E_1^c = \frac{1}{2\pi i}\int_\Gamma (\lambda I - B)^{-1}A(\lambda I - B)^{-1}\exp(\lambda\Delta t)d\lambda - \Delta t \exp\left(\frac{\Delta t}{2}B\right)A\exp\left(\frac{\Delta t}{2}B\right), \qquad (3.12)$$

A consequence of the next theorem is that E_1^c is the principal error for sufficiently small steplengths Δt.

Theorem 3.1 *For sufficiently small time steps Δt, it is true that $\|E^c - E_1^c\| = O(\Delta t^2)$.*

Proof: The operator

$$E^c - E_1^c = \frac{1}{2\pi i}\int_\Gamma (\lambda I - A - B)^{-1}A(\lambda I - B)^{-1}A(\lambda I - B)^{-1}\exp(\lambda\Delta t)d\lambda$$
$$-\exp\left(\frac{\Delta t}{2}B\right)\sum_{k=2}^{\infty}\frac{\Delta t^k}{k!}A^k\exp\left(\frac{\Delta t}{2}B\right)$$

can be bounded as

$$\|E^c - E_1^c\| \le \left\|\frac{1}{2\pi i}\int_\Gamma (\lambda I - A - B)^{-1}A(\lambda I - B)^{-1}A(\lambda I - B)^{-1}\exp(\lambda\Delta t)d\lambda\right\|$$
$$+\left\|\exp\left(\frac{\Delta t}{2}B\right)\right\|^2\sum_{k=2}^{\infty}\frac{\Delta t^k}{k!}\|A\|^k$$

The rest of the terms of the exact solution (3.11) can be bounded as

$$\left\|\frac{1}{2\pi i}\int_\Gamma (\lambda I - A - B)^{-1}A(\lambda I - B)^{-1}A(\lambda I - B)^{-1}\exp(\lambda\Delta t)d\lambda\right\|$$
$$\le \frac{\|A\|^2}{2\pi}\int_\Gamma \|(\lambda I - A - B)^{-1}\|\|(\lambda I - B)^{-1}\|^2|\exp(\lambda\Delta t)||d\lambda|$$
$$\le \frac{\|A\|^2M_1^2M_2}{2\pi}\int_\Gamma \frac{|\exp(\lambda\Delta t)|}{|\lambda+a_1|^2|\lambda+a_2|}|d\lambda| \le \frac{\|A\|^2M_1^2M_2}{2\pi}\int_\Gamma\frac{|\exp(\lambda\Delta t)|}{|\lambda+a_1|^3}|d\lambda|$$
$$+\frac{\|A\|^2M_1^2M_2}{2\pi}\int_\Gamma\frac{|\exp(\lambda\Delta t)|}{|\lambda+a_2|^3}|d\lambda| = \frac{\|A\|^2M_1^2M_2}{2\pi}\int_{\Gamma_1}\frac{|\exp((\lambda-a_1)\Delta t)|}{|\lambda|^3}|d\lambda|$$
$$+\frac{\|A\|^2M_1^2M_2}{2\pi}\int_{\Gamma_2}\frac{|\exp((\lambda-a_2)\Delta t)|}{|\lambda|^3}|d\lambda|,$$

where the contours Γ_1 and Γ_2 are obtained from the contour Γ by shifts to the left on a_1 and a_2 correspondingly. Putting $\mu = \lambda\Delta t$, we continue to estimate the last integrals:

$$= \frac{\|A\|^2M_1^2M_2}{2\pi}\Delta t^2\left(\exp(-a_1\Delta t)\int_{\Gamma_1}\frac{|\exp(\mu)|}{|\mu|^3}|d\mu| + \exp(-a_2\Delta t)\int_{\Gamma_2}\frac{|\exp(\mu)|}{|\mu|^3}|d\mu|\right)$$

The terms of the splitting expansion can be bounded with the help of

$$\left\|\exp\left(\frac{\Delta t}{2}B\right)\right\| \le \frac{M_1}{2\pi}\int_{\Gamma_1}\frac{|\exp(\mu)|}{|\mu|}|d\mu|\exp\left(-a_1\frac{\Delta t}{2}\right).$$

and

$$\sum_{k=2}^{\infty}\frac{\Delta t^k}{k!}\|A\|^k = \Delta t^2\|A\|^2\sum_{k=0}^{\infty}\frac{\Delta t^k}{(k+2)!}\|A\|^k \le \Delta t^2\|A\|^2\exp(\Delta t\|A\|)$$

Combining the obtained estimates we get the statement of the theorem for sufficiently small time steps.

□

The theorem has the following important consequence.

Corollary 3.2 *For sufficiently small time steps Δt the error of the splitting* **BAB** *applied to the problem (3.1) $e = E^c u_0$ is dominated by the leading error term $e \approx e^1 = E_1^c u_0$ if the latter is nonzero.*

Thus we clearly see that for sufficiently small step sizes Δt the order behavior of the error E^c of the splitting must be that of the first error term E_1^c. Using

$$\left(\frac{1}{2\pi i}\int_\Gamma(\lambda I-B)^{-1}A(\lambda I-B)^{-1}\exp(\lambda\Delta t)d\lambda\right)\phi_l = \sum_k\frac{1}{2\pi i}\int_\Gamma\frac{\exp(\lambda\Delta t)d\lambda}{(\lambda-\lambda_k)(\lambda-\lambda_l)}\alpha_l^k\phi_k$$

$$= \Delta t\sum_{\lambda_k\neq\lambda_l}\frac{\exp(\lambda_k\Delta t)-\exp(\lambda_l\Delta t)}{\Delta t\lambda_k-\Delta t\lambda_l}\alpha_l^k\phi_k + \Delta t\sum_{\lambda_k=\lambda_l}\exp(\lambda_k\Delta t)\alpha_l^k\phi_k$$

and

$$\exp\left(\frac{\Delta t}{2}B\right)\phi_l = \frac{1}{2\pi i}\int_\Gamma\frac{1}{\lambda-\lambda_l}\exp\left(\lambda\frac{\Delta t}{2}\right)d\lambda\phi_l = \exp\left(\lambda_l\frac{\Delta t}{2}\right)\phi_l$$

that gives

$$\exp\left(\frac{\Delta t}{2}B\right)A\exp\left(\frac{\Delta t}{2}B\right)\phi_l = \sum_k\exp\left(\Delta t\frac{\lambda_k+\lambda_l}{2}\right)\alpha_l^k\phi_k$$

we obtain the presentation of the leading error term e^1 for the solution of the problem (3.1) with the initial data presented in the form (3.10):

$$e^1 = E_1^c u_0 = E_1^c\left(\sum_{l=1}^{\infty}a_l^0\phi_l\right) = \Delta t\sum_{k,l=1}^{\infty}\beta_{k,l}\alpha_l^k a_l^0\phi_k \tag{3.13}$$

or

$$e^1 = \sum_{k=1}^{\infty}e_k^1\phi_k, \qquad e_k^1 = \Delta t\sum_{l=1}^{\infty}\beta_{k,l}\alpha_l^k a_l^0,$$

with the same values $\beta_{k,l}$ as we obtained for the ODE case (see (2.18)). This expression for the error of the splitting is valid for both finite–dimensional and infinite–dimensional operators. In the finite dimensional case we will get a finite sum.

Our goal is to connect the leading error term of the splitting $e^1(\Delta t, x)$ with the initial data $u_0(x)$. For this purpose we will associate the initial data $u_0(x) \in L^2(\Omega)$ and the error $e^1(\Delta t, x) \in L^2(\Omega)$ with their coefficients in the expansion (3.4): $\{a_k^0\}, \{e_k^1\} \in l^2$. Let us introduce a map $\hat{E}_1^c$ from $\{a_k^0\}$ into $\{e_k^1\}$ for $\{a_k^0\} \in l^2$:

$$e^1 = \hat{E}_1^c a^0 \qquad (\hat{E}_1^c)_{k,l} = \Delta t\beta_{k,l}\alpha_l^k, \quad k, l = 1, \dots, \infty. \tag{3.14}$$

Transforming the analysis from the function space $L^2(\Omega)$ to the sequence space l^2, we see a possibility to estimate the order of the splitting. The infinite–dimensional "matrix" $\hat{E}_1^c$ can be chosen to give an order of the splitting method. We would like to choose some norm which essentially characterizes the error matrix $\hat{E}_1^c$ and at the same time can be used to estimate the error analytically. The 2-norm appears not to be practical. We prefer to use the Frobenius norm of $\hat{E}_1^c$

$$||\hat{E}_1^c||_F = \Delta t \left(\sum_{k,l=1}^{\infty} |\beta_{k,l}\alpha_l^k|^2 \right)^{1/2} \tag{3.15}$$

because it can be more easily computed and as we will show later for some cases it is equivalent to the Frobenius norm of the error of a semi–discretization of (3.1).

For splittings **BAB**, **ABA** and **AB+BA** we can exploit the symmetry of $\beta_{k,l}$ in order to simplify the sum

$$||\hat{E}_1^c||_F^2 = \Delta t^2 \sum_{k,l=1}^{\infty} |\beta_{k,l}\alpha_l^k|^2 = \Delta t^2 \sum_{l=1,k\geq l+1}^{\infty} |\beta_{k,l}|^2(|\alpha_l^k|^2 + |\alpha_k^l|^2) \tag{3.16}$$

Let us consider the finite-dimensional analogue of the error analysis presented in this section. It is clear that the leading error term E_1^c corresponds to the leading error term E_1 introduced for ODEs. In the ODE case we have a finite number of eigenvectors instead of the infinite set of eigenfunctions (3.3). The eigenvectors form the transformation matrix $S = [\phi_1, \phi_2, \cdots, \phi_N]$, which was used to simplify the ODE system (2.1). It is easy to check that the coefficients α_l^k form the matrix $\hat{A}$:

$$(\hat{A})_{k,l} = (S^{-1}AS)_{k,l} = \alpha_l^k.$$

This implies that the order behavior given by $\hat{E}_1^c$ corresponds to the reduced order behavior of the matrix $\hat{E}_1$, indicating the leading error term of the transformed problem (2.14) in the ODE case.

The next term of the error operator E^c is

$$E_2^c = \frac{1}{2\pi i}\int_\Gamma (\lambda I - B)^{-1}A(\lambda I - B)^{-1}A(\lambda I - B)^{-1}\exp(\lambda\Delta t)d\lambda - \frac{\Delta t^2}{2}\exp\left(\frac{\Delta t}{2}B\right)A^2\exp\left(\frac{\Delta t}{2}B\right)$$

Using

$$\frac{1}{2\pi i}\int_\Gamma (\lambda I - B)^{-1}A(\lambda I - B)^{-1}A(\lambda I - B)^{-1}\exp(\lambda\Delta t)d\lambda\phi_j$$
$$= \sum_{i,q=1}^{\infty} \frac{1}{2\pi i}\int_\Gamma \frac{\exp(\lambda\Delta t)}{(\lambda-\lambda_j)(\lambda-\lambda_q)(\lambda-\lambda_i)}d\lambda\alpha_j^q\alpha_q^i\phi_i$$

and

$$\frac{1}{2\pi i}\int_\Gamma \frac{\exp(\lambda\Delta t)}{(\lambda-a)(\lambda-b)(\lambda-c)}d\lambda$$
$$= \begin{cases} \Delta t^2 \dfrac{(c-b)\exp(a\Delta t) + (a-c)\exp(b\Delta t) + (b-a)\exp(c\Delta t)}{(a-b)(b-c)(c-a)}, & \text{if } a \neq b,\ a \neq c,\ b \neq c \\[2ex] \Delta t^2 \dfrac{\exp(c\Delta t) - \exp(a\Delta t)}{(a-c)^2} + \dfrac{\exp(a\Delta t)}{(a-c)}, & \text{if } \quad a = b \neq c; \\[2ex] \dfrac{\Delta t^2}{2}\exp(a\Delta t), & \text{if } \quad a = b = c; \end{cases}$$

to find the first term of E_2^c, we get the error expression

$$e^2 = E_2^c u_0 = \Delta t^2 \sum_{i,q,j=1}^{\infty} \gamma_{i,q,j}\alpha_j^q\alpha_q^i a_j^0\phi_i \tag{3.17}$$

with the same coefficients $\gamma_{i,q,j}$ as we obtained in the ODE case (see (2.21) and Lemma 2.2).

As for the first error term we can write the second term in the form

$$e^2 = \sum_{i=1}^{\infty} e_i^2 \phi_i, \qquad e_i^2 = \Delta t^2 \sum_{q,j=1}^{\infty} \gamma_{i,q,j} \alpha_j^q \alpha_q^i a_j^0$$

and introduce the infinite "matrix" which maps $\{a_k^0\}$ into $\{e_k^2\}$ for $\{a_k^0\} \in l^2$:

$$e^2 = \hat{E}_2^c a^0 \qquad (\hat{E}_1^c)_{i,j} = \Delta t^2 \sum_{q=1}^{\infty} \gamma_{i,q,j} \alpha_j^q \alpha_q^i, \quad i,j = 1, \dots, \infty. \tag{3.18}$$

The second error term obtained for the linear ODE case (2.20) is the finite–dimensional analogue of (3.18).

Considering the other cases **AB**, **BA**, **ABA** and **AB+BA** in the same manner, we obtain the expressions (3.13) and (3.17) for the first and second error terms correspondingly with the same β and γ coefficients as we found for these splitting methods in the ODE case. Theorem 3.1 can also be extended on the other splitting methods.

3.2 Examples

We now consider equation (3.1) with $A(x,\partial) = f(x)$ and $B(x,\partial) = \frac{\partial^2}{\partial x^2}$, i.e. the equation

$$u_t = u_{xx} + f(x)u, \qquad \left[\frac{\partial^2}{\partial x^2}, f(x)\right] \neq 0. \tag{3.19}$$

Example 3.1 We can take, for instance, $f(x) = \cos(\pi x)$ and examine the problem

$$u_t = u_{xx} + \cos(\pi x)u \qquad \text{in} \qquad [0,1] \tag{3.20}$$

with some initial condition $u(x,t) = u_0(x)$ and homogeneous boundary conditions

$$u(0,t) = u(1,t) = 0, \qquad u_0(0) = u_0(1) = 0$$

This problem fits the general frame we have described above. The "stiff" part operator $\frac{\partial^2}{\partial x^2}$ has a set of eigenvalues and eigenfunctions which forms a basis in $L^2([0,1]) \cap L_0^1([0,1])$:

$$\lambda_k = -\pi^2 k^2, \qquad \phi_k = \sin(\pi k x), \qquad k = 1, \dots, \infty. \tag{3.21}$$

Thus the solution of the problem (3.20) has the infinite series

$$u(t,x) = \sum_{k=1}^{\infty} a_k(t) \sin(\pi k x) \tag{3.22}$$

The action of the bounded operator $A = \cos(\pi x)$ on the eigenfunctions of B has a very simple form

$$A(x)\phi_k = \frac{1}{2}(\phi_{k+1} + \phi_{k-1}), \qquad k = 1, \dots, \infty,$$

where we introduced $\phi_0 \equiv 0$, thus we obtain

$$\alpha_k^i = \begin{cases} \frac{1}{2}, & \text{if} \quad |k-i| = 1; \\ 0, & \text{if} \quad |k-i| \neq 1. \end{cases}$$

Following our error analysis, we want to find the Frobenius norm of the leading error term:

$$\|\hat{E}_1^c\|_F^2 = \frac{\Delta t^2}{4}\sum_{i=1}^{\infty}(|\beta_{i,i+1}|^2 + |\beta_{i+1,i}|^2),$$

which for symmetric splitting methods can be simplified further

$$\|\hat{E}_1^c\|_F^2 = \frac{\Delta t^2}{2}\sum_{i=1}^{\infty}|\beta_{i,i+1}|^2$$

For the **BAB** case, using

$$\begin{aligned}\beta_{k,k+1} &= \frac{\exp(-\pi^2\Delta t k^2) - \exp(-\pi^2\Delta t(k+1)^2)}{-\pi^2\Delta t(k^2-(k+1)^2)} - \exp\left(-\pi^2\Delta t\left(k^2+k+\frac{1}{2}\right)\right)\\ &= \exp\left(-\pi^2\Delta t\left(k^2+k+\frac{1}{2}\right)\right)\left(\frac{\sinh\left(\pi^2\Delta t\left(k+\frac{1}{2}\right)\right)}{\pi^2\Delta t\left(k+\frac{1}{2}\right)} - 1\right)\end{aligned}$$

and approximating

$$\frac{\sinh(x)}{x} - 1 \approx \begin{cases} x^2/6, & \text{if} \quad 0 \le x \le 1.6 \\ \dfrac{\exp(x)}{2x} - 1, & \text{if} \quad 1.6 \le x < \infty \end{cases}$$

we find that

$$\sum_{k=1}^{\infty}\beta_{k,k+1}^2 \approx I_1 + I_2,$$

where

$$\begin{aligned} I_1 &= \int_1^{k^*}\exp\left(-2\pi^2\Delta t\left(k^2+k+\frac{1}{2}\right)\right)\frac{1}{36}\left(\pi^2\Delta t\left(k+\frac{1}{2}\right)\right)^4 dk,\\ I_2 &= \int_{k^*}^{\infty}\exp\left(-2\pi^2\Delta t\left(k^2+k+\frac{1}{2}\right)\right)\left(\frac{\exp\left(2\pi^2\Delta t\left(k+\frac{1}{2}\right)\right)}{2\left(\pi^2\Delta t\left(k+\frac{1}{2}\right)\right)^2}\right)dk \end{aligned}$$

and k^* is taken from the relation $\pi^2\Delta t\left(k^*+\frac{1}{2}\right) = 1.6$. The estimation of the integrals shows that for sufficiently small steplengths Δt, which give $k^* \gg 1$, we have

$$I_1 = O(\Delta t^{3/2}) \qquad \text{and} \qquad I_2 = o(I_1).$$

Finally, we get the order behavior of the form $\|\hat{E}_1^c\|_F = O(\Delta t^{1.75})$. The same result was obtained for **ABA** and **AB+BA** cases. For the cases **AB** and **BA** we find the order behavior $\|\hat{E}_1^c\|_F = O(\Delta t^{1.25})$.

◊

Example 3.2 Choosing $f(x) = x$, we get the equation

$$u_t = u_{xx} + xu \qquad \text{in} \qquad [0,1]. \tag{3.23}$$

This example is more complicated than Example 3.1. The action of the bounded operator $A = x$ on the eigenfunctions of B gives us

$$A(x)\phi_k = x\sin(\pi k x) = \sum_{i=1}^{\infty}\alpha_k^i\sin(\pi i x),$$

where

$$\alpha_k^i = \begin{cases} \frac{1}{2}, & \text{if} \quad k = i; \\ -\frac{8}{\pi^2}\frac{ki}{(k^2-i^2)^2}, & \text{if} \quad |i-k| \quad \text{is odd;} \\ 0, & \text{if} \quad |i-k| \quad \text{is even and nonzero;} \end{cases}$$

and the Frobenius norm of the leading error term matrix is a truly two–dimensional sum. The estimation of the double sum can be done in the same manner as we estimated one–dimensional sum in the Example 3.1, but it is more complicated and we shall not reproduce it here. We present only the final result: for sufficiently small steplengths Δt the cases **BAB**, **ABA** and **AB+BA** have the error order $\|\hat{E}_1^c\|_F = O(\Delta t^{1.5})$; and the cases **AB** and **BA** show the order $\|\hat{E}_1^c\|_F = O(\Delta t^{1.25})$. ◊

From these examples we can see that different factors contribute to the order reduction of the splittings. According to (3.13) the order behavior is determined by the joint contribution of the elements $\beta_{i,k}$, which are Δt dependent, and the coefficients α_k^i, which do not depend on Δt. It is perhaps not surprising that different values α_k^i can cause different types of order reduction. In fact, we observe different order reductions for the Examples 3.1 and 3.2.

The analytical order estimations for the Examples 3.1 and 3.2 will be verified by numerical tests in the next point.

3.3 Numerical experiments

We can model the order reductions shown for the PDE (3.1) using some semi–discretization of the PDE. Let us introduce a uniform mesh in the interval $[0, 1]$: for some natural integer N we have mesh points $x_i = ih$, $i = 0, \dots, N+1$, $h = \frac{1}{N+1}$. The space discretization of the equation (3.19) gives us the system of ordinary differential equations for the internal points of the mesh

$$u_i' = \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + f(x_i)u_i, \qquad i = 1, \dots, N; \tag{3.24}$$

and boundary conditions $u_0(t) = u_{N+1}(t) = 0$. In the matrix form the system (3.24) can be presented as

$$u' = A_h u + B_h u, \qquad u = (u_1, \dots, u_N)^T, \tag{3.25}$$

where $A_h = \text{diag}(f(x_i))$ and B_h is the tridiagonal matrix

$$B_h = \frac{1}{h^2}\begin{bmatrix} -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 1 & -2 \end{bmatrix}$$

Of course, we can apply the order behavior analysis developed in Section 2 directly to the system (3.24) in order to find its order reduction. The result will be the same as for the order behavior which we got the PDE (3.19).

The matrix B_h has the set of eigenvalues and eigenvectors

$$\lambda_s = \frac{1}{h^2}(-2 + 2\cos(s\phi)), \qquad x_k^{(s)} = \sin(ks\phi), \qquad \phi = \pi/(N+1)$$

The matrix S, constructed from eigenvectors $S = [x^{(1)}, x^{(2)}, \cdots, x^{(N)}]$, has the elements $S_{i,j} = \sin(ij\phi)$. For our purpose it is important that the matrix S is symmetric and that $S^{-1} = \frac{2}{N+1}S$ since it implies that $\|SQS^{-1}\|_F = \|Q\|_F$ for any matrix $Q \in \mathbb{R}^{N\times N}$.

Thus, we can conclude that the system (3.24) $\|E\|_F = \|\hat{E}\|_F$. In its turn $\|\hat{E}\|_F \approx \|\hat{E}_1\|_F$ is expected to have the same reduced order as the order of $\|\hat{E}_1^c\|_F$. Consequently, we should observe experimentally the order reduction results obtained in the Examples 3.1 and 3.2.

The results of numerical experiments fully support our theoretical predictions for the equation (3.20), see Figures 7 and 8, and the equation (3.23), see Figures 9 and 10. The problems were discretized in space with $N = 50$, providing us with the ODE systems of the form (3.24). The computations were performed with the help of MATLAB package. The MATLAB function **expm** was chosen to compute the exponentials of the matrices.

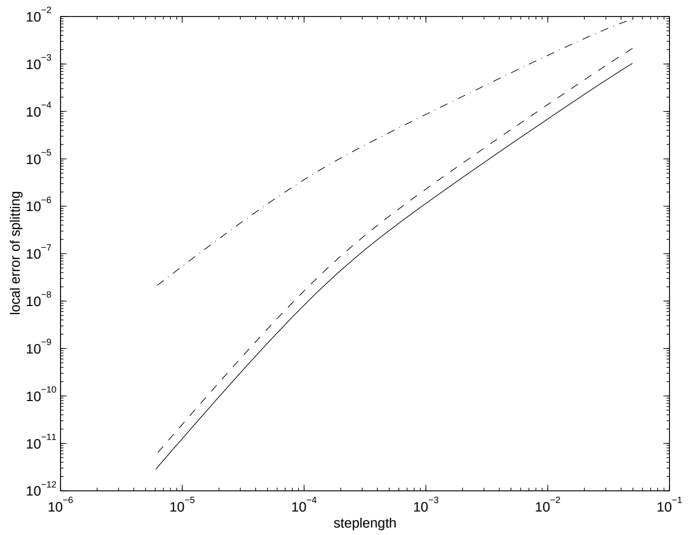

Figure 7: Frobenius norm of the error matrix E for the discretization of the problem (3.20). The solid line stands for **BAB** splitting, the dashed line for **ABA** (**AB+BA**) and the dash–dot line for the splittings **AB** and **BA**.

The figures show that for substantially small time steps the operator splitting methods show the classical orders 2 and 3. While the step size is getting larger, it reaches a break point. After this point we observe a reduced order until the step size gets too large to keep precision of the computations. The main difference between the PDE (3.1) and its semi–discretization is that for an applied splitting method the continuous equation is always "stiff" and the order is always reduced. For ODEs it is so. We have some range of step sizes for which the order of a splitting is reduced. If we decrease the step size, at some moment it will become small enough to provide the classical order of the splitting.

4 Conclusions

We have considered order reductions of the five splitting methods named **AB**, **BA**, **ABA**, **BAB** and **AB+BA** in the introduction. The splitting methods were applied to two types of problems: linear ODE systems (2.1) and linear PDEs (3.1). Using expansions for the exact solution and the splitting methods valid for moderate stepsizes, we obtained the principal error terms. These expressions explain the order reduction phenomena observed in numerical experiments.

The analysis given here does not apply to arbitrary problems, it is in fact hard to obtain precise results for the order behavior without making assumptions about the eigenspace structure of the given problem. Nevertheless, we believe that the general ideas can be used to analyze a number

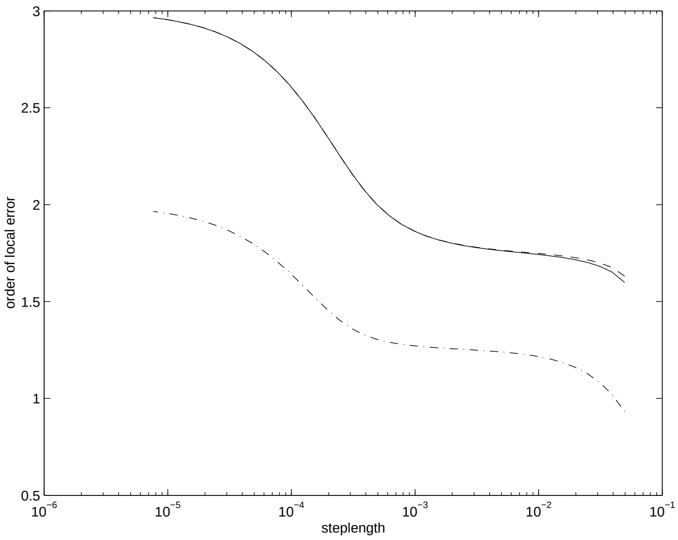

Figure 8: The order of the local error for the discretization of the problem (3.20). The same notations as in Figure 7 are used.

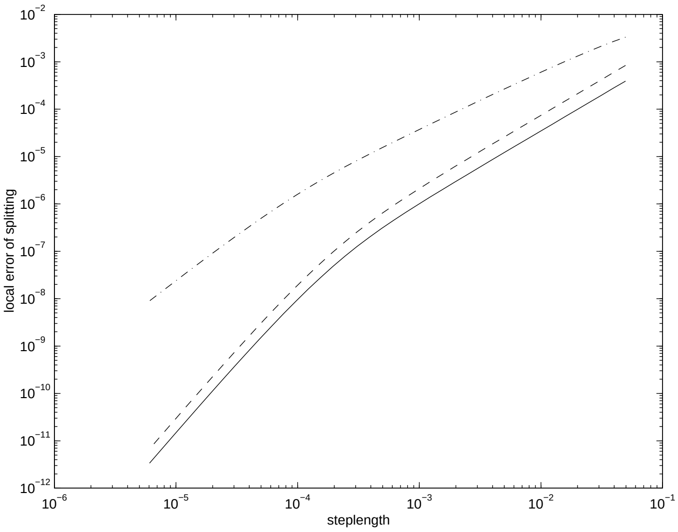

Figure 9: Discretized problem (3.23). The norm of the error matrix E. The notations are the same as in Figure 7.

of interesting applications like air pollution modeling, combustion, reactive flows, etc. (references can be found, for example, in [6]).

As a future extension of this work, we would like to consider also nonlinear problems, in which case, the expansions should be replaced by Lie series.

References

[1] G.STRANG, *On construction and comparison of difference schemes*, SIAM J. Numer. Anal. 5 (1968), pp. 506–517.

[2] N.N.YANENKO, *The Method of Fractional Steps*, New–York, 1971.

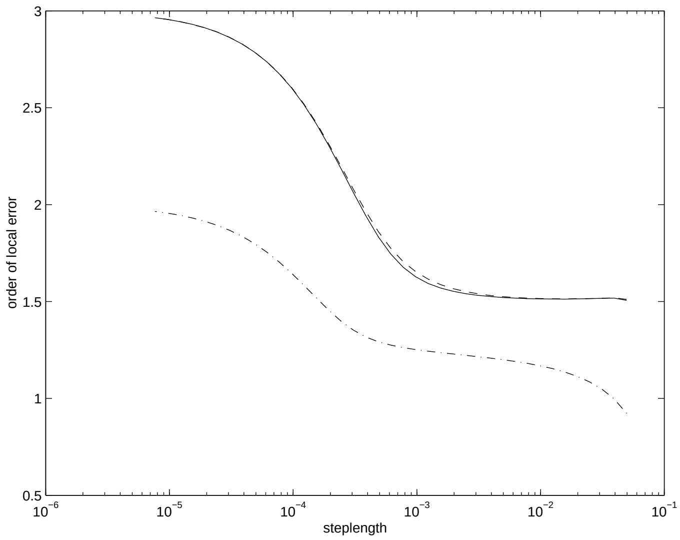

Figure 10: The local error order for the discretization of (3.23). The same notations as in Figure 7 are used.

[3] L.LANSER and J.G.VERWER, *Analysis of operator splitting for advection–diffusion–reaction problems in air pollution modeling*, In *Proceedings 2nd Meeting on Numerical methods for differential equations*, Coimbra, Portugal, February 1998.

[4] J.M.SANZ–SERNA, *The state of the art in numerical analysis*, chapter Geometrical integration, Clarendon Press, Oxford, 1997, pp. 121–143.

[5] H.YOSHIDA, *Construction of higher order symplectic integrators*, Physics Letters A 150, (1990), pp. 262–268.

[6] B.SPORTISSE, *An analysis of operator splitting techniques in the stiff case*, Technical Report 98–127, CERMICS, 1998.

[7] J.G.VERWER and B.SPORTISSE, *A note on operator splitting analysis in the stiff linear case*, Technical Report MAS–R9830, CWI, 1998.

[8] B.SPORTISSE, G.BENCTEUX and P.PLION, *Method of lines versus operator splitting for reaction–diffusion systems with fast chemistry*, In *Proceedings APMS'98*, ENPC–INRIA, October 26–29 1998.

[9] A.OSTERMANN and M.ROCHE, *Runge–Kutta methods for partial differential equations and fractional orders of convergence*, Math.Comp. 59, (1992), pp. 403–420.

[10] C.GONZÁLEZ and A.OSTERMANN, *Optimal convergence results for Runge–Kutta discretizations of linear nonautonomous parabolic problems*, BIT 39, (1999), pp. 79–95.

[11] A.OSTEMANN and M.THALHAMMER, *Non–smooth data error estimates for linearly implicit Runge–Kutta methods*, Preprint 1-1999, Institute of Mathematics and Geometry, University of Innsbruck, Austria.

[12] A.PASY, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer, New York, 1983.

[13] D.HENRY, *Geometric Theory of Semilinear Parabolic Equations*, LNM 840, Springer, Berlin, 1981.