Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Computational Physics 195 (2004) 576–593
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

The behaviour of the local error in splitting methods applied to stiff problems ☆

Roman Kozlov [a], Anne Kværnø [b], Brynjulf Owren [b,*]

[a] Department of Informatics, University of Oslo, Norway
[b] Department of Mathematical Sciences, NTNU, Norway

Received 17 February 2003; received in revised form 1 October 2003; accepted 3 October 2003

Abstract

Splitting methods are frequently used in solving stiff differential equations and it is common to split the system of equations into a stiff and a nonstiff part. The classical theory for the local order of consistency is valid only for stepsizes which are smaller than what one would typically prefer to use in the integration. Error control and stepsize selection devices based on classical local order theory may lead to unstable error behaviour and inefficient stepsize sequences. Here, the behaviour of the local error in the Strang and Godunov splitting methods is explained by using two different tools, Lie series and singular perturbation theory. The two approaches provide an understanding of the phenomena from different points of view, but both are consistent with what is observed in numerical experiments.

Keywords: Time integration; Geometric integration; Splitting methods; Numerical analysis; Order reduction; Singular perturbation

1. Introduction

Since the important early contributions by Marchuk [11] and Yanenko [18], splitting methods have steadily increased their popularity, and today they constitute an invaluable tool in several areas of computational mathematics. For instance, in the area of geometric integration, such splitting methods are frequently used to obtain structure preserving algorithms [7,12]. In some large scale engineering problems, operator splitting may be the only known practical way of carrying out time integration. Splitting is used in different ways, sometimes one applies splitting to the space dimensions, like in the original work of Strang [16]. Another much used possibility is to split according to some physical phenomenon, for instance by

☆ This work was in part sponsored by Center for Advanced Study, Oslo.
* Corresponding author. Present address: The Norwegian University of Science and Technology, 7491 Trondheim, Norway. Tel.: +47-7-3593518; fax: +47-7-3593524.
E-mail addresses: Roman.Kozlov@ifi.uio.no (R. Kozlov), Anne.Kvarno@math.ntnu.no (A. Kværnø), Brynjulf.Owren@math.ntnu.no (B. Owren).
URLs: http://www.math.ntnu.no/~anne/, http://www.math.ntnu.no/~bryn/.

0021-9991/$ - see front matter
doi:10.1016/j.jcp.2003.10.011

splitting linear stiff diffusion terms and nonlinear convection terms and integrating each of them separately. Recently, many authors have investigated splitting methods for PDEs and in particular studied the order of convergence of the local and global error [1,3,9].

In what follows, we shall study the behaviour of the local error in splitting methods used to integrate stiff ordinary differential equations. We claim and demonstrate through examples that this behaviour may have severe consequences for the performance of standard local stepsize control devices and so the design of variable stepsize integrators for such schemes should be given careful attention. Some ideas similar to those presented here have been used also by Lubich in [10].

Generally, we start from a system of autonomous ordinary differential equations, say

$$y' = \frac{\mathrm{d}y}{\mathrm{d}t} = F(y), \quad y(0) = y_0. \tag{1}$$

The solution space may well be a manifold or in many cases simply (some open subset of) Euclidean space. Splitting can be described by a decompositon of F into a sum of two or more terms, for simplicity, say

$$F(y) = A(y) + B(y). \tag{2}$$

Applying a splitting method to this problem means that we compose solutions of each of the two problems

$$y' = A(y) \quad \text{and} \quad y' = B(y), \tag{3}$$

over small time intervals. To ease the notation we introduce flow maps, e.g. we denote by $y(h) = \exp(hF)y_0$ the solution of (2) with initial condition $y(0) = y_0$. Thus, one simple splitting method is obtained by calculating $y_1 \approx y(h)$ as

$$y_1 = \exp(hA)\exp(hB)y_0.$$

It is well known and easy to prove that whenever the operators A and B are sufficiently smooth, the local error behave as

$$\exp(hF)y_0 - \exp(hA)\exp(hB)y_0 = \mathcal{O}(h^2) \quad \text{as } h \to 0. \tag{4}$$

In (1), one typically imposes a Lipschitz condition on the vector field F, and reasonable splitting methods therefore also have Lipschitz continuous A and B, but for stiff problems we often have one or more stiffness parameters such that F is not uniformly bounded in these parameters. Such a situation may for instance occur if, say

$$F(y) = A(y) + B(y) = A(y) + \frac{1}{\varepsilon}\tilde{B}(y),$$

where $\varepsilon > 0$. For small values of ε we typically observe the local error behaviour of (4) only when $h < \varepsilon$. This phenomenon is similar to what we see in the theory of order reduction in Runge–Kutta methods first discussed in [4,14] and further elaborated by Dekker and Verwer [2].

In the rest of the paper, we will always assume that the two terms of the splitting is a nonstiff vector field A and a stiff vector field B. We consider two first-order splitting methods BA, AB, and two second-order splitting methods BAB and ABA, defined as follows:

$$\begin{aligned}
&\text{BA} && y_1 = \exp(hB)\exp(hA)y_0,\\
&\text{AB} && y_1 = \exp(hA)\exp(hB)y_0,\\
&\text{BAB} && y_1 = \exp\left(\frac{h}{2}B\right)\exp(hA)\exp\left(\frac{h}{2}B\right)y_0,\\
&\text{ABA} && y_1 = \exp\left(\frac{h}{2}A\right)\exp(hB)\exp\left(\frac{h}{2}A\right)y_0.
\end{aligned} \tag{5}$$

In Section 2 we present an example of a stiff–nonstiff splitting which serves as a motivation for the further studies. We then analyse the local error in splitting methods in two different ways. First we use an approach based on tools from the theory of Lie series, and then we will repeat the analysis by means of singular perturbation theory. The two approaches provide an understanding of the problems from two different points of view, each having its strengths and weaknesses.

2. An example

A very popular test case for ODE solvers is the Van der Pol oscillator which can be formulated as a second-order ODE

$$x'' + \frac{1}{\varepsilon}(x^2 - 1)x' + x = 0, \quad x(0) = x_0, \;\; x'(0) = \dot{x}_0.$$

We rewrite this into a first-order system, setting $y := x, \; z := x'$, and split as follows:

$$\begin{bmatrix} y' \\ z' \end{bmatrix} = \begin{bmatrix} z \\ -y \end{bmatrix} + \frac{1}{\varepsilon}\begin{bmatrix} 0 \\ (1-y^2)z \end{bmatrix} = A(v) + \frac{1}{\varepsilon}\tilde{B}(v) \tag{6}$$

with $v = (y, z)$. The flows of each of the two vector fields on the right-hand side can be computed exactly, the first, A is a rotation in the yz-plane, the second, $B = (1/\varepsilon)\tilde{B}$, leaves y constant and decays z exponentially with time constant proportional to $1/\varepsilon$, thus the second vector field is stiff near the initial point for small values of ε.

This equation is now solved by a variable stepsize scheme based on splitting. The solution is advanced by the second-order scheme BAB, and the local error estimate is given by

$$\text{le} = \|v_{n+1} - \tilde{v}_{n+1}\|,$$

where $\tilde{v}_{n+1}$ is computed by the first-order splitting scheme BA. The stepsize is adjusted to ensure the local error to be less than the tolerance, however the stepsize is also restricted above by 0.1. The results of the simulation in terms of the solution and the step sizes chosen by the code are given in Fig. 1. There is a good

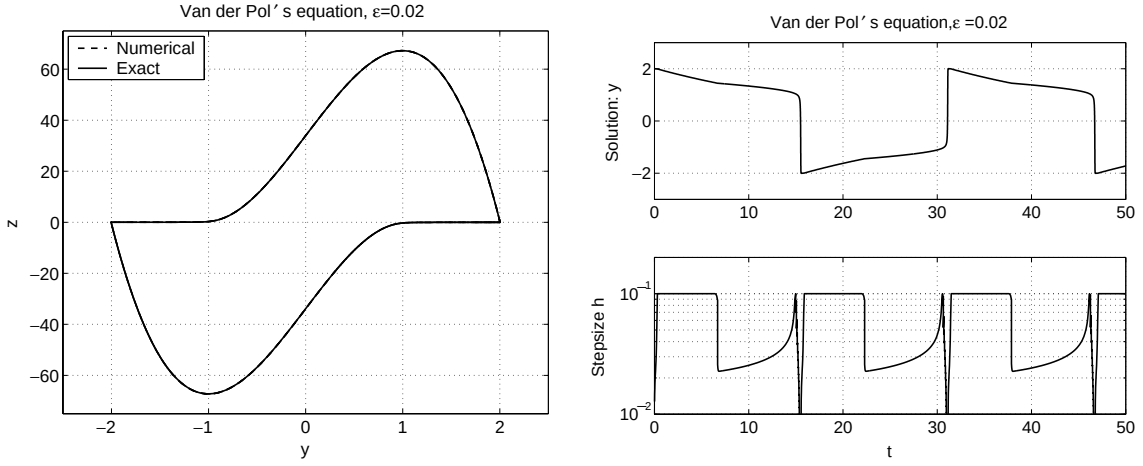

Fig. 1. To the left: The limit cycle of Van der Pol's equation. To the right: The numerical solution of y as well as the stepsize sequence.

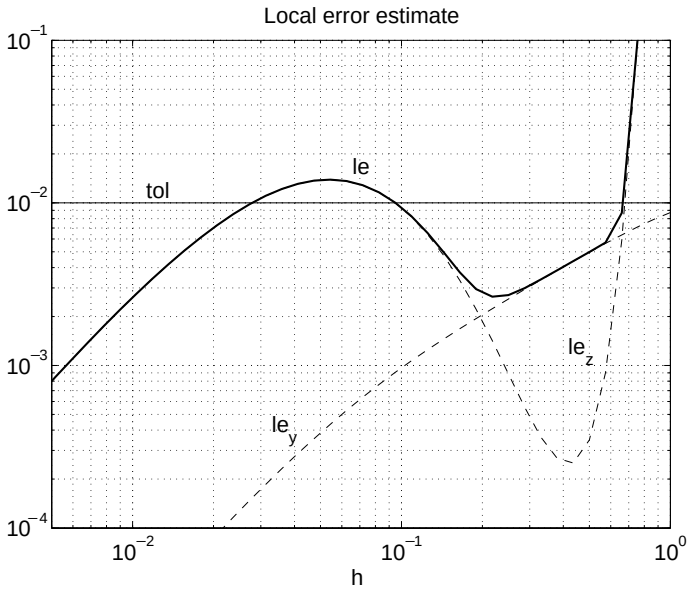

Fig. 2. Local error estimate at $t = 6.71$.

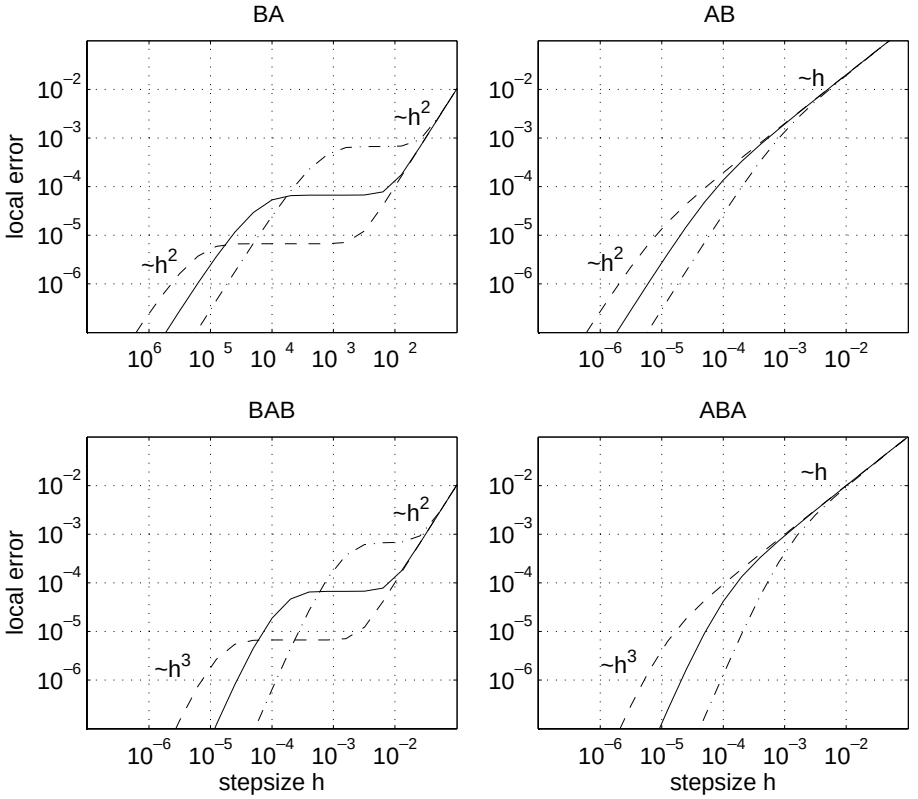

Fig. 3. Local error of van der Pol's equation, with $\varepsilon = 10^{-3}$ (— · —), $\varepsilon = 10^{-4}$ (—) and $\varepsilon = 10^{-5}$ (— —).

agreement between the phase plots of the numerical and the exact limit cycle, even if there is a considerable phase error present. The bottom right picture shows a more curious situation. There is a severe drop in the stepsize at $t \approx 6.7$ without any apparent reason. The cause of this drop is made apparent by considering the local error estimate as a function of the stepsize h, see Fig. 2.

This phenomenon is usually referred to as the "hump", e.g. [8, p. 113]. Another illustration of the situation is given in Fig. 3, giving a plot of the local errors for different values of ε. For $h \lesssim \varepsilon$ the error behaves as expected from the classical error analysis. For $h \gtrsim \varepsilon$ the order is reduced to 1 for the AB and the ABA schemes, while the error is constant at the size of ε for the BA and the BAB schemes. However, for $h \gtrsim \sqrt{\varepsilon}$ the order is 2 for both methods. By these pictures, it is apparent that an error estimator based on a combination of the BA and the BAB schemes is bound to fail. In fact, the dominant error terms for the BA and the BAB scheme are the same for $h \gtrsim \varepsilon$, causing the odd behaviour of the error estimate as can be seen in Fig. 2.

The "hump" phenomenon has been explained for Runge–Kutta and Rosenbrock methods by Hairer et al. [5,6] by consideration of the singular perturbation problems. A similar analysis for splitting methods is performed in Section 4. Before that however, we will perform a local error analysis by means of Lie series.

3. Local error analysis with Lie series

The analysis we present in this section is based on the use of Lie series, see for instance Olver [13]. Suppose that $\mathscr{D}$ is some open subset of $\mathbb{R}^m$. We denote by $C^\omega(\mathscr{D}, \mathbb{R})$ the analytic functions on $\mathscr{D}$. All differential equations we consider belong to the set of analytic vector fields on $\mathscr{D}$, denoted $\mathfrak{X}(\mathscr{D})$. One can also think of $\mathscr{D}$ as a local coordinate chart belonging to some manifold $\mathscr{M}$, most of the discussion that follows make only local considerations. Choosing coordinates $x_1, \ldots, x_m$, the vector field F can be written in the form $(F_1(x), \ldots, F_m(x))$, but it is useful to associate F with the differential operator

$$F_1(x)\frac{\partial}{\partial x_1} + \cdots + F_m(x)\frac{\partial}{\partial x_m}. \tag{7}$$

For instance, in (6) we would write

$$A(y) = y_2\frac{\partial}{\partial y_1} - y_1\frac{\partial}{\partial y_2}, \quad B(y) = \frac{1}{\varepsilon}(1 - y_1^2)y_2\frac{\partial}{\partial y_2}. \tag{8}$$

Thus, in this sense, vector fields are operators which act on functions defined on $\mathscr{D}$. In particular, the operators satisfy the important Leibniz' rule

$$F[\psi \cdot \varphi] = F[\psi] \cdot \varphi + \psi \cdot F[\varphi],$$

for any two analytic functions φ, ψ. Here we have used square brackets to signify that a vector field acts on a function, e.g. $F : \psi \mapsto F[\psi]$. The definition behaves naturally under coordinate transformations, so we define a vector field as a linear operator $F : C^\omega(\mathscr{D}, \mathbb{R}) \to C^\omega(\mathscr{D}, \mathbb{R})$.

The usefulness of this way of interpreting vector fields is evident when we for $\psi \in C^\omega(\mathscr{D}, \mathbb{R})$, $t \in \mathbb{R}$, $p \in \mathscr{D}$, and $F \in \mathfrak{X}(\mathscr{D})$ consider the expansion

$$\psi(\exp(tF)p) = \psi(p) + tF[\psi](p) + \frac{t^2}{2}F^2[\psi](p) + \cdots = \exp(tF)[\psi](p). \tag{9}$$

The powers of the vector fields is defined in the obvious way, $F^2[\psi] = F[F[\psi]]$, etc. Also in the sequel, we will sometimes make formal calculations with series without considering their convergence properties.

We now introduce the Lie Poisson bracket, $[\cdot,\cdot] : \mathfrak{X}(\mathscr{D}) \times \mathfrak{X}(\mathscr{D}) \to \mathfrak{X}(\mathscr{D})$, between two vector fields A and B as the commutator,

$$C = [A, B] = AB - BA. \tag{10}$$

In terms of local coordinates, $x_1, \ldots, x_m$, we have $C = \sum_i C_i(\partial/\partial x_i)$ where we easily calculate from Eqs. (7) and (10)

$$C_i = \sum_{j=1}^{m} \left(A_j \frac{\partial B_i}{\partial x_j} - B_j \frac{\partial A_i}{\partial x_j} \right).$$

It will be useful to define the operator $\mathrm{ad}_B : \mathfrak{X}(\mathscr{D}) \to \mathfrak{X}(\mathscr{D})$ as $\mathrm{ad}_B(A) = [B, A]$.

The main idea of the analysis lies in studying various parts of the Lie series expansions for the vector fields $F = A + B$, keeping in mind that the stiff vector field B and its powers do not give any useful information since no appropriate bounds can be obtained, whereas we will assume that $\exp(t(A + B))$ as well as $\exp(tB)$ can be uniformly bounded in $t \geqslant 0$.

3.1. Expanding the exact solution

In the analysis of the local error, we consider the difference

$$\psi(\exp(h(A + B))p) - \psi(\Phi_h p), \quad p \in \mathscr{D},$$

where Φ_h can be any of the splitting approximations (5). In what follows, we are going to make formal calculations with Lie series (9) and in manipulating these series we are not going to discuss convergence properties, but just refer to the a priori analyticity assumption.

We compute

$$\exp(h(A + B))[\psi](p) = \psi(p) + h(A + B)[\psi](p) + \cdots$$

Thus, we shall make formal calculations with the operator series $\exp(h(A + B))$ as well as series of the type

$$\sum_k \alpha_k \mathrm{ad}_B^k(A) = \phi(\mathrm{ad}_B)(A),$$

in terms of analytic functions $\phi(\zeta) = \sum_{k=0}^{\infty} \alpha_k \zeta^k$. In particular we shall make use of the fact that

$$\exp(hB)\phi(\mathrm{ad}_B)(A)[\psi](p) = \phi(\mathrm{ad}_B)(A)[\psi](\exp(hB)p).$$

We start by considering the following formula, easily proved by induction:

$$(A + B)^m = B^m + \sum_{\ell=0}^{m-1} B^\ell A (A + B)^{m-\ell-1}.$$

Substituting into the series for $\exp(t(A + B))$ we obtain

$$\exp(t(A + B)) = \exp(tB) + \sum_{m=1}^{\infty} \frac{t^m}{m!} \sum_{\ell=0}^{m-1} B^\ell A (A + B)^{m-\ell-1}.$$

An even more compact and convenient form can be obtained by using the identity

$$\int_0^t \frac{s^k}{k!} \frac{(t - s)^\ell}{\ell!} \mathrm{d}s = \frac{t^{k+\ell+1}}{(k + \ell + 1)!},$$

thus by substituting $m = k + \ell + 1$ in the summation above, one gets

$$\exp(t(A+B)) = \exp(tB) + \int_0^t \exp((t-s)B)A\exp(s(A+B))\mathrm{d}s.$$

This can be interpreted as the linear variation of constants formula, see e.g. [9], but as shown, it makes perfect sense also as formal calculation with Lie series. Applying the formula once again to the term $\exp(s(A+B))$, we get

$$\exp(t(A+B)) = \exp(tB) + \int_0^t \exp((t-s)B)A\exp(sB)\mathrm{d}s + R, \tag{11}$$

where

$$R = \int_0^t \int_0^s \exp((t-s)B)A\exp((t-\sigma)B)A\exp(\sigma(A+B))\mathrm{d}\sigma\,\mathrm{d}s.$$

We apply the well-known identity

$$\exp(-sB)A\exp(sB) = \mathrm{Ad}_{\exp(-sB)}(A) = \exp(-\mathrm{ad}_{sB})(A) \tag{12}$$

in (11), where Ad_C is the linear operator defined as

$$\mathrm{Ad}_C(A) = CAC^{-1}. \tag{13}$$

We get

$$\exp(t(A+B)) = \exp(tB) + \exp(tB)\int_0^t \exp(-s\,\mathrm{ad}_B)(A)\mathrm{d}s + R,$$

which we can write in the form

$$\exp(t(A+B)) = \exp(tB) + t\exp(tB)\phi(\mathrm{ad}_{tB})(A) + R, \tag{14}$$

where ϕ is the analytic function

$$\phi(\zeta) = \frac{1-\exp(-\zeta)}{\zeta}. \tag{15}$$

We observe that the assumed bounds on $\exp(tB)$, $\exp(t(A+B))$ for $t > 0$, and A immediately imply that $R = \mathcal{O}(h^2)$.

3.2. The numerical solution

We consider the splitting methods BAB, ABA, BA, AB. One should note that the ordering of exponentials is reversed when passing from the flow map to the formal series of operators. That is, given two vector fields A and B in $\mathfrak{X}(\mathscr{D})$, $\psi \in C^\omega(\mathscr{D},\mathbb{R})$ and $p \in \mathscr{D}$ we find

$$\begin{aligned}\psi(\exp(A)\exp(B)p) &= \exp(A)[\psi](\exp(B)p) = \exp(B)[\exp(A)[\psi]](p)\\ &:= \exp(B)\exp(A)[\psi](p).\end{aligned} \tag{16}$$

We get by formal calculations, expanding the flow of the nonstiff vector field A.

BAB

$$\exp\left(\frac{hB}{2}\right)\exp(hA)\exp\left(\frac{hB}{2}\right) = \exp(hB) + h\exp(hB)\mathrm{Ad}_{\exp(-\frac{h}{2}B)}(A) + \mathcal{O}(h^2),$$

ABA

$$\exp\left(\frac{hA}{2}\right)\exp(hB)\exp\left(\frac{hA}{2}\right) = \exp(hB) + \frac{h}{2}\exp(hB)\big(A + \mathrm{Ad}_{\exp(-hB)}(A)\big) + \mathcal{O}(h^2),$$

BA

$$\exp(hA)\exp(hB) = \exp(hB) + h\exp(hB)\mathrm{Ad}_{\exp(-hB)}(A) + \mathcal{O}(h^2),$$

AB

$$\exp(hB)\exp(hA) = \exp(hB) + h\exp(hB)(A) + \mathcal{O}(h^2).$$

We now use identity (12) together with the expansion for the exact flow (14) to obtain in all four cases an expression for the operators involved in the local truncation error of the form

$$E_{\mathrm{loc}}(x) = h\Phi(\mathrm{ad}_{hB})(A)(\mathrm{e}^{hB}x), \tag{17}$$

where Φ is an analytic function

$$\Phi(\zeta) = \phi(\zeta) - \tilde{\phi}(\zeta),$$

$\phi(\zeta)$ is given by (15), and $\tilde{\phi}(\zeta)$ is as in the following table:

Type	BAB	ABA	BA	AB
$\tilde{\phi}(\zeta)$	$\mathrm{e}^{-\zeta/2}$	$\frac{1}{2}(1+\mathrm{e}^{-\zeta})$	$\mathrm{e}^{-\zeta}$	1

The graphs of each function $\Phi(\zeta)$ are displayed in Fig. 4. In particular, we note the behaviour of $|\Phi(\zeta)|$ near $\zeta = \infty$:

Type	BAB	ABA	BA	AB
$\lvert\Phi(\zeta \to \infty)\rvert$	$\frac{1}{\zeta}$	$\frac{1}{2}$	$\frac{1}{\zeta}$	1

3.3. Linear stiff and polynomial nonstiff vector fields

We consider the important case when the stiff vector field B is linear with constant coefficients. In this section the stiffness parameter ε is not present in the problem, we do not assume that the vector field is of the form $(1/\varepsilon)\tilde{B}$, but rather that the stiffness is reflected in the size of the eigenvalues of B.

This case is important for instance when one uses a linearization of a general vector field and thereby extracts the stiff part as a linear term and leaves the nonlinear part nonstiff. In PDEs one frequently has the situation that there is a stiff linear part which is integrated by an implicit scheme whereas the somewhat less stiff nonlinear part is integrated with an explicit scheme.

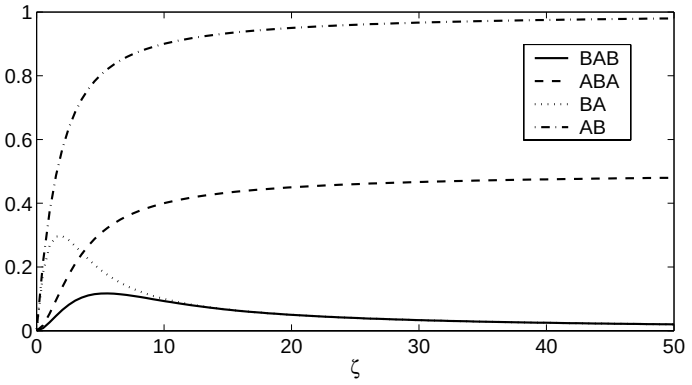

Fig. 4. $|\Phi(\zeta)|$ for various splitting methods.

Let $\mathscr{D} \subset \mathbf{R}^m$ have coordinates $x = (x_1, \ldots, x_m)$. It is well known [17, p. 13] that the tangent space at each point in $\mathscr{D}$ is naturally isomorphic to $\mathscr{H}^*_{m,1}$, the dual of the space of linear polynomials with vanishing constant term. In the above coordinates, this space has as basis $(\partial/\partial x_1, \ldots, \partial/\partial x_m)$. Locally, we can model the tangent bundle as $\mathscr{D} \times \mathscr{H}^*_{m,1}$. We consider the vector field $B : x \mapsto \sum_{i,j} b_{ij} x_j (\partial/\partial x_i)$. The nonstiff polynomial vector field A is of the form

$$A(x) = \sum_i a_i(x) \frac{\partial}{\partial x_i},$$

where each $a_i(x)$ is a polynomial in $x_1, \ldots, x_m$ of degree q. Denote this space by $\mathscr{P}_{m,q}$, and we denote by $\mathscr{H}^*_{m,1} \otimes \mathscr{P}_{m,q} \subset \mathfrak{X}(\mathscr{D})$ the corresponding space of polynomial vector fields. It is convenient for us to further divide the space $\mathscr{P}_{m,q}$ into a direct sum of its homogeneous components according to the degree of the polynomials

$$\mathscr{P}_{m,q} = \mathscr{H}_{m,0} \oplus \cdots \oplus \mathscr{H}_{m,q}.$$

Similarly, $\mathscr{H}^*_{m,1} \otimes \mathscr{H}_{m,q}$ is the corresponding space of vector fields whose components are homogeneous polynomials. The crucial observation now is that for linear vector fields, the operator $\mathrm{ad}_B : \mathfrak{X}(\mathscr{D}) \to \mathfrak{X}(\mathscr{D})$ is invariant on every subspace $\mathscr{H}^*_{m,1} \otimes \mathscr{H}_{m,k}$. The dimension of $\mathscr{H}_{m,q}$ is $\binom{m+q-1}{m-1}$ and one can for instance use a basis of monomials indexed by $(i_1, \ldots, i_q)$ where $1 \leqslant i_1 \leqslant \cdots \leqslant i_q \leqslant m$, of the form $x_{i_1} \cdot x_{i_2} \cdots x_{i_q}$.

To begin with, note that the linear vector field B, acting as a derivation operator, is an endomorphism of $\mathscr{H}_{m,q}$.

Lemma 3.1. *Suppose that B, acting as a derivation on $\mathscr{H}_{m,1}$, has m linearly independent eigenvectors $p_1, \ldots, p_m$ corresponding to eigenvalues $\lambda_1, \ldots, \lambda_m$, so that $B[p_i] = \lambda_i p_i$. Then B, acting as a derivation on $\mathscr{H}_{m,q}$ has $\binom{m+q-1}{m-1}$ eigenvalues with corresponding linearly independent eigenvectors, indexed by $1 \leqslant i_1 \leqslant \cdots \leqslant i_q \leqslant m$ such that*

$$\lambda_{(i_1,\ldots,i_q)} = \sum_{r=1}^{q} \lambda_{i_r}, \quad p_{(i_1,\ldots,i_q)} = \prod_{r=1}^{m} p_{i_r}.$$

Proof. That the listed vectors are linearly independent is clear, since the change of coordinates $(y_1, \ldots, y_m) = (p_1(x), \ldots, p_m(x))$ turns the polynomials $p_{(i_1,\ldots,i_q)}$ into the standard monomial basis for $\mathscr{H}_{m,q}$ in the new variables $y_1, \ldots, y_m$. The derivation property of B yields

$$B[p_{(i_1,\ldots,i_q)}] = B[p_{i_1}\cdots p_{i_q}] = \sum_{\ell=1}^{q} B[p_{i_\ell}] \prod_{\substack{r=1\\ r\neq\ell}}^{q} p_{i_r} = \left(\sum_{\ell=1}^{q} \lambda_{i_\ell}\right) p_{(i_1,\ldots,i_q)}. \qquad \square$$

The operator ad_B as an endomorphism of the space $\mathscr{H}^*_{m,1} \otimes \mathscr{H}_{m,q}$ is now constructed as

$$\mathrm{ad}_B = I^* \otimes B - B^* \otimes I,$$

I^* and I being the identity operators on $\mathscr{H}^*_{m,1}$ and $\mathscr{H}_{m,q}$, respectively. Note also that here B^* is the dual operator of B.

Clearly the two terms in this expression for ad_B commute so they share a common set of eigenvectors. We have:

Theorem 3.2. *The operator* ad_B *acts as an endomorphism on* $\mathscr{H}^*_{m,1} \otimes \mathscr{H}_{m,q}$. *Suppose that B is nondefective as an endomorphism of* $\mathscr{H}_{m,1}$, *with right eigenvectors* $p_i \in \mathscr{H}_{m,1}$ *and left eigenvectors* $\alpha_i \in \mathscr{H}^*_{1,m}$, $1 \leqslant i \leqslant m$. *Then its eigenvalues indexed by* $1 \leqslant j \leqslant m$, $1 \leqslant i_1 \leqslant \cdots \leqslant i_q \leqslant m$ *are*

$$\lambda_{\iota,j} := \lambda_{i_1} + \lambda_{i_2} + \cdots + \lambda_{i_q} - \lambda_j$$

with corresponding eigenvectors

$$\alpha_j \otimes p_{(i_1,\ldots,i_q)},$$

$p_{(i_1,\ldots,i_q)} = \prod_{r=1}^{q} p_{i_r} \in \mathscr{H}_{m,q}$.

Proof. Writing $\mathbf{i} = (i_1,\ldots,i_q)$ and $\lambda_{\mathbf{i}} = \sum_{r=1}^{q} \lambda_{i_r}$ we get

$$\begin{aligned}(I^* \otimes B - B^* \otimes I)(\alpha_j \otimes p_{\mathbf{i}}) &= (\alpha_j \otimes B[p_{\mathbf{i}}]) - (B^*\alpha_j \otimes p_{\mathbf{i}}) = (\alpha_j \otimes \lambda_{\mathbf{i}} p_{\mathbf{i}}) - (\lambda_j \alpha_j \otimes p_{\mathbf{i}}) \\ &= (\lambda_{\mathbf{i}} - \lambda_j)(\alpha_j \otimes p_{\mathbf{i}}). \qquad \square\end{aligned}$$

Introducing the projectors $\Pi_{\mathbf{i},j} = \Pi^*_j \otimes \Pi_{\mathbf{i}}$ onto the invariant subspace corresponding to the eigenvalue $\lambda_{\mathbf{i},j}$ of ad_B, we write

$$\mathrm{ad}_B = \sum_{\mathbf{i},j} \lambda_{\mathbf{i},j} \Pi^*_j \otimes \Pi_{\mathbf{i}}.$$

Since the $\Pi_{\mathbf{i},j}$ have the properties

$$\sum_{\mathbf{i},j} \Pi_{\mathbf{i},j} = I, \quad \Pi_{\mathbf{i},j} \bigcirc \Pi_{\mathbf{i}',j'} = \delta_{(\mathbf{i},j),(\mathbf{i}',j')} \Pi_{\mathbf{i},j},$$

we get for any analytic function Φ that

$$\Phi(\mathrm{ad}_B) = \sum_{\mathbf{i},j} \Phi(\lambda_{\mathbf{i},j}) \Pi_{\mathbf{i},j},$$

so that in terms of the eigenvector decomposition of A,

$$A = \sum_{\mathbf{i},j} a_{\mathbf{i},j} (\alpha_j \otimes p_{\mathbf{i}}).$$

From (17) we get that for any nonstiff polynomial vector field A the principal term of the local error is

$$E_{\mathrm{loc}}(x) = h\sum_{\mathbf{i},j} a_{\mathbf{i},j}\Phi(h\lambda_{\mathbf{i},j})(\alpha_j \otimes p_{\mathbf{i}}(\exp(hB)x)) = h\sum_{\mathbf{i},j} a_{\mathbf{i},j}\Phi(h\lambda_{\mathbf{i},j})\mathrm{e}^{h\lambda_{\mathbf{i}}}(\alpha_j \otimes p_{\mathbf{i}}(x)).$$

If we split the eigenvalues as $\lambda_{\mathbf{i},j} = \lambda_{\mathbf{i}} - \lambda_j$, we can further simplify the expression and we find that

$$E_{\mathrm{loc}}(x) = h\sum_{\mathbf{i},j} a_{\mathbf{i},j}\Phi_1(h\lambda_{\mathbf{i}}, h\lambda_j)(\alpha_h \otimes p_{\mathbf{i}}(x)), \tag{18}$$

$$\Phi_1(u,v) = \phi_1(u,v) - \tilde{\phi}_1(u,v),$$

where

$$\phi_1(u,v) = (\exp(u) - \exp(v))/(u-v),$$

and $\tilde{\phi}_1(u,v)$ depends on the splitting methods as follows:

Type	BAB	ABA	BA	AB
$\tilde{\phi}_1(u,v)$	$\mathrm{e}^{(u+v)/2}$	$\frac{1}{2}(\mathrm{e}^u + \mathrm{e}^v)$	e^v	e^u

To further analyse the behaviour of E_{loc}, one may study one mode at the time of the decomposition above. Suppose that the spectrum of B can be separated into two nonempty sets, $\sigma(B) = \sigma_s(B) \cup \sigma_{ns}(B)$, as described in [2, p. 9]. We exclude the possibility of B having large positive real parts or with large imaginary parts. The set $\sigma_s := \sigma_s(B)$ consists of eigenvalues λ with large negative real parts. In particular, we are interested in situations where the stepsize h satisfies $-\sqrt{\mathrm{Re}\,\lambda} \lessapprox 1/h \lessapprox -\mathrm{Re}\,\lambda$. The set $\sigma_{ns} := \sigma_{ns}(B)$ consists of eigenvalues belonging to some moderately sized disc centered at the origin. More to the point, we assume that with the stepsizes of interest, the function Φ_1 is always well approximated by its truncated Maclaurin series when the arguments are of the form $h\sum_k \lambda_{i_k}$ where each $\lambda_{i_k} \in \sigma_{ns}$. In terms of the stiffness parameter ε one would assume that each $\lambda_\varepsilon \in \sigma_s$ is such that $h\mathrm{Re}(\lambda_\varepsilon) \to -\infty$ as $\varepsilon \to 0^+$. For an eigenvalue $\lambda_{\mathbf{i}} = \lambda_{i_1} + \cdots + \lambda_{i_q}$ we will also say that $\lambda_{\mathbf{i}} \in \sigma_s$ if at least one $\lambda_{i_r} \in \sigma_s$ and that otherwise $\lambda_{\mathbf{i}} \in \sigma_{ns}$.

Considering the decomposition (19) of $E_{\mathrm{loc}}(x)$, we see that the modes can be divided into four classes. We discuss each of the cases, and we will slightly abuse the big-oh notation in what follows.

1. $\lambda_{\mathbf{i}} \in \sigma_s$, $\lambda_j \in \sigma_s$. In this case, $\Phi_1(u,v) = \mathcal{O}(\exp(\lambda))$ for all four splitting methods where $\lambda \in \sigma_s$ so the contribution from such modes is negligible.
2. $\lambda_{\mathbf{i}} \in \sigma_s$, $\lambda_j \in \sigma_{ns}$. Here we get

$$\phi_1(h\lambda_{\mathbf{i}}, h\lambda_j) \approx -\frac{\exp(h\lambda_j)}{h\lambda_{\mathbf{i}}}.$$

   For the for splitting methods we see that the cases BAB and AB will all have exponentially small contributions from $\tilde{\phi}_1(h\lambda_{\mathbf{i}}, h\lambda_j)$ whereas in ABA and BA it would add contributions of size $\exp(h\lambda_j) = \mathcal{O}(1)$.
3. $\lambda_{\mathbf{i}} \in \sigma_{ns}$, $\lambda_j \in \sigma_s$. Now

$$\phi_1(h\lambda_{\mathbf{i}}, h\lambda_j) \approx -\frac{\exp(h\lambda_{\mathbf{i}})}{h\lambda_j}.$$

   The situation for the 4 splitting methods is similar to the previous case except that the role of BA and AB is reversed: We can neglect the contribution of $\tilde{\phi}$ in the BAB and BA whereas both ABA and AB will have contributions of size $\exp(h\lambda_{\mathbf{i}}) = \mathcal{O}(1)$.

4. $\lambda_{\mathbf{i}} \in \sigma_{ns}$, $\lambda_j \in \sigma_{ns}$. Here, both arguments to Φ_1 are "small", so that Taylor series can be used and we consider the first nonzero term in the Maclaurin expansion of $\Phi_1(u,v)$:

Type	BAB	ABA	BA	AB
	$\frac{1}{24}(u-v)^2$	$-\frac{1}{12}(u-v)^2$	$\frac{1}{2}(u-v)$	$-\frac{1}{2}(u-v)$

These expressions correspond to the classical order results for the various splitting methods.

Summing up, one can now clearly see that the local error in the four splitting methods is composed from terms in the decomposition (18) of four different types 1–4 above. Type 1 can be ignored, and we assume the stepsize to belong to an interval where type 4 is small compared to types 2 and 3. The local error in modes of type 2 and 3 consists of two terms, corresponding to the exact and numerical solution. The exact solution always contributes with terms of type $1/\lambda_{\mathbf{i}}$ with $\lambda_{\mathbf{i}} \in \sigma_s$. The numerical solution contributes with exponentially small terms in the BAB case, but in the other three cases there will always be terms of type $h\exp(h\lambda_j) = \mathcal{O}(h)$, $\lambda_j \in \sigma_{ns}$. Note that with the range of stepsizes we consider, the term $h\exp(h\lambda_j)$, $\lambda \in \sigma_{ns}$ will dominate the term $1/\lambda_{\mathbf{i}}$, $\lambda_{\mathbf{i}} \in \sigma_s$. These results are summarized in Table 1.

3.3.1. The steady case

The results derived above are valid for arbitrary initial values, so they include also the transient phase of the integration when the problem is not necessarily considered to be stiff. In order to understand what happens after the transient has died out, one may assume that the initial value is of the form $x = \exp(tB)y$ and y is chosen arbitrarily. This just causes $\Phi_1(h\lambda_{\mathbf{i}}, h\lambda_j)$ in (18) to be replaced by $\Phi_1(h\lambda_{\mathbf{i}}, h\lambda_j)\exp(t\lambda_{\mathbf{i}})$ and x by y. The analysis differ only in the type 2 case above for the BA splitting. Rather than getting the $h\exp(h\lambda_j) = \mathcal{O}(h)$, $\lambda_j \in \sigma_{ns}$ contribution, one gets the term $h\exp(t\lambda_{\mathbf{i}} + h\lambda_j)$ with $\lambda_{\mathbf{i}} \in \sigma_s$, which can be neglected. As a consequence, the dominating term from type 2 nodes in the BA splitting is now $\mathcal{O}(1)$. See also Table 1.

One may easily apply these results to the case of general polynomial vector fields, simply by adding up the contributions from all of the homogeneous components. The approach presented here also suggests a way to consider arbitrary analytic vector fields A.

Finally, we note that the degree of commutativity between the stiff and nonstiff vector fields is here measured in terms of the size of the coefficients $a_{\mathbf{i},j}$.

3.4. Application to the van der Pol equation

In this case, the stiff vector field B is nonlinear, so the results of the previous subsection do not apply, but the general expression (18) for the local error can still be used. It is necessary to obtain information about the vector field $\Phi(\mathrm{ad}_{hB})(A)$ where A and $B = (1/\varepsilon)\tilde{B}$ are given by (8), and since Φ is an analytic function, we begin by calculating arbitrary powers $\mathrm{ad}^k_{hB}(A) = \mathrm{ad}^k_{\theta\tilde{B}}(A)$, where $\theta = h/\varepsilon$. It easily proved by induction that

$$\mathrm{ad}^k_{\tilde{B}}(A) = (1-y^2)^k z\frac{\partial}{\partial y} + y\Big(-(y^2-1)^k + 2kz^2(1-y^2)^{k-1}\Big)\frac{\partial}{\partial z}.$$

Table 1
The behaviour of the local error in general for linear B and polynomial A

Type	BAB	ABA	BA	AB
General	$\mathcal{O}(1)$	$\mathcal{O}(h)$	$\mathcal{O}(h)$	$\mathcal{O}(h)$
Steady	$\mathcal{O}(1)$	$\mathcal{O}(h)$	$\mathcal{O}(1)$	$\mathcal{O}(h)$

By substituting this expression into the analytic function Φ, we get

$$\Phi(\mathrm{ad}_{\theta\tilde{B}})(A) = \Phi(\theta(1-y^2)z)\frac{\partial}{\partial y} + \Big(2yz^2\theta\Phi'(\theta(1-y^2)) - y\Phi(\theta(y^2-1))\Big)\frac{\partial}{\partial z}.$$

This must be composed with the stiff flow map $\exp(\theta\tilde{B})$, that is, we set $y = \bar{y}$ and $z = \exp(-\alpha\theta)\bar{z}$, where $\alpha = \bar{y}^2 - 1 > 0$. Thus, stepping from the point $\bar{y} = (\bar{y}, \bar{z})$ we get

$$E_{\mathrm{loc}} = h\Phi(-\alpha\theta \mathrm{e}^{-\alpha\theta}\bar{z})\frac{\partial}{\partial y} + h\Big(2\bar{y}\bar{z}^2\theta \mathrm{e}^{-2\alpha\theta}\Phi'(-\alpha\theta) - \bar{y}\Phi(\alpha\theta)\Big)\frac{\partial}{\partial z}.$$

In the first term on the right hand side, the argument to Φ tends rapidly to zero as $\alpha\cdot\theta$ increases. In the second term, the unbounded terms in $\Phi'(-\alpha\theta)$ are killed by the premultiplication with $\exp(-2\alpha\theta)$ when $\alpha\cdot\theta$ increases. This happens in all the 4 cases of splitting methods we have considered. Finally, from the preceding discussion of the behaviour of $\Phi(\zeta)$ when ζ tends to infinity, we may for instance look at the BAB case, and we find that the third term behaves as $h\bar{y}(1/\alpha\theta)(\partial/\partial z)$ so we may conclude that

$$E_{\mathrm{loc}} \approx -\frac{\bar{y}}{\bar{y}^2-1}\cdot\varepsilon\frac{\partial}{\partial z},$$

when $\alpha\cdot\theta = (\bar{y}^2-1)(h/\varepsilon)$ is large.

4. Singular perturbation approach

In this section the local error analysis is done for singular perturbation problems. To use this approach, we will assume the vector fields to be of the form $A = (f_A, g_A)$, $B = (f_B, (1/\varepsilon)g_B)$, thus the ODE system under consideration can be written as

$$\begin{aligned} y' &= f_A(y,z) + f_B(y,z),\\ \varepsilon z' &= \varepsilon g_A(y,z) + g_B(y,z), \quad 0 < \varepsilon \ll 1. \end{aligned} \tag{19}$$

We seek solutions of the form

$$\begin{aligned} y(t) &= y_s(t) + \eta(\tau),\\ z(t) &= z_s(t) + \zeta(\tau), \quad \tau = t/\varepsilon, \end{aligned} \tag{20}$$

where $y_s(t), z_s(t)$ represents the smooth solutions and $\eta(\tau), \zeta(\tau)$ the transients. These solutions are written as power series in ε:

$$y_s(t) = y_0(t) + \varepsilon y_1(t) + \varepsilon^2 y_2(t) + \cdots, \quad z_s(t) = z_0(t) + \varepsilon z_1(t) + \varepsilon^2 z_2(t) + \cdots, \tag{21}$$

$$\eta(\tau) = \eta_0(\tau) + \varepsilon\eta_1(\tau) + \varepsilon^2\eta_2(\tau) + \cdots, \quad \zeta(\tau) = \zeta_0(\tau) + \varepsilon\zeta_1(\tau) + \varepsilon^2\zeta_2(\tau) + \cdots. \tag{22}$$

We will assume that the logarithmic norm of the Jacobian $g_{B,z}$ satisfies the condition $\mu(g_{B,z}) < -1$ in an ε-independent neighbourhood of the solution. Thus, the transients will satisfy

$$\|\eta_j(\tau)\| \leqslant \mathrm{e}^{-\kappa\tau}, \quad \|\zeta_j(\tau)\| \leqslant \mathrm{e}^{-\kappa\tau}, \qquad j = 1,2,3,\ldots,$$

for some $\kappa > 0$. See Hairer and Wanner [8] for a detailed discussion of the transients.

For $t \gtrsim \varepsilon$, which is the timescale of interest here, the transients will be damped out. Thus, we are only looking for the smooth solution. To do so, insert (21) into (19), expand the functions into power series of ε,

and then collect equal terms of ε. This procedure results in a series of differential-algebraic equations (DAEs):

$$\left.\begin{array}{l} \left.\begin{array}{l} y_0' = f(y_0, z_0) \\ 0 = g_B(y_0, z_0) \end{array}\right\}\text{Index 1} \\ \left.\begin{array}{l} y_1' = f_y(y_0, z_0)y_1 + f_z(y_0, z_0)z_1 \\ z_0' = g_A(y_0, z_0) + g_{B,y}(y_0, z_0)y_1 + g_{B,z}(y_0, z_0)z_1 \end{array}\right\}\text{Index 2,} \\ \quad\vdots \end{array}\right.$$

where $f = f_A + f_B$. In the following, we will refer to $(y_0(t), z_0(t))$ as the index 1 solution, $(y_1(t), z_1(t))$ as the index 2 solution, and so on. Since $g_{B,z}$ is nonsingular by assumption, the equation $g_B(y_0, z_0) = 0$ can be solved with respect to z_0. Inserting this solution into the first equation yields an ODE in y_0. Similarly, the fourth equation can be solved with respect to z_1, which inserted into the third equation gives an ODE for y_1, and so on. Thus, the initial values for the y_i's can be chosen freely, while the z_i's have to satisfy some algebraic constraints. For the given problem, these constraints are

$$\begin{array}{l} g_B = 0 \quad \text{Index 1 constraint,} \\ g_{B,y}f + g_{B,z}(g_A + g_{B,y}y_1 + g_{B,z}z_1) = 0 \quad \text{Index 2 constraint,} \\ \quad\vdots \end{array} \tag{23}$$

where the functions f, g_A and g_B and their derivatives are all evaluated at y_0, z_0.

To find the order of the splitting methods, we will need the power series in h of the exact smooth solutions. For the index 1 variables y_0, z_0, these series can be expressed in terms of trees, see e.g. Roche [15]. We will only need the first few terms, which are given by

$$\begin{aligned} y_0(t+h) &= y_0(t) + hf + \frac{h^2}{2}(f_y f + f_z(-g_{B,z})^{-1}g_{B,y}f) + \cdots, \\ z_0(t+h) &= z_0(t) + h(-g_{B,z})^{-1}g_{B,y}f + \frac{h^2}{2}(-g_{B,z})^{-1}\Big(2g_{B,yz}(f, (-g_{B,z})^{-1}g_{B,y}f) \\ &\quad + g_{B,zz}((-g_{B,z})^{-1}g_{B,y}f, (-g_{B,z})^{-1}g_{B,y}f) + g_{B,y}f_y f + g_{B,yy}(f,f) + g_{B,y}f_z(-g_{B,z})^{-1}g_{B,y}f\Big) + \cdots \end{aligned} \tag{24}$$

All functions and their derivatives are calculated in $y_0(t), z_0(t)$. Similar series can also be found for the higher index variables.

To analyse the numerical schemes, we will first have to discuss the flow of the two vector fields separately. Let us start with the stiff part, given by

$$\begin{array}{ll} y' = f_B(y,z), & y(t_0) = y^0, \\ \varepsilon z' = g_B(y,z), & z(t_0) = z^0. \end{array} \tag{25}$$

This equation is of the form (19), thus the smooth solution is given by (21). For the index 1 variables y_0, z_0 the power series are given by (24), using $f_A = 0$ and $g_A = 0$. The algebraic variables z_0 and z_1 have to satisfy the constraints

$$\begin{array}{l} g_B = 0 \quad \text{Index 1 constraint,} \\ g_{B,y}f_B + g_{B,z}(g_{B,y}y_1 + g_{B,z}z_1) = 0 \quad \text{Index 2 constraint,} \\ \quad\vdots \end{array} \tag{26}$$

We notice that z_0 has to satisfy the same index 1 constraint both for the full problem (19) and the stiff part (25). However, the index 2 constraint for B differs from the index 2 constraint (23) for the full problem, causing a discrepancy in the algebraic variable of size ε for all splitting methods concluding the step with the flow of B.

The nonstiff flow is the solution of the system

$$\begin{aligned} y' &= f_A(y,z), \quad y(t_0) = y^0, \\ z' &= g_A(y,z), \quad z(t_0) = z^0. \end{aligned} \tag{27}$$

The Taylor-expansion of this problem is

$$\begin{aligned} y(t_0+h) &= y^0 + hf_A + \frac{h^2}{2}(f_{A,y}f_A + f_{A,z}g_A) + \cdots, \\ z(t_0+h) &= z^0 + hg_A + \frac{h^2}{2}(g_{A,y}f_A + g_{A,z}g_A) + \cdots, \end{aligned} \tag{28}$$

where all functions and their derivatives are evaluated in y^0, z^0.

Fig. 5 illustrates this process. For the full problem $A+B$ the transient will rapidly take the solution to the manifold $\mathscr{M}_{B+A}$, given by (23). Similarly, the transient of B takes the solution to the manifold $\mathscr{M}_B$, given by the constraints (26). The flow of A moves the solution away from both manifolds. Thus we might expect an error of $\mathscr{O}(\varepsilon)$ for all methods concluding their step by B, and an error of order $\mathscr{O}(h)$ for methods concluding with A. This is consistent with the numerical results given in Fig. 3, as well as those given in Table 1, the steady case. In the following, a more refined analysis will confirm this.

Since the numerical solution is alternating between the flow of B and the flow of A, it is reasonable to assume that the initial values of the nonstiff problem A can be expressed as a power series in ε, like

$$\begin{aligned} y(t_0) &= y^0 = y_0^0 + \varepsilon y_1^0 + \varepsilon^2 y_2^0 + \cdots, \\ z(t_0) &= z^0 = z_0^0 + \Delta z_0 + \varepsilon(z_1^0 + \Delta z_1) + \varepsilon^2(z_2^0 + \Delta z_2) + \cdots. \end{aligned} \tag{29}$$

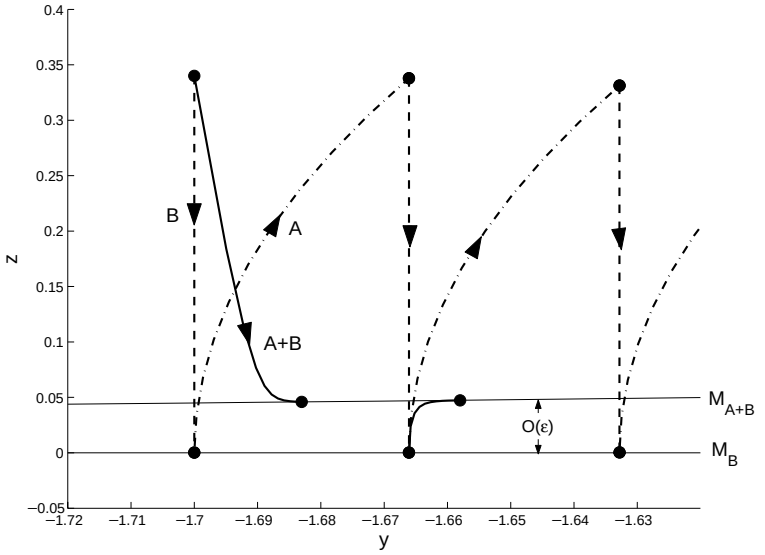

Fig. 5. The flows of $A(\text{—}\cdot\text{—}\cdot\text{—}), B(\text{——})$ and $A+B$ (—) for van der Pol's equation.

In the series of z^0, z_j^0 denote the variables satisfying the algebraic constraints (26). Thus, if the step is composed such that a step given by A is proceeding a step of B, then $\Delta z_j = 0$. This is not the case for the first step, neither will it be for the ABA-scheme. We will make the assumption that $\Delta z_j = \mathcal{O}(h)$. By inserting (29) into (28), expanding the functions around y_0^0, z_0^0, and collecting equal terms of ε we find that the solution of (27) is of the form (21), with

$$\begin{aligned}
y_0(t_0+h) &= y_0^0 + hf_A + hf_{A,z}\Delta z_0 + \frac{h^2}{2}(f_{A,y}f_A + f_{A,z}g_A) + \mathcal{O}(h^3),\\
z_0(t_0+h) &= z_0^0 + \Delta z_0 + hg_A + hg_{A,z}\Delta z_0 + \frac{h^2}{2}(g_{A,y}f_A + g_{A,z}g_A) + \mathcal{O}(h^3),\\
y_1(t_0+h) &= y_1^0 + h(f_{A,y}y_1^0 + f_{A,z}(z_1^0 + \Delta z_1)) + \mathcal{O}(h^2),\\
z_1(t_0+h) &= z_1^0 + \Delta z_1 + h(g_{A,y}y_1^0 + g_{A,z}(z_1^0 + \Delta z_1)) + \mathcal{O}(h^2),\\
&\vdots
\end{aligned} \tag{30}$$

This can now be used to find the series of the numerical solutions of the splitting schemes. Let us start with the BA-scheme. Let $(y_0^A, \tilde{z}_0^A)$ be the solution after one step following the flow of A, given by (30). This will be the inital values for the solution given by the stiff flow B. However, the inital values of the smooth solution have to satisfy $g_B(y_0^A, z_0^A) = 0$, thus

$$z_0^A = z_0^0 + h(-g_{B,z})^{-1}g_{B,y}f_A + \mathcal{O}(h^2).$$

The solution from the vector field B using (y_0^A, z_0^A) as initial values is given by

$$\begin{aligned}
y_0^1 &= y_0^A + hf_B + \frac{h^2}{2}\left(f_{B,y}f_B + f_{B,z}(-g_{B,z})^{-1}g_{B,y}f_B\right) + \mathcal{O}(h^3)\\
&= y_0^0 + hf + hf_{A,z}\Delta z_0 + \frac{h^2}{2}(f_{A,y}f_A + f_{A,z}g_A + 2f_{B,y}f_A + f_{B,y}f_B + 2f_{B,z}(-g_{B,z})^{-1}g_{B,y}f_A\\
&\quad + f_{B,z}(-g_{B,z})^{-1}g_{B,y}f_B) + \mathcal{O}(h^3),
\end{aligned}$$

the expressions being evaluated at y_0^A, z_0^A. Comparing this with the exact solution, the local error will be

$$y_0(t_0+h) - y_0^1 = -hf_z\Delta z_0 + \mathcal{O}(h^2).$$

The local error of z_0^1 will be of the same order, since both the exact and the numerical solution of this component are obtained from the algebraic constraint $g_B(y_0, z_0) = 0$.

Comparing the second constraints of Eqs. (23) and (26) yields the error in z_1^1:

$$z_1(t_0+h) - z_1^1 = (-g_{B,z})^{-1}(g_A - (-g_{B,z})^{-1}g_{B,y}f_A) + \mathcal{O}(h).$$

Thus, to conclude the analysis of the BA-scheme, the local error after one step will be

$$\begin{aligned}
y(t_0+h) - y^1 &= -hf_{A,z}\Delta z_0 + \frac{h^2}{2}(f_{A,y}f_B - f_{B,y}f_A + f_{A,z}(-g_{B,z})^{-1}g_{B,y}(f_A + f_B)\\
&\qquad - f_{B,z}(-g_{B,z})^{-1}g_{B,y}f_A - f_{A,z}g_A) + \mathcal{O}(h^3 + \varepsilon h + \varepsilon^2),\\
z(t_0+h) - z^1 &= \varepsilon(-g_{B,z})^{-1}(g_A - (-g_{B,z})^{-1}g_{B,y}f_A) + \mathcal{O}(h^2 + \varepsilon h + \varepsilon^2).
\end{aligned}$$

Thus, the local order of the y-components is 2, but it might drop to 1 if the initial value z_0 is not properly chosen. The z-components have a constant error of size ε.

A similar analysis for the remaining schemes shows that:

AB

$$y(t_0+h)-y^1=\frac{h^2}{2}(f_{B,y}f_A-f_{A,y}f_B-f_{A,z}g_A+f_{A,z}(-g_{B,z})^{-1}g_{B,y}(f_A-f_B)+f_{B,z}(-g_{B,z})^{-1}g_{B,y}f_A)+\mathcal{O}(h^3+\varepsilon h),$$
$$z(t_0+h)-z^1=h((-g_{B,z})^{-1}g_{B,y}f_A-g_A)+\mathcal{O}(h^2+\varepsilon).$$

BAB

$$y(t_0+h)-y^1=\frac{h^2}{2}(f_{A,z}(-g_{B,z})^{-1}g_{B,y}f_A-f_{A,z}g_A)+\mathcal{O}(h^3+\varepsilon h+\varepsilon^2),$$
$$z(t_0+h)-z^1=\varepsilon(-g_{B,z})^{-1}(g_A-(-g_{B,z})^{-1}g_{B,y}f_A)+\mathcal{O}(h^2+\varepsilon h+\varepsilon^2).$$

ABA

$$y(t_0+h)-y^1=\frac{h^2}{4}(f_{A,z}(-g_{B,z})^{-1}g_{B,y}f_A-f_{A,z}g_A)-\frac{h}{2}f_{A,z}\Delta z_0+\mathcal{O}(h^2+\varepsilon h+\varepsilon^2),$$
$$z(t_0+h)-z^1=\frac{h}{2}((-g_{B,z})^{-1}g_{B,y}f_A-g_A)+\mathcal{O}(h^2+\varepsilon h).$$

In general, since the ABA-step is usually followed by another ABA-step, the initial values will normally be inconsistent, and $\Delta z_0=-(h/2)(-g_{B,z})^{-1}g_{B,y}f_A$.

These results are in coincidence with the results given in Table 1. The steady case is equivalent to the situation where the initial values are satisfying the constraints (26). For the general case, Δz_j might differ from zero, causing an error $\mathcal{O}(h)$ for the BA-scheme.

In the case of van der Pol's equation (6), the index 1 constraint is simply $(1-y_0^2)z_0=0$, thus $z_0=0$. Under this condition the discrepancy between the index 2 manifolds of the full and the stiff problem will be $\varepsilon\cdot y_0/(1-y_0^2)$, resulting in a constant error of this size in the z-component for the BA and the BAB schemes. This is consistent with the results obtained when the same problem was analysed by the use of Lie-series.

5. Conclusion

We have analysed the principal part of the local error in various types of splitting methods where the vector field has been split into a stiff and a nonstiff part. Two approaches have been used, one based on Lie series and the other on singular perturbation. In general, the presented analysis holds for general nonlinear vector fields, but some specific examples that appear more commonly in applications have been given particular attention.

One may ask what the theory presented here says about which of the four splitting methods which should be used in practice. We believe that the smallest local error is obtained with the BAB and BA schemes for stepsize larger than the asymptotic regime. However in selecting the scheme to be used, one should look closer at the global error. Here, we have considered instead the local error and we have argued that it is important to study by itself. This is because it allows us to study error and stepsize control and eventually it might aid us in designing robust and efficient new splitting methods.

The analysis shows that there are stepsize intervals for which the local error behaves very differently from what the classical theory based on Taylor series expansion predicts. We see that with certain choices of splitting methods, the local error can be almost constant for fairly large stepsize intervals. One may be led to think that this behaviour contradicts other known results from the literature, e.g. [9] on the order of the

local error in splitting methods. However, this is not so, because the order zero behaviour reported here happens only for finitely small stepsizes, and thus the results can still be reconciled with those of Jahnke and Lubich. The difference is that the analysis we present here is somewhat more detailed.

Finally, we believe that an interesting open problem is to conduct a similarly detailed analysis of the global error in the situations described in this paper, indeed, some initial numerical results show that the global error also behaves differently from what the classical theory predicts.

References

[1] C. Besse, B. Bidégaray, S. Descombes, Order estimates in time of splitting methods for the nonlinear Schr´odinger equation, SIAM J. Numer. Anal. 40 (2002) 26–40.
[2] K. Dekker, J.G. Verwer, Stability of Runge–Kutta Methods for Stiff Initial Value Problems, North-Holland, Amsterdam, 1984.
[3] S. Descombes, Convergence of a splitting method of high order for reaction–diffusion systems, Math. Comput. 70 (2001) 1481–1501.
[4] R. Frank, J. Schneid, C.W. Ueberhuber, The concept of *B*-convergence, SIAM J. Numer. Anal. 18 (1981) 753–780.
[5] E. Hairer, Ch. Lubich, M. Roche, Error of Runge–Kutta methods for stiff problems studied via differential algebraic equations, BIT 28 (1988) 678–700.
[6] E. Hairer, Ch. Lubich, M. Roche, Error of Rosenbrock methods for stiff problems studied via differential algebraic equations, BIT 29 (1989) 77–90.
[7] E. Hairer, Ch. Lubich, G. Wanner, Geometric Numerical Integration, Number 31 in Springer Series in Computational Mathematics, Springer, Berlin, 2002.
[8] E. Hairer, G. Wanner, Solving Ordinary Differential Equations II, Stiff and Differential-Algebra, Problems, second ed., Springer, Berlin, 1996.
[9] T. Jahnke, Ch. Lubich, Error bounds for exponential operator splittings, BIT 40 (4) (2000) 735–744.
[10] Ch. Lubich, Variational splitting for multiconfiguration quantum dynamics, Preprint, Univ Tübingen, January 2003.
[11] G.I. Marchuk, On the theory of the splitting-up method, in: Numerical Solution of Partial Differential Equations, II, Academic Press, New York, 1971, pp. 469–500.
[12] R.I. McLachlan, G.R.W. Quispel, Splitting methods, Acta Numer. 11 (2002) 341–434.
[13] P.J. Olver, Applications of Lie Groups to Differential Equations, GTM 107, second ed., Springer, Berlin, 1993.
[14] A. Prothero, A. Robinson, On the stability and accuracy of one-step methods for solving stiff systems of ordinary differential equations, Math. Comput. 28 (1974) 145–162.
[15] Michel Roche, Rosenbrock methods for differential algebraic equations, Numer. Math. 52 (1988) 45–63.
[16] G. Strang, On the construction and comparison of difference schemes, SIAM J. Numer. Anal. 5 (1968) 506–517.
[17] F.W. Warner, Foundations of Differentiable Manifolds and Lie Groups, GTM 94, Springer, Berlin, 1983.
[18] N.N. Yanenko, Splitting methods for partial differential equations, In: Proc. IFIP Congr., vol. 71 of Information processing, IFIP, North-Holland, Amsterdam, 1972, pp. 1206–1213.