Centrum voor Wiskunde en Informatica

REPORT*RAPPORT*

A Note on Operator Splitting in a Stiff Linear Case

J.G. Verwer, B. Sportisse

Modelling, Analysis and Simulation (MAS)

MAS-R9830 December 1998

Report MAS-R9830
ISSN 1386-3703

CWI
P.O. Box 94079
1090 GB Amsterdam
The Netherlands

A Note on Operator Splitting in a Stiff Linear Case

J.G. Verwer *
CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

B. Sportisse †
Centre d'Enseignement et de Recherche en Mathématiques,
Informatique et Calcul Scientifique (CERMICS), 6 et 8 avenue Blaise Pascal,
Cité Descartes - Champs-sur-Marne, 77455 - Marne-la-Vallée CEDEX2, France

Abstract

This note is concerned with the numerical technique of operator splitting for initial value problems. Using a stiff linear ODE system as model problem, error bounds are derived for standard 1st- and 2nd-order splitting methods. The analysis focuses on deriving bounds independent of stiffness. The aim is to study the influence of stiffness on accuracy. Attention is paid to the influence of the splitting sequence on the splitting error and to the order reduction phenomenon.

1991 Mathematics Subject Classification: Primary: 65L05, 65L70. Secondary: 65M15.
1991 Computing Reviews Classification System: G.1.7 and G.1.8.
Keywords and Phrases: Numerical time integration, operator splitting, error analysis.
Note: Work carried out under project MAS1.4 - Discretization of Initial Value Problems.

1 Introduction

Operator splitting is a popular numerical technique in many large-scale PDE applications involving multiple time scales. One such application is atmospheric air pollution modelling where transport-chemistry equations of the advection-diffusion-reaction type are solved. These equations model the long range transport and chemical transformations of natural and anthropogenic trace gases [16]. The chemistry is extremely stiff, resulting in multiple time scales. Compared to many chemical reactions the transport process is extremely slow. In the integration process, transport and chemistry are normally splitted, both by 1st-order splitting and 2nd-order Strang splitting [14]. Splitting gives rise to errors which come on top of all numerical errors induced by discretizing and integrating the transport and stiff chemical kinetics equations. Insight in the splitting errors is of obvious interest, in particular since the splitting step size is adjusted to the slow transport process so that the fast chemical reactions are underresolved.

Having the application of air pollution modelling in mind, Sportisse [12] presents a splitting error analysis for a stiff linear ODE system in $\mathbb{R}^m$,

$$\dot{w} = Aw + Bw, \quad t > 0, \quad w(0) = w^0, \tag{1}$$

where the matrix A is supposed to be stiff and B nonstiff. The matrix A thus represents the stiff chemical kinetics part and B the transport part. A restriction is that transport usually gives rise to mildly stiff problems, whereas here B is assumed truly nonstiff. However, studying this linear

* Jan.Verwer@cwi.nl
† sportiss@cermics.enpc.fr

model system is interesting and helpful in getting insight in the real application. The analysis in [12] has led to two interesting conclusions. First, 2nd-order Strang splitting suffers from order reduction from two to one. Second, putting the stiff computation after the nonstiff one may enhance accuracy. Let us notice that the order reduction phenomenon for splitting methods and Strang splitting due to stiffness has been discussed before in the literature, see e.g. [3, 4, 5, 11].

The approach adopted in [12] has been inspired by numerical work on dynamical systems of singularly perturbed type [9, 10]. The results are obtained for reduced solutions projected on certain smooth manifolds. The purpose of the current note is to provide an alternative splitting error analysis, leading to results very similar to those in [12]. The analysis presented here compares the splitting approximations directly to the sought solutions. Like in the quoted papers, and many others, the aim is to study the influence of stiffness on accuracy. For that purpose we focus on error bounds which are independent of stiffness. Our approach is reminiscent of the B-convergence theory for Runge-Kutta methods (see e.g. [1, 2]).

The contents is as follows. In Section 2 we define our class of model problems. Section 3 describes the splitting methods, five in total. Section 4 is the main one. Here a consistency and convergence analysis is presented directed at splitting a stiff and a nonstiff linear problem. In Section 5 we illustrate the theory by means of two simple numerical examples. Section 6 concludes the note with final remarks.

2 The linear problem class

We consider constant coefficient linear ODE systems (1) where the matrix A is supposed to be stiff and B nonstiff. Stiffness means that τA is huge in norm for the range of step sizes τ considered [1, 2]. In this note the step size τ represents a splitting step size. We thus assume

$$\|\tau A\| \gg 1, \quad \|\tau B\| = O(\tau). \tag{2}$$

The matrix A is supposed to be diagonizable with a well-conditioned eigensystem:

$$A = XDX^{-1}, \quad D = \mathrm{diag}\,(d_k), \quad \mathrm{Re}\,(d_k) \le C, \quad \|X\| \le C, \quad \|X^{-1}\| \le C. \tag{3}$$

Throughout, C denotes an appropriate constant of moderate size independent of stiffness taking on different values when used in different instances. Independent of stiffness means that the constants do not grow without bound with $\|A\|$.

Generally, A has stiff eigenvalues and nonstiff ones. For convenience of notation, in the remainder we will denote stiff eigenvalues d_k $(k = 1, \ldots, m)$ by λ and nonstiff ones by μ. For the eigenvalues, the notions stiff and nonstiff are associated with the properties

$$\mathrm{Re}(\tau\lambda) \ll -1, \quad |\tau\mu| = O(\tau). \tag{4}$$

Hence stiffness is associated with eigenvalues with a large negative real part.

Let us express the solution of (1) in the exponential matrix form

$$w(t+\tau) = e^{\tau(A+B)}w(t). \tag{5}$$

Introduce the eigencomponent vector z of w defined by X, i.e.

$$w = Xz. \tag{6}$$

Trivially,

$$z(t+\tau) = e^{\tau(D+E)}z(t), \quad E = X^{-1}BX, \tag{7}$$

which solves

$$\dot{z} = Dz + Ez, \quad t > 0, \quad z(0) = z^0. \tag{8}$$

In analogy with (2) there holds $\|\tau D\| \gg 1$ and $\|\tau E\| = O(\tau)$. The transformed linear problem (8) will be used in the consistency analysis of the splitting methods. In particular, we will distinguish between stiff and nonstiff components z_k of the solution vector $z(t) \in \mathbb{C}^m$. We call z_k a *stiff* component, associated to a stiff eigenvalue λ, if z_k is a solution of

$$\dot{z}_k(t) = \lambda z_k(t) + (Ez(t))_k. \tag{9}$$

Likewise, z_k is called *nonstiff* if z_k is a solution of one of the equations

$$\dot{z}_k(t) = \mu z_k(t) + (Ez(t))_k. \tag{10}$$

From (9) we obtain the inequality

$$|z_k(t+\tau)| \le |e^{\tau\lambda}|\,|z_k(t)| + C\left|\frac{1-e^{\tau\lambda}}{\lambda}\right|. \tag{11}$$

In view of assumption (4), the exponential is negligibly small so that up to an exponentially small term, $z_k(t+\tau)$ is proportional to the reciprocal of the stiff eigenvalue. We denote this in the usual way by

$$z_k(t+\tau) = O(\frac{1}{\lambda}). \tag{12}$$

This holds for any $t \ge 0$, $\tau > 0$ under the stiffness assumption (4). For initial values $z_k(0) \ne 0$, the normal situation, we encounter a fast initial transient at $t = 0$.

3 The splitting methods

Following [12] we consider five different splitting methods: two 1st-order methods, two 2nd-order methods based on Strang splitting, and a 1st-order method based on the idea of source splitting. It is emphasized that in this section order means order of consistency in the classical sense for $\tau \to 0$, without giving notice to stiffness. We also stress that in the whole of the paper we assume exact integration at substeps and thus examine only splitting errors.

a) $\mathcal{AB}$ Method $\mathcal{AB}$ is defined by first integrating the stiff problem $\dot{w} = Aw$ and then the nonstiff one $\dot{w} = Bw$, both over $[t, t+\tau]$. Denoting $\tilde{w}$ as the approximation, we have

$$\tilde{w}(t+\tau) = e^{\tau B} e^{\tau A}\,\tilde{w}(t). \tag{13}$$

This method is 1st-order consistent in the classical sense. Taylor expansion of the local error

$$le = w(t+\tau) - e^{\tau B} e^{\tau A}\,w(t), \tag{14}$$

yields

$$le = \tfrac{1}{2}\tau^2(BA - AB)w(t) + O(\tau^3), \quad \tau \to 0, \tag{15}$$

revealing the 1st-order consistency. However, first order consistency is meaningful only when the $\mathrm{O}(\tau^2)$ term is negligible compared to the $\mathrm{O}(\tau^3)$ remainder. Since $\|\tau A\| \gg 1$, this may not be true and hence a straightforward Taylor expansion not giving notice to stiffness is not meaningful.

b) $\mathcal{BA}$ Method $\mathcal{BA}$ differs from $\mathcal{AB}$ only in the sequence of the integrations. Now the stiff integration comes last. Method $\mathcal{BA}$ yields the 1st-order approximations

$$\tilde{w}(t+\tau) = e^{\tau A} e^{\tau B}\,\tilde{w}(t). \tag{16}$$

c) $\mathcal{ABA}$ Method $\mathcal{ABA}$ is based on symmetric Strang splitting [14]. First $\dot{w} = Aw$ is solved using half the step size, then $\dot{w} = Bw$ using the whole step size, followed by $\dot{w} = Aw$ using half the step size. So again the stiff integration comes last. The result is the approximation

$$\tilde{w}(t+\tau) = e^{\frac{1}{2}\tau A} e^{\tau B} e^{\frac{1}{2}\tau A}\,\tilde{w}(t). \tag{17}$$

Taylor expansion will show 2nd-order consistency in the classical sense.

d) $\mathcal{BAB}$ Method $\mathcal{BAB}$ is also based on symmetric Strang splitting and differs from $\mathcal{ABA}$ only in the sequence in which the splitted subproblems are solved. Here the nonstiff integration comes last. The resulting 2nd-order approximation reads

$$\tilde{w}(t+\tau) = e^{\frac{1}{2}\tau B} e^{\tau A} e^{\frac{1}{2}\tau B}\, \tilde{w}(t). \tag{18}$$

e) $\mathcal{SP}$ In the above splitting methods, initial values for substeps are defined by the final result of a preceding substep. At any split step we thus introduce a transient for the stiff integration (initial values differ from the final result from a preceding stiff integration). Because stiff solvers often encounter difficulties in the transient phase, it makes sense to avoid the artificial transients. The experiments in [7] illustrate this for applications in air pollution modelling. Method $\mathcal{SP}$ (Source Splitting) is designed with this goal. In this method, the result of the nonstiff integration is treated as a piecewise constant source for the stiff integration, so that no discontinuities in solution values arise. For the present linear problem, $\mathcal{SP}$ thus solves the linear system $\dot{w} = Aw + s$, where s is the piecewise constant source generated by the nonstiff integration. Consequently, $\mathcal{SP}$ is defined by

$$\tilde{w}(t+\tau) = e^{\tau A}\left(\tilde{w}(t) + (I - e^{-\tau A})A^{-1}s(t)\right), \tag{19}$$

where

$$s(t) = \left(e^{\tau B} - I\right)\tau^{-1}\tilde{w}(t). \tag{20}$$

Method $\mathcal{SP}$ is called $\mathcal{NTS}$ in [12]. The order of consistency is equal to one. See [16] and references therein for nonlinear applications.

Except for $\mathcal{SP}$, all methods are exact if A and B commute. This result is well known. Interesting similar results on commutativity for nonlinear problems can be found in [8]. In that paper the emphasis also lies on systems of advection-diffusion-reaction problems encountered in air pollution modelling. $\mathcal{SP}$ does not gain advantage from commutativity.

4 Consistency and convergence in the stiff case

All five splitting methods fit in a standard linear recurrence format $\tilde{w}(t+\tau) = L(\tau A, \tau B)\,\tilde{w}(t)$ which is invariant under the transformation (6). Hence we might as well apply the methods directly to the transformed problem (8) and analyse instead $\tilde{z}(t+\tau) = L(\tau D, \tau E)\,\tilde{z}(t)$. Convergence and consistency estimates for the approximations $\tilde{z}(t)$ immediately carry over to $\tilde{w}(t) = X\tilde{z}(t)$ in view of the boundedness of X. This transformation couples the stiff and nonstiff approximations $\tilde{z}_k$.

Introduce the global error $ge(t) = z(t) - \tilde{z}(t)$ and the local (truncation) error

$$le(t+\tau) = z(t+\tau) - L(\tau D, \tau E)z(t). \tag{21}$$

These errors are connected by the inhomogeneous error recurrence

$$ge(t+\tau) = L(\tau D, \tau E)\, ge(t) + le(t+\tau). \tag{22}$$

In the analysis we will focus on the local errors. Stability is not an issue here since the subproblems $\dot{z} = Dz$ and $\dot{z} = Ez$ are stable, see (2)-(3). Hence the linear operators $L(\tau D, \tau E)$ are stable under the same conditions as used in the consistency analysis. Consistency then implies convergence under the same conditions.

For the consistency analysis we thus employ the transformed problem (8). Hereby we distinguish between the equations

$$\dot{z}_k(t) = \mu z_k(t) + (Ez(t))_k$$

and

$$\dot{z}_k(t) = \lambda z_k(t) + (Ez(t))_k,$$

defining *nonstiff* and *stiff* solution components z_k, respectively. The error analysis should lead to estimates which are independent of eigenvalues λ. Error bounds not giving notice to stiffness, may involve error constants which grow without bound for increasing λ. This means that the convergence order deduced from such bounds may not reflect actual error behaviour. In the following sections we will derive local and global error relations reflecting what is observed in an actual computation. To this end we employ the asymptotic assumption

$$\mathrm{Re}(\tau\lambda) \ll -1, \ |\tau\mu| = O(\tau), \ \|\tau E\| = O(\tau) \quad \text{for} \quad \tau \to 0. \tag{23}$$

In particular, we allow infinite stiffness, i.e. $\mathrm{Re}(\tau\lambda) \to -\infty$. When appropriate we will use the notions *stiff consistency* and *stiff convergence* in connection with error relations derived under this assumption.

4.1 Method $\mathcal{BA}$

4.1.1 Consistency

Stiff components Applied to the transformed problem $\dot{z} = Dz + Ez$, method $\mathcal{BA}$ reads

$$\tilde{z}(t+\tau) = e^{\tau D} e^{\tau E} \tilde{z}(t). \tag{24}$$

Let z_k be a stiff component with stiff eigenvalue λ. Because $e^{\tau E}$ is bounded and D is diagonal, (23) induces that up to exponentially small values, $ge_k(t+\tau) = le_k(t+\tau) = z_k(t+\tau)$ for any $t \geq 0$. As the exact solution $z_k(t+\tau)$ is proportional to the reciprocal of the stiff eigenvalue (see (12)), both the global and local error satisfy

$$ge_k(t+\tau) = le_k(t+\tau) = O(\frac{1}{\lambda}), \quad t \geq 0. \tag{25}$$

Putting $\mathrm{Re}(\tau\lambda) \ll -1$ we get $O(\tau)$. But as long as $\mathrm{Re}(\tau\lambda) \ll -1$, step size reduction will be of no influence. Consequently, in the stiff case the stiff components have relative errors of size $O(1)$.

Nonstiff components Here we distinguish between $t = 0$ (transient phase) and $t > 0$. Let z_k be a nonstiff component with a nonstiff eigenvalue μ. At t = 0 we have

$$le_k(\tau) = z_k(\tau) - e^{\tau\mu}(e^{\tau E} z(0))_k = z_k(\tau) - z_k(0) + O(\tau),$$

where the $O(\tau)$ term is independent of stiffness. For all $t \geq 0$, the first derivative of all nonstiff components $z_k(t)$ is bounded with respect to stiff eigenvalues λ. This follows trivially from the differential equation $\dot{z}_k(t) = \mu z_k(t) + (Ez(t))_k$. Consequently, we may apply Taylors theorem with remainder term to obtain that $z_k(\tau) - z_k(0) = O(\tau)$, where the constant involved is again independent of stiffness. This proves that $le_k(\tau) = O(\tau)$, i.e., stiff concistency of order zero in the first time step.

Next consider the local error for $t > 0$,

$$le_k(t+\tau) = z_k(t+\tau) - e^{\tau\mu}(e^{\tau E} z(t))_k.$$

We may expand

$$e^{\tau\mu}(e^{\tau E} z(t))_k = z_k(t) + \tau\mu z_k(t) + \tau(Ez(t))_k + O(\tau^2) = z_k(t) + \tau\dot{z}_k(t) + O(\tau^2),$$

up to an $O(\tau^2)$ term independent of stiffness. If we also may expand $z_k(t+\tau)$ up to an $O(\tau^2)$ term independent of stiffness, stiff consistency of order one is proved. Sufficient is boundedness of the second derivative with respect to stiff eigenvalues λ (Taylor's theorem with remainder term). We have

$$\ddot{z}(t) = (D^2 + DE + ED + E^2)z(t).$$

All entries in the k-th row of D^2, DE and E^2 are independent of stiff eigenvalues λ. Stiff eigenvalues λ do enter the k-th row of ED. However, all components of $Dz(t)$ connected with stiff eigenvalues are bounded in λ due to the proportionality relation (12). Because E is bounded, we may conclude that for $t > 0$ the nonstiff components of the second derivative vector $\ddot{z}(t)$ are bounded too. This completes the proof of stiff consistency of 1-st order at times $t > 0$.

4.1.2 Convergence

We have shown that approximations to stiff components z_k do not converge for $\tau \to 0$ as long as $\mathrm{Re}(\tau\lambda) \ll -1$. Their global errors ge_k satisfy

$$ge_k(t) = O(\frac{1}{\lambda}). \tag{26}$$

In the relative sense the error is even $O(1)$. For nonstiff approximations $\tilde{z}_k$ we have shown zero order stiff consistency at the first step point and first order stiff consistency at all later step points. The zero order at the first step is allowed as this error adds up only once in the global error recurrence

$$ge_k(t+\tau) = e^{\tau\mu}(e^{\tau E}ge(t))_k + le_k(t+\tau).$$

This recurrence couples all remaining stiff and nonstiff error components to the k-th one through the term $e^{\tau E}e(t)$. Because $\|\tau E\| = O(\tau)$, for nonstiff components z_k the global errors satisfy

$$ge_k(t) = O(\tau) + O(\frac{1}{\lambda}) = O(\tau). \tag{27}$$

For nonstiff components the term $O(1/\lambda)$ can be neglected compared to $O(\tau)$ as long as $\mathrm{Re}(\tau\lambda) \ll -1$.

4.2 Method $\mathcal{AB}$

4.2.1 Consistency

Applied to the transformed problem $\dot{z} = Dz + Ez$, method $\mathcal{AB}$ defines the approximations

$$\tilde{z}(t+\tau) = e^{\tau E}e^{\tau D}\,\tilde{z}(t). \tag{28}$$

Now the stiff computation comes first, followed by the nonstiff one. The local errors read

$$le(t+\tau) = z(t+\tau) - e^{\tau E}e^{\tau D}z(t). \tag{29}$$

Like for method $\mathcal{BA}$ we distinguish between stiff and nonstiff components.

Stiff components Since $\|\tau E\| = O(\tau)$ and $\|e^{\tau D}\| = O(1)$ independent of stiffness, we may write

$$le(t+\tau) = z(t+\tau) - (I+\tau E)e^{\tau D}z(t) + O(\tau^2).$$

Neglecting exponentially small values associated to stiff eigenvalues λ, we write

$$le_k(t+\tau) = O(\frac{1}{\lambda}) - \tau(Ee^{\tau D}z(t))_k + O(\tau^2).$$

The vector $e^{\tau D}z(t)$ is composed of exponentially small values for stiff components and values $e^{\tau\mu}z_k(t)$ for nonstiff components z_k. The latter are $O(1)$. Hence, again neglecting exponentially small values, we may write

$$le_k(t+\tau) = O(\frac{1}{\lambda}) + O(\tau), \tag{30}$$

where the $O(\tau)$ term arises from the coupling nonstiff to stiff through the matrix E. In the relative sense this coupling is error prone, as nonstiff $O(\tau)$ components are much larger than the stiff ones. Hence, for step sizes of practical interest we will encounter large relative local errors, even of magnitude

$$O(\tau\lambda) \gg 1.$$

This peculiar error behaviour is a consequence of the fact that the nonstiff computation comes after the stiff one. However, these large local errors for stiff components do not accumulate. Neither do they affect the global errors for nonstiff components. To see this we inspect the global error recursion

$$ge(t+\tau) = e^{\tau E} e^{\tau D} ge(t) + le(t+\tau). \tag{31}$$

In the multiplication $e^{\tau D} ge(t)$, the stiff exponential entries $e^{\tau\lambda}$ effectively eliminate all stiff global error components $ge_k(t)$. Hence, when stepping from time t to time $t+\tau$, these errors do not accumulate, nor are they coupled to the nonstiff components of $ge(t+\tau)$.

Nonstiff components Introduce the auxiliary local error

$$l(t) = e^{-\tau E}\, le(t). \tag{32}$$

We expand this auxiliary error in a manner independent of stiffness as follows:

$$\begin{aligned} l(t+\tau) &= e^{-\tau E} z(t+\tau) - e^{\tau D} e^{\tau E} e^{-\tau E} z(t) \\ &= (I - \tau E)\, z(t+\tau) - e^{\tau D} e^{\tau E} (I - \tau E)\, z(t) + O(\tau^2) \\ &= z(t+\tau) - e^{\tau D} e^{\tau E} z(t) - \tau E z(t+\tau) + e^{\tau D} e^{\tau E}\, \tau E z(t) + O(\tau^2) \\ &= le_{ED}(t+\tau) - \tau E z(t+\tau) + e^{\tau D} \tau E z(t) + O(\tau^2), \end{aligned} \tag{33}$$

where $le_{ED}(t+\tau)$ is the local error of method $\mathcal{BA}$ (see Section 4.1.1). So far we have only expanded $e^{\tau E}$ so that the $O(\tau^2)$ terms are truly independent of stiffness. Also recall that $e^{\tau D}$ is bounded.

Let z_k be a nonstiff component associated to a nonstiff eigenvalue μ. Then we can write

$$(e^{\tau D} \tau E z(t))_k = (\tau E z(t))_k + O(\tau^2).$$

Inserting this into the found expression for $l(t+\tau)$ and using the differential equation, we get

$$l_k(t+\tau) = (le_{ED})_k(t+\tau) - \tau(\dot z_k(t+\tau) - \mu z_k(t+\tau)) + \tau(\dot z_k(t) - \mu z_k(t)) + O(\tau^2).$$

At the initial time $\dot z_k$ is bounded and for $t > 0$ also the second derivative. Hence at the first step $l_k(\tau) = O(\tau)$ and at all later steps $l_k(t+\tau) = O(\tau^2)$, independent of stiffness (see also Section 4.1.1).

Because of the transformation $le(t+\tau) = e^{\tau E}\, l(t+\tau)$, we also have to examine components $l_k(t+\tau)$ belonging to stiff components z_k. Consider again (33) for a stiff components z_k. Neglecting the exponentially small term, it follows immediately that

$$l_k(t+\tau) = (le_{ED})_k(t+\tau) + O(\tau) = O(\frac{1}{\lambda}) + O(\tau), \quad t \geq 0.$$

Next, computing $le_k(t+\tau)$ for a nonstiff component z_k from

$$le(t+\tau) = (I + \tau E)\, l(t+\tau) + O(\tau^2),$$

yields $le_k(\tau) = O(\tau)$ and $le_k(t+\tau) = O(\tau^2)$, $t \geq 0$. Observe that all stiffness terms $O(\tau/\lambda)$ may be neglected compared to nonstiff components of $l(t+\tau)$. In conclusion, for nonstiff components we have zero order stiff consistency at the first step and at all later steps stiff consistency of order one. In this respect method $\mathcal{AB}$ does not differ from its counterpart $\mathcal{AB}$.

4.2.2 Convergence

The global error recursion (31) couples errors for all stiff and nonstiff components z_k. As observed above, the exponential $e^{\tau\lambda}$ eliminates all errors $e_k(t)$ for stiff components as long as $\mathrm{Re}(\tau\lambda) \ll -1$. So in the stiff case only errors associated to nonstiff components are accumulated. Assuming stability, and using the above consistency results for the stiff and nonstiff components, we get

$$ge_k(t) = O(\tau), \tag{34}$$

for all components z_k. But the stiff ones suffer from large relative errors of size $O(\tau\lambda) \gg 1$.

We should like to point out that convergence for method $\mathcal{AB}$ can also be derived in a more direct way from the convergence of $\mathcal{BA}$ by using the transformation $\hat{w}(t) = e^{-\tau B}\tilde{w}(t)$ (cf. (32)). These auxiliary approximations satisfy the $\mathcal{BA}$ sequence, i.e.,

$$\hat{w}(t+\tau) = e^{\tau A} e^{\tau B} \hat{w}(t). \tag{35}$$

The proof goes in three steps. First, the new approximation $\hat{w}(t)$ converges to the exact solution defined by the initial value $\hat{w}(0) = e^{-\tau B} w(0)$. Second, because $\|\tau B\| = O(\tau)$ and $\dot{w} = (A+B)w$ is supposed to be stable, $\hat{w}(t)$ also converges to the sought solution with initial value $w(0)$. Third, in view of $\|\tau B\| = O(\tau)$ the actual approximations then also converge to the sought solution with initial value $w(0)$.

4.3 Methods $\mathcal{ABA}$ and $\mathcal{BAB}$

Applied to the transformed problem $\dot{z} = Dz + Ez$, method $\mathcal{ABA}$ defines the approximations

$$\tilde{z}(t+\tau) = e^{\frac{1}{2}\tau D} e^{\tau E} e^{\frac{1}{2}\tau D}\, \tilde{z}(t). \tag{36}$$

According to the alternate application of methods $\mathcal{AB}$ and $\mathcal{BA}$, we may also write

$$\tilde{z}(t+\tau) = e^{\frac{1}{2}\tau D} e^{\frac{1}{2}\tau E} \hat{z}(t+\tfrac{1}{2}\tau), \quad \hat{z}(t+\tfrac{1}{2}\tau) = e^{\frac{1}{2}\tau E} e^{\frac{1}{2}\tau D}\, \tilde{z}(t). \tag{37}$$

Stiff consistency and convergence thus can be deduced from the results for $\mathcal{AB}$ and $\mathcal{BA}$. Unfortunately, the method lacks second order stiff consistency and convergence [12]. Simple counterexamples can be constructed to show this analytically. Also the numerical tests given in Section 5 exemplify this order reduction. Most likely the reduction occurs generally, as 2nd-order consistency for nonstiff components z_k requires the retrieval of the second derivative $\ddot{z}_k(t)$. The derivations for methods $\mathcal{AB}$ and $\mathcal{BA}$ indicate that this is not possible, at least in general, simply because the stiff entries in the diagonal exponential operator cannot be expanded.

Method $\mathcal{BAB}$ also suffers from order reduction from two to one, for the same reason. The two methods do approximate stiff components z_k completely different. From (36) we see that $\mathcal{ABA}$ treats stiff components z_k in the same way as method $\mathcal{BA}$, due to the fact that the stiff computation comes last. Likewise, method $\mathcal{BAB}$ treats stiff components z_k in the same way as $\mathcal{AB}$, where the nonstiff computation comes last.

4.4 Method $\mathcal{ST}$

4.4.1 Consistency

Applied to the transformed problem $\dot{z} = Dz + Ez$, method $\mathcal{ST}$ defines the approximations

$$\tilde{z}(t+\tau) = e^{\tau D}\tilde{z}(t) + (e^{\tau D} - I)(\tau D)^{-1}\,(e^{\tau E} - I)\tilde{z}(t). \tag{38}$$

It is useful to compare (38) with the true solution derived from the variation of constants formula

$$z(t+\tau) = e^{\tau D} z(t) + e^{\tau D} \int_0^\tau e^{-sD} E z(t+s)\, ds. \tag{39}$$

Inserting

$$\int_0^\tau e^{-sD} ds = (I - e^{-\tau D}) D^{-1}$$

in (38) gives

$$\tilde{z}(t+\tau) = e^{\tau D}\tilde{z}(t) + e^{\tau D}\int_0^\tau e^{-sD}(e^{\tau E} - I)\tau^{-1}\tilde{z}(t)\,ds. \tag{40}$$

Hence the local truncation error satisfies

$$le(t+\tau) = e^{\tau D}\int_0^\tau e^{-sD}\left(E(z(t+s) - z(t)) + O(\tau E)z(t)\right)\,ds,$$

or

$$le(t+\tau) = e^{\tau D}\int_0^\tau e^{-sD}E(z(t+s) - z(t))\,ds + (e^{\tau D} - I)D^{-1}O(\tau E)z(t). \tag{41}$$

Stiff components Neglecting exponentially small values, for stiff components (41) yields

$$le_k(t+\tau) = \int_0^\tau e^{(\tau-s)\lambda}E(z(t+s) - z(t))\,ds + O(\frac{\tau}{\lambda}). \tag{42}$$

At this stage we distinguish between $t = 0$ and $t > 0$. At $t = 0$ we impose $E(z(t+s)-z(t)) = O(1)$, resulting in $le_k(\tau) = O(1/\lambda)$. For $t > 0$ the first derivative $\dot{z}$ is bounded in λ so that we may invoke $E(z(t+s) - z(t)) = O(\tau)$, resulting in $le_k(t+\tau) = O(\tau/\lambda)$. Hence for $t > 0$ the local error is proportional to $\tau z_k(t) = O(\tau/\lambda)$. This means that for $t > 0$ the local error has the right 'stiffness' size. Recall that none of the other splitting methods has this property.

Nonstiff components For nonstiff components we can rewrite (41) to

$$le_k(t+\tau) = \int_0^\tau e^{(\tau-s)\mu}E(z(t+s) - z(t))\,ds + O(\tau^2). \tag{43}$$

Because $e^{(\tau-s)\mu}$ is bounded at $t = 0$ we have $le_k(\tau) = O(\tau)$, i.e. stiff consistency of order zero. Likewise, for $t > 0$ we have $le_k(t+\tau) = O(\tau^2)$, i.e. stiff consistency of order one.

4.4.2 Convergence

Consider the global error recurrence

$$ge(t+\tau) = \left(e^{\tau D} + (e^{\tau D} - I)(\tau D)^{-1}\left(e^{\tau E} - I\right)\right)ge(t) + le(t+\tau).$$

Again neglecting exponentially small values, for stiff components we can write

$$ge_k(t+\tau) = \frac{-1}{\tau\lambda}(O(\tau E)ge(t))_k + le_k(t+\tau)$$

and observe a multiplication by $\tau E/\tau\lambda$. As $|\tau\lambda| \gg 1$, in first approximation $ge_k(t+\tau) = le_k(t+\tau)$. Consequently, for stiff components we deduce stiff convergence of order one and this also holds in the relative sense. When considering nonstiff components we may write $le(\tau) = O(\tau)$ at $t = 0$ and $O(\tau^2)$ for $t > 0$. Assuming stability, we thus again conclude stiff convergence of order one also for nonstiff components.

4.5 Summary

For ease of survey we summarize all convergence orders derived:

		$\mathcal{BA}$	$\mathcal{AB}$	$\mathcal{ABA}$	$\mathcal{BAB}$	$\mathcal{ST}$
Nonstiff errors	ge_k	$O(\tau)$	$O(\tau)$	$O(\tau)$	$O(\tau)$	$O(\tau)$
Stiff errors	ge_k	$O(\frac{1}{\lambda})$	$O(\tau)$	$O(\frac{1}{\lambda})$	$O(\tau)$	$O(\frac{\tau}{\lambda})$
Stiff errors	ge_k/z_k	$O(1)$	$O(\tau\lambda)$	$O(1)$	$O(\tau\lambda)$	$O(\tau)$

5 Numerical illustration

We next illustrate the analysis numerically with two simple example problems (1).

5.1 Example 1

The first is defined by

$$A = \begin{pmatrix} -10^4 & 10^4 & 1 \\ 10^4 & -10^4 & 2 \\ 1 & 1 & -2 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 0.5 & 0.25 \\ 0.1 & 0 & 0.1 \\ 0.2 & 0.4 & -1 \end{pmatrix}, \quad w(0) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \tag{44}$$

With this definition A is stiff and B nonstiff for stepsizes τ between 10^{-3} and 1, say. Matrix A satisfies condition (3), has one eigenvalue $\lambda = -20000$ and two eigenvalues $\mu = -3, 1$. According to our definition, component z_1 of the transformed problem (8) is the stiff one and z_2, z_3 are nonstiff. The choice of B is of minor importance, although for the test we wish to avoid commutativity. The five splitting methods are applied over the time interval $[0, 1]$, using $\tau = 1(10^{-1})10^{-3}$.

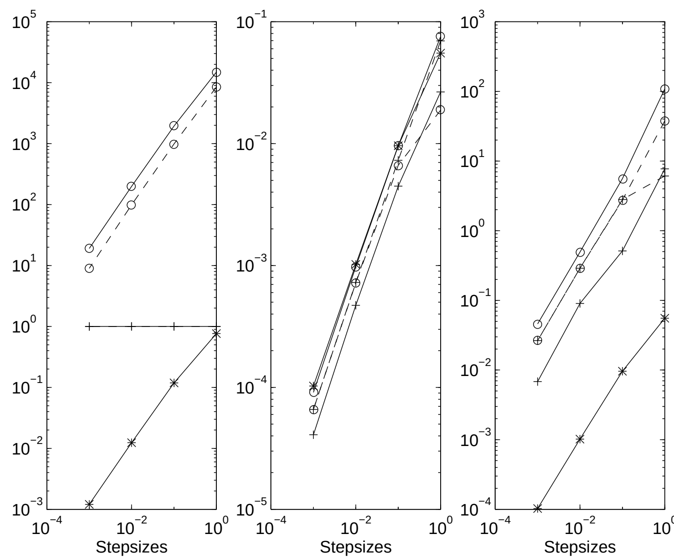

Figure 1: Relative splitting errors in z for matrices (44). Component z_1 left, z_2 middle and z_3 right. Solid -+- refers to method $\mathcal{BA}$, solid -o- to $\mathcal{AB}$ and solid -*- to $\mathcal{ST}$. Dashed -+- refers to $\mathcal{ABA}$ and dashed -o- to $\mathcal{BAB}$.

Figure 1 shows the relative errors for all three components of the transformed problem (8). The errors are in excellent agreement with the theory. For the stiff component z_1 we observe huge differences. Splittings $\mathcal{AB}$ and $\mathcal{BAB}$ show 1-st order convergence but the errors are large. The

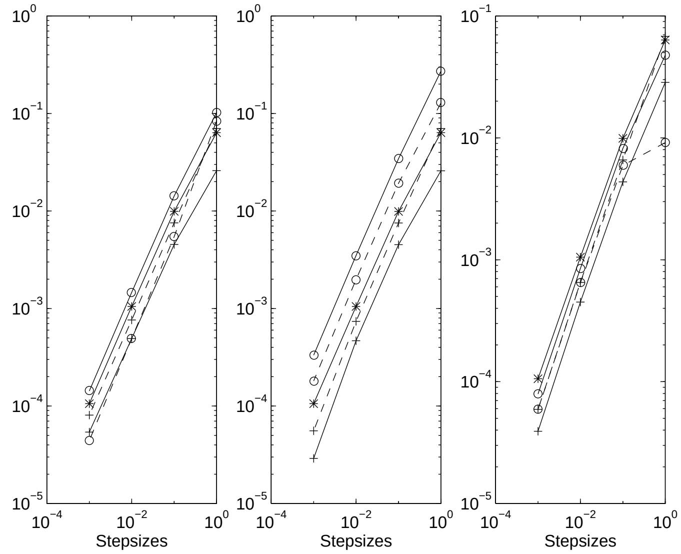

Figure 2: Relative splitting errors in w for matrices (44). Component w_1 left, w_2 middle and w_3 right. Solid -+- refers to method $\mathcal{BA}$, solid -o- to $\mathcal{AB}$ and solid -*- to $\mathcal{ST}$. Dashed -+- refers to $\mathcal{ABA}$ and dashed -o- to $\mathcal{BAB}$.

error prone coupling from nonstiff to stiff caused by the nonstiff computation at the end of the split step is evident. Splittings $\mathcal{BA}$ and $\mathcal{ABA}$ do not converge and their relative errors are exactly equal to one (in this step size range), which is caused by the strong exponential decay in the stiff computation at the end of the split step. The 1st-order of the source splitting $\mathcal{ST}$ is also nicely shown and the $\mathcal{ST}$ approximations are of the right size. For the nonstiff component z_2 the relative errors are nearly equal for all splittings. All give order one, clearly illustrating the order reduction for the Strang splittings. We also see this for the nonstiff component z_3, for which $\mathcal{ST}$ provides the highest accuracy (benefitting from the eigenvalue $\mu = -3$).

Figure 2 shows the relative errors for all three components of the original problem. Needless to say that these errors are the ones which count. All methods are confirmed to be stiffly convergent of order one, again exemplifying the order reduction of the Strang splittings. The source splitting method $\mathcal{ST}$ no longer shows advantage. Also the errors for the two Strang splitting methods are somewhat disappointing.

5.2 Example 2

The second example is defined by

$$A = \begin{pmatrix} -10^6 & -10^6 & 1 \\ 10^6 & -10^6 & 2 \\ 1 & 1 & -2 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 0.5 & 0.25 \\ 0.1 & 0 & 0.1 \\ 0.2 & 0.4 & -1 \end{pmatrix}, \quad w(0) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \tag{45}$$

The large entries of A are made a factor 100 larger than in the first example and the (1,2)-entry is negative now. Matrix A again satisfies condition (3), has a pair of complex-valued stiff eigenvalues $\lambda = -10^6\,(1 \pm i)$ and one real nonstiff eigenvalue $\mu = -2$. According to our definition, components z_1 and z_2 of the transformed problem (8) are the stiff ones and z_3 is nonstiff. The matrix B and the initial condition $w(0)$ are the same as in the first example. Noteworthy is that the negative (1,2)-entry of A changes the solution behaviour drastically. The large negative diagonal of A now result in a rapid decay for component w_1 and w_2, quite similar to the decay for z_1 and z_2. At $t = 1$ we have $w_1(1) \approx -2.1\,10^{-8}$, $w_2(1) \approx 8.3\,10^{-8}$ and $w_3(1) \approx 5.0\,10^{-2}$.

The rapid decay for w_1 and w_2 does reveal itself in huge differences in relative accuracies, quite similar as what we encounter for z_1 and z_2. Figure 3 shows the relative accuracies for the w-components at t = 1 for the step sizes $\tau = 10^{-1}(10^{-1})10^{-4}$. The results are again in excellent agreement with the theory. All methods compute w_3, z_3 with order one. For w_1, z_1 and w_2, z_2 the same huge differences are observed as for the stiff component z_1 in the first example.

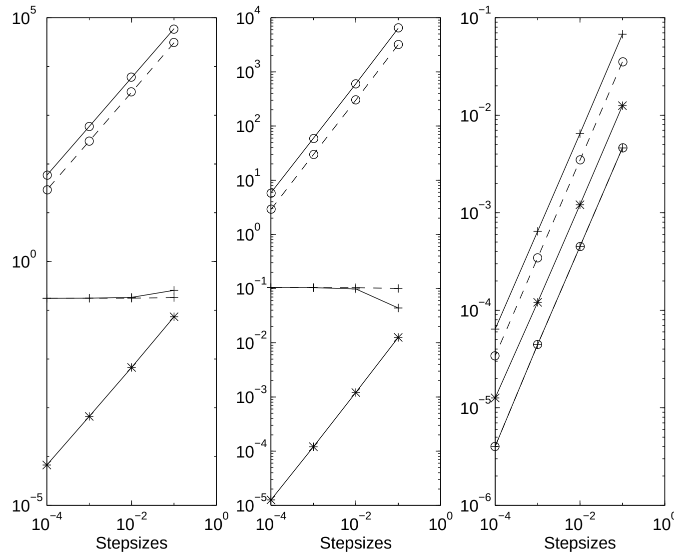

Figure 3: Relative splitting errors in w for matrices (45). Component w_1 left, w_2 middle and w_3 right. Solid -+- refers to method $\mathcal{BA}$, solid -o- to $\mathcal{AB}$ and solid -*- to $\mathcal{ST}$. Dashed -+- refers to $\mathcal{ABA}$ and dashed -o- to $\mathcal{BAB}$.

6 Conclusions

The analysis has revealed that for nonstiff solution components, as defined in Section 2, all five splitting methods are comparable in the sense that they all converge with $O(\tau)$ independent of the level of stiffness. For actual problems their accuracies of course differ and there seems to be no winner, as our two numerical examples indicate. On the other hand, large differences exist in the computation of the stiff components. For these the source splitting method shows a clear advantage. For stiff components the analysis has also confirmed the findings of [12]: putting the stiff computation at the end of the split step, as in method $\mathcal{BA}$ and $\mathcal{ABA}$, enhances accuracy for stiff components.

Interestingly, a similar conclusion is drawn in [6] in a comparison of two related dimensional splitting methods based on the trapezoidal and midpoint rule. The trapezoidal splitting rule turns out to be the better one as it ends with implicit Euler steps (stiff computations), whereas the midpoint splitting rule ends with explicit Euler steps (nonstiff computations). The precise arguments differ from the ones presented here, but they tell the same story: combining stiff and nonstiff computations may cause loss of accuracy which cannot be explained by classical consistency analysis.

The order reduction for Strang splitting gives rise to the question how to obtain stiff convergence of order two with a splitting method, see also [11]. Probably the most simple possibility is to apply Richardson extrapolation to method $\mathcal{BA}$ or $\mathcal{ST}$. This can be done locally, so that a new splitting method results, or globally at output points, like in [15].

The linear ODE system (1) serves as a model problem, useful to study the effect of stiffness on accuracy. Our main interest lies in understanding splitting methods for stiff, nonlinear problems.

In [12, 13] one finds interesting numerical results for a stiff, nonlinear reaction-diffusion problem. These confirm the order reduction of Strang splitting and again show an advantage of putting the stiff computation at the end of the split step. For this stiff nonlinear case a satisfactory theory does not yet exist.

References

[1] K. Dekker and J.G. Verwer, *Stability of Runge-Kutta methods for stiff non-linear differential equations*, North-Holland Elsevier Science Publishers, 1984.

[2] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems*, Springer-Verlag, Berlin, 2nd edition, 1996.

[3] W. Hundsdorfer and J.G. Verwer, *Stability and convergence of the Peaceman-Rachford ADI method for initial-boundary value problems*, Math. Comp. 53, 81–101, 1989.

[4] W. Hundsdorfer, *Unconditional convergence of some Crank-Nicolson LOD methods for initial-boundary value problems*, Math. Comp. 58, 35-53, 1992.

[5] W. Hundsdorfer and J.G. Verwer, *A note on splitting errors for advection-reaction equations*, Appl. Numer. Math. 18, 191–199, 1994.

[6] W. Hundsdorfer, *Trapezoidal and midpoint splittings for initial-boundary value problems*, Math. Comp. 67, 1047–1062, 1998.

[7] O. Knoth and R. Wolke, *A comparison of fast chemical kinetic solvers in a simple vertical diffusion model*, In Procs. 20th International Meeting on Air Pollution Modelling and its Applications, Valencia, Spain, 1993.

[8] D. Lanser and J.G. Verwer, *Analysis of operator splitting for advection-diffusion-reaction problems from air pollution modelling*, CWI Report MAS-R9805, 1998.

[9] Ch. Lubich, K. Nipp and D. Stoffer, *Runge-Kutta solutions of stiff differential equations near stationary points*, SIAM J. Numer. Anal. 32, 1296–1307, 1995.

[10] K. Nipp and D. Stoffer, *Invariant manifolds and global error estimates of numerical integration schemes applied to stiff systems of singular perturbation type - Part I: RK methods*, Numer. Math. 70, 245–257, 1995.

[11] Shi Jin, *Runge-Kutta methods for hyperbolic conservation laws with stiff relaxation terms*, J. Comput. Phys. 122, 51–67, 1995.

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

[13] B. Sportisse, G. Bencteux and P. Plion, *Method of lines versus operator splitting methods for reaction diffusion systems in air pollution modelling*, In: Proceedings APMS'98, ENPC-INRIA, October 26-29, 1998.

[14] G. Strang, *On the construction and comparison of difference schemes*, SIAM J. Numer. Anal. 5, 506–517, 1968.

[15] J.G. Verwer and H.B. de Vries, *Global extrapolation of a first order splitting method*, SIAM J. Sci. Stat. Comput. 6, 771–780, 1985.

[16] J.G. Verwer, W. Hundsdorfer and J.G. Blom, *Numerical time integration for air pollution models*, CWI Report MAS-R9825, 1998.