Properties of two additive splitting procedures

Istvan Farago[1], Per Grove Thomsen[2] and Zahari Zlatev[3]

[1] Eötvös Loránd University
Pázmány P. s. 1/c, 1117 Budapest, HUNGARY
[2] Institute for Informatics and Mathematical Modelling
Technical University of Denmark DK-2800 Lyngby, DENMARK
[3] National Environmental Research Institute
Frederiksborgvej 399, P. O. Box 358, DK-4000 Roskilde, DENMARK

Abstract. Two additive splitting procedures are defined and studied in this paper. It is shown that these splitting procedures have good stability properties. Some other splitting procedures, which are traditionally used in mathematical models used in many scientific and engineering fields, are sketched. Some conclusions, which are related to the comparison of the additive splitting procedures with the other splitting procedures, are drawn.
Key Words: operator splitting, numerical methods, time discretization, stability
AMS Classification: 65N06, 65N30, 65L05.

1 Introduction

In our investigation we will assume that there are only two operators, i.e. we will demonstrate our methods on the Cauchy problem of the form:

$$\left.\begin{aligned} \frac{dw(t)}{dt} = (A+B)w(t), \quad t \in (0,T] \\ w(0) = w_0. \end{aligned}\right\} \tag{1}$$

We assume that the operators are bounded linear operators, and hence the exact solution is $w(t) = \exp(t(A+B))w(0)$.

2 Traditional operator splitting methods

In this section we give a short overview of the different known operator splitting methods which which are used in many different applications. For more details, see [8], [1], [3].

2.1 Sequential splitting

If the sub-problems involving operators A and B from (1) are treated one after the other, then the resulting algorithm is called a sequential splitting procedure.

It is very often worthwhile to describe how the different splitting procedures are to be applied in practice by

– giving the order in which the simple operators A and B, are applied, and
– indicating the splitting time-step size τ, which is actually used.

In our particular case, see again (1), the application of the sequential splitting procedure at a given splitting time-step can be described by sequence

$$(A)_\tau,\ (B)_\tau. \tag{2}$$

It is necessary to explain how the two sub-models are coupled. Assume that n splitting time-steps have successfully been performed and the next splitting time-step, time-step $n+1$, has to be carried out. The approximation obtained at time-step n is used as a starting approximation when the first sub-model is treated. The approximation obtained at the end of computations related to the first sub-model is used as a starting value for the second sub-model. The approximation obtained when the computations related to the second sub-model are accomplished is accepted as an approximation of the solution of problem (1) at time-step $n+1$. In this way everything is prepared to start the computations related to time-step $n+2$. It is necessary to explain how to start the computations at time-step 1, but this is not causing problems, because it is assumed that $w(0) = w_0$ is given; see again (1).

It should be noted here that if we change the order of the application of the operators, then the results will normally not be the same, i.e. the sequence $(A)_\tau,\ (B)_\tau$ is in general different from the sequence $(B)_\tau,\ (A)_\tau$.

The sequential splitting is in general leading to a numerical approximation of order one. The implication of this fact is that as a rule it is not advisable to use numerical algorithms of order higher than one in the treatment of the sub-problems involving the simpler operators A and B when a sequential splitting procedure is to be used.

2.2 Marchuk–Strang splitting

Sometimes it is desirable to apply more accurate splitting procedures. Accuracy of order two can be achieved in the following way. Consider an arbitrary splitting time-step, say step n (i.e. the computations are to be carried out from $t = t_n$ to $t = t_{n+1} = t_n + \tau$). Assume that the sub-models are treated as follows:

– Carry out computations by using the first operator from $t = t_n$ to $t = t_n + 0.5\tau$
– Use the second operator to perform computations from $t = t_n$ to $t = t_n + \tau$.
– Perform computations from $t = t_n$ to $t = t_n + \tau$ by applying again the first operator.

This splitting procedure has been proposed in 1968 simultaneously by Marchuk and Strang (see [5] and [6]). It is also called symmetric splitting

In the notation used in the previous sub-section the calculations at an arbitrary splitting time-step can be described by the sequence:

$$(A)_{0.5\tau},\ (B)_{\tau},\ (A)_{0.5\tau}. \tag{3}$$

The same idea, as in the previous section, can be used to couple the used operators: (i) we start with the approximation obtained at time-step n, (ii) the approximation obtained when the computations related to $(A)_{0.5\tau}$ are accomplished is taken as a starting value for the second sub-model $(B)_{\tau}$, (iii) the approximation obtained when the computations related to $(B)_{\tau}$ are accomplished is taken as a starting value for the third sub-model and (iv) the approximation obtained when the computations related to $(B)_{\tau}$ are accomplished is taken as an approximation of the problem (1) at time-step $n+1$. It is clear that after (iv) we are ready to proceed by performing the computations at splitting time-step $n+2$.

As in the previous section it should be noted here that if we change the order of the application of the operators, then the results will normally not be the same, i.e. the sequence $(A)_{0.5\tau},\ (B)_{\tau}\ (A)_{0.5\tau}$ is in general different from the sequence $(B)_{0.5\tau},\ (A)_{\tau}\ (B)_{0.5\tau}$.

2.3 Weighted sequential splitting

Consider again (1). Assume that the computations at the splitting time-step under consideration are carried out by applying successively the sub-problems as shown by the following two sequences of operators:

$$(A)_{\tau},\ (B)_{\tau} \tag{4}$$

and

$$(B)_{\tau},\ (A)_{\tau}, \tag{5}$$

i.e. we perform first a sequential splitting time-step starting with the sub-problem containing operator A and after that proceed with a sequential splitting time-step starting with operator B.

Denote by $w_{forward}(t_n+\tau)$ the result obtained when the splitting time-step (4) is finished and by $w_{backward}(t_n+\tau)$ the corresponding result obtained when the computations involved in (5) are accomplished. Then an approximation:

$$w_{n+1} = \theta\ w_{forward}(t_n+\tau) + (1-\theta)\ w_{backward}(t_n+\tau) \tag{6}$$

of the solution $w(t_n+\tau)$ of (1) at $t = t_{n+1} = t_n+\tau$ is calculated using some weighting parameter θ and one can proceed with the calculations for the next splitting time-step.

The splitting procedures based on the performance of the calculations at each splitting time-step using the formulae (4), (5) and (6) are called weighted sequential splitting procedures. Very often these splitting procedures are used with $\theta = 0.5$. The order of accuracy of the weighted sequential procedures is in general one, but experimental results indicate that these procedures are more stable than the simple sequential procedure.

It is clear that the sub-problems involved in (4) and (5) can be coupled in the same way as this was done for the sequential splitting procedure in Sub-section 2.

2.4 Weighted Marchuk–Strang splitting

A weighted Marchuk–Strang splitting procedure can be obtained from the ordinary Marchuk–Strang splitting procedure in the same way as the weighted sequential procedure was obtained from the ordinary sequential procedure in the previous sub-section. Assume that the computations at the splitting time-step under consideration are carried out by applying successively the sub-problems as shown by the following two sequences of operators:

$$(A)_{0.5\tau},\ (B)_{\tau}\ (A)_{0.5\tau} \tag{7}$$

and

$$(B)_{0.5\tau},\ (A)_{\tau}\ (B)_{0.5\tau} \tag{8}$$

i.e. we perform first a Marchuk–Strang splitting time-step starting with the sub-problem containing operator A and after that proceed with a Marchuk–Strang splitting time-step starting with operator B.

Denote again by $w_{forward}(t_n+\tau)$ the result obtained when the splitting time-step (7) is finished and by $w_{backward}(t_n+\tau)$ the corresponding result obtained when the computations involved in (8) are accomplished. Then an approximation:

$$w_{n+1} = \theta\ w_{forward}(t_n+\tau) + (1-\theta)\ w_{backward}(t_n+\tau) \tag{9}$$

of the solution $w(t_n+\tau)$ of (1) at $t = t_{n+1} = t_n + \tau$ is calculated using some weighting parameter θ and one can proceed with the calculations for the next splitting time-step.

The splitting procedures based on the performance of the calculations at each splitting time-step using the formulae (7), (8) and (9) are called weighted Marchuk–Strang splitting procedures or weighted symmetric splitting procedures. Very often these splitting procedures are used with $\theta = 0.5$. The order of accuracy of the weighed sequential procedures is in general the same as the order of the ordinary Marchuk–Strang splitting procedures (i.e. two), but experimental results indicate that these procedures are more stable than the simple sequential procedure.

It is clear that the sub-problems involved in (7) and (8) can be coupled in the same way as this was done for the sequential splitting procedure in Sub-section 2.2.

2.5 Extension for the case of more than two operators

The splitting procedures sketched in the previous four sub-sections have been derived for the case where the right-hand-side of the original problems contains as in (1) a sum of two operators. In principle at least, all these procedures can be defined for the case where the right-hand-side of the original problem is a sum of more than two operators (see, for example, [8]).

3 Additive splitting

This method is based on a simple idea: we solve the different sub-problems by using the same initial function. We obtain the split solution by the use of these results and the initial condition. The algorithm is the following:

$$\begin{aligned} &\frac{dw_1^n}{dt}(t) = Aw_1^n(t), \qquad (n-1)\tau < t \leq n\tau, \\ &w_1^n((n-1)\tau) = w_{\mathrm{sp}}^N((n-1)\tau), \end{aligned} \tag{10}$$

and

$$\begin{aligned} &\frac{dw_2^n}{dt}(t) = Bw_2^n(t), \qquad (n-1)\tau < t \leq n\tau, \\ &w_2^n((n-1)\tau) = w_{\mathrm{sp}}^N((n-1)\tau). \end{aligned} \tag{11}$$

Then the split solution at the mesh-points is defined as

$$w_{\mathrm{sp}}^N(n\tau) = w_1^n(n\tau) + w_2^n(n\tau) - w_{\mathrm{sp}}^N((n-1)\tau). \tag{12}$$

Here $n = 1, 2, \ldots N$, where $w_{\mathrm{sp}}^N(0) = w_0$.

One can see the main advantage of this method at first sight: it can be parallelized on the operator level in a natural way (like the symmetrically weighted sequential splitting).
The local splitting error (for bounded operators) can be investigated directly.

Theorem 1. *The additive splitting is a first order accurate splitting method for the bounded operators.*

Proof. The solution of the additive splitting at $t = \tau$ is defined as

$$w_{\mathrm{sp}}^N(\tau) = [\exp(A\tau) + \exp(B\tau) - I]w_0, \tag{13}$$

where I denotes the identity operator. Hence, we get

$$\begin{aligned} w_{sp}^N(\tau) &= \Big(I + A\tau + \frac{1}{2}A^2\tau^2 + I + B\tau + \frac{1}{2}B^2\tau^2 - I\Big)w_0 + \mathcal{O}(\tau^3) = \\ &= \Big(I + (A+B)\tau + \tfrac{1}{2}(A^2+B^2)\tau^2\Big)w_0 + \mathcal{O}(\tau^3). \end{aligned} \tag{14}$$

Comparing this expression with the similar Taylor expansion of the exact solution, we get the local splitting error

$$Err_{\mathrm{sp}}(\tau) = \frac{1}{2}\big((AB + BA)\tau^2\big)w_0 + \mathcal{O}(\tau^3), \tag{15}$$

which proves the statement. ■

The following statement shows that the additive splitting approximates the exact solution not only at the mesh-points.

Corollary 2. *The additive splitting approximates the exact solution on the whole split time interval* $[0, \tau]$.

Proof. The exact solution of the additive splitting at $t \in [0, \tau]$ is

$$w_{\mathrm{sp}}^N(t) = [\exp(At) + \exp(Bt) - I]w_0. \tag{16}$$

Hence, at any point of the time interval we have

$$Err_{\mathrm{sp}}(t) = \frac{1}{2}\big((AB + BA)t^2\big)w_0 + \mathcal{O}(t^3), \tag{17}$$

which proves the statement.

4 Modified additive splitting method

For the additive splitting introduced by (10)- (12) the critical point is the proof of the stability. The problem is based on the definition $w_{\mathrm{sp}}^N(t)$ in (12): in fact, we should estimate in the norm the operator $\exp(At)+\exp(Bt)-I$ and to show that in norm it is bounded by one. Due to the presence of the subtraction in the formula. the triangle inequality doesn't useful for this aim even in the case when the sub-problems are contractive,

Therefore we modify the method. We execute the separated splitting steps with double operators, and then we compute the splitting approximation as the arithmetical mean of the results. Consequently, the algorithm reads as follows:

$$\begin{aligned} &\frac{dw_1^n}{dt}(t) = 2Aw_1^n(t), \qquad (n-1)\tau < t \leq n\tau, \\ &w_1^n((n-1)\tau) = w_{\mathrm{sp}}^N((n-1)\tau), \end{aligned} \tag{18}$$

and

$$\begin{aligned} &\frac{dw_2^n}{dt}(t) = 2Bw_2^n(t), \qquad (n-1)\tau < t \leq n\tau, \\ &w_2^n((n-1)\tau) = w_{\mathrm{sp}}^N((n-1)\tau). \end{aligned} \tag{19}$$

Then the split solution at the mesh-points is defined as

$$w_{\mathrm{sp}}^N(n\tau) = \frac{1}{2}\left(w_1^n(n\tau) + w_2^n(n\tau)\right). \tag{20}$$

Here, as before $n = 1, 2, \dots N$, where $w_{\rm sp}^N(0) = w_0$.
The investigation of the local splitting error (for bounded operators) for the modified additive splitting (18)- (20) is similar as it was done in Theorem 1.

Theorem 3. *For the bounded operators the modified additive splitting is a first order accurate splitting method.*

Proof. The solution of the modified additive splitting at $t = \tau$ is defined as

$$w_{\rm sp}^N(\tau) = \frac{1}{2}[\exp(2A\tau) + \exp(2B\tau)]w_0, \tag{21}$$

Hence, we get

$$\begin{aligned} w_{sp}^N(\tau) = \frac{1}{2}\Big(I + 2A\tau + \frac{1}{2}(2A\tau)^2 + I + 2B\tau + \frac{1}{2}(2B\tau)^2\Big)w_0 + \mathcal{O}(\tau^3) = \\ = \Big(I + (A+B)\tau + (A^2 + B^2)\tau^2\Big)w_0 + \mathcal{O}(\tau^3). \end{aligned} \tag{22}$$

Comparing this expression with the similar Taylor expansion of the exact solution, we get the local splitting error

$$Err_{\rm sp}(\tau) = -\frac{1}{2}\big((A-B)^2\tau^2\big)w_0 + \mathcal{O}(\tau^3), \tag{23}$$

which proves the statement. ■

I the following we analyze the approximation property of the method on the whole time interval.

Corollary 4. *The additive splitting approximates the exact solution on the whole split time interval* $[0, \tau]$.

Proof. The exact solution of the modified additive splitting at $t \in [0, \tau]$ is

$$w_{\rm sp}^N(t) = \frac{1}{2}[\exp(2At) + \exp(2Bt)]w_0. \tag{24}$$

Hence, at any point of the time interval we have

$$Err_{\rm sp}(t) = -\frac{1}{2}\big((A-B)^2t^2\big)w_0 + \mathcal{O}(t^3), \tag{25}$$

which proves the statement.

5 Convergence of the modified additive splitting for contractive bounded operators

The introduced additive splittings, as any operator splitting method, can be considered as a numerical time-discretization method. Hence, the question of its convergence is a crucial question during the application. This leads to the problem of the investigation of its consistency and stability. In the following we assume that the exponential of the operators tA and tB are contractive, i.e the relations

$$\| \exp(tA)\| \leq 1 \text{ and } \| \exp(tB)\| \leq 1 \tag{26}$$

are valid.

The notion of *consistency* is closely related to the local splitting error. Let us denote by $T_{\mathrm{sp}}(\tau)$ the operator which transforms $w_{\mathrm{sp}}^N(t_n)$ into $w_{\mathrm{sp}}^N(t_{n+1})$, i.e.

$$T_{\mathrm{sp}}(\tau) w_{\mathrm{sp}}^N(t_n) = w_{\mathrm{sp}}^N(t_{n+1}). \tag{27}$$

Obviously, for the additive splitting

$$T_{\mathrm{sp}}(\tau) = \exp(\tau A) + \exp(\tau B) - I, \tag{28}$$

while for the modified additive splitting it is

$$T_{\mathrm{sp}}(\tau) = \frac{1}{2}(\exp(2\tau A) + \exp(2\tau B)). \tag{29}$$

Definition 5. A splitting method called consistent numerical method to the problem (1), when

$$\lim_{\tau\to 0} \| \left[\frac{T_{\mathrm{sp}}(\tau) - I}{\tau} - (A+B) \right] w(t) \| = 0 \tag{30}$$

for all $t \in [0, T]$ and the convergence is uniform in t. (Here $w(t)$ denotes the solution of the problem (1).)

Remark. For the solution of the well-posed abstract Cauchy problem (1) the relation

$$\lim_{\tau\to 0} \| \left[\frac{w(t+\tau) - w(t)}{\tau} - (A+B) \right] w(t) \| = 0 \tag{31}$$

holds for all $t \in [0, T]$ and the convergence is uniform in t. Therefore, using (30), a splitting method is consistent if and only if the relation

$$\lim_{\tau\to 0} \| \frac{w(t+\tau) - T_{\mathrm{sp}}(\tau) w(t)}{\tau} \| = 0 \tag{32}$$

for all $t \in [0, T]$ and the convergence is uniform in t. Substituting $t = 0$ into (32), we get the relation

$$\lim_{\tau\to 0} \| \frac{Err_{\mathrm{sp}}(\tau)}{\tau} \| = 0 \tag{33}$$

which coincides with the condition

$$Err_{\mathrm{sp}}(\tau) = \mathcal{O}(\tau^{p+1}) \tag{34}$$

with some $p > 0$. Consequently, when an operator splitting is a consistent then its order can be defined by the rate of the convergence in (30). However, the converse isn't true: the relation (34) doesn't imply automatically the consistency.

The notion of *stability* means that the powers of the operator $T_{\mathrm{sp}}(\tau)$ are bounded on the initial element, i.e.

Definition 6. A splitting method called stable numerical method if

$$\|(T_{\mathrm{sp}}(\tau))^n w_0\| < \infty \tag{35}$$

for all $n\tau \in [0, T]$.

For the well-posed problem (1) the Lax theorem states the following: for a consistent numerical method the stability is the necessary and sufficient condition of the convergence.

Theorem 7. *Under the contractivity condition (26) both the additive and the modified splittings are consistent.*

Proof. First we show the validity of (30) for $T_{\mathrm{sp}}(\tau)$ defined in (29). We have

$$\begin{aligned}
&\left[\frac{T_{\mathrm{sp}}(\tau) - I}{\tau} - (A+B)\right] w(t) = \\
&= \left[\frac{\frac{1}{2}(\exp(2\tau A) + \exp(2\tau B)) - I}{\tau} - (A+B)\right] \exp(t(A+B))w_0 = \\
&= \left[\tau(A^2 + B^2) + \mathcal{O}(\tau^2)\right] \exp(t(A+B))w_0
\end{aligned} \tag{36}$$

According to (26), $\|\exp(t(A+B))\| \leq 1$. Hence, the right side of (36) converges to zero as τ tends to zero, uniformly in t. This results in the consistency of the modified additive splitting.
The proof of the statement for the additive splitting is similar. ■

Due to the consistency property, we have to show the stability. However, as it was already mentioned, for the additive splitting (10)- (12) to do it even for contractive operators, it is a difficult task. But for the modified additive splitting, in the contractive case we can show it directly.

Theorem 8. *Under the contractivity condition (26) the modified additive splitting is contractive and hence it is stable.*

Proof. The statement is an obvious consequence of the assumptions and of the formula (20). ■

Hence, using the Lax theorem, we obtained

Theorem 9. *Under the contractivity condition (26) the modified additive splitting is convergent.*

6 General discussion of the splitting procedures

6.1 Computing time

Assume here that all splitting procedures are used with the same splitting time-stepsize τ) and that the computing time is the most important factor. Under these assumptions the following conclusions can be drawn.

- From a computational point of view, the sequential splitting is the best one (i.e. it is the cheapest one with regard to computing time).
- If the computational work for the treatment of the sub-problem involving operator B is much larger than the computational work for treating the sub-problem involving operator A, then the Marchuk–Strang splitting procedure is only slightly more expensive computationally than the sequential splitting procedure. In some environmental models this is precisely the case. In advection-chemistry models (see, for example, [7], [8]) the chemical part often takes about 95% of the computing time. It is clear that for such models the Marchuk–Strang splitting procedure is quite competitive with the sequential splitting procedure with regard to the computing time used.
- The remaining splitting procedures are much more expensive with regards to the computing time than the sequential splitting procedure. The computing time will be increased by a factor at least equal to two when any of these splitting procedures is applied instead of the sequential splitting procedure.

6.2 Accuracy of the results

Assume that the accuracy of the results is the most important issue. Then it is perhaps a good idea to apply the Marchuk–Strang splitting procedure or the weighted Marchuk–Strang splitting procedure, because second order accuracy can be achieved when these splitting procedures are used. It should be emphasized, however, that the use of these two splitting procedures imposes some requirements to the numerical methods that are to be used in the three sub-problems arising when the splitting procedure selected is applied to (1): in general the numerical methods for solving differential equations that are to be applied in each sub-problems must be of order at least equal to two.

The remaining splitting procedures are of order one. This means, let us reiterate, that these splitting procedures should not be used when the accuracy required is a critical factor.

6.3 Robustness of the results

Some experimental results indicate that the weighted splitting procedures are sometimes more robust than the remaining splitting procedures (i.e. in some difficult situations when the solution changes very quickly these two splitting procedures are giving rather good results; such situations occur during sun-sets and sun-rises when some chemical models are run, see [8]).

6.4 Proving theoretical properties

It is important to prove some theoretical results (as, for example, convergence results, stability results, etc.). If such results are proved, then it will be clear for which classes of problems it is possible to guarantee that the splitting procedures will ensure good results. Some properties of the additive splitting procedures were proved in this paper. Much more similar results are highly desirable.

6.5 General conclusion

The short discussion given in the previous four sub-section shows very clearly that the choice of the best splitting procedure is by far not an easy task. In the different situations arise different requirement and very often these requirement are working in opposite directions. Therefore, in each particular situation one has first to find out what is the most important requirement and to select the splitting procedure according to this most important requirement. It is also necessary to remember that it is necessary to select a good splitting procedure, but it might turn out that this is not sufficient. The numerical methods for solving differential equations, which have to be used during the numerical treatment of every sub-problem, are also very important and one must be very careful when such methods are to be used in combination with splitting procedures.

7 Conclusions and closing remarks

We have introduced new consistent splittings, namely, the additive splitting and the modified additive splitting. The main benefit of these methods is their algorithmic simplicity, more precisely its natural parallelization on the operator level. However, these methods have some drawbacks, too.

1. Stability of the additive splitting is a crucial problem and to prove it seems to be very difficult. In practice, we have obtained a stable (and hence convergent) solution only for small time steps.
2. The additive splitting has always first order accuracy, even for commuting operators. Hence, it is not accurate enough for many practical applications.

References

1. Csomós, P., Faragó, I., Havasi , Á.: *Weighted sequential splittings and their analysis*, Comput. Math. Appl., **50** (2005) 1017-1031.
2. Faragó, I., Geiser, J. *Iterative operator- splitting methods for linear problems*, Weierstass Institute für Angewandte Analysis, **1043** (2005) 1-18. International Journal of Computational Science and Engineering (to appear)
3. Hundsdorfer, W., Verwer, J. G.: *Numerical solution of time-dependent advection-diffusion-reaction equations*, Springer, Berlin, 2003.
4. Horváth, R.: *Operator splittings for the numerical solution of the Maxwell's equation*, Lecture Notes in Computer Science, **3743** (2006) 363-371.

5. Marchuk, G. I. (1968). *Some application of splitting-up methods to the solution of mathematical physics problems.* Applik. Mat., 13(2).
6. Strang, G. (1968). *On the construction and comparison of difference schemes.* SIAM J. Numer. Anal., 5: 505-517.
7. Zlatev, Z.: *Computer treatment of large air pollution models*, Kluwer Academic Publishers, Dordrecht-Boston-London, 1995.
8. Zlatev, Z. and Dimov, I.: *Computational and numerical challenges in environmental modelling*, Studies in Computational Mathematics 13, Elsevier, Amsterdam, 2006.