UDC 004.75

ON THE EFFICIENCY OF THE OPERATOR SPLITTING METHOD

Faragó István

Department of Applied Analysis, Eötvös Loránd University
H-1117 Budapest, Pázmány P. s. 1/c, Hungary
Email: faragois@cs.elte.hu

Operator splitting is a widely used procedure in the numerical solution of complex problems. The point in operator splitting is the replacement of the original model with one in which appropriately chosen groups of the sub-processes, described by the model, take place successively in time. This de-coupling procedure allows us to solve a few simpler problems instead of the whole one. In this paper we investigate the application of the operator splitting method to real-life problems from a practical point of view, with great emphasis on long-range air pollution transport. We discuss the main benefits and drawbacks of this approach.
Key-words: air pollution modelling, operator splitting, numerical solution.

Introduction

The operator splitting method (OSM) is a widely used method for solving real-life problems. It can be applied in the numerical modelling process of many different time-dependent complex physical phenomena. The original continuous mathematical model of such phenomena can be described in the form of an abstract Cauchy problem (ACP) as follows:

$$\begin{cases} \dfrac{du(t)}{dt} = \sum_{i=1}^{n} A_i u(t), & t \in (0,T], \\ u(0) = u_0, \end{cases} \qquad (1)$$

where u denotes the unknown function, u_0 is a given element which describes the initial state of the process; A_i (for i = 1, 2,…n) are given (usually densely defined linear) operators.

Frequently, the OSM is applied to the so called semi-discretized problem. This means that the original continuous problem has already been discretized w.r.t. the space variables. Then (1) denotes the system of ordinary differential equations obtained in this way. Clearly, for linear problems A_i are matrices.

1. An important example: air pollution modelling

The transport of air pollutants is one of the most widely investigated phenomena for which (1) can serve as a mathematical model.

Let $c_j = c_j(\mathbf{x},t)$ denote the concentration of the j-th air pollutant, and c the vector function of these functions. Then the time evolution of the vector c can be described mathematically by the system of partial differential equations (Zlatev, 1995, Prussov and Dorosenko, 2003)

$$\left. \begin{aligned} \frac{\partial c}{\partial t} &= -\nabla(\mathbf{u}c) + \nabla(K\nabla c) + E - \sigma c + R(c) \qquad t \in (0,T] \\ c(\mathbf{x},0) &= c_0(\mathbf{x}) \end{aligned} \right\} \qquad (2)$$

where $\mathbf{u} = \mathbf{u}(\mathbf{x},t)$ is a vector-function describing the wind velocity, $K = K(\mathbf{x},t)$ is the diffusion coefficient function, $E = E(\mathbf{x},t)$ is the function of emission, $\sigma = \sigma(\mathbf{x},t)$ describes the deposition and R defines the chemical reactions of the pollutants. The initial function $c_0(x)$ is given. Using these notations, the terms in equation (10) have the following physical meaning. The first term on the right-hand side describes the transportation due to the velocity field, which is called *advection*. The second term expresses the *turbulent diffusion*, the third term the *emission*, the fourth term describes the *deposition* and the last term defines the *chemistry* of the pollutants.

Then, using the notation

- ♦ $A_1 c = -\nabla(\mathbf{u}c)$ - the advection operator
- ♦ $A_2 c = \nabla(K\nabla c)$ - the diffusion operator
- ♦ $A_3 c = -\sigma c$ - the deposition operator

♦ $A_4c = E$ - the emission operator

♦ $A_5c = R(c)$ - the chemistry operator

the problem (2) can be written in the form (1).

The number of chemical species involved in a modern air pollution model sometimes reaches 200, or even more, which results in a huge system of partial differential equations. The analytical solution of such a problem is obviously impossible to find. Hence we have to treat it numerically. We note that in case of semi-discretization usually the number of spatial grid points equals many millions. This means that the system of ordinary differential equations obtained after spatial discretization is extremely big, hence the use of any numerical method developed for systems of ODE's is rather complicated. Moreover, the model equations contain terms that have different physical meanings and so different mathematical properties (e.g. linear, non-linear, stiff and non-stiff). Therefore, it is impossible to find such a universal numerical method which would perform well when applied directly to the original system. The application of operator splitting allows us to treat the different physical terms separately.

Operator splitting method (OSM) is a kind of problem decomposition: we divide the spatial differential operator of the global system into a few simpler operators and solve the corresponding problems one after the other, by connecting them through their initial conditions.

The simpler systems which are obtained in this manner, and are sometimes called sub-systems, might have some special properties that can be exploited in the numerical solution. The sub-systems are usually easier to treat numerically than the whole system.

Splitting can be performed in several ways. We expect the method to be accurate as well as efficient enough. The latter property depends on the number of computations and the possibility of performing the computations in parallel. Taking into account the latter requirement, we made attempts to construct a new splitting scheme which does not require a lot of computational work, and is parallelizable on the operator level.

There are different kinds of OSM which can be applied to such problems. For their description and study we refer to the literature. (See. e.g. Hundsdorfer and Verwer, 2003, Zlatev, 1995, and Dimov et al., 2003 and 2005.) In the following we give an overview about the general benefits and drawbacks of the OSM. We emphasize that these properties are independent of the concrete OSM which was chosen.

2. Benefits of the operator splitting process

In this section we list those properties of the OSM which make it attractive during the mathematical modelling process of complex physical phenomena.

2.1 Easier theoretical investigation of the convergence. In fact, the OSM can be viewed as a one-step time-discretization method. Hence, it is quite natural to raise the question: if the different sub-problems are solved exactly, under which conditions is the split (discretized) solution convergent to the exact solution when the discretization parameter tends to zero? The answer is based on the famous Lax equivalence theorem, which roughly speaking means that the consistency and stability together imply the convergence. This means that we should check only the above two simpler properties.

To do it for the globally discretized unsplit problem is a very difficult (usually hopeless) task.

Since the consistency is very close to the local splitting error order with p > 0, therefore in practice we only have to check the stability. This latter property means that the norm of the split solution at any time level cannot be bigger than the norm of the initial function multiplied by some constant independent of τ. (C.f. Havasi (2001).) This constant is called stability constant and it clearly cannot be less than one. When it is equal to one, then the method is called contractive.

The stability: if both sub-problems are stable, it does not yield the stability of the whole split discretization method. However, when the sub-problems are contractive, then the total OSM is also contractive, and hence stable.

2.2 Choice of a suitable numerical method. The use of the OSM results in a sequence of sub-problems, which also represent Cauchy problems, but with simpler operators. However, these problems cannot be solved exactly, either. So, we should apply numerical methods to the sub-problems. The benefit of the OSM is that we can use different numerical methods to the different sub-problems according to the special features of the problems. After defining the numerical methods to each sub-problems, we can consider the global numerical algorithm as a discretization method to the original unsplit continuous model (1). This approach gives possibility to derive some well-known numerical schemes or create new methods. (See Faragó, (2005).) This approach makes it possible to increase the efficiency of the global algorithm.

2.3 Applicability of the existing software products. When we divide the original problem into a sequence of sub-problems, it is fairly reasonable to do it in such a way that the split tasks were standard problems to which the already existing software products (e.g. MATLAB library tool-boxes) are directly applicable. For instance, in the air pollution modelling the sub-problems, defined by the operators given in Section 3, are standard and can be used by using library program routines. We note that the OSM gives high flexibility in choosing the sub-operators. In the

choosing process the "standardization" requirements formulated above should coincide with some other requirements, too (see later).

2.4 Increase of the computational efficiency. The choice of the time-discretization parameter τ plays a crucial role in the efficiency of the applied numerical method. Usually, due to some stability conditions, we cannot choose it arbitrarily and there is an upper bound. However, our aim is to avoid this restriction because a too small τ usually results in a lot of difficulties. (In order to get the numerical solution on some fixed time level T we should solve the sub-problems on T/τ time levels, which may get extremely big.) But this latter fact causes some troubles:

- ♦ the computational work increases significantly,
- ♦ for stiff problems the numerical implementation is almost impossible,
- ♦ due to the big number of the arithmetic operations, the computational errors may increase dramatically.

Hence, our aim is to get rid of the too strict restriction and to allow to choose τ as large as possible.

We show that the OSM is a powerful method to increase the discretization parameter τ.

In the following we consider the following example:

$$\frac{\partial u(t)}{\partial t} = (10^6 + \sin xt)\frac{\partial^2 u(t)}{\partial x^2}, \quad t \in (0,T)$$

When we use the above problem with the explicit Euler method, then the condition of stability is

$$\frac{\tau}{h^2} \le \frac{1}{\max 2(10^6 + \sin xt)} \approx 0.5 \cdot 10^{-6},$$

which means that $\tau \approx 0.5 \cdot 10^{-6} h^2$. This yields that the choice h = 0.01 results in $\tau \approx 0.5 \cdot 10^{-10}$ for the upper bound. (When T = 1, then $2 \cdot 10^{11}$ time steps are required.) If we use the sequential splitting

$$\frac{\partial u_1(t)}{\partial t} = 10^6 \frac{\partial^2 u_1(t)}{\partial x^2}, \quad t \in ((k-1)\tau, k\tau])$$

$$\frac{\partial u_2(t)}{\partial t} = \sin(xt)\frac{\partial^2 u_2(t)}{\partial x^2}, \quad t \in ((k-1)\tau, k\tau])$$

which are connected via the initial conditions, then we can observe that the first sub-problem can be solved by spectral method, which does not require any mesh. Hence, we need to construct the mesh only for the second sub-problem. Applying the explicit Euler method to this problem, the stability bound turns into

$$\frac{\tau}{h^2} \le \frac{1}{2 \cdot \max \sin(xt)} \le 0.5.$$

This means that for the same fixed space-discretization parameter h we can select a 10^6-times bigger splitting step size τ within the stability condition.

2.5 Use of numerical-analytical methods. In certain cases, even if the original problem cannot be solved analytically, after splitting one of the sub-problems can be solved analytically. Hence the possible choice of the time-discretization parameter τ can be enlarged again. This is demonstrated in the following example.

$$\frac{\partial u(t)}{\partial t} = (10^6 + \sin xt)\frac{\partial u(t)}{\partial x},$$

$$u(x,0) = u_0(x)$$

for t > 0, $x \in \Re$, where $u_0(x)$ is a given function. Using the explicit Euler finite difference scheme, the well-known CFL condition implies the bound

$$\frac{\tau}{h} \le \frac{1}{10^6}.$$

Applying the sequential splitting we get the sub-problems

$$\frac{\partial u_1(t)}{\partial t} = 10^6 \frac{\partial u_1(t)}{\partial x}, \quad t \in ((k-1)\tau, k\tau])$$

$$\frac{\partial u_2(t)}{\partial t} = \sin(xt)\frac{\partial u_2(t)}{\partial x}, \quad t \in ((k-1)\tau, k\tau])$$

connected again via the initial conditions.

As one can see, the first sub-problem can be solved analytically by using the D'Alambert formula, because the coefficient in the elliptic part is constant. At the same time, using the explicit Euler method to the second sub-problem, for the bound of the method we obtain

$$\frac{\tau}{h} \le 1.$$

This means that for some fixed *h* we could increase the possible choice of τ 10^6-times bigger.

2.6 Preservation of the main qualitative properties. As we have already mentioned, the original physical phenomenon has a lot of basic qualitative properties, which are inherent to the physical process and originate from the physics of the process. The adequate continuous mathematical model also has these properties. Therefore, it is natural to require that the "good" discrete model also has the discrete analogue of the above qualitative properties. (For instance, such properties are the non-negativity preservation of the initial data, contractivity in time, maximum principle, shape preservation of the initial function, etc.) For some standard models and discretization methods (e.g. heat equation, wave propagation with finite difference and finite element method) this theory has been developed, and we know the conditions under which the models are qualitative property preserving.

However, when we discretize the original unsplit problem with some numerical method, then the obtained discrete model is not standard, and hence to check the validity of the discrete qualitative properties is a very difficult task.

Clearly, when we use some operator splitting method and such numerical methods for the sub-problems which preserve the qualitative properties, then the global discretization method is also qualitative property preserving. Therefore, when we split the unsplit problem into a sequence of "standard" sub-problems, we can give some sufficient conditions under which the qualitative properties are preserved.

3. Drawbacks of the operator splitting process

In the previous section we listed some advantages of the operator splitting approach. However, in our opinion, this method has some disadvantages, too. In the following we list those problems which arise by use of the OSM.

3.1. The suitable choice of the sub-operators. The original complex physical problem consists of different - usually simpler- processes. In fact, when we consider the problem (1), then our mathematical model is

$$\begin{cases} \dfrac{du(t)}{dt} = Au(t), & t \in (0,T], \\ u(0) = u_0, \end{cases} \tag{1a}$$

where the operator A describes the whole complex physical process. The partition of the operator A into the sum A = ΣA_i is not always natural. The choice of suitable sub-operators might be difficult. For instance, for the air pollution modelling the choice of the sub-operators, given in Section 2 is not unique. We can decompose the operator

$$Au = -\nabla(\mathbf{u}c) + \nabla(K\nabla c) + E - \sigma c + R(c)$$

also as the sum $Au = \sum_{i=1}^{5} A_i u(t)$, where now

♦ $A_1 c = -\sum_{i=1}^{2} \partial_i (u_i c)$ - the horizontal advection operator,

♦ $A_2 c = \sum_{i=1}^{2} \partial_i (k_i \partial_i c)$ - the horizontal diffusion operator,

♦ $A_3 c = -\sigma c$ - the deposition operator,

♦ $A_4 c = E + R(c)$ - the emission and chemistry operator,

♦ $A_5 c = \partial_3 (u_3 c) + \partial_3 (k_3 \partial_3 c)$ - the vertical transport operator.

This kind of decomposition is used in the Danish Eulerian Model (DEM) and is called DEM decomposition. The decomposition given in Section 2 is called physical decomposition.
The main advantage of the DEM decomposition is its high flexibility for 2D problems because only the last operator contains the vertical part.

3.2 Local splitting error. Replacing the original problem (1) with one of the above listed split models usually results in a new kind of error.
Let us denote the exact solution of the problem (1) by *u*(t), and by $u_{sp}(n\tau)$ the exact solution of the chosen split problem, respectively. Their difference at the point t = τ is called local splitting error, i.e.,

$$Err_{sp}(\tau) = u(\tau) - u_{sp}(\tau). \tag{4}$$

When $Err_{sp}(\tau) = O(\tau^{p+1})$, then the splitting method is called p-th order. Generally, the increase of the order of the splitting results in a more accurate split solution. Therefore the quality of a splitting can be characterised by its order, too. For the above discussed splitting methods the orders can be computed directly and they are

- ♦ for the sequential splitting one,
- ♦ for the Strang-Marchuk splittig two;
- ♦ for the symmetrically weighted sequential splitting two,
- ♦ for the additive splitting one;
- ♦ for the iterated splitting 2m+1.

Since the split sub-problems cannot be solved exactly, we should apply numerical methods to their solution. Hence, the obtained numerical result includes two kinds of errors: the local splitting error and the error of the numerical methods. The analysis of the interaction of these two errors is usually a very complicated task. Hence, the error analysis for the numerical solution obtained by use of some numerical method to the unsplit problem, is simpler and we can control it more easily than for the split models.

We note that under some condition the splitting error may disappear (for some splittings in case of commutativity of the operators) but these conditions are mostly unrealistic in real-life applications.

3.3. Handling the boundary conditions.
When using the OSM, a serious problem is handling the boundary conditions for the split problems. More precisely, the question is: how to describe the boundary conditions for the different sub-problems of different type? E.g. the simplified diffusion-advection model of (1) in 1D has the form

$$\left.\begin{aligned}
\frac{\partial c}{\partial t} &= -\nabla(\mathbf{u}c) + \nabla(K\nabla c), \quad x \in (0,1), \qquad t \in (0,T]; \\
c(x,0) &= c_0(x), \quad x \in (0,1).
\end{aligned}\right\} \tag{21}$$

This is a parabolic problem and hence we should define two boundary conditions, namely, at the point x = 0 and x = 1. However, using e.g. the sequential splitting, the first sub-problem (advection part) is a first order hyperbolic problem. For such a problem we can use only one boundary condition, i.e., one of them. Hence the boundary condition at the other point will not be satisfied which may cause some difficulties.

1. Dimov, I., Faragó, I., Zlatev, Z., 2003, Parallel computations with large-scale air pollution models, *Problems in Programming* 3, 44-52.
2. Dimov, I., Faragó, I., Havasi, Á., Zlatev, Z., 2006, Different splitting techniques with application to air pollution models, *Int. J. Environmental Pollution* (to appear).
3. Faragó, I., 2005, Splitting methods for abstract Cauchy problems, *Lect. Notes Comp. Sci.* 3401, Springer Verlag, Berlin, 35-45.
4. Havasi, Á., Bartholy, J., Faragó, I., 2001, Splitting method and its application in air pollution modeling. *Időjárás* 105, 39–58.
5. Hundsdorfer, W., Verwer, J. G., 2003, Numerical solution of time-dependent advection-diffusion-reaction equations. Springer, Berlin.
6. Prusov, V., Dorosenko, A., 2003, Modelling and forecasting atmospheric pollution over region, *Annales Univ. Sci. ELTE*, 46, 27-45.
7. Zlatev, Z., 1995, Computer treatment of large air pollution models. Kluwer, Amsterdam.
1. 456 – 1480