

## MODIFIED DIFFERENTIAL EQUATIONS \*

PHILIPPE CHARTIER<sup>1</sup>, ERNST HAIRER<sup>2</sup> AND GILLES VILMART<sup>1,2</sup>

**Abstract.** Motivated by the theory of modified differential equations (backward error analysis) an approach for the construction of high order numerical integrators that preserve geometric properties of the exact flow is developed. This summarises a talk presented in honour of Michel Crouzeix.

**Résumé.** Motivé par la théorie des équations modifiées (analyse rétrograde de l'erreur), une approche pour la construction de méthodes numériques d'ordre élevé préservant des propriétés géométriques du flot exact est développée. Ceci résume une présentation donnée en honneur de Michel Crouzeix.

### INTRODUCTION

Modified differential equations in combination with backward error analysis (cf. the monographs [3], [5]) form an important tool for studying the long-time behaviour of numerical integrators for ordinary differential equations. The main idea of this theory is sketched and, by inverting the roles of the exact and numerical flows, a new approach for the construction of high order numerical integrators for ordinary differential equations is developed [1]. As an application, a computationally efficient and highly accurate modification of the Discrete Moser–Veselov algorithm for the simulation of the free rigid body is presented [4].

#### 1. MODIFIED EQUATIONS FOR BACKWARD ERROR ANALYSIS

Consider an initial value problem

$$\dot{y} = f(y), \quad y(0) = y_0 \tag{1}$$

with sufficiently smooth vector field  $f(y)$ , and a numerical one-step integrator  $y_{n+1} = \Phi_{f,h}(y_n)$ . The idea of backward error analysis is to search for a modified differential equation

$$\dot{z} = f_h(z) = f(z) + hf_2(z) + h^2f_3(z) + \dots, \quad z(0) = y_0, \tag{2}$$

which is a formal series in powers of the step size  $h$ , such that the numerical solution  $\{y_n\}$  is formally equal to the exact solution of (2),

$$y_n = z(nh) \quad \text{for } n = 0, 1, 2, \dots, \tag{3}$$

see the left picture of Figure 1.

The idea of backward error analysis was originally introduced by Wilkinson (1960) in the context of numerical linear algebra. For the integration of ordinary differential equations it was not used until one became interested in the long-time behaviour of numerical solutions. Without considering it as a theory, Ruth [9] uses the

---

\* This work was partially supported by the Fonds National Suisse, project No. 200020-109158.

<sup>1</sup> INRIA Rennes, Campus de Beaulieu, 35042 Rennes-Cedex, France

<sup>2</sup> Section de Mathématiques, Université de Genève, 2-4 rue du Lièvre, CH-1211 Genève 4, Switzerland

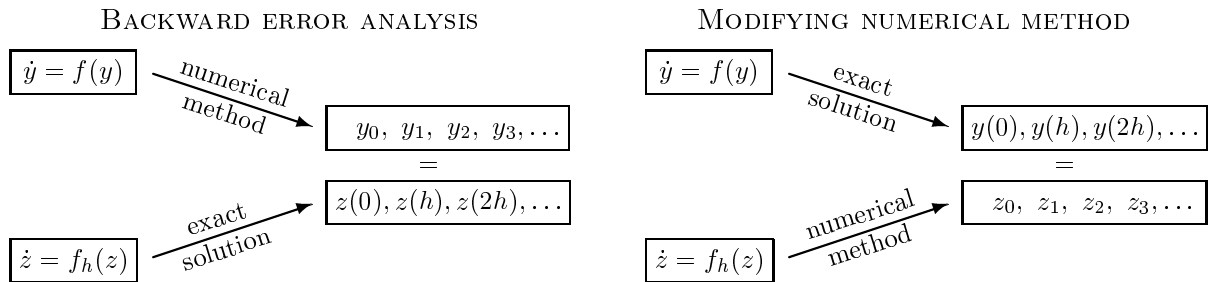


FIGURE 1. Backward error analysis opposed to modifying numerical integrators

idea of backward error analysis to motivate symplectic integrators for Hamiltonian systems. In fact, applying a symplectic numerical method to a Hamiltonian system  $\dot{y} = J^{-1}\nabla H(y)$  gives rise to a modified differential equation that is Hamiltonian. This permits to transfer known properties of perturbed Hamiltonian systems (e.g., conservation of energy, KAM theory for integrable systems) to properties of symplectic numerical integrators. One became soon aware that this kind of reasoning is not restricted to Hamiltonian systems, and new insight can be obtained with the same techniques also for reversible differential equations, for Poisson systems, for divergence-free problems, etc. A rigorous analysis has been developed in the nineties. We refer the interested reader to [3, Chapter IX], where backward error analysis and its applications are explained in detail.

## 2. MODIFYING NUMERICAL INTEGRATORS

Backward error analysis is a purely theoretical tool that gives much insight into the long-term integration with geometric numerical methods. We shall show that by simply exchanging the roles of the “numerical method” and the “exact solution” (cf. the two pictures in Figure 1), it can be turned into a means for constructing high order integrators that conserve geometric properties. They will be useful for integrations over long times.

Let us be more precise. As before, we consider an initial value problem (1) and a numerical integrator. But now we search for a modified differential equation, again of the form (2), such that the numerical solution  $\{z_n\}$  of the method applied with step size  $h$  to (2) yields formally the exact solution of the original differential equation (1), i.e.,

$$z_n = y(nh) \quad \text{for } n = 0, 1, 2, \dots, \quad (4)$$

see the right picture of Figure 1. Notice that this modified equation is different from the one considered before. However, due to the close connection with backward error analysis, all theoretical and practical results have their analogue in this new context. The modified differential equation is again an asymptotic series that usually diverges, and its truncation inherits geometric properties of the exact flow if a suitable integrator is applied. The coefficient functions  $f_j(z)$  can be computed recursively by using a formula manipulation program like MAPLE. This can be done by developing both sides of  $z(t+h) = \Phi_{f_h, h}(z(t))$  into a series in powers of  $h$ , and by comparing their coefficients. Once a few functions  $f_j(z)$  are known, the following algorithm suggests itself.

**Algorithm 2.1** (modifying integrator). *Consider the truncation*

$$\dot{z} = f_h^{[r]}(z) = f(z) + hf_2(z) + \dots + h^{r-1}f_r(z) \quad (5)$$

of the modified differential equation corresponding to  $\Phi_{f_h, h}(y)$ . Then,

$$z_{n+1} = \Psi_{f_h, h}(z_n) := \Phi_{f_h^{[r]}, h}(z_n)$$

defines a numerical method of order  $r$  that approximates the solution of (1). We call it *modifying integrator*, because the vector field  $f(y)$  of (1) is modified into  $f_h^{[r]}$  before the basic integrator is applied.

This is an alternative approach for constructing high order numerical integrators for ordinary differential equations (classical approaches are multistep, Runge–Kutta, Taylor series, extrapolation, composition, and splitting methods). It is particularly interesting in the context of geometric integration because, as known from backward error analysis, the modified differential equation inherits the same structural properties as (1) if a suitable integrator is applied.

A few known methods can be cast into the framework of modifying integrators although they have not been constructed in this way. The most important are the generating function methods as introduced by Feng [2]. These are high order symplectic integrators obtained by applying a simple symplectic method to a modified Hamiltonian system. The corresponding Hamiltonian is the solution of a Hamilton–Jacobi partial differential equation. Another special case is a modification of the discrete Moser–Veselov algorithm for the Euler equations of the rigid body, proposed by McLachlan and Zanna [7]. The general approach of Algorithm 2.1 is introduced and discussed in [1].

**Example 2.2.** For the numerical integration of (1) we consider the implicit midpoint rule

$$y_{n+1} = y_n + h f\left(\frac{y_n + y_{n+1}}{2}\right). \quad (6)$$

The truncated modified vector field corresponding to this method is

$$\begin{aligned} f_h^{[5]} &= f + \frac{h^2}{12} \left( -f' f' f + \frac{1}{2} f''(f, f) \right) + \frac{h^4}{120} \left( f' f' f' f' f - f''(f, f' f' f) + \frac{1}{2} f''(f' f, f' f) \right) \\ &+ \frac{h^4}{240} \left( -\frac{1}{2} f' f' f''(f, f) + f' f''(f, f' f) + \frac{1}{2} f''(f, f''(f, f)) - \frac{1}{2} f^{(3)}(f, f, f' f) \right) \\ &+ \frac{h^4}{80} \left( -\frac{1}{6} f' f^{(3)}(f, f, f) + \frac{1}{24} f^{(4)}(f, f, f, f) \right) \end{aligned} \quad (7)$$

and applying the midpoint rule to  $\dot{z} = f_h^{[5]}(z)$  yields a numerical approximation of order 6 for (1). At first glance this modified equation looks extremely complicated and it is hard to imagine that the modifying midpoint rule can compete with other methods of the same order. This is true in general, but there are important differential equations for which the evaluation of  $f_h^{[r]}(y)$  is not much more expensive than that of  $f(y)$ , so that the modifying integrators of Algorithm 2.1 can become efficient. A first example is the equations of motion for the full dynamics of a rigid body (see [1, 4] and Section 3 below).

As another example, consider the  $N$ -body problem, which is Hamiltonian  $\dot{q} = p$ ,  $\dot{p} = -\nabla U(q)$  and has potential

$$U(q) = \sum_{1 \leq j < k \leq N} U_{jk}(\|q_j - q_k\|)$$

(the sum is over  $j$  and  $k$ ), where  $U_{jk}(r)$  is a scalar function ( $U_{jk}(r) = -1/r$  for the gravitational potential) and  $\|\cdot\|$  stands for the Euclidean norm. Here,  $q \in \mathbb{R}^{3N}$  is composed by the position vectors  $q_j \in \mathbb{R}^3$ . The vector field requires the computation of

$$\frac{\partial U(q)}{\partial q_j} = \sum_{1 \leq k \neq j \leq N} V_{jk}(\|q_j - q_k\|)(q_j - q_k), \quad V_{jk}(r) = \frac{U'_{jk}(r)}{r} \quad (8)$$

the sum being only over  $k$ . The main observation is now that every summand in this expression depends only on two variables  $q_j$  and  $q_k$ . Therefore, many mixed higher derivatives vanish, and all expressions appearing in the modified equation (7) reduce to a sum over only one summation index. Typically, the computation of the square root in  $\|q_j - q_k\|$  is the most time consuming part, and derivatives of  $U_{jk}(r)$  can often be obtained with negligible cost when computed together with the value  $U_{jk}(r)$ . In this situation, modifying numerical integrators can be implemented efficiently. As noted by McLachlan [6], this feature of  $N$ -body problems can be exploited also in an efficient implementation of implicit Runge–Kutta methods.

### 3. ACCURATE RIGID-BODY INTEGRATOR BASED ON THE DMV ALGORITHM

As illustration of how efficient modifying integrators can be, we consider the equations of motion for a rigid body,

$$\dot{y} = \hat{y} I^{-1} y, \quad \dot{Q} = Q \widehat{I^{-1} y}, \quad \text{where} \quad \hat{a} = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix} \quad (9)$$

for a vector  $a = (a_1, a_2, a_3)^T$ . Here,  $I = \text{diag}(I_1, I_2, I_3)$  is the matrix formed by the moments of inertia,  $y$  is the vector of the angular momenta, and  $Q$  is the orthogonal matrix that describes the rotation relative to a fixed coordinate system. As numerical integrator we choose the Discrete Moser–Veselov algorithm (DMV) [8],

$$\hat{y}_{n+1} = \Omega_n \hat{y}_n \Omega_n^T, \quad Q_{n+1} = Q_n \Omega_n^T, \quad (10)$$

where the orthogonal matrix  $\Omega_n$  is computed from

$$\Omega_n^T D - D \Omega_n = h \hat{y}_n. \quad (11)$$

Here, the diagonal matrix  $D = \text{diag}(d_1, d_2, d_3)$  is determined by  $d_1 + d_2 = I_3$ ,  $d_2 + d_3 = I_1$ , and  $d_3 + d_1 = I_2$ . This algorithm is an excellent geometric integrator and shares many geometric properties with the exact flow. It is symplectic, it exactly preserves the Hamiltonian, the Casimir and the angular momentum  $Qy$  (in the fixed frame), and it keeps the orthogonality of  $Q$ . Its only disadvantage is the low order two.

The technique of modifying integrators cannot be directly applied to increase the order of this method, because the algorithm (10) is not defined for general problems (1). It is, however, defined for arbitrary  $I_j$ , and therefore we look for modified moments of inertia  $\tilde{I}_j$  such that the DMV algorithm applied with  $\tilde{I}_j$  yields the exact solution of (9). It is shown in [4] that this is possible with

$$\frac{1}{\tilde{I}_j} = \frac{1}{I_j} \left( 1 + h^2 s_3(y_n) + h^4 s_5(y_n) + \dots \right) + h^2 d_3(y_n) + h^4 d_5(y_n) + \dots \quad (12)$$

The expressions  $s_k(y)$  and  $d_k(y)$  can be computed by a formula manipulation package similar as the modified differential equation is obtained. The first of them are

$$\begin{aligned} s_3(y_n) &= -\frac{1}{3} \left( \frac{1}{I_1} + \frac{1}{I_2} + \frac{1}{I_3} \right) H(y_n) + \frac{I_1 + I_2 + I_3}{6 I_1 I_2 I_3} C(y_n), \\ d_3(y_n) &= \frac{I_1 + I_2 + I_3}{6 I_1 I_2 I_3} H(y_n) - \frac{1}{3 I_1 I_2 I_3} C(y_n), \end{aligned}$$

where

$$C(y) = \frac{1}{2} (y_1^2 + y_2^2 + y_3^2) \quad \text{and} \quad H(y) = \frac{1}{2} \left( \frac{y_1^2}{I_1} + \frac{y_2^2}{I_2} + \frac{y_3^2}{I_3} \right) \quad (13)$$

are the Casimir and the Hamiltonian of the system. The physical interpretation of this result is the following: after perturbing suitably the form of the body, an application of the DMV algorithm yields the exact motion of the body. Truncating the series in (12) after the  $h^{2r-2}$  terms, yields a modifying DMV algorithm of order  $2r$ .

**Example 3.1.** We consider an asymmetric rigid body with moments of inertia  $I_1 = 0.6$ ,  $I_2 = 0.8$ , and  $I_3 = 1.0$  on the interval  $[0, 10]$ . Initial values are  $y(0) = (1.8, 0.4, -0.9)^T$  and  $Q(0)$  is the identity matrix. The implementation of the modifying DMV algorithm is done using quaternions as explained in [4]. Although  $H(y)$  and  $C(y)$  are constant along the numerical solution, we recompute the values of  $\tilde{I}_j$  in every step to simulate the presence of an external potential.

We apply the DMV algorithm and its extensions to order 4, 6, 8, and 10 with many different step sizes, and we plot in Figure 2 the global error at the endpoint as a function of the cpu times. The execution times are

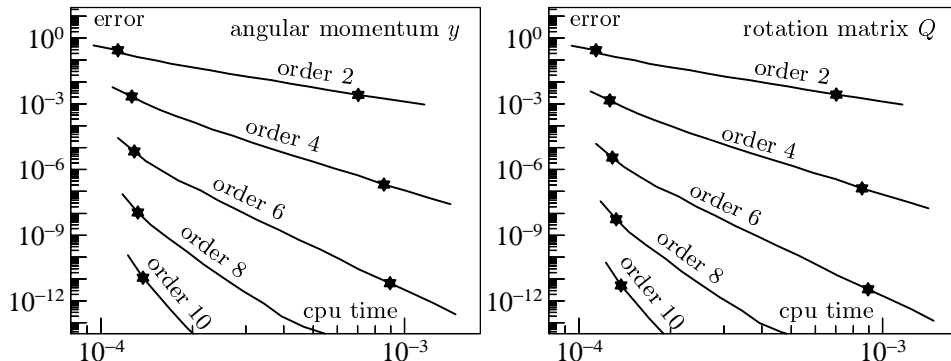


FIGURE 2. Work-precision diagram for the DMV algorithm (order 2) and for the modifying DMV integrators of orders 4, 6, 8, and 10.

the average of 1000 experiments. The symbols indicate the values obtained with the step sizes  $h = 0.1$  and  $h = 0.01$ , respectively.

The pictures nicely illustrate the expected orders of the algorithms (order  $p$  corresponds to a straight line with slope  $-p$ ). Much more interesting is the fact that high accuracy is obtained more or less for free. Consider the results obtained with step size  $h = 0.1$ . The error for the DMV algorithm (order 2) is more than 20%. With very little extra work, the modification of order 10 gives an accuracy of more than 11 digits with the same step size.

## REFERENCES

- [1] P. Chartier, E. Hairer, and G. Vilmart. Numerical integrators based on modified differential equations. *Submitted for publication*, 2006.
- [2] K. Feng. Difference schemes for Hamiltonian formalism and symplectic geometry. *J. Comp. Math.*, 4:279–289, 1986.
- [3] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer Series in Computational Mathematics 31. Springer-Verlag, Berlin, second edition, 2006.
- [4] E. Hairer and G. Vilmart. Preprocessed Discrete Moser–Veselov algorithm for the full dynamics of the free rigid body. *Submitted for publication*, 2006.
- [5] B. Leimkuhler and S. Reich. *Simulating Hamiltonian Dynamics*. Cambridge Monographs on Applied and Computational Mathematics 14. Cambridge University Press, Cambridge, 2004.
- [6] R. I. McLachlan. A new implementation of symplectic Runge–Kutta methods. *Submitted for publication*, 2006.
- [7] R. I. McLachlan and A. Zanna. The discrete Moser–Veselov algorithm for the free rigid body, revisited. *Found. Comput. Math.*, 5:87–123, 2005.
- [8] J. Moser and A. P. Veselov. Discrete versions of some classical integrable systems and factorization of matrix polynomials. *Comm. Math. Phys.*, 139:217–243, 1991.
- [9] R. D. Ruth. A canonical integration technique. *IEEE Trans. Nuclear Science*, NS-30:2669–2671, 1983.