Performance of Taylor Model Methods for Validated Integration of ODEs

Martin Berz and Kyoko Makino

Department of Physics and Astronomy
Michigan State University
East Lansing, MI 48824, USA

Abstract. The performance of various Taylor model (TM)-based methods for the validated integration of ODEs is studied for some representative computational problems. For nonlinear problems, the advantage of the method lies in the ability to retain dependencies of final conditions on initial conditions to high order, leading to the ability to treat large boxes of initial conditions for extended periods of time. For linear problems, the asymptotic behavior of the error of the methods is seen to be similar to that of non-validated integrators.

1 Introduction

Taylor model methods provide functional inclusions of a function f of the form

$$f(x) \in P(x) + I \text{ for } x \in B$$

where B is the domain box, P is the nth order Taylor polynomial of f around a point $x_0 \in B \subset R^m$ expressed by floating point coefficients, and I is an interval remainder bound including errors associated to the floating point representation of P. Taylor models of a given function can be obtained from its code list by use of Taylor model arithmetic. Compared to other validated methods, the approach has the following advantages:

- Reduction of the dependency problem [1], since the bulk of the functional dependency is represented by the polynomial P
- High-order approximation, since the width of I scales with order $(n + 1)$ if n is the order of P [2]
- Simplification of global optimization, since original function is replaced by P, which has little dependency and is easily amenable to domain reduction
- Availability of advanced tools for solutions of implicit equations

The methods also allow the development of validated integration schemes [3] [4] that represent the dependence of the solution at the time t_k after the kth step in terms of the initial conditions x_i as a Taylor model

$$x(x_i, t_k) \in P_k(x_i) + I_k.$$

Frequently the dynamics is also represented in a pre-conditioned form by factoring a co-moving coordinate frame as

$$x(x_i, t_k) \in (C_k + J_k) \circ (P_{rk}(x_i) + I_{rk})$$

where C_k represents a change of variables. Frequently used choices in the study of dynamics are linear transformations based on curvilinear coordinates (PC-CV) [5] [4] [6] [7] [8] and blunted parallelepiped coordinates (PC-BL), which prevents the occurring parallelepipeds from becoming ill-conditioned [9]. For purposes of comparison, we also utilize the QR coordinate system used for the description of the error term by Lohner in his code AWA [10] [11] [12], which determines a linear relationship between final conditions and initial conditions.

Finally it is also possible to give up the direct connection between final conditions and initial conditions and merely ask that

$$x(x_i, t_k) \in \bigcup_{x_i \in B} P_k^* (x_i, t_k)$$

where P_k^* is a polynomial obtained in the so-called shrink wrap approach [13] [9], of which we use the blunted version (SW-BL). All these features are implemented in the code COSY-VI.

2   Nonlinear Problems

As an example for the behavior for a nonlinear problem, we utilize a classical example from the literature [14] [15] of validated integration of ODEs, the Volterra equations

$$\frac{dx_1}{dt} = 2x_1(1 - x_2), \quad \frac{dx_2}{dt} = -x_2(1 - x_1). \tag{2.1}$$

The ODEs admit an invariant which has the form

$$C(x_1, x_2) = x_1 x_2^2 e^{-x_1 - 2x_2} = \text{Constant}, \tag{2.2}$$

which is useful for the practical study of the methods under consideration. In the quadrant characterized by $x_{1,2} > 0$, the constant is positive, which implies that contour lines are restricted to this quadrant and even form closed curves. Figure 1 illustrates the shape of C and a few of its contour lines. The period of one cycle of the solution depends on the initial condition, and outer orbits take longer.

We study the ODEs for the initial conditions

$$x_{01} \in 1 + [-0.05, 0.05], \quad x_{02} \in 3 + [-0.05, 0.05].$$

In the contour line plot of the invariant in figure 1, the center of these initial conditions lies on the outermost of the three shown contour lines. Within the Taylor model framework of COSY-VI, the initial conditions are represented by the initial Taylor models
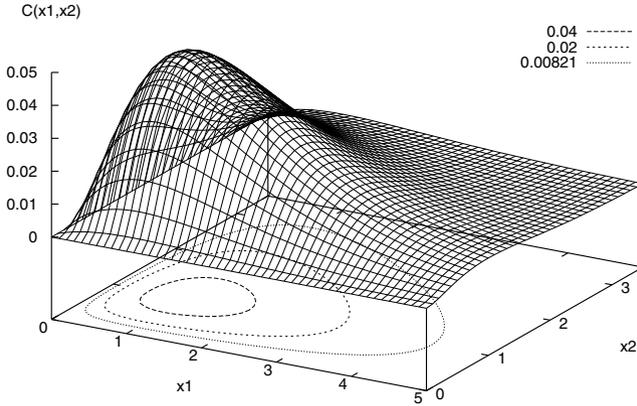
Fig. 1. The invariant of the Volterra equations, which has the form $C(x_1, x_2) = x_1 x_2^2 e^{-x_1 - 2x_2}$. A few contour lines are shown, including the one passing through the initial condition $(x_1, x_2) = (1, 3)$ being studied

```
I   COEFFICIENT                ORDER EXPONENTS
1    1.000000000000000           0     0 0  0
2   0.5000000000000000E-01       1     1 0  0
R   [-.5884182030513415E-015,0.5884182030513415E-015]


1    3.000000000000000           0     0 0  0
2   0.5000000000000000E-01       1     0 1  0
R   [-.1476596622751470E-014,0.1476596622751470E-014]
```

We use the ODEs to compare the behavior of the code COSY-VI and the code AWA by Lohner [12] and to study the longer term behavior of the validated integration process. As a first step, in figure 2 we show the overall width of the validated enclosure of the solution for one full revolution of the validated enclosure produced by AWA and COSY-VI. It is seen that shortly before completion of the revolution, the enclosures produced by AWA grow rapidly.

We now study the Taylor model representation of the solution obtained by COSY-VI after one revolution; the x_1 component has the form

```
I   COEFFICIENT                ORDER EXPONENTS
1   0.9999999999999984          0     0 0  0
2   0.4999999999999905E-01      1     1 0  0
3   0.1593548596541794          1     0 1  0
4   0.2987903618516347E-02      2     2 0  0
5   0.7967742982712876E-02      2     1 1  0
6   0.1745863785260356E-01      2     0 2  0
7   0.4979839364191599E-04      3     3 0  0
8   0.5551021321878651E-03      3     2 1  0
9   0.6348634117324201E-03      3     1 2  0
```
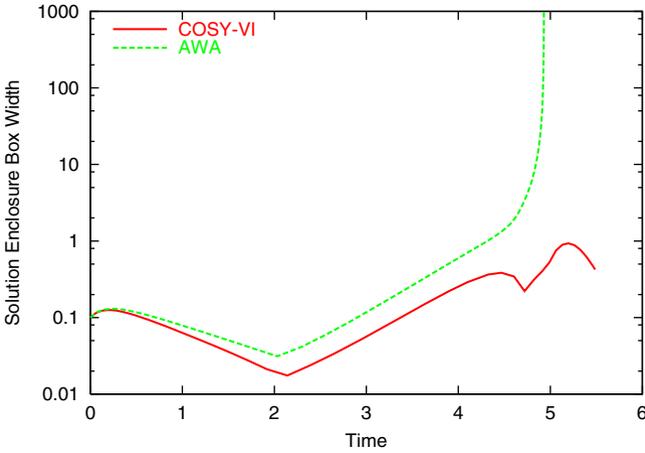
Fig. 2. The width of validated enclosure for solution of the Volterra equation determined by AWA and COSY-VI

```
 10   0.1191291278992926E-02    3     0 3   0
 11   0.3258832737620100E-05    4     4 0   0
 12   0.3241341695678232E-06    4     3 1   0
 13   0.3862783715688610E-04    4     2 2   0
 14   0.2689662978922477E-05    4     1 3   0
 15   0.3564904362283420E-04    4     0 4   0
            .....
171   0.1136167325983013E-18   18     117   0
  R   [-.4707095747144810E-010,0.4699004714805186E-010]
```

It can be seen that the zeroth order term is reproduced nearly exactly, as necessary after one revolution. Also the dependence on the first variable is nearly as in the original Taylor model. However, there is now an additional dependence on the second variable, which is significantly larger than the dependence on the first variable, and which induces a significant shearing of the solution. There are also higher order dependencies on initial variables; up to machine precision, terms of up to order 18 contribute, and some of the second order contributions indeed have a magnitude similar to the first order contribution, an indication of strongly nonlinear behavior.

The dependence on the second variable and the higher order contributions are the reason why the box enclosure produced by COSY-VI shown in figure 2 is larger at the end of the integration than it was in the beginning. To determine how much of this is actual overestimation, we insert the Taylor models representing the flow at time t into the invariant in eq. (2.2) of the ODE and subtract from it the value of the invariant at time 0. To the extent the Taylor models represent the true solution, the coefficients of the resulting polynomial should vanish. The bound of the resulting Taylor model is a measure for the sharpness of the approximation.
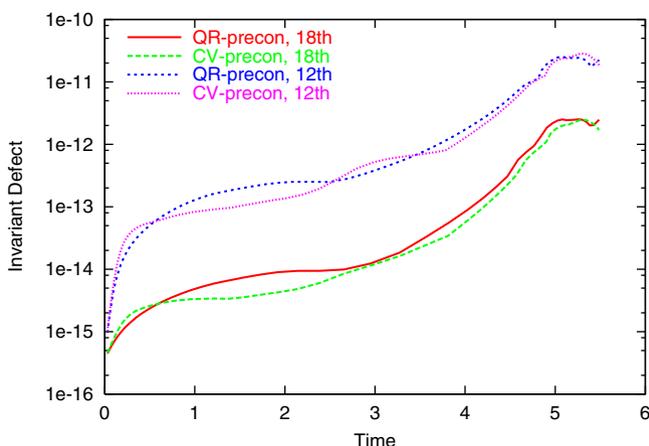
Fig. 3. Invariant defect of the Taylor model integration of the Volterra equation for one revolution for oders 12 and 18 and curvilinear (CV) and QR preconditioning

As a specific example, we show the resulting Taylor model of the invariant defect after one full revolution.

```
    I   COEFFICIENT              ORDER  EXPONENTS
    1   0.1214306433183765E-16     0     0 0  0
    2   -.1100465205072787E-16     1     1 0  0
    3   -.4109126233720062E-16     1     0 1  0
    4   0.3169597288625592E-17     2     2 0  0
    5   0.2035589578841535E-16     2     1 1  0
    6   0.2318159770045569E-16     2     0 2  0
    7   -.3702063375093326E-18     3     3 0  0
    8   -.2109853192492055E-17     3     2 1  0
    9   0.7358175212798107E-17     3     1 2  0
   10   0.2956745849914467E-16     3     0 3  0
             .....
   76   0.2469510337886316E-19    15     6 9  0
    R   [-.1689839352652954E-011,0.1691154903051651E-011]
```

Indeed all remaining coefficients are very small, and the remaining terms are just of the magnitude of machine precision. The main contribution is in fact the remainder term of the Taylor model evaluation of the invariant of magnitude around 10^{-12}, which is similar to that of the Taylor model solution after one revolution. Overall, this study shows that the original domain box of width 10^{-1} could be transported for one revolution with an actual overestimation of only around 10^{-12}.

Figure 3 shows in detail the size of the invariant defect as a function of integration time for up to one revolution. Shown are computation orders 18 and 12 and curvilinear (CV) as well as QR preconditioning. It can be seen that the method of order 18 produces an overestimation of around 10^{-12} after one revolution; after a fast ramp-up away from

the floating point error floor, a plateau is reached, until the error again increases because the system of ODEs enters a region of strong nonlinearity. On the other hand, figure 2 shows that AWA already early on exhibits relative overestimation of about 2 and then fails before $t = 5$.

In order to assess the long-term behavior of the integration, it is important to first consider some of the specific properties of the Volterra equations. As can be seen from the one-revolution Taylor model enclosure of the solution, one of the important features of the ODEs is that the revolution period strongly depends on the original position on the invariant curves of the Volterra equations. This entails that when integrating the flow of an initial condition box of significant size, some of its corners will lag more and more behind some of the others, and the box will become more and more elongated. Analyzing the one-revolution Taylor model, one sees that within only about five revolutions, the image of the original box has a size similar to the entire parameter space reached during the revolution; thus simplistic long-term integration of this ODE is not possible without further modifications.

One way to treat this problem is to assess the dynamics in terms of a Poincare section, a method frequently employed in the study of long-term behavior (see for example [5]). Here however, we will restrict our attention to a more immediate tool for assessing the long-term behavior, namely repeated forward-backward integration. This approach maintains the nonlinear effects of the problem while away from initial conditions, but avoids the "lag" problem because after one forward-backward cycle, all initial conditions return to their original values.

In the following we assess the long-term forward-backward integration of the Volterra equation using the shrink wrap approach utilized in COSY-VI [13] [9]. Roughly speaking, in this approach the interval remainder bounds are "absorbed" back into the range of the Taylor polynomial of the flow by slightly enlarging the corresponding coefficients. Thus the remaining interval dependencies and the associated risk of eventual blowup disappear. If the intervals to be absorbed into the range are sufficiently small in each step, the increase in the size of the coefficients will also be small. The quality of the invariant over suitable integration times suggests that this is indeed the case.

In figure 4 we show the results of some longer term integration of 100 forward-backwards cycles. The pictures show the width of the solution enclosure box as a function of cycles. The left picture shows the situation for five full cycles; while the box width varies greatly, namely by nearly two orders of magnitude, over one forward-backward cycle it returns to nearly the previous status. The repeated pattern of five very similar looking box widths is visible; furthermore, within each of the five cycles, the widths are mirror symmetric around the middle, which corresponds with the turn-around point.

The right picture in figure 4 shows the situation from cycle 95 to cycle 100. The remarkable fact is that while the curves have significant fine structure, there is no difference discernible to the naked eye; hence the transport after nearly 100 cycles looks almost the same as in the beginning, although the box itself was relatively large and got enhanced to a width of nearly one in each of the forward and backward passes.
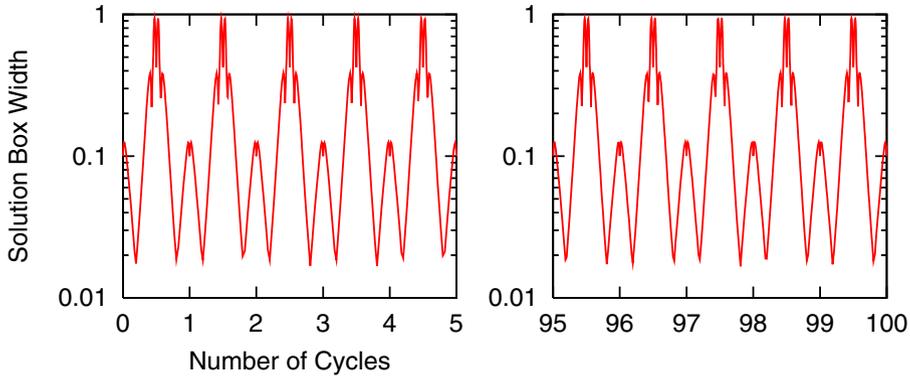
Fig. 4. Evolution of box widths during forward-backward cycles of the Volterra equation. Shown are the situation for the first five cycles as well as cycles 95 to 100

3 Autonomous Linear Problems

While of rather limited practical significance, linear autonomous problems are of theoretical interest in the study of integration schemes because much is known about the asymptotic behavior of error growth of common methods. Thus it is interesting to study the behavior of validated integration schemes for such cases. Of particular interest are cases that have eigenvalues of varying magnitudes, since for several validated ODE integration schemes asymptotically this leads to difficulties with ill-conditioned linear algebra operations.

To asses the behavior of the methods, we study three test cases for linear ODEs $x' = Bx$ originally proposed by Markus Neher. We compare the performance of COSY-VI with that of the code AWA. This comparison is interesting because AWA can easily handle these cases since different from the situation in the previous section, the dependence on initial conditions is always linear, and thus COSY-VI's advantage to handle higher order dependence on initial condition is irrelevant. But they are challenging for COSY-VI because only first order terms in initial conditions appear, resulting in extreme sparsity in the Taylor model data structure on computers.

We show results of various computation modes with COSY-VI, namely QR preconditioning (PC-QR), curvilinear preconditioning (PC-CV), blunted parallelepiped preconditioning (PC-BL), and blunted shrink wrapping (SW-BL). Both codes use automatic step size control. COSY-VI uses order 17. AWA uses order 20, and the modes 1 through 4; frequently the mode 4, the "intersection of QR decomposition and interval-vector" performs best. All runs were performed in the arithmetic environment of a 450MHz Pentium III processor. Integration was performed until $t = 1000$ or until failure for the initial box

$$(1, 1, 1) + 10^{-6} \cdot [-1, 1]^3.$$

As a first example, we study a pure contraction with three distinct eigenvalues and obtained the following result of performance.

$$B_1 = \begin{pmatrix} -0.6875 & -0.1875 & 0.08838834762 \\ -0.1875 & -0.6875 & 0.08838834762 \\ 0.08838834762 & 0.08838834762 & -0.875 \end{pmatrix} \approx \begin{pmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & -\frac{3}{4} & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Mode	t max	Steps	Width
AWA	1000	1216	1.4e-35
VI PC-QR	1000	1633	3.1e-38
VI PC-CV	1000	1463	4.1e-38
VI PC-BL	1000	1620	3.8e-36
VI SW-BL	1000	1726	6.3e-36

We note that the sweeping variable in COSY's Taylor model arithmetic [2], which controls the size of terms retained as significant, has been set to 10^{-40} for this case. Running with a larger value for the sweeping variable leads to a solution set with a size of roughly that larger size.

As the next example, we study a pure rotation

$$B_2 = \begin{pmatrix} 0 & -0.7071067810 & -0.5 \\ 0.7071067810 & 0 & 0.5 \\ 0.5 & -0.5 & 0 \end{pmatrix} \approx \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Mode	t max	Steps	Width
AWA	1000	2549	3.5e-6
VI PC-QR	1000	2021	3.5e-6
VI PC-CV	1000	2046	3.5e-6
VI PC-BL	1000	2021	3.5e-6
VI SW-BL	1000	2030	3.5e-6

Finally we study a combination of a contraction with a rotation

$$B_3 = \begin{pmatrix} -0.125 & -0.8321067810 & -0.3232233048 \\ 0.5821067810 & -0.125 & 0.6767766952 \\ 0.6767766952 & -0.3232233048 & -0.25 \end{pmatrix} \approx \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} \end{pmatrix}$$

Mode	t max	Steps	Width
AWA	1000	3501	3.0e-6
VI PC-QR	1000	2772	3.0e-6
VI PC-CV	1000	2769	3.0e-6
VI PC-BL	1000	2746	4.7e-6
VI SW-BL	1000	2728	1.2e-5

4 Conclusion

Summarizing the results of the numerical experiments for various nonlinear and linear problems shows the following results:

- As expected, the ability to treat higher order dependences on initial conditions leads to a significant performance advantage for COSY-VI for nonlinear problems and larger initial condition domain boxes
- For extended computation times in the Volterra equations, curvilinear preconditioning of Taylor model integration behaves similar to QR preconditioning, and both of them behave significantly better than the AWA approach
- Shrink wrapping allows extended integration periods; over 100 forward-backward cycles through the Volterra equation, growth of box width is not discernible within printer resolution even for rather large boxes where AWA cannot complete a single forward integration
- For linear autonomous problems, the COSY-VI preconditioning methods based on QR, curvilinear, and blunted parallelepiped, all show qualitatively the same behavior as the QR mode of AWA. The latter is known to have error growth similar to the non-validated integration for autonomous linear ODEs. Thus we observe that the three modes of COSY-VI achieve the same type of error growth.
- The number of integration steps of all methods are rather similar.

References

1. K. Makino and M. Berz. Efficient control of the dependency problem based on Taylor model methods. *Reliable Computing*, 5(1):3–12, 1999.
2. K. Makino and M. Berz. Taylor models and other validated functional inclusion methods. *International Journal of Pure and Applied Mathematics*, 6,3:239–316, 2003. available at http://bt.pa.msu.edu/pub.
3. M. Berz and K. Makino. Verified integration of ODEs and flows using differential algebraic methods on high-order Taylor models. *Reliable Computing*, 4(4):361–369, 1998.
4. K. Makino. *Rigorous Analysis of Nonlinear Motion in Particle Accelerators*. PhD thesis, Michigan State University, East Lansing, Michigan, USA, 1998. Also MSUCL-1093.
5. M. Berz. *Modern Map Methods in Particle Beam Physics*. Academic Press, San Diego, 1999. Also available at http://bt.pa.msu.edu/pub.
6. K. Makino and M. Berz. Perturbative equations of motion and differential operators in non-planar curvilinear coordinates. *International Journal of Applied Mathematics*, 3,4:421–440, 2000.
7. M. Berz and K. Makino. Preservation of canonical structure in nonplanar curvilinear coordinates. *International Journal of Applied Mathematics*, 3,4:401–419, 2000.
8. M. Berz. Arbitrary order description of arbitrary particle optical systems. *Nuclear Instruments and Methods*, A298:426, 1990.
9. K. Makino and M. Berz. Suppression of the wrapping effect by Taylor model based validated integrators. submitted. Also MSUHEP-40910, available at http://bt.pa.msu.edu/pub.
10. R. J. Lohner. Enclosing the solutions of ordinary initial and boundary value problems. In E. Kaucher, U. Kulisch, and C. Ullrich, editors, *Computer Arithmetic: Scientific Computation and Programming Languages*, pages 255–286. Teubner, Stuttgart, 1987.
11. R. J. Lohner. *Einschliessung der Losung gewohnlicher Anfangs- und Randwertaufgaben und Anwendungen*. Dissertation, Fakultat fur Mathematik, Universitat Karlsruhe, 1988.
12. R. J. Lohner. AWA - Software for the computation of guaranteed bounds for solutions of ordinary initial value problems.

13. K. Makino and M. Berz. The method of shrink wrapping for the validated solution of ODEs. Technical Report MSUHEP-20510, Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, 2002.

14. W. F. Ames and E. Adams. Monotonically convergent numerical two-sided bounds for a differential birth and death process. In K. Nickel, editor, *Interval Mathematics*, volume 29 of *Lecture Notes in Computer Science*, pages 135–140, Berlin; New York, 1975. Springer-Verlag.

15. R. E. Moore. *Methods and Applications of Interval Analysis*. SIAM, 1979.