From Interval Analysis to Taylor Models - An Overview

Markus Neher
Universität Karlsruhe, Institut für Angewandte Mathematik
76128 Karlsruhe, Germany
markus.neher@math.uni-karlsruhe.de

Abstract

Interval arithmetic has been widely used in enclosure methods for almost 40 years. Today, it is a well established tool for the calculation of rigorous error bounds for many problems in numerical analysis.

Despite its overall success, interval arithmetic suffers from two drawbacks: the dependency problem and the so-called wrapping effect, which both may overestimate the true error of some computation.

To reduce overestimation, Taylor models have been developed as a symbiosis of a computer algebra method and interval arithmetic by M. Berz and his group since the 1990s. Software implementations of Taylor models have been applied to a variety of problems, such as global optimization problems, validated multidimensional integration, or the solution of ODEs and DAEs.

The validated solution of ODEs is used for exemplifying the distinction of interval methods and Taylor model methods.

1 Interval Computations

Interval arithmetic attracted wide attention with the publication of the pioneering work of R. E. Moore [24] in 1966. In his book, not only the foundation of interval arithmetic has been laid; it also contains interval methods for various fundamental problems, such as a discussion of inclusion functions for the computation of range bounds for real functions, an interval version of Newton's method, a chapter on automatic differentiation, and algorithms for the validated solution of ODEs. Extensive introductions to interval computations have also been given in the monographs [1, 11, 28].

Real interval arithmetic is based on closed bounded intervals. Arithmetic operations between two intervals $X = [\underline{x}, \overline{x}]$ and $Y = [\underline{y}, \overline{y}]$ are defined in such a way, that the result contains all values that are obtainable by relating the real numbers in X and Y. For example, the sum of two intervals is defined as

$$[\underline{x}, \overline{x}] + [\underline{y}, \overline{y}] := [\underline{x} + \underline{y}, \overline{x} + \overline{y}].$$

Interval operations between interval vectors and interval matrices are constructed by applying scalar interval arithmetic componentwise. Interval functions are defined in the same manner. Interval evaluation of a rational expression $r(x)$ for some interval X means inserting X for x and evaluating according to the rules of interval arithmetic. Other real functions are extended via range enclosures. For example, the interval exponential function is given by

$$e^X := [e^{\underline{x}}, e^{\overline{x}}].$$

Interval enclosures of solutions to continuous problems, such as integral equations, ODEs or PDEs, are defined in a similar way.

Interest in interval arithmetic has been primarily aroused by the limitations of the floating point number format used on digital computers. In floating point arithmetic, real numbers are approximated by a finite screen of machine representable numbers. Whenever a real number x that is not exactly representable is stored on a computer, or when intermediate results of a calculation are approximated by machine numbers, roundoff errors occur. It was early observed (e.g., see [9]) how these errors may accumulate until the computed result is no longer meaningful.

Interval arithmetic provides a tool for controlling roundoff errors automatically. Instead of approximating x by a machine number, the real value x is enclosed into a real interval $X = [\underline{x}, \overline{x}]$ with machine representable bounds $\underline{x} \le \overline{x}$. In all subsequent calculations, x is replaced by X. The operations between real numbers are then replaced by the respective interval operations. For restoring mathematical rigor in a floating point number system, the rules of interval arithmetic are extended by directed rounding operations, such that the exact result of any arithmetic operation is automatically enclosed in a floating-point interval, including roundoff errors. For example, sum and difference of two floating point intervals

$$
\begin{aligned}
X + Y &:= [\underline{x} \mathbin{\overset{+}{\bigtriangledown}} \underline{y}, \overline{x} \mathbin{\overset{+}{\bigtriangleup}} \overline{y}] \supseteq \{x + y \mid x \in X,\ y \in Y\}, \\
X - Y &:= [\underline{x} \mathbin{\overset{-}{\bigtriangledown}} \overline{y}, \overline{x} \mathbin{\overset{-}{\bigtriangleup}} \underline{y}] \supseteq \{x - y \mid x \in X,\ y \in Y\}
\end{aligned}
$$

are each calculated with two basic arithmetic operations between the endpoints of the intervals X and Y, where $\bigtriangledown$, $\bigtriangleup$ denote operations with downward or upward rounding, respectively. Definitions of validated computer arithmetic have been given in [15, 16].

Interval analysis is also used for bounding all kinds of truncation errors: the truncation error of an infinite iteration, as in the interval version of Newton's method [24, Chap. 7], the remainder term of a convergent series [27], discretization errors in the numerical solution of differential equations [24, Chap. 10], etc.

2 Dependency Problem and Wrapping Effect

Interval arithmetic is sometimes affected by overestimation, such that computed error bounds are over-pessimistic. Overestimation is often caused by the *dependency problem*, which is the lack of interval arithmetic to identify different occurrences of the same variable. For example, $x - x = 0$ holds for each $x \in [1, 2]$, but $X - X$ for $X = [1, 2]$ yields $[-1, 1]$. A second source of overestimation is the *wrapping effect*, which appears when intermediate results of a computation are enclosed into intervals. The wrapping effect was first observed by Moore in 1965 [23]; a recent analysis has been given by Lohner [18].

Example 1. Wrapping effect.
Consider the function

$$f : (x, y) \to \frac{\sqrt{2}}{2}(x + y, y - x).$$

The image of the square $[0, \sqrt{2}]^2$ is the rotated square with corners $(0, 0)$, $(1, -1)$, $(2, 0)$, $(1, 1)$. On the other hand, interval computation yields

$$f([0, \sqrt{2}], [0, \sqrt{2}]) = ([0, 2], [-1, 1]).$$

Note that the observed overestimation (the area of the interval enclosure is twice the area of the range) is not a result of dependency, but of the enclosure of a non-interval shaped range into an interval. Overestimations of this kind are one of the major problems in the interval arithmetic treatment of ODEs.

3 Taylor Models

For reducing both the dependency problem and the wrapping effect, interval arithmetic has been endorsed with symbolic extensions. Symbolic-numeric computations have been proposed under various names since the 1980s [7, 13, 20]. Early implementations in software were also given [7, 12], but to the author's knowledge, these packages have not been widely distributed and are not available today.

A rigorous multivariate Taylor arithmetic has been developed by Berz and his group since the 1990s [2, 20]. A *Taylor model* of a function f on some interval X consists of the Taylor polynomial p_n of order n of f and an interval remainder term I_n, which encloses the approximation error $|f - p_n|$ on X. In computations that involve f, the function is replaced by $p_n + I_n$. The polynomial part is propagated by symbolic calculations where possible. The interval remainder term is processed according to the rules of interval arithmetic. All truncation and roundoff errors in intermediate operations are also enclosed into the remainder interval of the final result.

Example 2. Multiplication of two univariate Taylor models of order 2.

Let $J := [-1, 1]$ and

$$T_1(x) := 1 + \tfrac{1}{2}x + \tfrac{1}{4}x^2 + [0, 0.2], \quad T_2(x) := 1 - \tfrac{1}{5}x + [-0.1, 0.1], \quad \text{where} \quad x \in J.$$

For all $x \in J$, it holds that

$$\begin{aligned}
T_1(x)\,T_2(x) &\in (1 + \tfrac{1}{2}x + \tfrac{1}{4}x^2)(1 - \tfrac{1}{5}x) + (1 + \tfrac{1}{2}x + \tfrac{1}{4}x^2)[-0.1, 0.1] \\
&\quad + (1 - \tfrac{1}{5}x)[0, 0.2] + [0, 0.2] \cdot [-0.1, 0.1] \\
&\subset (1 + \tfrac{3}{10}x + \tfrac{3}{20}x^2) - \tfrac{1}{20}x^3 + (1 + \tfrac{1}{2}J + \tfrac{1}{4}J^2)[-0.1, 0.1] \\
&\quad + (1 - \tfrac{1}{5}J)[0, 0.2] + [-0.02, 0.02] \\
&\subset 1 + \tfrac{3}{10}x + \tfrac{3}{20}x^2 + [-0.05, 0.05] + [-0.175, 0.175] + [0, 0.24] + [-0.02, 0.02] \\
&= 1 + \tfrac{3}{10}x + \tfrac{3}{20}x^2 + [-0.245, 0.485],
\end{aligned}$$

so that we may define

$$T_1(x) \cdot T_2(x) := 1 + \tfrac{3}{10}x + \tfrac{3}{20}x^2 + [-0.245, 0.485].$$

In Example 2, straightforward interval evaluation for computing the remainder interval of the product has been used for simplicity. In general, it does not yield optimal bounds, and it should be replaced by some more accurate estimation scheme in a practical implementation of Taylor model arithmetic. A software implementation of Taylor model arithmetic has been given by Berz and his co-workers [3, 21] in the COSY Infinity package. The software is available free of charge to non-commercial users [4]. Using COSY Infinity, Taylor models have been applied with great success to a large variety of problems, such as global optimization [25], validated multidimensional integration [6], or the validated solution of ODEs and DAEs [5, 10].

4 Interval Methods and Taylor Model Methods for ODEs

The distinction between interval methods and Taylor model methods is now illustrated for the validated solution of initial value problems for ODEs. We consider the smooth interval IVP

$$u' = f(t,u), \quad u(t_0) \ \in \ [u_0], \quad t \in T = [t_0, t_{\text{end}}] \tag{1}$$

where $f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ is a sufficiently smooth function, $u_0 \in \mathbb{IR}^n$ is a given interval in the space variables and $t_{\text{end}} > t_0$ is a given endpoint of the time interval. For each $u_0 \in [u_0]$, the point IVP

$$u' = f(t,u), \quad u(t_0) \ = \ u_0 \tag{2}$$

has a classical solution, which is denoted by $u(t; t_0, u_0)$. In the following, we assume that $u(t; t_0, u_0)$ is bounded for all $t \in T$ and for all $u_0 \in [u_0]$.

Our goal when solving (1) is to calculate bounds on the flow of the interval IVP. For each $t \in T$, we wish to calculate an interval $[u(t)]$ such that

$$u(t; t_0, u_0) \in [u(t)]$$

holds for all $u_0 \in [u_0]$. The tube $[u(t)]$, $t \in T$, then contains all solutions of $u' = f(t,u)$ that emerge from $[u_0]$.

All enclosure methods for ODEs that we are aware of subdivide the domain of integration into subintervals. On each grid point, the flow of the given ODE is enclosed into a set with a certain geometric structure, for example an n-dimensional rectangle. In the general case, the shape of the flow has a different geometric design, so that the flow is wrapped into some larger set, which serves as initial set for the next time step. To maintain the validity of the method, all solutions of the ODE emerging from the increased initial set must be enclosed in subsequent time steps. Thus, the method picks up additional solutions of the ODE (that is, solutions not emerging from the original initial set) during the integration process. If the accumulated flow becomes too large, the method may break down because it can no longer compute a sufficiently tight enclosure. A thorough discussion of the wrapping effect in interval methods for ODES is found in [26].

Various methods have been proposed to fight the wrapping effect, and there are several techniques which are effective in reducing overestimation of the flow for some problem classes [8, 14, 17, 23, 24]. Nevertheless, the ability of interval methods in minimizing wrapping is limited by the fact that interval based enclosure sets are convex. If the flow is a nonconvex set, any interval wrap must be at least as large as the convex hull of the flow.

In contrast, Taylor model methods use multivariate polynomials in the initial values plus a small interval remainder term for representing the flow. Thus it is possible to work with nonlinear boundary curves, including non-convex enclosure sets for crescent-shaped or twisted flows. For nonlinear ODEs, this increased flexibility in admissible boundary curves is an intrinsic advantage over traditional interval methods.

We refer to the recent paper of Makino and Berz [22] for the description of Taylor model methods for ODES. Our intention here is to explain the fundamental difference between interval methods and Taylor model methods with a simple, but illuminating nonlinear example.

Example 3: Taylor model method for nonlinear model IVP.
We consider the IVP

$$\begin{aligned} u' &= v, \quad u(0) = 1 + a, \\ v' &= u^2, \quad v(0) = -1 + b, \end{aligned} \tag{3}$$

where differentiation is with respect to t. The symbolic parameters a and b (modeling uncertainties in initial conditions) are assumed to vary in the interval $[-0.05, 0.05]$. In an interval method, one would use interval initial values $u_0 = [0.95, 1.05]$, $v_0 = [-1.05, -0.95]$ instead. The initial flow of the IVP (3) at $t = t_0$ is thus given by

$$u_0(a, b) := 1 + a,$$
$$v_0(a, b) := -1 + b,$$

where a, $b \in [-0.05, 0.05]$.

In the following Taylor model integration of (3), order $p = 3$ and step size $h = 0.1$ are used. In each integration step, the multivariate Taylor series (with respect to a, b and t) of the solution of (3) is employed. The third order Taylor polynomial serves as an approximate solution. The truncation error of the series is enclosed into a suitable remainder interval (for the computation of the remainder interval, we refer to [19]):

Taylor model of order 3 at $t_0 = 0$ in 3 variables:

$$\tilde{u}_1(t, a, b) := 1 + a - t + bt + t^2/2 + at^2 - t^3/3 + [-5.1\text{E}{-5}, 7.9\text{E}{-5}],$$
$$\tilde{v}_1(t, a, b) := -1 + b + t + 2at - t^2 + a^2t - at^2 + bt^2 + 2t^3/3 + [-1.8\text{E}{-4}, 1.7\text{E}{-4}]$$

where $t \in [0, 0.1]$ and a, $b \in [-0.05, 0.05]$. Evaluating at $t = h$, we obtain the

Flow at $t_1 = 0.1$ (Taylor model of order 2 in space variables):

$$u_1(a, b) := \tilde{u}_1(0.1, a, b) = 0.905 + 1.01a + 0.1b + [-5.1\text{E}{-5}, 7.9\text{E}{-5}],$$
$$v_1(a, b) := \tilde{v}_1(0.1, a, b) = -0.909 + 0.19a + 1.01b + 0.1a^2 + [-1.8\text{E}{-4}, 1.7\text{E}{-4}],$$

which serves as initial set for the second integration step. Proceeding as before, we obtain the

Taylor model of order 3 at $t_1 = 0.1$ in 3 variables:

$$\begin{aligned} \tilde{u}_2(t, a, b) := {} & 0.905 + 1.01a + 0.1b - 0.909t + 0.19at + 1.01bt + 0.409t^2 \\ & +0.1a^2t + 0.914at^2 + 0.0905bt^2 - 0.274t^3 + [-1.2\text{E}{-4}, 1.7\text{E}{-4}], \\ \tilde{v}_2(t, a, b) := {} & -0.909 + 0.19a + 1.01b + 0.818t + 0.1a^2 + 1.83at + 0.181bt - 0.823t^2 \\ & +1.02a^2t + 0.202abt + 0.01b^2t - 0.747at^2 + 0.823bt^2 + 0.522t^3 \\ & +[-3.4\text{E}{-4}, 3.3\text{E}{-4}] \end{aligned}$$

(where again $t \in [0, 0.1]$ and a, $b \in [-0.05, 0.05]$) and the

Flow at $t_2 = 0.2$ (Taylor model of order 2 in space variables):

$$\begin{aligned} u_2(a, b) &:= \tilde{u}_2(0.1, a, b) = 0.818 + 1.04a + 0.202b + 0.01a^2 + [-1.2\text{E}{-4}, 1.7\text{E}{-4}], \\ v_2(a, b) &:= \tilde{v}_2(0.1, a, b) = -0.835 + 0.365a + 1.04b + 0.202a^2 + 0.0202ab + 0.001b^2 \\ & \qquad +[-3.4\text{E}{-4}, 3.3\text{E}{-4}]. \end{aligned}$$

The flow of the IVP at t_j, $j = 0, 1, 2$, is contained in the set

$$\left\{ \left(u_j(a,b) + I_j, u_j(a,b) + J_j \right) \mid a, b \in [-0.05, 0.05] \right\},$$

where I_j and J_j denote the respective remainder intervals. Such sets need not be convex and thus are more flexible in enclosing the flow than enclosures derived from intervals (cf. Figure 1).

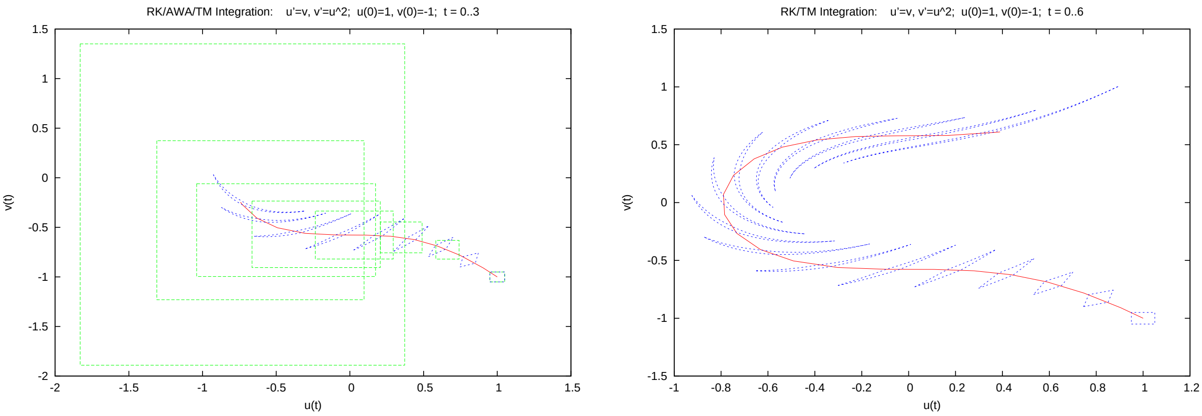

Figure 1: Quadratic model IVP.

Finally, we use the quadratic model IVP (3) to compare the performance of Lohner's software AWA [17] with the COSY Infinity integrator written by Makino. For the computation, Taylor expansions of order 18 were used in both programs. The computed enclosure sets are shown in Figure 1. In the left picture, integration is performed in the time interval $[0, 3]$. In the beginning, the enclosures from AWA (rectangular boxes) and COSY Infinity (nonlinear sets) are of similar quality. Near the end of the integration domain, the enclosures from AWA start exploding. While AWA aborts integration at $t = 3.75$, COSY Infinity is able to continue the integration much longer (right picture; enclosures of AWA are not shown). We attribute this to the ability of Taylor model methods to use nonconvex enclosure sets of the flow.

Conclusion

We have compared traditional interval arithmetic with Taylor model arithmetic. For the validated solution of initial value problems for ODEs, we have shown how Taylor model methods benefit from symbolic computations. Increased flexibility in admissible boundary curves of enclosures is an intrinsic advantage over traditional interval methods, not only for the solution of ODEs. In future research, we hope to contribute to the further development and spreading of Taylor model methods.

References

[1] G. Alefeld and J. Herzberger. *Introduction to Interval Computations.* Academic Press, New York, 1983.

[2] M. Berz. From Taylor series to Taylor models. In *AIP Conference Proceedings 405*, pages 1–23, 1997.

[3] M. Berz. Cosy Infinity Version 8 reference manual. NSCL Technical Report MSUCL-1088, Michigan State University, 1998.

[4] M. Berz. COSY INFINITY. `http://bt.pa.msu.edu/index_files/cosy.htm`, March 2005.

[5] M. Berz and K. Makino. Verified integration of ODEs and flows using differential algebraic methods on high-order Taylor models. *Reliable Computing*, 4:361–369, 1998.

[6] M. Berz and K. Makino. New methods for high-dimensional verified quadrature. *Reliable Computing*, 5:13–22, 1999.

[7] J.-P. Eckmann, H. Koch, and P. Wittwer. A computer-assisted proof of universality in area-preserving maps. *Memoirs of the AMS*, 47(289), 1984.

[8] P. Eijgenraam. *The Solution of Initial Value Problems Using Interval Arithmetic.* Mathematical Centre Tracts No. 144, Stichting Mathematisch Centrum, 1981.

[9] G. E. Forsythe. Pitfalls in computation, or why a math book is not enough. *Amer. Math. Monthly*, 77:931–956, 1970.

[10] J. Hoefkens, M. Berz, and K. Makino. Verified high-order integration of DAEs and higher-order ODEs. In W. Kraemer and J. Wolff von Gudenberg, editors, *Scientific Computing, Validated Numerics and Interval Methods*, pages 281–292. Kluwer, Dordrecht, Netherlands, 2001.

[11] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter. *Applied Interval Analysis.* Springer, London, 2001.

[12] E. Kaucher. Self-validating computation of ordinary and partial differential equations. In E. Kaucher, U. Kulisch, and Ch. Ullrich, editors, *Computerarithmetic: Scientific Computation and Programming Languages*, pages 221–254. Teubner, Stuttgart, 1987.

[13] E. W. Kaucher and W. L. Miranker. *Self-Validating Numerics for Function Space Problems.* Academic Press, New York, 1984.

[14] W. Kühn. Rigorously computed orbits of dynamical systems without the wrapping effect. *Computing*, 61:47–67, 1998.

[15] U. Kulisch. *Advanced Arithmetic for the Digital Computer.* Springer, Wien, 2002.

[16] U. Kulisch and W. L. Miranker. *Computer Arithmetic in Theory and Practice.* Academic Press, New York, 1981.

[17] R. Lohner. *Einschließung der Lösung gewöhnlicher Anfangs- und Randwertaufgaben und Anwendungen*. PhD thesis, Universität Karlsruhe, 1988.

[18] R. Lohner. On the ubiquity of the wrapping effect in the computation of error bounds. In U. Kulisch, R. Lohner, and A. Facius, editors, *Perspectives of Enclosure Methods*, pages 201–217. Springer, Wien, 2001.

[19] K. Makino. *Rigorous analysis of nonlinear motion in particle accelerators*. PhD thesis, Michigan State University, 1998.

[20] K. Makino and M. Berz. Remainder differential algebras and their applications. In M. Berz, C. Bischof, G. Corliss, and A. Griewank, editors, *Computational Differentiation: Techniques, Applications and Tools*, pages 63–74. SIAM, Philadelphia, 1996.

[21] K. Makino and M. Berz. COSY INFINITY version 8. *Nuclear Instruments & Methods in Physics Research A*, 427:338–343, 1999.

[22] K. Makino and M. Berz. Suppression of the wrapping effect by Taylor model-based validated integrators. MSU Report MSUHEP 40910, Michigan State University, 2004.

[23] R. E. Moore. Automatic local coordinate transformations to reduce the growth of error bounds in interval computation of solutions of ordinary differential equations. In L. B. Rall, editor, *Error in Digital Computation, Vol. II*, pages 103–140. John Wiley and Sons, New York, 1965.

[24] R. E. Moore. *Interval Analysis*. Prentice Hall, Englewood Cliffs, N.J., 1966.

[25] P. S. V. Nataraj and K. Kotecha. Global optimization with higher order inclusion function forms. Part 1: A combined Taylor-Bernstein form. *Reliable Computing*, 10:27–44, 2004.

[26] N. Nedialkov and K. R. Jackson. A new perspective on the wrapping effect in interval methods for initial value problems for ordinary differential equations. In U. Kulisch, R. Lohner, and A. Facius, editors, *Perspectives of Enclosure Methods*, pages 219–264. Springer, Wien, 2001.

[27] M. Neher. Improved validated bounds for Taylor coefficients and for Taylor remainder series. *J. Comput. Appl. Math.*, 152:393–404, 2003.

[28] A. Neumaier. *Interval Methods for Systems of Equations*. Cambridge University Press, Cambridge, 1990.