ON TAYLOR MODEL BASED INTEGRATION OF ODES

M. NEHER*, K. R. JACKSON†, AND N. S. NEDIALKOV‡

Abstract. Interval methods for verified integration of initial value problems (IVPs) for ODEs have been used for more than 40 years. For many classes of IVPs, these methods are able to compute guaranteed error bounds for the flow of an ODE, where traditional methods provide only approximations to a solution. Overestimation, however, is a potential drawback of verified methods. For some problems, the computed error bounds become overly pessimistic, or the integration even breaks down. The dependency problem and the wrapping effect are particular sources of overestimations in interval computations.

Berz and his co-workers have developed Taylor model methods, which extend interval arithmetic with symbolic computations. The latter is an effective tool for reducing both the dependency problem and the wrapping effect. By construction, Taylor model methods appear particularly suitable for integrating nonlinear ODEs. We analyze Taylor model based integration of ODEs and compare Taylor model methods with traditional enclosure methods for IVPs for ODEs.

AMS subject classifications. 65G40, 65L05, 65L70.

Key words. Taylor model methods, verified integration, ODEs, IVPs.

1. Introduction. The numerical solution of initial value problems (IVPs) for ODEs is one of the fundamental problems in scientific computation. Today, there are many well-established algorithms for approximate solution of IVPs. However, traditional integration methods usually provide only approximate values for the solution. Precise error bounds are rarely available. The error estimates, which are sometimes delivered, are not guaranteed to be accurate and are sometimes unreliable.

In contrast, reliable integration computes guaranteed bounds for the flow of an ODE, including all discretization and roundoff errors in the computation. Originated by Moore in the 1960s [33], interval computations are a particularly useful tool for this purpose. There is a vast literature on interval methods for verified integration [6, 8, 9, 10, 12, 19, 21, 22, 24, 29, 31, 32, 33, 35, 36, 37, 38, 39, 40, 44, 45, 46, 47], but there are still many open questions. The results of interval arithmetic computations are often impaired by overestimation caused by the dependency problem and by the wrapping effect. In verified integration, overestimation may degrade the computed enclosure of the flow, enforce miniscule step sizes, or even bring about premature abortion of an integration.

Berz and his co-workers have developed Taylor model methods, which combine interval arithmetic with symbolic computations [2, 5, 25, 27, 28]. In Taylor model methods, the basic data type is not a single interval, but a *Taylor model*,

$$\mathcal{U} := p_n(x) + \boldsymbol{i}$$

consisting of a multivariate polynomial $p_n(x)$ of order n in m variables, and a remainder interval $\boldsymbol{i}$. In computations that involve $\mathcal{U}$, the polynomial part is propagated by symbolic calculations wherever possible, and thus not significantly affected by the dependency problem or the wrapping effect. Only the interval remainder term and polynomial terms of order higher than n, which are usually small, are bounded using interval arithmetic.

Taylor model arithmetic is an extension of interval arithmetic with a comprehensive variety of applicable enclosure sets. Nevertheless, there has been some debate about the usefulness and the limitations of Taylor model methods [42]. To some extent, this may be due to the sometimes cursory description of technical details of Taylor model arithmetic, which may be obvious to the experts of Taylor models, but which are less trivial to others.

The motivation of this paper is to analyze Taylor model methods for the verified integration of ODEs and to compare these methods with existing interval methods. Taylor models are better suited for integrating ODEs than interval methods whenever richness in available enclosure sets and reduction of the dependency problem is an advantage. This is usually the case for IVPs for nonlinear ODEs,

*Institut für Angewandte Mathematik, Universität Karlsruhe (TH), 76128 Karlsruhe, Germany
†Computer Science Department, University of Toronto, 10 King's College Rd, Toronto, ON, M5S 3G4, Canada
‡Department of Computing and Software, McMaster University, Hamilton, ON, L8S 4L7, Canada

especially in combination with large initial sets or with large integration domains. Although parameter intervals or initial sets can be handled by subdivision, this approach is only practical in low dimensions.

The advantage of Taylor model methods is less obvious for linear ODEs, where interval methods should perform equally well. Nevertheless, we include a discussion of Taylor model methods for linear ODEs in this paper for two reasons. First, the discussion is simpler for linear ODEs than for nonlinear ones. Second, if Taylor model methods failed on linear ODEs, they would likely fail on nonlinear ODEs as well. However, some of the most advantageous properties of Taylor models only take effect on nonlinear problems. We use a simple nonlinear model problem to illustrate these advantages.

The paper is structured as follows. In the next section, basic concepts of interval arithmetic and Taylor model methods are reviewed. Interval methods for ODEs are presented in Section 3. The naive Taylor model method is described in Section 4, which is followed by a discussion of Taylor model methods for linear ODEs. A nonlinear model problem is used to explain preconditioned Taylor model methods for ODEs in Section 6. In the last section, numerical examples for linear ODEs are given.

2. Preliminaries.

2.1. Interval Arithmetic. Interval arithmetic [1, 14, 33, 41] is a powerful tool for verified computations. In interval arithmetic, operations between intervals are employed to calculate guaranteed bounds for continuous problems with a finite number of basic arithmetic operations. We assume that the reader is familiar with real interval arithmetic and floating point interval arithmetic. The latter is based on a screen of floating-point numbers. Rigor of a computation is achieved by enclosing real numbers by floating-point intervals (that is, intervals with floating-point upper and lower bounds), and by performing all calculations with directed rounding according to the rules of interval arithmetic [20]. Successful software implementations of floating point interval arithmetic have for example been given in [3, 17, 18].

The set of compact real intervals is denoted by

$$\mathbb{IR} = \{\, \boldsymbol{x} = [\underline{x}, \overline{x}] \mid \underline{x}, \overline{x} \in \mathbb{R},\ \underline{x} \leq \overline{x} \,\}.$$

A real number x is identified with a point interval $\boldsymbol{x} = [x, x]$. The *midpoint* and the *width* of an interval $\boldsymbol{x}$ are denoted by $\mathrm{m}(\boldsymbol{x}) := (\overline{x} + \underline{x})/2$ and $\mathrm{w}(\boldsymbol{x}) := \overline{x} - \underline{x}$, respectively. The set of all m-dimensional interval vectors is denoted by $\mathbb{IR}^m$. In this paper, intervals are denoted by boldface. Lower-case letters are used for denoting scalars and vectors. Matrices are denoted by upper-case letters.

2.2. Dependency Problem and Wrapping Effect. Interval methods are sometimes affected by overestimation, whence the computed error bounds may be overly pessimistic. Overestimation is often caused by the *dependency problem*, that is the failure of interval arithmetic to identify different occurrences of the same variable. For example, the range of $f(x) := x/(1+x)$ on $\boldsymbol{x} = [1, 2]$ is $[1/2, 2/3]$, but interval-arithmetic evaluation yields

$$\frac{\boldsymbol{x}}{1+\boldsymbol{x}} = \frac{[1,2]}{[2,3]} = \left[\frac{1}{3}, 1\right].$$

In general, the dependency problem is not easily removed. To diminish overestimation, alternative evaluation schemes, such as centered forms [33], have been developed. A discussion of computer methods for the range of functions is given in [43].

A second source of overestimation is the *wrapping effect*, which appears when intermediate results of a computation are enclosed by intervals. The wrapping effect was first observed by Moore in 1965 [32]; a recent analysis has been given by Lohner [23].

2.3. Taylor Model Arithmetic. For reducing both the dependency problem and the wrapping effect, interval arithmetic has been extended with symbolic computations. Symbolic-numeric computations have been proposed under various names since the 1980s [11, 16, 25]. Early implementations in software were also given [11, 15], but to the authors' knowledge, these packages have not been widely distributed and are not available today.

Starting in the 1990s, Berz and his group developed a rigorous multivariate Taylor arithmetic [2, 25, 28]. In these references, a *Taylor model* is defined in the following way. Let $f : D \subset \mathbb{R}^m \to \mathbb{R}$ be a function that is $(n+1)$ times continuously differentiable in an open set containing the box $\boldsymbol{x}$. Let x_0 be a point in $\boldsymbol{x}$, let p_n denote the nth order Taylor polynomial of f around x_0, and let $\boldsymbol{i}$ be an interval such that

$$f(x) \in p_n(x - x_0) + \boldsymbol{i} \quad \text{for all} \quad x \in \boldsymbol{x}. \tag{2.1}$$

Then the pair $(p_n, \boldsymbol{i})$ is called an nth order Taylor model of f around x_0 on $\boldsymbol{x}$.

This original definition of a Taylor model is useful for computations in exact arithmetic, but it must be extended for floating point computations. For example, there is no Taylor model of $e^x \approx 1 + x + (1/2)x^2 + (1/6)x^3 + \dots$ of order $n \geq 3$ in IEEE 754 floating point arithmetic, since the coefficient of x^3 is not exactly representable as a floating point number. In [29], instead of the Taylor polynomial of f, an arbitrary polynomial p_n with floating point coefficients is used in (2.1), but the definition of a Taylor model in [29] assumes that the width of $\boldsymbol{i}$ is of order $O\big(\|\mathrm{w}(\boldsymbol{x})\|^n\big)$. In this paper, such an assumption on the width of $\boldsymbol{i}$ is not required.

We use calligraphy letters for denoting Taylor models:

$$\mathcal{U} := p_n(x) + \boldsymbol{i}, \quad x \in \boldsymbol{x},$$

where $\boldsymbol{x} \in \mathbb{IR}^m$, $\boldsymbol{i} \in \mathbb{IR}$ are intervals, and p_n is an m-variate polynomial of order n. $\boldsymbol{x}$ is called the *domain interval* of $\mathcal{U}$, and $\boldsymbol{i}$ is its *remainder interval.* A Taylor model is the set of all m-variate continuous functions f such that

$$f(x) \in p_n(x) + \boldsymbol{i}$$

holds for all $x \in \boldsymbol{x}$. Evaluating $\mathcal{U}$ for all $x \in \boldsymbol{x}$, we obtain the *range* of $\mathcal{U}$:

$$\mathrm{Rg}\,(\mathcal{U}) := \{z = p(x) + \iota \mid x \in \boldsymbol{x}, \iota \in \boldsymbol{i}\}.$$

Example 2.1. Taylor models of e^x *and* $\cos x$. Let $\boldsymbol{x} := [-\frac{1}{2}, \frac{1}{2}]$ and $x_0 := 0$. Then Taylor's theorem is a natural starting point for constructing Taylor models. We have

$$e^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 e^{\xi}, \quad \cos x = 1 - \frac{1}{2}x^2 + \frac{1}{6}x^3 \sin \xi, \quad x, \xi \in \boldsymbol{x},$$

from which we derive Taylor models for $f_1(x) := e^x$ and $f_2(x) := \cos x$:

$$\mathcal{U}_1(x) := 1 + x + \tfrac{1}{2}x^2 + [-0.035, 0.035], \quad \mathcal{U}_2(x) := 1 - \tfrac{1}{2}x^2 + [-0.010, 0.010], \quad x \in \boldsymbol{x},$$

respectively.

Taylor model arithmetic has been defined in [2, 25, 28]. We use the same arithmetic rules, even though our Taylor models differ slightly from the Taylor models defined in these references. The difference only affects the function set that is defined by a Taylor model.

In computations that involve a Taylor model $\mathcal{U}$, the polynomial part is propagated by symbolic calculations wherever possible. In floating point computations, the roundoff errors of the symbolic operations are rigorously estimated and the estimate is added to the remainder interval of the final result. This part of the computation is hardly affected by the dependency problem or the wrapping effect. Only the interval remainder term and polynomial terms of order higher than n (which in applications are usually small) are processed according to the rules of interval arithmetic.

Example 2.2. Multiplication of two univariate Taylor models of order 2. Let $\boldsymbol{x} := [-\frac{1}{2}, \frac{1}{2}]$ and

$$\mathcal{U}_1(x) := 1 + x + \tfrac{1}{2}x^2 + [-0.035, 0.035], \;\; \mathcal{U}_2(x) := 1 - \tfrac{1}{2}x^2 + [-0.010, 0.010], \quad \text{where} \quad x \in \boldsymbol{x}.$$

For all $x \in \boldsymbol{x}$, it holds that

$$\begin{aligned}
\mathcal{U}_1(x) \cdot \mathcal{U}_2(x) &\subseteq (1 + x + \tfrac{1}{2}x^2)(1 - \tfrac{1}{2}x^2) + \big(\tfrac{1}{2} + \tfrac{1}{2}(1+x)^2\big)[-0.010, 0.010] \\
&\quad + (1 - \tfrac{1}{2}x^2)[-0.035, 0.035] + [-0.035, 0.035] \cdot [-0.010, 0.010] \\
&\subseteq (1 + x) - \tfrac{1}{2}x^3 - \tfrac{1}{4}x^4 + [0.625, 1.625] \cdot [-0.010, 0.010] + [0.875, 1] \cdot [-0.035, 0.035] + [-0.004, 0.004] \\
&\subseteq 1 + x - [-0.063, 0.063] - [-0.016, 0.016] + [-0.202, 0.202] = 1 + x + [-0.281, 0.281],
\end{aligned}$$

so we may define

$$\mathcal{U}_1(x) \cdot \mathcal{U}_2(x) := 1 + x + [-0.281, 0.281].$$

This product is a Taylor model for the function $e^x \cos x$, $x \in \boldsymbol{x}$:

$$e^x \cos x \in 1 + x + [-0.281, 0.281], \quad x \in \boldsymbol{x}.$$

In Example 2.2, direct interval evaluation for computing the remainder interval of the product has been used for simplicity. Due to the dependency problem, this does not always yield optimal bounds. More accurate estimation schemes have been proposed in [30].

Compositions $\mathcal{U}_1 \circ \mathcal{U}_2$ of Taylor models are evaluated in a similar way as products; $\circ$ denotes the composition operator for functions, namely

$$(f \circ g)(x) = f\big(g(x)\big).$$

Example 2.3. Composition of two univariate Taylor models of order 2. Let $\boldsymbol{x} := [-\frac{1}{2}, \frac{1}{2}]$ and

$$\mathcal{U}_1(x) := 1 + x + \tfrac{1}{2}x^2 + [-0.035, 0.035], \quad \mathcal{U}_2(x) := 1 - \tfrac{1}{2}x^2 + [-0.010, 0.010], \quad \text{where} \quad x \in \boldsymbol{x}.$$

It is tempting to compute the composition $\mathcal{U}_1 \circ \mathcal{U}_2$ in the following manner.

$$\begin{aligned}
\mathcal{U}_1(x) \circ \mathcal{U}_2(x) &\subseteq 1 + (1 - \tfrac{1}{2}x^2 + [-0.010, 0.010]) + \tfrac{1}{2}(1 - \tfrac{1}{2}x^2 + [-0.010, 0.010])^2 + [-0.035, 0.035] \\
&\subseteq 2 - \tfrac{1}{2}x^2 + [-0.045, 0.045] + \tfrac{1}{2}(1 - x^2 + \tfrac{1}{4}x^4 + [-0.020, 0.020] - x^2[-0.010, 0.010] + [-0.001, 0.001]) \\
&\subseteq \tfrac{5}{2} - x^2 + \tfrac{1}{8}x^4 - x^2[-0.005, 0.005] + [-0.056, 0.056] \\
&\subseteq \tfrac{5}{2} - x^2 + [0, 0.008] - [-0.002, 0.002] + [-0.056, 0.056] = \tfrac{5}{2} - x^2 + [-0.058, 0.066].
\end{aligned}$$

Hence, we may define

$$\mathcal{U}_1(x) \circ \mathcal{U}_2(x) := \frac{5}{2} - x^2 + [-0.058, 0.066]. \tag{2.2}$$

However, the above computation does not yield a Taylor model for $e^{\cos x}$ for all $x \in \boldsymbol{x}$. Evaluating (2.2) at $x = 0$, we obtain

$$\mathcal{U}_1(0) \circ \mathcal{U}_2(0) = [2.442, 2.566] \not\ni e = e^{\cos 0}.$$

The reason for this failure lies in the range of $\mathcal{U}_2$, which is not contained in $\boldsymbol{x}$. Compositions of Taylor models are indeed computed as above, but it is required that the domain of $\mathcal{U}_1$ contains the range of $\mathcal{U}_2$.

In our example, it suffices to compute the remainder term for the exponential function on the interval $[-1, 1]$. Using Lagrange's representation of the remainder term, we have

$$\frac{e^\xi}{3!}x^3 \in [-\frac{e}{6}, \frac{e}{6}] \subseteq [-0.454, 0.454] \quad \text{for all } \xi \in [-1, 1] \text{ and all } x \in [-1, 1].$$

Using $[-0.454, 0.454]$ instead of $[-0.035, 0.035]$ in the derivation of (2.2) yields

$$\mathcal{U}_1(x) \circ \mathcal{U}_2(x) := \frac{5}{2} - x^2 + [-0.477, 0.485],$$

which is a verified enclosure of $\mathcal{U}_1(x) \circ \mathcal{U}_2(x)$ for all $x \in \boldsymbol{x}$. Note that it is still not a verified enclosure for all $x \in [-1, 1]$. The latter requires that the interval term of $\mathcal{U}_2$ is also computed for $x \in [-1, 1]$.

A *Taylor model vector* is a vector with Taylor model components. When no ambiguity arises, we call a Taylor model vector simply a Taylor model. Arithmetic operations for Taylor model vectors are defined componentwise.

2.3.1. Floating-Point Taylor Model Arithmetic. On a computer with floating-point arithmetic, a Taylor model is defined by a polynomial with machine representable coefficients and a suitable remainder interval that takes account for the roundoff errors. These roundoff errors can occur

- when a function is represented by a Taylor model, or
- when operations between Taylor models are executed.

Example 2.4. Addition of two univariate floating-point Taylor models. For simplicity, we use Taylor models of order 1 and a floating-point number system with a mantissa of four decimal digits. Let

$$\boldsymbol{x} := [-1, 1], \quad f_1(x) := 1 + x + \frac{1}{8}x^2, \ x \in \boldsymbol{x}, \quad f_2(x) := 1 + \frac{1}{3}x, \ x \in \boldsymbol{x}.$$

Then linear Taylor models for f_1 and f_2 are given by

$$\mathcal{U}_1(x) := 1 + x + [0, 0.125], \quad \mathcal{U}_2(x) := 1 + 0.3333x + [-0.0001, 0.0001], \quad x \in \boldsymbol{x}.$$

For $j = 1, 2$, the inclusion condition

$$f_j(x) \in \mathcal{U}_j(x) \quad \text{for all} \quad x \in \boldsymbol{x}$$

does not define $\mathcal{U}_1$ and $\mathcal{U}_2$ uniquely. For example,

$$\widetilde{\mathcal{U}}_1(x) := 1 + x + [-0.125, 0.125], \quad x \in \boldsymbol{x}$$

is also a valid, but less accurate, Taylor model for f_1.

A Taylor model for $f_1 + f_2$ is obtained by performing $\mathcal{U}_1 + \mathcal{U}_2$ with suitable outward rounding. The interval bound for the roundoff error in $x + 0.3333x$ depends of the domain $\boldsymbol{x}$.

$$\begin{aligned}\mathcal{U}_1(x) + \mathcal{U}_2(x) &\subseteq 2 + (x + 0.3333x) + [-0.0001, 0.1251]\\ &\subseteq 2 + (1.333x + [-0.0003, 0.0003]) + [-0.0001, 0.1251] = 2 + 1.333x + [-0.0004, 0.1254].\end{aligned}$$

A software implementation of Taylor model arithmetic has been developed by Berz and Makino [3, 26] in the COSY Infinity package [4]. Using COSY Infinity, Taylor models have been applied with success to a variety of problems, including global optimization [34], verified multidimensional integration [7], and the verified solution of ODEs and DAEs [6, 13].

2.4. Representation of Intervals by Taylor Models. For a given vector $c \in \mathbb{R}^m$ and a given diagonal matrix $C \in \mathbb{R}^{m \times m}$ with nonnegative diagonal elements, the range of the Taylor model vector

$$\mathcal{U} := c + Cx, \quad x \in \boldsymbol{x} \tag{2.3}$$

is an m-dimensional interval vector. Vice versa, each interval vector $\boldsymbol{z} \in \mathbb{IR}^m$ can be represented by a Taylor model vector of the form (2.3). There is freedom of choice in selecting c, C, and $\boldsymbol{x}$. A convenient choice is

$$c = \mathrm{m}(\boldsymbol{z}), \quad C = \mathrm{diag}\left(\frac{1}{2}\mathrm{w}(\boldsymbol{z})\right), \quad \boldsymbol{x} = [-1, 1]^m,$$

where $[-1, 1]^m$ denotes an interval vector with $[-1, 1]$ in each component.

Example 2.5. Let $\boldsymbol{z} = ([1, 2], [-2, 2])^T$. Then we have

$$\boldsymbol{z} = \mathrm{Rg}\left(\begin{pmatrix} \frac{3}{2} \\ 0 \end{pmatrix} + \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}\right), \quad \begin{pmatrix} x \\ y \end{pmatrix} \in [-1, 1]^2.$$

3. Interval Methods for ODEs.

3.1. Interval Initial Value Problems. We consider the smooth interval IVP

$$u' = f(t, u), \quad u(t_0) \in \boldsymbol{u}_0, \quad t \in \boldsymbol{t} = [t_0, t_{\mathrm{end}}], \tag{3.1}$$

where $f : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m$ is a sufficiently smooth function, $\boldsymbol{u}_0 \in \mathbb{IR}^m$ is a given interval vector in the space variables, and $t_{\mathrm{end}} > t_0$ is a given endpoint of the time interval. (The case $t_{\mathrm{end}} < t_0$ is handled similarly).

While the ODE is defined in the traditional way, the initial value is allowed to vary in the interval $\boldsymbol{u}_0$. In applications, this variability is used for modeling uncertainties in initial conditions. For each $u_0 \in \boldsymbol{u}_0$, the point IVP

$$u' = f(t, u), \quad u(t_0) = u_0$$

has a classical solution, which is denoted by $u(t; t_0, u_0)$. In the following, we assume that $u(t; t_0, u_0)$ exists and is bounded for all $t \in \boldsymbol{t}$ and for all $u_0 \in \boldsymbol{u}_0$.

Our goal when solving (3.1) is to calculate bounds on the flow of the interval IVP. For each $t \in \boldsymbol{t}$, we wish to calculate an interval $\boldsymbol{u}(t)$ such that

$$u(t; t_0, u_0) \in \boldsymbol{u}(t)$$

holds for all $u_0 \in \boldsymbol{u}_0$. The tube $\boldsymbol{u}(t)$, $t \in \boldsymbol{t}$, then contains all solutions of $u' = f(t, u)$ that emerge from $\boldsymbol{u}_0$.

3.2. Interval Methods for IVPs. All enclosure methods for ODEs that we are aware of subdivide the domain of integration into subintervals. At each grid point, the flow of the given ODE is enclosed by a set with a certain geometric structure, for example an m-dimensional rectangle. In the general case, the shape of the flow has a different geometry, so that the flow is wrapped by some larger set, which serves as the 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. The method thus 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. It is essential for any verified integration method to minimize the excess introduced by the wrapping of intermediate enclosures of the flow.

In Moore's *direct interval method* [31, 32, 33], the widths of the enclosures at subsequent time steps are always increasing, even for shrinking flows. For linear autonomous ODEs, the direct interval method is only suited for pure contractions. If the flow is rotated, the rotation of the initial set usually provokes exponential growth of the widths of the computed interval enclosures.

In the *parallelepiped method* [32, 33, 12, 21], the flow of the ODE at intermediate time steps is enclosed by parallelepipeds instead of rectangular boxes. This choice is motivated by the shape of the flow of a linear ODE with interval initial values, which is a parallelepiped at any time. For this problem, the only source of overestimation is the remainder interval accounting for the discretization error and the accumulated roundoff errors, if the computation is performed in floating-point arithmetic. These quantities must be enclosed by the final parallelepiped enclosure, but the wrapping only affects small quantities. The algebraic crux of the parallelepiped method is the verified inversion of certain matrices A_j [21, 36], which often tend to become singular after some time steps, so that the method breaks down either due to excessive wrapping or because the verified matrix inversion is no longer feasible. Hence, breakdown of the parallelepiped method is a rule rather than an exception.

To preserve good condition numbers in the matrices A_j, Lohner [21] developed the *QR method*. His idea was to stabilize the iteration by orthogonalization of the matrices, so that the algebraic problem of inverting the matrices is reduced to taking the transpose.

Various other interval 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 [12, 19, 21, 32, 33]. Nevertheless, the ability of interval methods to minimize wrapping is limited by the fact that interval-based enclosure sets are convex. If the flow is a non-convex set, as may arise for nonlinear ODEs, any interval wrap must be at least as large as the convex hull of the flow.

4. Taylor Model Methods for ODEs. Taylor model methods use multivariate polynomials in the initial values plus a small interval remainder term to represent the flow of an IVP. 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 of Taylor model methods over traditional interval methods, making Taylor model methods very effective in some cases in reducing the wrapping effect.

We refer to the recent paper of Makino and Berz [29] for the general 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 nonlinear example.

4.1. Quadratic Model Problem. We consider the quadratic model problem

$$\begin{aligned} u' &= v, \qquad u(0) \in [0.95, 1.05], \\ v' &= u^2, \qquad v(0) \in [-1.05, -0.95], \end{aligned} \tag{4.1}$$

where the differentiation is with respect to t. In an interval method, one would use interval initial values $\boldsymbol{u}_0 = [0.95, 1.05]$ and $\boldsymbol{v}_0 = [-1.05, -0.95]$. In the Taylor model method, the initial set is described by parameters, which we call a and b, and which we choose in the interval $[-0.05, 0.05]$. The initial conditions of the IVP (4.1) at $t = t_0$ are thus given by

$$\begin{aligned} u_0(a, b) &:= 1 + a, \qquad a \in \boldsymbol{a} := [-0.05, 0.05], \\ v_0(a, b) &:= -1 + b, \qquad b \in \boldsymbol{b} := [-0.05, 0.05]. \end{aligned}$$

For illustration, we use order $n = 3$ and step size $h = 0.1$ in the Taylor model integration of (4.1). All numbers are displayed here rounded to six decimal digits. In each integration step, the multivariate Taylor series (with respect to t, a, and b) of the solution of (4.1) is employed. The third-order Taylor polynomial serves as an approximate solution. The truncation error of the series is enclosed by a suitable remainder interval.

The first integration step consists of integrating the IVP

$$\begin{aligned} u' &= v, \qquad u(0) = 1 + a, \\ v' &= u^2, \qquad v(0) = -1 + b \end{aligned} \tag{4.2}$$

for $0 \le t \le h$. We use the Picard iteration to calculate a multivariate Taylor polynomial approximation of the solution to (4.2). Using the initial approximations

$$\begin{aligned} u^{(0)}(\tau, a, b) &= 1 + a, \\ v^{(0)}(\tau, a, b) &= -1 + b \end{aligned}$$

(τ is time), the first step of the Picard iteration yields

$$\begin{aligned} u^{(1)}(\tau, a, b) &= u_0(a, b) + \int_0^\tau v^{(0)}(s, a, b)\, ds = 1 + a - \tau + b\tau, \\ v^{(1)}(\tau, a, b) &= v_0(a, b) + \int_0^\tau \left(u^{(0)}(s, a, b)\right)^2 ds = -1 + b + \tau + 2a\tau + a^2\tau. \end{aligned}$$

After two more Picard iterations (and omitting the higher order terms), we obtain the third order Taylor polynomials

$$\begin{aligned} u^{(3)}(\tau, a, b) &= 1 + a - \tau + b\tau + \frac{1}{2}\tau^2 + a\tau^2 - \frac{1}{3}\tau^3, \\ v^{(3)}(\tau, a, b) &= -1 + b + \tau + 2a\tau - \tau^2 + a^2\tau - a\tau^2 + b\tau^2 + \frac{2}{3}\tau^3, \end{aligned}$$

as multivariate approximations to the solution of (4.2). For a verified enclosure of the flow, the Taylor polynomials have to be furnished with suitable remainder bounds. Their derivation is based on a fixed point iteration [24]. Intervals $\boldsymbol{i}_0$ and $\boldsymbol{j}_0$ are sought such that the inclusions

$$\begin{aligned} u_0 + \int_0^\tau \left(v^{(3)}(s, a, b) + \boldsymbol{j}_0\right) ds &\subseteq u^{(3)}(\tau, a, b) + \boldsymbol{i}_0, \\ v_0 + \int_0^\tau \left(u^{(3)}(s, a, b) + \boldsymbol{i}_0\right)^2 ds &\subseteq v^{(3)}(\tau, a, b) + \boldsymbol{j}_0 \end{aligned}$$

simultaneously hold for all $a \in \boldsymbol{a}$, for all $b \in \boldsymbol{b}$, and for all $\tau \in [0, 0.1]$. For the details of the computation of the remainder interval, we refer to [24]. In our example, these inclusions are fulfilled, for example, for

$$\boldsymbol{i}_0 = [-5.09307\text{E-}5, 7.86167\text{E-}5] \quad \text{and} \quad \boldsymbol{j}_0 = [-1.75707\text{E-}4, 1.60933\text{E-}4].$$

An enclosure of the flow of the IVP (4.2) for $t \in [0, 0.1]$ is given by the Taylor models

$$\begin{aligned} \widetilde{\mathcal{U}}_1(\tau, a, b) &:= 1 + a - \tau + b\tau + \frac{1}{2}\tau^2 + a\tau^2 - \frac{1}{3}\tau^3 + \boldsymbol{i}_0, \\ \widetilde{\mathcal{V}}_1(\tau, a, b) &:= -1 + b + \tau + 2a\tau - \tau^2 + a^2\tau - a\tau^2 + b\tau^2 + \frac{2}{3}\tau^3 + \boldsymbol{j}_0, \end{aligned}$$

where $a, b \in [-0.05, 0.05]$, $\tau \in [0, 0.1]$, and $t = \tau$.

Evaluating $\widetilde{\mathcal{U}}_1$ and $\widetilde{\mathcal{V}}_1$ at $\tau = h = 0.1$, we obtain the enclosure of the flow at $t_1 = 0.1$ (Taylor models of order at most 2 in the space variables):

$$\begin{aligned} \mathcal{U}_1(a, b) &:= \widetilde{\mathcal{U}}_1(0.1, a, b) = 0.904667 + 1.01a + 0.1b + \boldsymbol{i}_0, \\ \mathcal{V}_1(a, b) &:= \widetilde{\mathcal{V}}_1(0.1, a, b) = -0.909333 + 0.19a + 1.01b + 0.1a^2 + \boldsymbol{j}_0, \end{aligned} \tag{4.3}$$

which is the initial set for the second integration step. The latter is performed with a slight modification. We do not use the interval remainder terms in $\mathcal{U}_1$ and $\mathcal{V}_1$ when computing the polynomial part of the Taylor model in the space and time variables. The Picard iteration is again performed for $\tau \in [0, 0.1]$, with initial approximations

$$
\begin{aligned}
u^{(0)}(\tau, a, b) &= 0.904667 + 1.01a + 0.1b, \\
v^{(0)}(\tau, a, b) &= -0.909333 + 0.19a + 1.01b + 0.1a^2.
\end{aligned}
$$

After three iterations (and again omitting higher order terms), we obtain

$$
\begin{aligned}
u^{(3)}(\tau, a, b) = {} & 0.904667 + 1.01a + 0.1b - 0.909333\tau + 0.19a\tau + 1.01b\tau + 0.409211\tau^2 \\
& +0.1a^2\tau + 0.913713a\tau^2 + 0.0904667b\tau^2 - 0.274215\tau^3, \\
v^{(3)}(\tau, a, b) = {} & -0.909333 + 0.19a + 1.01b + 0.818422\tau + 0.1a^2 + 1.82743a\tau + 0.180933b\tau - 0.822644\tau^2 \\
& +1.0201a^2\tau + 0.202ab\tau + 0.01b^2\tau - 0.74654a\tau^2 + 0.82278b\tau^2 + 0.522429\tau^3.
\end{aligned}
$$

To compute the interval remainder term, we must find intervals $\boldsymbol{i}_1$ and $\boldsymbol{j}_1$ fulfilling the inclusions

$$
\begin{aligned}
\mathcal{U}_1(a, b) + \int_0^\tau \left(v^{(3)}(s, a, b) + \boldsymbol{j}_1\right) ds &\subseteq u^{(3)}(\tau, a, b) + \boldsymbol{i}_1, \\
\mathcal{V}_1(a, b) + \int_0^\tau \left(u^{(3)}(s, a, b) + \boldsymbol{i}_1\right)^2 ds &\subseteq v^{(3)}(\tau, a, b) + \boldsymbol{j}_1
\end{aligned}
\tag{4.4}
$$

for all a, $b \in [-0.05, 0.05]$ and for all $\tau \in [0, 0.1]$. (Note that $\boldsymbol{i}_0$ and $\boldsymbol{j}_0$ are contained in $\mathcal{U}_1$ and $\mathcal{V}_1$, respectively, from (4.3)). Suitable remainder intervals are, for example

$$
\boldsymbol{i}_1 = [-1.12850\text{E-}4, 1.65751\text{E-}4], \quad \boldsymbol{j}_1 = [-3.31917\text{E-}4, 3.24724\text{E-}4].
$$

Thus, the flow of the IVP (4.2) for $t \in [0.1, 0.2]$ is contained in the Taylor models

$$
\begin{aligned}
\widetilde{\mathcal{U}}_2(\tau, a, b) &= u^{(3)}(\tau, a, b) + \boldsymbol{i}_1, \\
\widetilde{\mathcal{V}}_2(\tau, a, b) &= v^{(3)}(\tau, a, b) + \boldsymbol{j}_1
\end{aligned}
$$

where a, $b \in [-0.05, 0.05]$, $\tau \in [0, 0.1]$, $t = \tau + 0.1$.

Evaluating at $\tau = 0.1$, we obtain the enclosure of the flow at $t_2 = 0.2$ (Taylor models of order at most 2 in the space variables):

$$
\begin{aligned}
\mathcal{U}_2(a, b) := \widetilde{\mathcal{U}}_2(0.1, a, b) &= 0.817551 + 1.03814a + 0.201905b + 0.01a^2 + \boldsymbol{i}_1, \\
\mathcal{V}_2(a, b) := \widetilde{\mathcal{V}}_2(0.1, a, b) &= -0.835195 + 0.365277a + 1.03632b \\
& \quad +0.20201a^2 + 0.0202ab + 0.001b^2 + \boldsymbol{j}_1.
\end{aligned}
$$

For larger values of t, the integration can be continued as in the second integration step described above.

Remark 4.1.

1. The sets $(\mathcal{U}_j, \mathcal{V}_j)$ containing the flow of the IVP (4.2) generally become more and more irregular for increasing j. Integration over a larger domain is shown in Figure 6.1.
2. In the above calculations, the polynomial parts of the Taylor models are independent of the initial domain intervals for a and b and independent of the step size h, but the interval remainder bounds are not.
3. The order of the method refers to the order of the multivariate Taylor polynomials with respect to space and time variables that are calculated in the integration step. When the initial sets are defined by linear functions in a and b, then it follows by induction that the maximum order of the polynomials representing the flow at the grid points (obtained after evaluating t) is always at least one less than the order of the method.

In the above example, we have used the so-called *naive* Taylor model integration method to illustrate the qualitative difference of interval methods and Taylor model methods for solving IVPs. For practical computations, the naive Taylor model method is not very useful. The interval remainder terms are propagated as in the direct interval method. The inclusion (4.4) implies that the diameters of the interval remainder terms are nondecreasing. Often, these diameters grow exponentially, and the method breaks down early. More advanced Taylor model integration methods are discussed in the next section. For clarity, we summarize the major steps of the naive Taylor model method as Algorithm 4.1.

Algorithm 4.1 (naive Taylor model method)

Let the initial set be given as a Taylor model vector in m space variables.

For $j := 0, 1 \dots, j_{\max} - 1$:

1. Compute the Taylor polynomial p_n (of dimension m in $m+1$ variables) of the solution of the $j+1$st time step, using Picard iteration.
2. Compute a remainder interval vector $\boldsymbol{i}$, using Schauder's fixed point theorem (via interval iteration based on Picard iteration).
3. Evaluate $\widetilde{\mathcal{U}} = p_n + \boldsymbol{i}$ at t_{j+1}. The resulting m-dimensional Taylor model $\mathcal{U}$ contains the flow of the IVP and serves as initial set for the next time step.

4.2. Shrink Wrapping and Preconditioning. For successful integration over long time spans, sophisticated treatment of the interval terms is required. For this purpose, Berz and Makino invented two schemes which they call *shrink wrapping* and *preconditioning*. Shrink wrapping is a method to absorb the interval remainder term into the symbolic part of the Taylor model. From a geometric viewpoint, it resembles the parallelepiped method. Shrink wrapping uses the same linear map as the parallelepiped method, so that it has the same limitations when this map becomes ill-conditioned. Preconditioning aims at maintaining a small condition number for the shrink wrapping map. Thus it stabilizes the integration process, like the QR interval method does.

For clarity of the presentation, we describe shrink wrapping and preconditioning for the special case of linear autonomous ODEs. The generalization to nonlinear ODEs is straightforward. We refer to [29] for the details.

5. Taylor Model Methods for Linear ODEs. For a linear ODE, the flow of an interval IVP is a parallelepiped for all time, so Taylor models seem to have no obvious advantage over interval methods. On the other hand, if Taylor model methods failed on linear ODEs, they would probably not be effective for nonlinear ODEs. The purpose of this section is to show that they can be as good as interval methods for linear ODEs.

We consider the linear autonomous ODE

$$\begin{aligned} u' &= B\,u \\ u(0) &= \mathcal{U}_0, \end{aligned} \tag{5.1}$$

where B is a given real matrix, $\boldsymbol{x}$ is a given interval vector, and $\mathcal{U}_0 = p_n(x)$, $x \in \boldsymbol{x}$, is a Taylor model vector with zero remainder interval describing the initial set. x is used to denote the vector of the space variables. We assume that the enclosure step in the Taylor model method is feasible with some constant step size $h > 0$ and some order $n \in \mathbb{N}$.

5.1. Naive Taylor Model Method. In the first integration step, Picard iteration of order n is used to compute the multivariate Taylor polynomial

$$u_{1,n} := P_n(tB)\,p_n(x), \quad \text{where} \;\; P_n(tB) := \sum_{k=0}^{n} \frac{(tB)^k}{k!}.$$

Introducing $T := P_n(hB)$, the verification step consists of finding an interval vector $\boldsymbol{i}_1$ such that

$$p_n(x) + \int_0^h B\big(P_n(\tau B)\,p_n(x) \;+ \boldsymbol{i}_1\big)\,d\tau \;\subseteq\; P_n(hB)\,p_n(x) \;+ \boldsymbol{i}_1 = Tp_n(x) \;+ \boldsymbol{i}_1$$

holds for all $x \in \boldsymbol{x}$ (see for example [24, Ch. 6]). At $t_1 = h$, the flow of the IVP (5.1) is then enclosed by the Taylor model

$$\mathcal{U}_1 := T\,p_n(x) + \boldsymbol{i}_1.$$

Subsequent integration steps are performed in the same manner, but with a slight modification in the verification step. In the jth integration step, $j \geq 2$, $\boldsymbol{i}_j$ is sought such that the inclusion

$$T^{j-1}p_n(x) + \boldsymbol{i}_{j-1} + \int_0^h B\left(P_n(\tau B)\, T^{j-1}p_n(x) + \boldsymbol{i}_j\right) d\tau \subseteq T^j p_n(x) \,+ \boldsymbol{i}_j$$

is fulfilled for all $x \in \boldsymbol{x}$. Letting

$$\mathcal{U}_j := T\,\mathcal{U}_{j-1} + \boldsymbol{i}_j, \quad j = 1, 2, \ldots,$$

the *naive Taylor model method* for (5.1) consists of the iteration

$$\mathcal{U}_j = T^j\,\mathcal{U}_0 + \sum_{k=1}^{j} (T\circ)^{j-k}\boldsymbol{i}_k, \quad j = 1, 2, \ldots, \tag{5.2}$$

where

$$(T\circ)^0\boldsymbol{x} := \boldsymbol{x}, \quad (T\circ)^k\boldsymbol{x} := T \cdot \left((T\circ)^{k-1}\boldsymbol{x}\right), \quad k \in \mathbb{N}.$$

Apart from the different computation of the remainder interval, for the initial value problem (5.1), the naive Taylor model method (5.2) coincides with the direct interval method that occurs in [36]. Hence, the naive Taylor model method (5.2) has the same divergence property as the direct interval method, for which it was shown in [36] that after j steps we have

$$\mathrm{w}\left((T\circ)^{j-1}\boldsymbol{i}_1\right) \;=\; |T|^{\,j-1}\,\mathrm{w}(\boldsymbol{i}_1)$$

(for $A = (a_{ij})$, we denote by $|A|$ the matrix with components $|a_{ij}|$). The key point here is that the spectral radius of $|T|^{\,j-1}$ may be much larger than the spectral radius of T^{j-1}, which describes the natural error growth of a point method. If this is the case, the error bounds for the naive Taylor model method may be much larger than the true error.

5.2. Naive Taylor Model Method with Shrink Wrapping. Berz and Makino [29] defined shrink wrapping as a method for absorbing the interval part of the Taylor model into the polynomial part by modifying the polynomial coefficients. The set defined by the sum of the given polynomial and interval is wrapped by a set defined by a pure polynomial. The new set may be larger than the initial set, but it is less prone to the dependency problem and to the wrapping effect in succeeding calculations.

In the verified integration of ODEs, shrink wrapping is usually applied to the Taylor model enclosures of the flow at the grid points, before continuing the integration. In practical computations, shrink wrapping is performed when the size of the interval remainder term exceeds some heuristically chosen bound. After shrink wrapping, the initial set of the subsequent integration step is purely symbolic, which removes the dependency problem and simplifies the verification step. The success of the Taylor model based integration method depends on the successful reduction of the excess introduced in the shrink wrapping process.

The process of applying shrink wrapping to a Taylor model vector

$$\mathcal{U} := p(x) + \boldsymbol{i}, \quad x \in \boldsymbol{x},$$

is described in [29]. Here, we only outline its four basic steps. First, let $\widetilde{\mathcal{U}}$ denote the Taylor model that is obtained when the constant part of p is removed. Second, multiply $\widetilde{\mathcal{U}}$ by the inverse of the matrix associated with its linear part and obtain the Taylor model $\widehat{\mathcal{U}}$. Third, estimate the nonlinear part of $\widehat{\mathcal{U}}$, its Jacobian, and the interval term of $\widehat{\mathcal{U}}$, to obtain the shrink wrap factor $q \geq 1$. Fourth, multiply the polynomial part of $\widetilde{\mathcal{U}}$ with q and add the constant part of $\mathcal{U}$.

We illustrate shrink wrapping with the following nonlinear example. For clarity, we use two scalar Taylor models $\mathcal{U}$ and $\mathcal{V}$ instead of a Taylor model vector. The symbolic variables are denoted by a and b (instead of the vector x).

Example 5.1. Absorption of the interval part into the symbolic part of a Taylor model. We consider the Taylor model vector $(\mathcal{U}, \mathcal{V})^T$, where

$$\left.\begin{aligned} \mathcal{U}(a,b) &:= 2 + 4a + \tfrac{1}{2}a^2 + [-0.2, 0.2], \\ \mathcal{V}(a,b) &:= 1 + 3b + ab + [-0.1, 0.1], \end{aligned}\right\} \quad a,\, b \in [-1, 1]. \tag{5.3}$$

The set defined by (5.3) is shown in Figure 5.1. Following the above outline, we obtain

$$\begin{aligned}\widetilde{\mathcal{U}}(a,b) &= 4a + \tfrac{1}{2}a^2 + [-0.2, 0.2],\\ \widetilde{\mathcal{V}}(a,b) &= 3b + ab + [-0.1, 0.1].\end{aligned} \tag{5.4}$$

The matrix associated with the linear part of the Taylor model (5.4) is

$$C := \begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix}.$$

Multiplying (5.4) with C^{-1}, we have

$$\begin{aligned}\widehat{\mathcal{U}}(a,b) &= a + \tfrac{1}{8}a^2 + [-0.05, 0.05],\\ \widehat{\mathcal{V}}(a,b) &= b + \tfrac{1}{3}ab + [-0.034, 0.034].\end{aligned}$$

Estimating the nonlinear part and the interval terms as described in [29], we compute numbers s, t, and d satisfying

$$\begin{aligned}&s \geq |\frac{1}{8}a^2|\,,\ s \geq |\frac{1}{3}ab| \quad \text{for all } a, b \in [-1,1],\\ &t \geq |\frac{1}{4}a|\,,\ t \geq |\frac{1}{3}b|\,,\ t \geq |\frac{1}{3}a| \quad \text{for all } a, b \in [-1,1],\\ &d \geq 0.05,\ d \geq 0.034.\end{aligned}$$

These conditions are fulfilled for $s = t = \frac{1}{3}$ and $d = 0.05$, from which we deduce the shrink wrap factor [29]

$$q = 1 + d \cdot \frac{1}{(1-t)(1-s)} = \frac{89}{80}.$$

The final Taylor model after shrink wrapping is

$$\begin{aligned}\mathcal{U}_{\mathrm{sw}}(a,b) &:= 2 + \frac{89}{20}a + \frac{89}{160}a^2,\\ \mathcal{V}_{\mathrm{sw}}(a,b) &:= 1 + \frac{287}{80}b + \frac{89}{80}ab.\end{aligned} \tag{5.5}$$

As Figure 5.1 shows, the set defined by (5.3) is contained in the set defined by (5.5).

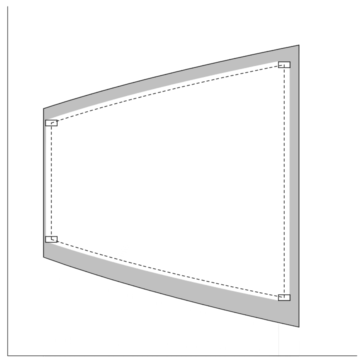

FIG. 5.1. *Sets of the Taylor models before (Eq. (5.3)) and after shrink wrapping (Eq. (5.5)). The dotted line is the boundary of the set that is described by the polynomial of the original Taylor model. The white area is the set described by the original Taylor model, including the interval term. The excess area introduced by shrink wrapping is shaded in grey.*

Applying shrink wrapping in the linear model problem (5.1) is rather simple. For simplicity, let us assume that shrink wrapping is performed in every integration step. Then we must compute [29] $q_j := 1 + d_j/2$, where

$$d_j := \|\mathrm{w}\big((T^j)^{-1}\boldsymbol{i}_j\big)\|_\infty\,.$$

If T is sufficiently well-conditioned, and if the interval terms are sufficiently small, then the factors d_j are almost zero, and shrink wrapping is feasible for many integration steps.

The naive Taylor model method with shrink wrapping resembles the parallelepiped method. By multiplying the non-constant coefficients of the Taylor polynomial, for linear autonomous ODEs the interval term is absorbed as in the parallelepiped method. While T^j is well-conditioned, d_j is small, and so is the excess area. On the other hand, q_j (and the excess area) becomes large if T^j becomes ill conditioned, which is eventually the case if T has eigenvalues of different magnitude. In this case the integration breaks down due to the growth of the Taylor polynomial coefficients.

The naive TM method with shrink wrapping is outlined as Algorithm 5.1.

Algorithm 5.1 (naive TM method with shrink wrapping)

Let the initial set be given as a Taylor model vector in m space variables.

For $j := 0, 1 \ldots, j_{\max} - 1$:

1. Compute the m-dimensional Taylor model $\mathcal{U} = p_n + \boldsymbol{i}$ (containing the flow of the IVP at t_{j+1}) as in the naive Taylor model method.
2. Absorb $\boldsymbol{i}$ into p_n by shrink wrapping.
3. Continue the integration with the modified polynomial as the initial set for the next time step.

5.3. Preconditioned Taylor Models. We showed in the previous section that shrink wrapping has the same limitations as the parallelepiped method in traditional interval arithmetic. To make Taylor model based integration successful for a larger class of IVPs, some stabilization process similar to the QR interval method is required. For restoring good condition numbers of the maps defined by the linear parts of the Taylor models in the integration process, Berz and Makino developed preconditioned Taylor models [29].

In the naive Taylor model method with or without shrink wrapping, the flow of the ODE $u' = f(t, u)$ is represented by a single Taylor model at each grid point. In the preconditioned Taylor model method, the flow of the ODE at $t = t_j$ is represented by a composition of a left and a right Taylor model

$$\mathcal{U}_l \circ \mathcal{U}_r = (p_{l,j} + \boldsymbol{i}_{l,j}) \circ (p_{r,j} + \boldsymbol{i}_{r,j}).$$

Definition 5.2. *The composition*

$$\mathcal{U}(x) := \big(p_l(x) + \boldsymbol{i}_l\big) \circ \big(p_r(x) + \boldsymbol{i}_r\big) \tag{5.6}$$

of two Taylor models

$$\begin{aligned} \mathcal{U}_l(x) &:= p_l(x) + \boldsymbol{i}_l, \;\; x \in \boldsymbol{x}_l, \\ \mathcal{U}_r(x) &:= p_r(x) + \boldsymbol{i}_r, \; x \in \boldsymbol{x}_r, \end{aligned}$$

is called a preconditioned Taylor model *if*

$$\mathrm{Rg}\,(\mathcal{U}_r) \subseteq \boldsymbol{x}_l. \tag{5.7}$$

The range enclosure condition (5.7) is essential in verified integration with preconditioned Taylor models (see discussion below). The factorization into a left and a right Taylor model is not unique. Two preconditioned Taylor models of the form (5.6) can have the same domain $\boldsymbol{z}$ and the same range, but different polynomials and remainder intervals. In verified integration, preconditioning is used to replace some representation of the flow at an intermediate grid point by a different set of initial values that is more suitable for continuing the integration. Here preconditioning is essentially a substitution in space variables. In the continuation of the integration, the right Taylor model is not involved at all. The following theorem is a reformulation of a proposition given without a proof by Makino and Berz [29].

THEOREM 5.3. *If the initial set of an IVP is given by a preconditioned Taylor model, then integrating the flow of the ODE only acts on the left Taylor model.*

For better understanding of this theorem, which is the key point of the preconditioned integration method, we present first a formal proof, then an example with symbolic integration, and finally a numerical example.

Proof. The space variables are parameters in the integration with respect to time. If $F(x,t)$ is a primitive of $f(x,t)$, that is if

$$\int f(x,t)\,dt = F(x,t),$$

then substituting $x = g(u)$ does not affect F:

$$\int f(g(u),t)\,dt = F(g(u),t).$$

Preconditioned integration uses $x = (p_{l,j} + \boldsymbol{i}_{l,j})$ and $g(u) = (p_{r,j} + \boldsymbol{i}_{r,j})$. □

Example 5.4. Preconditioned symbolic integration over two time steps. We consider the IVP

$$\begin{aligned} x' &= x(x+y), \quad x(0) = 1+a, \\ y' &= -x(x+y), \quad y(0) = -1+b. \end{aligned}$$

Its unique solution is

$$\begin{aligned} x(t) &= (1+a)e^{(a+b)t}, \\ y(t) &= a+b-(1+a)e^{(a+b)t}, \end{aligned}$$

so that at $t = 1$,

$$x(1) = (1+a)e^{a+b}, \quad y(1) = a+b-(1+a)e^{a+b}.$$

To continue the integration, we use the IVP

$$\begin{aligned} u' &= u(u+v), \qquad u(0) = \alpha, \\ v' &= -u(u+v), \quad v(0) = \beta \end{aligned}$$

and obtain

$$u(1) = \alpha e^{\alpha+\beta}, \quad v(1) = \alpha+\beta-\alpha e^{\alpha+\beta}.$$

Due to the substitution rule, $u(1) = x(2)$ and $v(1) = y(2)$. Indeed, letting

$$\begin{aligned} \alpha &= (1+a)e^{a+b}, \\ \beta &= a+b-(1+a)e^{a+b}, \end{aligned}$$

we obtain

$$\begin{aligned} u(1) &= (1+a)e^{2(a+b)} = x(2), \\ v(1) &= (a+b)-(1+a)e^{2(a+b)} = y(2). \end{aligned}$$

The same variable substitution as in Example 5.4 is applied when the initial set for an ODE is given by some preconditioned Taylor model $\mathcal{U}_l \circ \mathcal{U}_r$. To compute an enclosure of the flow, it suffices to integrate the given ODE for the initial values defined by $\mathrm{Rg}\,(\mathcal{U}_l)$, and to compose the integrated Taylor model with $\mathcal{U}_r$. If higher order terms appear in the composition process, they are included in the remainder interval of the result, as in Example 2.2.

In practice, preconditioning is used to replace the integrated preconditioned flow at the end of the j-th integration step,

$$\left(\oint \mathcal{U}_{l,j}\right) \circ \mathcal{U}_{r,j},$$

(where $\oint \mathcal{U}$ denotes integrated flow with respect to the given ODE) by a different preconditioned Taylor model

$$\mathcal{U}_{l,j+1} \circ \mathcal{U}_{r,j+1}.$$

The initial set for the $(j+1)$-st integration step is defined by $\mathrm{Rg}\,(\mathcal{U}_{l,j+1})$. The method is successful if

- the amount of overestimation in the wrapping of $\left(\oint \mathcal{U}_{l,j}\right) \circ \mathcal{U}_{r,j}$ by $\mathcal{U}_{l,j+1} \circ \mathcal{U}_{r,j+1}$ is sufficiently small, and if
- $\mathrm{Rg}\,(\mathcal{U}_{l,j+1})$ is better suited for continuing the integration than $\oint \mathcal{U}_{l,j}$. For example, preconditioning can be used to reduce the condition number of certain matrices that control the propagation of the global error (see example below), or to reduce the number of nonzero elements in the polynomial part of the left Taylor model.

In Lohner's QR-method, an ill-conditioned parallelepiped is wrapped by some well-conditioned m-dimensional rectangle. For preconditioning Taylor models, a large variety of well-conditioned wraps are conceivable. The optimal choice is still an open question for future research.

One important aspect of preconditioned integration is the computation of the remainder bounds in the Picard iteration. If the initial set is given by (5.6), the validity of the enclosure is already guaranteed if the remainder intervals hold for $x \in \mathrm{Rg}\,(\mathcal{U}_r)$. In practice, the remainder bounds are calculated for $x \in \boldsymbol{x}$, a larger set and a potential source of overestimation. In practical computations, overestimation (loss of accuracy) is usually converted to costs (increase of computation time). A common strategy is to limit the admissible size of the remainder intervals by some prescribed bound. Using a larger initial set then has the effect of reducing step sizes and increasing overall computation time.

A simple choice for the left Taylor model (the initial set) in each integration step is a well-conditioned linear map (a parallelepiped). The following description of preconditioned integration is a simplified version of the presentation in [29]. We consider the linear autonomous IVP

$$\begin{aligned} u' &= B\,u \\ u(0) &= u_0 = c_0 + C_0 x, \end{aligned} \tag{5.8}$$

where B is a real matrix, c_0 is a real vector, C_0 is a diagonal matrix, and x is contained in $[-1,1]^m$. The initial set is given by a Taylor model vector of the form (2.3). A suitable preconditioned Taylor model for this initial set is

$$p_{l,0}(x) = c_0 + C_0 x, \quad \boldsymbol{i}_{l,0} = 0, \quad p_{r,0}(x) = x, \quad \boldsymbol{i}_{r,0} = 0.$$

We assume that the flow at t_j is given by the preconditioned Taylor model

$$\mathcal{U}_j := (p_{l,j} + \boldsymbol{i}_{l,j}) \circ (p_{r,j} + \boldsymbol{i}_{r,j}) = (c_{l,j} + C_{l,j}\,x + \boldsymbol{i}_{l,j}) \circ (c_{r,j} + C_{r,j}\,x + \boldsymbol{i}_{r,j}),$$

where $c_{l,j}$ and $c_{r,j}$ are real vectors, $C_{l,j}$ and $C_{r,j}$ are real matrices. Using the matrix T from Section 5.1, the flow after integration is given by

$$\mathcal{U}_{j+1} := (T c_{l,j} + T C_{l,j}\,x + \boldsymbol{i}_{l,j+1}) \circ (p_{r,j} + \boldsymbol{i}_{r,j}).$$

For $c_{l,j+1} := T c_{l,j}$ and any nonsingular matrix $C_{l,j+1}$, the preconditioned Taylor model $\mathcal{U}_{j+1}$ can be rewritten as

$$\begin{aligned} \mathcal{U}_{j+1} &= (T c_{l,j} + C_{l,j+1}\,x + [0,0]) \circ \left\{ \left[C_{l,j+1}^{-1} T C_{l,j}\,x + C_{l,j+1}^{-1} \boldsymbol{i}_{l,j+1} \right] \circ (p_{r,j} + \boldsymbol{i}_{r,j}) \right\} \\ &= (c_{l,j+1} + C_{l,j+1}\,x + [0,0]) \circ \left\{ \left[C_{l,j+1}^{-1} T C_{l,j}\,x + C_{l,j+1}^{-1} \boldsymbol{i}_{l,j+1} \right] \circ (c_{r,j} + C_{r,j}\,x + \boldsymbol{i}_{r,j}) \right\} \\ &= (c_{l,j+1} + C_{l,j+1}\,x + [0,0]) \circ \left\{ C_{l,j+1}^{-1} T C_{l,j}\; (c_{r,j} + C_{r,j}\,x + \boldsymbol{i}_{r,j}) + C_{l,j+1}^{-1} \boldsymbol{i}_{l,j+1} \right\} \\ &= (c_{l,j+1} + C_{l,j+1}\,x + [0,0]) \\ &\quad \circ \left\{ C_{l,j+1}^{-1} T C_{l,j}\, c_{r,j} + C_{l,j+1}^{-1} T C_{l,j} C_{r,j}\,x + C_{l,j+1}^{-1} T C_{l,j}\, \boldsymbol{i}_{r,j} + C_{l,j+1}^{-1} \boldsymbol{i}_{l,j+1} \right\} \\ &=: (c_{l,j+1} + C_{l,j+1}\,x + [0,0]) \circ (c_{r,j+1} + C_{r,j+1}\,x + \boldsymbol{i}_{r,j+1}). \end{aligned}$$

The interval term $\boldsymbol{i}_{r,j}$ in the preconditioned Taylor model integration of (5.8) is propagated as the interval term in the parallelepiped and QR interval iteration, if $C_{l,j+1}$ is chosen as in those methods. For $C_{l,j+1} = TC_{l,j}$, the parallelepiped method is obtained, for $TC_{l,j}P_j = Q_jR_j$ (where P_j is a permutation matrix for sorting the columns of $TC_{l,j}$) and $C_{l,j+1} = Q_j$, the QR method. Numerical examples confirming these relations are presented in Section 7.

For nonlinear ODEs, the nonlinear terms in the left Taylor model can be shifted to the right Taylor model in the same manner [29]. However, the resulting Taylor model methods then differ from the corresponding interval methods. First, the symbolic parts of the composed Taylor models describe nonlinear enclosures sets of the flow, which need not be convex, in contrast to interval methods. Second, the nonlinear terms in the left Taylor models then also act on the interval terms in the right Taylor models. An analysis of the resulting interval propagation will be the subject of future research.

6. Preconditioned Quadratic Example. We now demonstrate QR preconditioned Taylor model integration for the quadratic model problem of Section 4.1, namely

$$\begin{aligned} u' &= v, \qquad u(0) \in [0.95, 1.05], \\ v' &= u^2, \qquad v(0) \in [-1.05, -0.95]. \end{aligned}$$

In each integration step, the left Taylor models are constructed via a QR factorization of the linear parts of the integrated Taylor models of the previous integration step. As in the naive integration of this IVP in Section 4.1, order $n = 3$ and step size $h = 0.1$ are used, and all numbers are displayed rounded to six decimal digits.

In the first integration step, the initial set is described by the left Taylor model in space variables at t_0. The right Taylor model at t_0 is the identity map in space variables. Hence, the first integration step is performed as in the naive Taylor model method (cf. Section 4.1), and we obtain the integrated left Taylor models (4.3), namely

$$\left.\begin{aligned} \widetilde{\mathcal{U}}_{l,1}(a,b) &:= 0.904667 + 1.01a + 0.1b + \widetilde{\boldsymbol{i}}_0, \\ \widetilde{\mathcal{V}}_{l,1}(a,b) &:= -0.909333 + 0.19a + 1.01b + 0.1a^2 + \widetilde{\boldsymbol{j}}_0, \end{aligned}\right\} \quad a,\ b \in [-0.05, 0.05],$$

where

$$\widetilde{\boldsymbol{i}}_0 = [-5.09307\text{E-}5, 7.86167\text{E-}5], \quad \widetilde{\boldsymbol{j}}_0 = [-1.75707\text{E-}4, 1.60933\text{E-}4].$$

For reasons that will soon become clear, we normalize the domain such that a and b are contained in $[-1, 1]$. Doing so (without changing the names of the variables), we have

$$\left.\begin{aligned} \widetilde{\mathcal{U}}_{l,1}(a,b) &:= 0.904667 + 0.0505a + 0.005b + \widetilde{\boldsymbol{i}}_0, \\ \widetilde{\mathcal{V}}_{l,1}(a,b) &:= -0.909333 + 0.0095a + 0.0505b + 0.00025a^2 + \widetilde{\boldsymbol{j}}_0, \end{aligned}\right\} \quad a,\ b \in [-1, 1].$$

So far, the right Taylor models have been unaffected by the integration process. Before continuing the integration, however, we precondition the left Taylor models. We extract the linear parts of $\widetilde{\mathcal{U}}_{l,1}$ and $\widetilde{\mathcal{V}}_{l,1}$, and obtain the matrix $C_{l,1}$, from which we compute a QR factorization.

$$C_{l,1} := \begin{pmatrix} 0.0505 & 0.005 \\ 0.0095 & 0.0505 \end{pmatrix} = \begin{pmatrix} 0.982762 & -0.184876 \\ 0.184876 & 0.982762 \end{pmatrix} \cdot \begin{pmatrix} 0.0513858 & 0.0142500 \\ 0 & 0.0487051 \end{pmatrix} =: QR.$$

The left Taylor models in the second integration step are built from the constant terms of $\widetilde{\mathcal{U}}_{l,1}$ and $\widetilde{\mathcal{V}}_{l,1}$ and from Q. Thus we get

$$\overline{\mathcal{U}}_{l,1}(a,b) := 0.904667 + 0.982762a - 0.184876b,$$

$$\overline{\mathcal{V}}_{l,1}(a,b) := -0.909333 + 0.184876a + 0.982762b.$$

The nonlinear term $0.00025a^2$ in $\widetilde{\mathcal{V}}_{l,1}$ and the interval terms $\widetilde{\boldsymbol{i}}_0$, $\widetilde{\boldsymbol{j}}_0$ are collected in the right Taylor models, which are multiplied by Q^T. We obtain

$$Q^T \cdot \begin{pmatrix} 0 \\ 0.00025a^2 \end{pmatrix} = \begin{pmatrix} 0.0000462190a^2 \\ 0.000245691a^2 \end{pmatrix}$$

and

$$\begin{pmatrix} \bar{\boldsymbol{\imath}}_0 \\ \bar{\boldsymbol{\jmath}}_0 \end{pmatrix} := Q^T \cdot \begin{pmatrix} \tilde{\boldsymbol{\imath}}_0 \\ \tilde{\boldsymbol{\jmath}}_0 \end{pmatrix} = \begin{pmatrix} [-8.25368\text{E-}5, 1.07014\text{E-}4] \\ [-1.87213\text{E-}4, 1.67575\text{E-}4] \end{pmatrix},$$

which yields

$$\left.\begin{aligned} \overline{\mathcal{U}}_{r,1}(a,b) &:= 0.0513858a + 0.0142500b + 0.0000462190a^2 + \bar{\boldsymbol{\imath}}_0, \\ \overline{\mathcal{V}}_{r,1}(a,b) &:= 0.0487051b + 0.000245691a^2 + \bar{\boldsymbol{\jmath}}_0, \end{aligned}\right\} \quad a,\ b \in [-1,1].$$

Before we can continue the integration, we must further modify the preconditioned Taylor models. This is probably the most surprising part of the algorithm. It is also crucial for the validity of the method. After the first time step, the flow of the IVP is contained in the composition of the left and right Taylor models. For continuing the integration, we want to drop the right Taylor model. On one hand, this is only feasible if the left Taylor model contains the flow of the IVP. On the other hand, the set defined by the left Taylor model should not be much larger than the current flow, because that would mean large overestimation. There are two potential solutions for ensuring the desired inclusion property. We can either modify the domain of the independent variables, or we may modify the left Taylor model by an additional transformation. We describe both alternatives in the following.

The starting point of the transformation is the range of the right Taylor model. We have

$$\begin{aligned} \operatorname{Rg}\left(\overline{\mathcal{U}}_{r,1}\right) &\subseteq 0.0513858 \cdot [-1,1] + 0.0142500 \cdot [-1,1] + 0.0000462190 \cdot [0,1] + [-8.25368\text{E-}5, 1.07014\text{E-}4] \\ &= [-0.0657183368, 0.065789033] \subseteq [-0.0657183, 0.0657890], \\ \operatorname{Rg}\left(\overline{\mathcal{V}}_{r,1}\right) &\subseteq 0.0487051 \cdot [-1,1] + 0.000245691 \cdot [0,1] + [-1.87213\text{E-}4, 1.67575\text{E-}4] \\ &= [-0.048892151, 0.049118366] \subseteq [-0.0488922, 0.0491184]. \end{aligned}$$

Thus we may continue the integration with the initial set for the second time step given by

$$\left.\begin{aligned} \widehat{\mathcal{U}}_{l,1}(a,b) &:= 0.904667 + 0.982762a - 0.184876b, \\ \widehat{\mathcal{V}}_{l,1}(a,b) &:= -0.909333 + 0.184876a + 0.982762b, \end{aligned}\right\} \quad \begin{aligned} &a \in [-0.0657183, 0.0657890], \\ &b \in [-0.0488922, 0.0491184] \end{aligned}$$

(unchanged polynomials, but modified domain).

Alternatively, we can apply a linear transformation on the left and the right Taylor models by a scaling matrix [29]. It is convenient here to denote the linear map (that is, a linear Taylor model $\mathcal{S}$ with zero constant part and zero interval remainder term) associated with a matrix S by the matrix itself. First note that for any nonsingular matrix S,

$$(\overline{\mathcal{U}}_{l,1}, \overline{\mathcal{V}}_{l,1}) \circ (\overline{\mathcal{U}}_{r,1}, \overline{\mathcal{V}}_{r,1}) = (\overline{\mathcal{U}}_{l,1}, \overline{\mathcal{V}}_{l,1}) \circ (S \circ S^{-1}) \circ (\overline{\mathcal{U}}_{r,1}, \overline{\mathcal{V}}_{r,1}) \subseteq ((\overline{\mathcal{U}}_{l,1}, \overline{\mathcal{V}}_{l,1}) \circ S) \circ (S^{-1} \circ (\overline{\mathcal{U}}_{r,1}, \overline{\mathcal{V}}_{r,1})),$$

where the subset property is induced by the subdistributivity law of interval arithmetic [1, p. 3]. Letting

$$S := \begin{pmatrix} 0.0657890 & 0 \\ 0 & 0.0491184 \end{pmatrix},$$

we obtain

$$\begin{aligned} (\overline{\mathcal{U}}_{l,1}, \overline{\mathcal{V}}_{l,1}) \circ S &= \begin{pmatrix} 0.904667 \\ -0.909333 \end{pmatrix} + \begin{pmatrix} 0.982762 & -0.184876 \\ 0.184876 & 0.982762 \end{pmatrix} \begin{pmatrix} 0.0657890 & 0 \\ 0 & 0.0491184 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \\ &= \begin{pmatrix} 0.904667 \\ -0.909333 \end{pmatrix} + \begin{pmatrix} 0.0646550 & -0.00908081 \\ 0.0121628 & 0.0482716 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}. \end{aligned}$$

Since S has been determined such that the range of each component of $S^{-1} \circ (\overline{\mathcal{U}}_{r,1}, \overline{\mathcal{V}}_{r,1})$ is contained in $[-1,1]$, it is feasible to continue the integration with the left Taylor models

$$\left.\begin{aligned} \mathcal{U}_{l,1}(a,b) &:= 0.904667 + 0.0646550a - 0.00908081b, \\ \mathcal{V}_{l,1}(a,b) &:= -0.909333 + 0.0121628a + 0.0482716b, \end{aligned}\right\} \quad a,\ b \in [-1,1]$$

as initial set for the second time step (modified polynomials, but original domain). The corresponding right Taylor models are

$$\begin{pmatrix} \mathcal{U}_{r,1} \\ \mathcal{V}_{r,1} \end{pmatrix} := S^{-1} \circ (\overline{\mathcal{U}}_{r,1}, \overline{\mathcal{V}}_{r,1}) = \begin{pmatrix} 15.2001 & 0 \\ 0 & 20.3590 \end{pmatrix} \begin{pmatrix} 0.0513858a + 0.01425b + 0.000046219a^2 + \bar{\boldsymbol{i}}_0 \\ 0.0487051b + 0.000245691a^2 + \bar{\boldsymbol{j}}_0 \end{pmatrix}$$

$$= \begin{pmatrix} 0.781070a + 0.216602b + 0.000702534a^2 + [-0.00125457, 0.00162662] \\ 0.991586b + 0.00500202a^2 + [-0.00381146, 0.00341165] \end{pmatrix}.$$

Remark 6.1. From a mathematical viewpoint, modification of the domain or of the polynomials are equivalent approaches for factorizing preconditioned Taylor models, but maintaining the integration domain via the scaling matrices is advantageous for the software implementation of the method, because it simplifies the estimation of the higher order terms in the integration step.

In the second integration step, we use the initial set defined by $\mathcal{U}_{l,1}$ and $\mathcal{V}_{l,1}$. Proceeding as before, we obtain the integrated left Taylor models (for a, $b \in [-1, 1]$)

$$\widetilde{\mathcal{U}}_{l,2}(a, b) := 0.817551 + 0.0664561a - 0.00433580b + \widetilde{\boldsymbol{i}}_1,$$

$$\widetilde{\mathcal{V}}_{l,2}(a, b) := -0.835195 + 0.0233831a + 0.0471479b \\ +0.000418026a^2 - 0.000117424ab + 0.00000824612b^2 + \widetilde{\boldsymbol{j}}_1,$$

where

$$\widetilde{\boldsymbol{i}}_1 = [-5.72276\text{E-}5, 9.15947\text{E-}5], \quad \widetilde{\boldsymbol{j}}_1 = [-1.80914\text{E-}4, 1.80850\text{E-}4].$$

Finally, the flow at t_2 is made up by the composition of the integrated left Taylor models and the previous right Taylor models. We have

$$\mathcal{U}_2(a, b) := \widetilde{\mathcal{U}}_{l,2}(\mathcal{U}_{r,1}(a, b), \mathcal{V}_{r,1}(a, b)) = 0.817551 + 0.0519069a + 0.0100952b + 0.000025a^2 \\ + [-3.48708\text{E-}4, 4.09534\text{E-}4],$$

$$\mathcal{V}_2(a, b) := \widetilde{\mathcal{V}}_{l,2}(\mathcal{U}_{r,1}(a, b), \mathcal{V}_{r,1}(a, b)) = -0.835195 + 0.0182638a + 0.0518160b + 0.000507287a^2 \\ -0.0000505ab - 0.0000025b^2 + [-4.38606\text{E-}4, 4.28392\text{E-}4],$$

where a, $b \in [-1, 1]$.

Algorithm 6.1 (QR preconditioned Taylor model method)

Let the initial set be given as a preconditioned Taylor model vector $\mathcal{U}_{l,0} \circ \mathcal{U}_{r,0}$ in m space variables, with $\mathcal{U}_{r,0}$ the identity map and symbolic variables in $[-1, 1]$.

For $j := 0, 1 \ldots, j_{\max} - 1$:

1. Integrate $\mathcal{U}_{l,j}$ (containing the flow of the IVP at t_j) as in the naive Taylor model method. Denote the integrated left Taylor model (containing the flow of the IVP at t_{j+1}) by $\widetilde{\mathcal{U}}_{l,j+1}$. The flow is also contained in $\widetilde{\mathcal{U}}_{l,l+1} \circ \mathcal{U}_{r,j}$.
2. Replace $\widetilde{\mathcal{U}}_{l,j+1} \circ \mathcal{U}_{r,j}$ by $\mathcal{U}_{l,j+1} \circ \mathcal{U}_{r,j+1}$:
   - (i) Compute the QR factorization of the linear part of $\widetilde{\mathcal{U}}_{l,j+1}$.
   - (ii) Shift all but the constant part of $\widetilde{\mathcal{U}}_{l,j+1}$ to $\mathcal{U}_{r,j}$. Make Q the linear part of $\widetilde{\mathcal{U}}_{l,j+1}$. Apply Q^{-1} on $\mathcal{U}_{r,j}$.
   - (iii) Bound the range of the new $\mathcal{U}_{r,j}$.
   - (iv) Apply a scaling matrix S_{j+1} on $\mathcal{U}_{r,j}$ such that each component of the range of $\mathcal{U}_{r,j+1} := S_{j+1}^{-1} \circ \mathcal{U}_{r,j}$ is contained in $[-1, 1]$ and spans $[-1, 1]$ approximately.
   - (v) Set $\mathcal{U}_{l,j+1} := \widetilde{\mathcal{U}}_{l,j+1} \circ S_{j+1}$.

Compared with the naive Taylor model integration performed in Section 4.1, the polynomial coefficients are identical except for roundoff errors. This does not invalidate the computations, since all roundoff errors are rigorously bounded by the interval terms. Even though preconditioned integration is the superior method with respect to accuracy in the long run, the interval terms after two integration steps are larger here. The advantage of preconditioning becomes only apparent after several integration steps (see Section 6.1). Algorithm 6.1 summarizes the preconditioned Taylor model method with domain normalization.

6.1. Numerical Comparison with the QR Interval Method. Finally, we compare the performance of Lohner's software AWA [21] with the COSY Infinity integrator written by Makino. We use the quadratic model IVP (4.1) for the comparison. For the computation, Taylor expansions of order 18 were used in both programs. In both programs, the QR method (QR preconditioning) is used. The computed enclosure sets are shown in Figure 6.1.

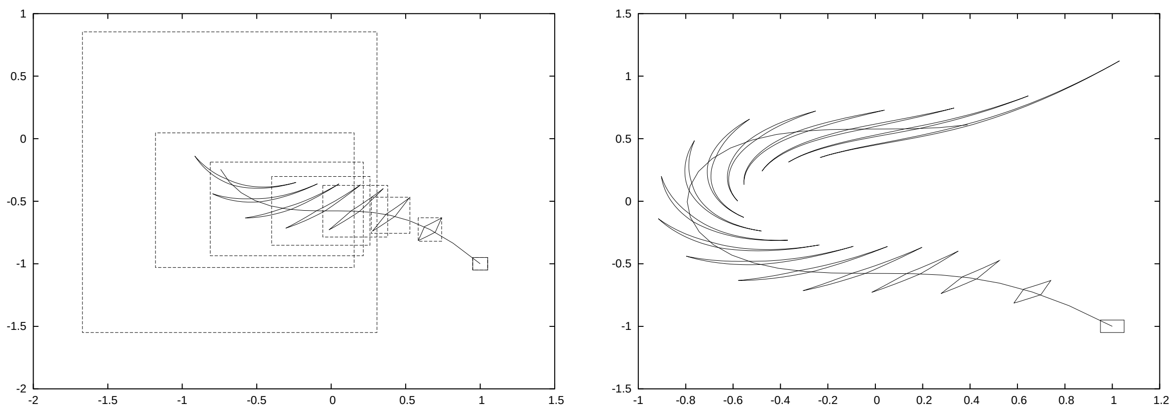

FIG. 6.1. *Integration of quadratic model IVP with AWA and COSY Infinity for* $t \in [0, 2.8]$ *(left), and with COSY Infinity for* $t \in [0, 6]$ *(right). Enclosures of the flow are shown for* $t_k = 0.4k$, $k = 0, 1, \ldots$. *The solid line in each picture belongs to the approximate solution that was computed with a Runge-Kutta method (for the model ODE with point initial values).*

In the left picture, integration is performed in the time interval $[0, 2.8]$. 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 non-convex enclosure sets of the flow.

This example shows that Taylor model methods may perform much better than interval methods on some problems, but this is not always the case. For some problems, interval methods can be as effective. Moreover, if they succeed, interval methods are often faster than Taylor model methods, because symbolic computations with multivariate polynomials are expensive.

7. Linear Numerical Examples. We compare interval methods and Taylor model methods for the linear autonomous ODE

$$u' = B\,u,$$

where B is a real 3×3 matrix. Numerical results are displayed for three different choices of B. In all examples, the initial values

$$\boldsymbol{u}_0 = \begin{pmatrix} [0.999, 1.001] \\ [0.999, 1.001] \\ [0.999, 1.001] \end{pmatrix}.$$

were used. The computations were performed with AWA and with the COSY Infinity integrator. In all examples, order 12 was chosen for the Taylor polynomial. Using lower orders (6 and 9 were tested) gave less accurate results, using higher orders (15 was tested) increased the computation times, but not the accuracy of the results. For integration with COSY Infinity, the minimal step size was set to 0.25.

In the tables, the following notation is used.

- AWA iv/AWA pe/AWA QR denote the direct interval method, the parallelepiped method and the QR method, respectively.
- TM na/TM sw/TM QR denote the naive Taylor model method without shrink wrapping, the naive Taylor model method with shrink wrapping, and the Taylor model method with QR preconditioning, respectively.

The observed performance of the methods is in agreement with the theoretical considerations in this paper. Naive Taylor model integration without shrink wrapping performs as the direct interval method (except for Example 1), naive Taylor model integration with shrink wrapping like the parallelepiped method, and QR preconditioned Taylor model integration similar to the QR method.

We call two matrices A and B *floating-point similar*, if A is obtained from B by a similarity transform executed in floating-point arithmetic. Floating-point similar matrices are denoted by $A \approx B$. Intervals are sometimes displayed using a short notation with upper and lower indexes. For example, 1.4^{7301}_{5593}E-001 denotes the interval [0.145593,0.147301].

Example 7.1. Pure Contraction.

$$B = \begin{pmatrix} -0.4375 & 0.0625 & -0.2651650429 \\ 0.0625 & -0.4375 & -0.2651650429 \\ -0.2651650429 & -0.2651650429 & -0.375 \end{pmatrix} \approx \begin{pmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & -\frac{3}{4} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

B has three distinct real eigenvalues, so that B describes a contraction without rotation. For such problems, the parallelepiped method is not well suited, because the matrices A_j, which have to be inverted, become nearly singular. The interval method breaks down, and the corresponding naive Taylor model method with shrink wrapping computes a practically useless enclosure of the solution.

Method	t_{end}	Steps	$y_1(t_{\text{end}})$
AWA iv	100	216	1.4^{7301}_{5593}E-001
AWA pe	52.6	131	aborted
AWA QR	100	216	1.4^{7301}_{5593}E-001
TM na	100	391	[−2.378E+26, 2.378E+26]
TM sw	100	272	[−2.282E+112, 2.282E+112]
TM QR	100	122	1.4^{7301}_{5593}E-001

Table 7.1. Numerical results for Example 7.1.

The direct interval method succeeds here. We also would have expected the naive Taylor model method without shrink wrapping to succeed. While the reason for its failure is not clear, it provides further evidence for our judgement that this method is not very effective. Both the QR interval method and the QR preconditioned Taylor model method succeed here.

Method	t	Steps	$y_1(t_{\text{end}})$
AWA iv	76.5	393	aborted
AWA pe	100	449	1.49^{522}_{222}E+000
AWA QR	100	449	1.49^{522}_{222}E+000
TM na	100	396	[−1.517E+45, 1.517E+45]
TM sw	100	343	1.49^{522}_{222}E+000
TM QR	100	343	1.49^{522}_{222}E+000

Table 7.2. Numerical results for Example 7.2.

Example 7.2. Pure Rotation.

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

B has eigenvalues $\pm i$ and 0. The flow of this IVP is a rotating interval box. As expected, the direct interval method and the naive Taylor model method fail, whereas the parallelepiped method and the naive Taylor model method with shrink wrapping (and also the QR based methods) succeed.

Example 7.3. Contraction and Rotation.

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

In our last example, B has eigenvalues $\pm i$ and $-1/2$, so contraction and rotation are combined. Here, the direct interval method and the naive Taylor model method are bound to fail because of the rotation, whereas the contraction causes the parallelepiped method and the Taylor model method with shrink wrapping to fail.

Method	t	Steps	$y_1(t_{\text{end}})$
AWA iv	85.5	507	aborted
AWA pe	75.2	404	aborted
AWA QR	100	516	1.34^{862}_{592}E+000
TM na	100	397	$[-1.605\text{E}+55, 1.605\text{E}+55]$
TM sw	100	357	$[-3.566\text{E}+106, 3.566\text{E}+106]$
TM QR	100	362	1.34^{862}_{592}E+000

Table 7.3. Numerical results for Example 7.3.

Only the QR based methods can successfully deal with both contraction and rotation of the initial set. For these methods, the overestimation of the final flow is hardly noticeable. This agrees with the general observation that the QR decomposition is a very effective tool in fighting the wrapping effect, both for the interval method and for the preconditioned Taylor model method.

Conclusion. We have compared traditional enclosure methods with Taylor model based integration. For the verified 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 increased use of Taylor model methods.

Acknowledgements. The authors thank Martin Berz and Kyoko Makino for several very helpful clarifying discussions on Taylor models, and for making the COSY Infinity integrator available. Our thanks also go to the referees for helpful comments.

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, 1997, pp. 1–23.
[3] ———, *Cosy Infinity Version 8 reference manual*, NSCL Technical Report MSUCL-1088, Michigan State University, 1998.
[4] ———, *COSY INFINITY*. `http://bt.pa.msu.edu/index_files/cosy.htm`, August 2006.
[5] M. Berz and G. Hoffstätter, *Computation and application of Taylor polynomials with interval remainder bounds*, Reliable Computing, 4 (1998), pp. 83–97.
[6] M. Berz and K. Makino, *Verified integration of ODEs and flows using differential algebraic methods on high-order Taylor models*, Reliable Computing, 4 (1998), pp. 361–369.
[7] ———, *New methods for high-dimensional verified quadrature*, Reliable Computing, 5 (1999), pp. 13–22.
[8] ———, *Performance of Taylor model methods for validated integration of ODEs*, Lecture Notes in Computer Science, 3732 (2006), pp. 65–73.
[9] G. F. Corliss, *Survey of interval algorithms for ordinary differential equations*, Appl. Math. Comput., 31 (1989), pp. 112–120.
[10] ———, *Guaranteed error bounds for ordinary differential equations*, in Theory and Numerics of Ordinary and Partial Differential Equations, M. Ainsworth, J. Levesley, W. A. Light, and M. Marletta, eds., Clarendon Press, Oxford, 1995.
[11] J.-P. Eckmann, H. Koch, and P. Wittwer, *A computer-assisted proof of universality in area-preserving maps*, Memoirs of the AMS, 47 (1984).
[12] P. Eijgenraam, *The Solution of Initial Value Problems Using Interval Arithmetic*, Mathematical Centre Tracts No. 144, Stichting Mathematisch Centrum, 1981.

[13] J. Hoefkens, M. Berz, and K. Makino, *Verified high-order integration of DAEs and higher-order ODEs*, in Scientific Computing, Validated Numerics and Interval Methods, W. Krämer and J. Wolff von Gudenberg, eds., Kluwer, Dordrecht, Netherlands, 2001, pp. 281–292.

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

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

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

[17] R. Klatte, U. Kulisch, Ch. Lawo, M. Rauch, and A. Wiethoff, *C-XSC: A C++ Class Library for Extended Scientific Computing*, Springer, Berlin, 1993.

[18] R. Klatte, U. Kulisch, M. Neaga, D. Ratz, and Ch. Ullrich, *Pascal-XSC – Language Reference with Examples*, Springer, Berlin, 1992.

[19] W. Kühn, *Rigorously computed orbits of dynamical systems without the wrapping effect*, Computing, 61 (1998), pp. 47–67.

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

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

[22] ———, *Computation of guaranteed solutions of ordinary initial and boundary value problems*, in Computational Ordinary Differential Equations, J. R. Cash and I. Gladwell, eds., Clarendon Press, Oxford, 1992, pp. 425–435.

[23] ———, *On the ubiquity of the wrapping effect in the computation of error bounds*, in Perspectives of Enclosure Methods, U. Kulisch, R. Lohner, and A. Facius, eds., Springer, Wien, 2001, pp. 201–217.

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

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

[26] ———, *COSY INFINITY version 8*, Nuclear Instruments & Methods in Physics Research A, 427 (1999), pp. 338–343.

[27] ———, *Efficient control of the dependency problem based on Taylor model methods*, Reliable Computing, 5 (1999), pp. 3–12.

[28] ———, *Taylor models and other validated functional inclusion methods*, Int. J. Pure Appl. Math., 4 (2003), pp. 379–456.

[29] ———, *Suppression of the wrapping effect by Taylor model-based validated integrators*, MSU Report MSUHEP 40910, Michigan State University, 2004.

[30] K. Makino, M. Berz, and Y. Kim, *Range bounding with Taylor models - some case studies*, WSEAS Transactions on Mathematics, 3 (2004), pp. 137–145.

[31] R. E. Moore, *The automatic analysis and control of error in digital computation based on the use of interval numbers*, in Error in Digital Computation, Vol. I, L. B. Rall, ed., John Wiley and Sons, New York, 1965, pp. 61–130.

[32] ———, *Automatic local coordinate transformations to reduce the growth of error bounds in interval computation of solutions of ordinary differential equations*, in Error in Digital Computation, Vol. II, L. B. Rall, ed., John Wiley and Sons, New York, 1965, pp. 103–140.

[33] ———, *Interval Analysis*, Prentice Hall, Englewood Cliffs, N.J., 1966.

[34] 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 (2004), pp. 27–44.

[35] N. S. Nedialkov, *Computing rigorous bounds on the solution of an IVP for an ODE*, PhD thesis, University of Toronto, 1999.

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

[37] ———, *Some recent advances in validated methods for IVPs for ODEs*, Appl. Numer. Math., 42 (2003), pp. 269–284.

[38] N. S. Nedialkov, K. R. Jackson, and G. F. Corliss, *Validated solutions of initial value problems for ordinary differential equations*, Appl. Math. Comput., 105 (1999), pp. 21–68.

[39] N. S. Nedialkov, K. R. Jackson, and J. Pryce, *An effective high-order interval method for validating existence and uniqueness of the solution of an IVP for an ODE*, Reliable Computing, 7 (2001), pp. 449–465.

[40] M. Neher, *Geometric series bounds for the local errors of Taylor methods for linear n-th order ODEs*, in Symbolic Algebraic Methods and Verification Methods, G. Alefeld, J. Rohn, S. Rump, and T. Yamamoto, eds., Springer, Wien, 2001, pp. 183–193.

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

[42] ———, *Taylor forms – use and limits*, Reliable Computing, 9 (2003), pp. 43–79.

[43] H. Ratschek and J. Rokne, *Computer Methods for the Range of Functions*, Ellis Horwood Limited, Chichester, 1984.

[44] R. Rihm, *Über Einschließungsverfahren für gewöhnliche Anfangswertprobleme und ihre Anwendung auf Differentialgleichungen mit unstetiger rechter Seite*, PhD thesis, Universität Karlsruhe, Germany, 1993.

[45] ———, *Interval methods for initial value problems in ODEs*, in Topics in Validated Computations, J. Herzberger, ed., Elsevier, Amsterdam, 1994, pp. 173–207.

[46] H. J. Stetter, *Validated solution of initial value problems for ODE*, Notes Rep. Math. Sci. Eng., 7 (1990), pp. 171–193.

[47] N. F. Stewart, *A heuristic to reduce the wrapping effect in the numerical solution of $x' = f(t, x)$*, BIT, 11 (1971), pp. 328–337.