Advances in Computational Mathematics **19:** 231–253, 2003.

Computing validated solutions of implicit differential equations *

Jens Hoefkens [a], Martin Berz [a] and Kyoko Makino [b]

[a] *Department of Physics and Astronomy, and National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, MI 48824-1321, USA*
E-mail: {hoefkens;berz}@msu.edu
[b] *Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801-3080, USA*
E-mail: makino@uiuc.edu

Received 16 November 2001; accepted 10 October 2002
Communicated by C.A. Micchelli

Ordinary differential equations (ODEs), including high-order implicit equations, describe important problems in mechanical and chemical engineering. However, the use of self-validated methods providing rigorous enclosures of the solution has mostly been limited to explicit and weakly nonlinear problems, and no general-purpose algorithm for the validated integration of general ODE initial value problems has been developed. Since most integration techniques for Differential Algebraic Equations (DAEs) are based on transformation to implicit ODEs, the integration of DAE initial value problems has traditionally been restricted to few hand-picked problems from the relatively small class of low-index systems. The recently developed Taylor model method combines high-order differential algebraic descriptions of functional dependencies with intervals for verification. It has proven its power in several applications, including verified integration of ODEs under avoidance of the wrapping effect. Recognizing antiderivation (integration) as a natural operation on Taylor models yields methods that treat DEs within a fully differential algebraic context as implicit equations made of conventional functions and antiderivation. This method has the potential to be applied to high-index DAE problems and allows the computation of guaranteed enclosures of final conditions from large initial regions for large classes of initial value problems. In the framework of this method, a Taylor model represents the highest derivative of the solution function occurring in the DE and all lower derivatives are treated as antiderivatives of this Taylor model. Consequently, one obtains a set of implicit equations involving only the highest derivative. Utilizing methods of verified inversion of functional dependencies described by Taylor models allows the computation of a guaranteed enclosure of the solution in the form of a Taylor model. The performance of the method is illustrated by detailed examples.

Keywords: differential algebraic equations, Taylor model, self-validated methods, interval methods

* Supported in part by the U.S. Department of Energy under contract DE-FG02-95ER40931 and the Alfred P. Sloan Foundation.

1. Introduction

Under certain conditions, the solutions of ordinary differential equations (ODEs) and differential algebraic equations (DAEs) can be expanded in Taylor series in both the independent variable and the initial conditions. In these cases, we can obtain good approximations of the solutions by computing the respective Taylor polynomials [14,15]. Utilizing the Taylor model approach as summarized in this paper often allows us to compute rigorous enclosures of the solutions of initial value problems [10,30].

By using index analysis [40,41] and computational differentiation, a given DAE can often be transformed into an equivalent system of implicit ODEs. If the derived system is described by a Taylor model, representing each derivative by an independent variable, verified inversion methods [7,22] can be utilized to solve for the highest derivatives as functions of lower order ones. The resulting Taylor model forms an enclosure of the right-hand side of an explicit ODE initial value problem that is equivalent to the original problem. While this explicit system is suitable for integration with Taylor model solvers [10], the intermediate inversion often requires a substantial increase in the dimensionality of the problem, limiting the approach to relatively small systems. An application of this inversion-based integration of differential equations has been discussed in [23].

In this paper we derive a method for the verified integration of general implicit ODEs that is based on the observation that solutions can be obtained as fixed points of a certain operator containing antiderivation. We show that this operator is particularly well suited for practical applications in Taylor model settings, since its restriction to Taylor polynomials is guaranteed to converge to the exact nth order expansion of the solution in at most $n + 1$ steps, where $n \in \mathbb{N}$ is the order of the Taylor models.

While sophisticated methods have been developed for the numerical integration of DAE problems [1,13,18,19], they are usually based on multistep methods and generally do not provide the possibility of verification and validation. However, since our proposed method of integrating implicit ODE initial value problems can also be used for the computation of the differentiation index ν_d of DAEs [1], it can be utilized for a verified index analysis. Combining this with methods of computational differentiation, we can obtain a scheme for transforming DAEs into implicit ODEs, which can be solved with the new method. Since this combination uses verified Taylor models at every stage of the computation, it can be used to compute Taylor model enclosures of the solutions of DAEs. By utilizing high order Taylor models ($n > 20$ is not uncommon), the scheme can even be applied to high-index problems that pose serious challenges to existing non-verified DAE integration methods.

2. Computational tools

In this section we summarize the theory behind the computational methods that form the basis of the new integration scheme. First we introduce an algebra with three operations, including the derivative operation (derivation). Such structures are generally called differential algebras [26,42,43] and we will focus mostly on ${}_nD_v$ which enables us to efficiently manipulate Taylor polynomials with floating point coefficients. Then, in section 2.2, we briefly discuss interval methods and summarize the Taylor model approach that combines high-order floating point polynomials with intervals. Taylor models provide self-validated computations with high-order convergence and can often outperform conventional intervals in terms of sharpness and computational cost.

2.1. The differential algebra ${}_nD_v$

Let $U \subset \mathbb{R}^v$ be an open set containing the origin and consider the space $\mathcal{C}^{n+1}(U, \mathbb{R}^w)$ of $(n+1)$-times continuously differentiable functions that map U into $\mathbb{R}^w$. We define the relation of *equality up to order n* as follows.

Definition 1. For $f, g \in \mathcal{C}^{n+1}(U, \mathbb{R}^w)$ we say that f equals g up to order n if $f(0) = g(0)$, and all partial derivatives of orders up to n agree at the origin. If f equals g up to order n, we denote that by $f =_n g$.

It is easy to see that equality up to order n establishes an equivalence relation on the space $\mathcal{C}^{n+1}(U, \mathbb{R}^w)$ [4]. The resulting equivalence classes are called *DA vectors*, and the class containing the function $f \in \mathcal{C}^{n+1}(U, \mathbb{R}^w)$ is denoted by $[f]_n$. The collection of these equivalence classes is called ${}_nD_v$. More details on this structure are given in [2,4].

Proposition 1. For $f \in \mathcal{C}^{n+1}(U, \mathbb{R}^w)$, the nth order Taylor polynomial $T_n(f)$ of f is contained in $[f]_n$.

This assertion follows easily from the basic definition of the equivalence classes. However, the fact that the nth order Taylor polynomial of f can be used as a representative for the class $[f]_n$ opens the door for a computer implementation of the structure ${}_nD_v$ by storing and manipulating the coefficients of Taylor polynomials.

2.1.1. Elementary operations and functions

Elementary operations like "$+$" and "$\times$" can be lifted from $\mathcal{C}^{n+1}(U, \mathbb{R}^w)$ in the usual way, and extend to the corresponding operations "$\oplus$" and "$\otimes$" on ${}_nD_v$ [2,3].

Definition 2. Let $f, g \in \mathcal{C}^{n+1}(U, \mathbb{R}^w)$ be two functions. Then the *sum* of the DA vectors $[f]_n$ and $[g]_n$ is given by

$$[f]_n \oplus [g]_n = [f + g]_n. \tag{1}$$

The *product* of the two DA vectors is defined by

$$[f]_n \otimes [g]_n = [f \times g]_n. \tag{2}$$

Together with the scalar multiplication $r \cdot [f]_n = [r \cdot f]_n$, this definition of the elementary operations makes ${}_nD_v$ an algebra [4] and provides a transparent extension to the equivalence classes of DA vectors; i.e., knowledge of the values and derivatives of f and g at the origin is sufficient to obtain Taylor polynomials of their sums and products. Moreover, in [2,3] the available operations on ${}_nD_v$ have been extended to include subtraction and, for a limited class of DA vectors, even the multiplicative inversion. From now on we omit the distinction between the operations on $\mathcal{C}^{n+1}(U, \mathbb{R}^w)$ and ${}_nD_v$ and will always use the same symbols "$+$" and "$\times$" for operations between numbers, functions, and DA vectors.

To fully utilize the differential algebra ${}_nD_v$, especially in numerical analysis and computer environments, it is necessary to extend the standard mathematical functions commonly available on computers to ${}_nD_v$: square root, exponential, logarithm, trigonometric and hyperbolic functions. Details on the exact definition and implementations of these functions have been presented in [4] and for the purposes of this paper it suffices to say that all computer functions have in fact been extended to ${}_nD_v$.

It has been shown that the derivative operation can be extended from $\mathcal{C}^{n+1}(U, \mathbb{R}^w)$ to the algebra ${}_nD_v$ in such a way that ${}_nD_v$ becomes a *differential algebra* [2,4]. While we will not use this intrinsic structure of ${}_nD_v$, we will make frequent use of the *antiderivation* of DA vectors to be presented in section 2.1.3.

2.1.2. Contracting operators and fixed points

In the previous section we have shown how elementary operations on the function space $\mathcal{C}^{n+1}(U, \mathbb{R}^w)$ are extended to the differential algebra ${}_nD_v$. Here we take a look at the more general concept of *operators* on ${}_nD_v$ and summarize an important fixed point theorem. The availability of this powerful fixed point theorem for operators on ${}_nD_v$ allows the use of the differential algebra ${}_nD_v$ in a large class of numerical applications, ranging from the analysis of dynamical systems [5,9,12] to global optimization [25,34].

Definition 3. For $[f]_n \in {}_nD_v$, the *depth* $\lambda([f]_n)$ is defined to be the order of the first non-vanishing derivative of f if $[f]_n \neq 0$, and $n + 1$ otherwise.

By definition of the equivalence classes, this definition is independent of the choice of $f \in [f]_n$. We note that any $a \in {}_nD_v$ with $\lambda(a) \geqslant 1$ satisfies the condition $a^{n+1} = 0$ and is therefore called *nilpotent*. Using the straightforward definition of the depth, contracting operators on ${}_nD_v$ are defined as follows.

Definition 4. Let $\mathcal{O}$ be an operator defined on $M \subset {}_nD_v$. $\mathcal{O}$ is *contracting* on M, if for any two $[f]_n, [g]_n \in M$,

$$\lambda\big(\mathcal{O}([f]_n) - \mathcal{O}([g]_n)\big) \geqslant \lambda\big([f]_n - [g]_n\big) \tag{3}$$

with equality if and only if $f =_n g$.

This definition has a striking similarity to the corresponding definitions on standard function spaces. Even more so, a theorem that resembles the Banach Fixed Point Theorem can be established on ${}_nD_v$. However, unlike in the case of the Banach Fixed

Point Theorem, in ${}_nD_v$ the sequence of iterates is guaranteed to converge in at most $n+1$ steps [4].

Theorem 1 (DA Fixed Point Theorem). Let $\mathcal{O}$ be a contracting operator defined on $M \subset {}_nD_v$ that maps M into itself (i. e., $\mathcal{O}(M) \subset M$). Then $\mathcal{O}$ has a unique fixed point $a \in M$. Moreover, for any $a_0 \in M$ it is $\mathcal{O}^{(n+1)}(a_0) = a$.

A proof and further discussion of the DA Fixed Point Theorem can be found in [4]. Here we just mention that, since $\lambda(a + b) \geqslant \min(\lambda(a), \lambda(b))$, it follows easily that the sum and composition of two contracting operators $\mathcal{O}_1$ and $\mathcal{O}_2$ defined on M is also a contracting operator.

2.1.3. Antiderivation

We conclude this section on functions on ${}_nD_v$ with an example of an operator that is unusual but, considering the structure of the differential algebra ${}_nD_v$, actually quite natural. The antiderivation; i.e., the integration with respect to any of the v variables, turns out to be a contracting operation on ${}_nD_v$.

Proposition 2 (Antiderivation is contracting). For $k \in \{1, \ldots, v\}$, the *antiderivation* $\partial_k^{-1} : {}_nD_v \to {}_nD_v$ is a contracting operator on ${}_nD_v$.

The proof of this important result is based on the fact that if $a, b \in {}_nD_v$ agree up to order l, the first non-vanishing derivative of $\partial_k^{-1}(a - b)$ is of order $l + 1$ [4,23,24]. It is important to realize that in the DA framework of ${}_nD_v$, antiderivation is easily computed and there is no fundamental difference between any of the standard mathematical functions and antiderivation. In fact, fully embracing antiderivation as a normal operation on DA vectors will enable us to develop a new and powerful method for the verified integration of ordinary differential equations (ODEs) and differential algebraic equations (DAEs) in the remainder of this paper.

2.2. Rigorous numerical analysis with Taylor models

The relative accuracy of floating point number representations on computers is limited, usually to around 16 decimal digits. Interval analysis takes this limitation into account and allows the computation of guaranteed enclosures for computational results: the result of every interval computation consists of two numbers; one is guaranteed to be smaller than the mathematically correct result, the other one is guaranteed to be larger than the correct result. Thus, if x_r is the analytical result of some computation $\mathcal{C}$, and $R = [x_l, x_u]$ is the corresponding result of interval analysis, then $x_r \in R$. On the other hand, interval analysis can also be seen as a computerization of set theory: starting with $I = [x_l, x_u]$ and some computation $\mathcal{C}$, the result $\mathcal{C}(I) = [y_l, y_u]$ contains the result of the computation $\mathcal{C}$ applied to *every* number contained in I. Further details on interval analysis can be found in [20,36,37].

Recently, Taylor models have been developed [30,31] as a combination of the high order Taylor polynomials described in the previous subsection and interval analysis to obtain guaranteed enclosures of functional descriptions. Details on applications of these methods in beam physics are given in [29,35]. It has been shown [32] that the Taylor model approach can often substantially alleviate the following problems inherent in conventional interval arithmetic:

- sharpness for large domain intervals,
- cancellation and dependency problem,
- dimensionality curse.

2.2.1. Taylor model methods

Commonly, we view Taylor models as sets of functions that are close to a reference polynomial as outlined below.

Definition 5. Let $\boldsymbol{D} \subset \mathbb{R}^v$ be a box with $x_0 \in \boldsymbol{D}$. Let $P : \boldsymbol{D} \to \mathbb{R}^w$ be a polynomial of order n and $R \subset \mathbb{R}^w$ be an non-empty convex compact set. Then $(P, x_0, \boldsymbol{D}, R)$ is called a Taylor model of order n with expansion point x_0 over $\boldsymbol{D}$.

Following these notations, P is called the *reference polynomial* and R is called the *remainder bound* of the Taylor model. We say that f is contained in a Taylor model $T = (P, x_0, \boldsymbol{D}, R)$ if $P(x) - f(x) \in R$ for all $x \in \boldsymbol{D}$ and the nth order Taylor series of f around x_0 equals P.

Methods have been developed to extend mathematical operations and functions to Taylor models such that the inclusion relationships are preserved. For example, for two given Taylor models, T_1 and T_2, the following theorem lays the foundation for the computation of Taylor models for the sum S, and the product P, of T_1 and T_2. In the context of this theorem $B(P, D)$ denotes a bound for the polynomial P over the compact set $\boldsymbol{D}$, $P_{(n)}$ stands for the truncation of P to order n, and $P_{(n+)}$ consists of the terms of degree larger than n in P.

Theorem 2. Let $T_1 = (P_1, x_0, \boldsymbol{D}, R_1)$ and $T_2 = (P_2, x_0, \boldsymbol{D}, R_2)$ be two Taylor models of order n and define

$$R_{\mathrm{P}} = R_1 \cdot R_2 + R_1 \cdot B(P_2, \boldsymbol{D}) + B(P_1, \boldsymbol{D}) \cdot R_2 + B\big((P_1 \cdot P_2)_{(n+)}, \boldsymbol{D}\big). \tag{4}$$

Obtain new Taylor models T_{S} and T_{P} by

$$T_{\mathrm{S}} = (P_1 + P_2, x_0, \boldsymbol{D}, R_1 + R_2), \tag{5}$$

$$T_{\mathrm{P}} = \big((P_1 \cdot P_2)_{(n)}, x_0, \boldsymbol{D}, R_{\mathrm{P}}\big). \tag{6}$$

Then, T_{S} and T_{P} are Taylor models for the sum T_{S} and product T_{P} of T_1 and T_2. In particular, for two functions $f_1 \in T_1$ and $f_2 \in T_2$, it is

$$(f_1 + f_2) \in T_{\mathrm{S}} \quad \text{and} \quad (f_1 \cdot f_2) \in T_{\mathrm{P}}. \tag{7}$$

Proof. If we define $\mathcal{C}^{n+1}$ functions $\delta_1 = f_1 - P_1$ and $\delta_2 = f_2 - P_2$, then $\delta_1(x) \in R_1$ and $\delta_2(x) \in R_2$ for any $x \in \boldsymbol{D}$. Thus, for a given $x \in \boldsymbol{D}$

$$\big((f_1 + f_2) - (P_1 + P_2)\big)(x) = \delta_1(x) + \delta_2(x) \in R_1 + R_2 = R_{\mathrm{S}}. \tag{8}$$

Since the nth order Taylor expansion of the sum $f_1 + f_2$ equals the sum of the Taylor series, T_{S} is indeed a Taylor model for the sum $T_1 + T_2$. Also, over the domain $\boldsymbol{D}$

$$\begin{aligned}\big((f_1 \cdot f_2) - (P_1 \cdot P_2)_{(n)}\big) &= \big((P_1 + \delta_1)(P_2 + \delta_2)\big) - \big((P_1 \cdot P_2) - (P_1 \cdot P_2)_{(n+)}\big) \\ &= P_1 \cdot \delta_2 + \delta_1 \cdot P_2 + \delta_1 \cdot \delta_2 + (P_1 \cdot P_2)_{(n+)}.\end{aligned} \tag{9}$$

Moreover, since the nth order Taylor polynomial of $f_1 \cdot f_2$ equals the polynomial product $(P_1 \cdot P_2)_{(n)}$, T_{P} is a Taylor model for the product $T_1 \cdot T_2$. □

Similarly, other elementary operations and intrinsic functions can been extended to Taylor models such that the fundamental inclusion properties of Taylor models are maintained. For example, the exponential function of Taylor models has been defined in such a way that for a given Taylor model T and its Taylor model exponential T_{E}

$$f \in T \Rightarrow \exp(f) \in T_{\mathrm{E}} \tag{10}$$

Further information on arithmetic and intrinsic functions on Taylor models can be found in [8,31,32].

Lastly, antiderivation, which is essentially the integration operation, extends naturally to Taylor models. It allows us to compute a Taylor model T_{I} from a given Taylor model T such that T_{I} contains primitives for each function f contained in T. Since the antiderivation does not fundamentally differ from other intrinsic functions on the set of Taylor models, it is often used in fundamental Taylor model algorithms. An important application of the antiderivation will be presented in section 2.2.2.

A consequence of the results summarized in this subsection is that we can compute Taylor models for any sufficiently smooth algorithm or computer function. This result follows directly by finite induction from theorem 2 and similar results for intrinsic functions and the antiderivation:

Theorem 3. Let $f : \boldsymbol{D} \subset \mathbb{R}^v \to \mathbb{R}^w$ be a sufficiently smooth function that is computable by a finite number of elementary operations, intrinsic functions, and antiderivations. Then, starting with Taylor models for the identity functions $x = (x_1, \ldots, x_v) \in \boldsymbol{D} \mapsto x_k$, a guaranteed enclosure of f can be obtained as a Taylor model by evaluating the code list of f with Taylor model operations.

We close this introduction to Taylor model methods by noting that a major advantage of Taylor model methods over regular interval analysis is that the sharpness of the enclosures obtained by Taylor models scales with the $(n + 1)$st order of the domain size [7]. Thus, the Taylor model approach is of particular advantage in the combination of Taylor models with high-order map codes like COSY Infinity [9,11] and gives small guaranteed enclosures even for complicated high-order maps in many variables.

2.2.2. Verified enclosures of flows

One of the most important applications of Taylor models stems from their applicability to numerical ODE integration. To illustrate the method, we consider the initial value problem

$$x' = f(t, x) \quad \text{and} \quad x(t_0) = x_0. \tag{11}$$

It is a well-known fact that the solution to (11) can be obtained as the fixed point of the Picard operator $\mathcal{O}$, defined by

$$\mathcal{O}(x) = x_0 + \int_{t_0}^{t} f(\tau, x)\,\mathrm{d}\tau. \tag{12}$$

Using the previously mentioned antiderivation, the operator $\mathcal{O}$ can be extended to Taylor models and yields an algorithm that allows the computation of verified *enclosures of flows* $\mathcal{M}(t, t_0, x_0)$ of (11). This approach has been presented in [10,30] and has recently been used successfully in a variety of applications ranging from solar system dynamics [12,21] to beam physics. Unlike methods that use conventional interval computations to enclose the final conditions, the Taylor model approach avoids the wrapping effect to very high order and is therefore capable of propagating extended initial regions over large integration intervals.

3. Validated integration of general differential equations

While sophisticated general-purpose methods for the verified integration of explicit first order ODEs have been developed [10,27,28,30,38,39], none of these can be readily used for the verified integration of implicit ODEs, let alone differential algebraic equations. As a first step toward the verified integration of general DAEs, we consider the problem of integrating the implicit ODE initial value problem $F(t, x, x') = 0$. Unlike in standard ODE integration methods, we will not limit ourselves to first order problems, but will eventually be able to integrate the arbitrary order problem

$$F\bigl(t, x, x', \ldots, x^{(p)}\bigr), \tag{13}$$

where the derivative of F with respect to the highest derivatives in each of the components is assumed to be nonsingular for all argument values in an appropriate domain.

Our integration method is based on an extended use of antiderivation. To motivate the combination of high order methods and antiderivation for the verified integration of general differential equations, consider the explicit second order ODE initial value problem

$$x'' = f(x, x', t), \quad x(t_0) = x_0,\ x'(t_0) = x'_0. \tag{14}$$

While the conventional approach to solving this system is based on order reduction to a two-dimensional first order problem, in the framework of Taylor models we can use the intrinsic antiderivation and substitute $\xi = x'$; i.e.,

$$x(t) = x_0 + \int_{t_0}^{t} \xi(\tau)\,\mathrm{d}\tau. \tag{15}$$

After inserting the expanded expression for $x(t)$ into (14), we obtain an explicit expression for ξ' in the form of a relatively simple integro-differential equation.

$$\xi' = F(\xi, t) = f\left(x_0 + \int_{t_0}^{t} \xi(\tau)\,\mathrm{d}\tau, \xi, t\right). \tag{16}$$

This system can readily be integrated with the existing Taylor model based integration scheme discussed in [10,30], and the solution $x(t)$ of (14) can be computed from the Taylor model for $\xi(t)$ by a final application of antiderivation.

$$x = x_0 + \int_{t_0}^{t} \xi(\tau)\,\mathrm{d}\tau. \tag{17}$$

While this approach seems obvious from a theoretical point of view, numerical analysis has traditionally avoided explicit references to antiderivation in its algorithms. And while symbolic tools make sophisticated use of antiderivation, they often suffer from large memory requirements and slow computations. However, differential algebraic frameworks [26,43] make the explicit use of antiderivation in numerical computations feasible by allowing fast computations with moderate memory requirements. From a practical point of view, the approach of using antiderivation has the advantage of reducing the computational complexity of the right-hand side of the ODE (14), often allowing for favorable computations.

3.1. High order Taylor model solutions

In this section we present a Taylor model-based algorithm for the self-validated integration of the general first order ODE initial value problem

$$F(t, x, x') = 0, \quad x(t_0) = x_0. \tag{18}$$

While we will later extend the algorithm to higher order ODEs, for now we assume that the problem is stated as an implicit first order system with an $(n+2)$-times continuously differentiable v-dimensional function F and regular Jacobian matrix

$$\frac{\partial F(t, u, v)}{\partial v} \tag{19}$$

in appropriate domains.

3.1.1. Taylor model integration of implicit ODEs

For notational convenience, we write the initial condition of the implicit ODE problem (18) as x_0, and denote functional dependence on the initial condition x_0 in a suitable domain by y. With these conventions, a single nth order integration step of the implicit first order ODE (18) consists of the following substeps:

1. Using a suitable numerical method like the Newton method discussed in [7,21], solve the implicit system
$$F(t_0, y, x') = 0 \tag{20}$$
for a consistent initial condition $x'(t_0, y) = x_0'(y)$.

2. Utilizing antiderivation, rewrite the original problem in a derivative-free form
$$\Phi(t, y, \xi) = F\left(t, y + \int_{t_0}^{t} \xi(\tau, y)\, \mathrm{d}\tau, \xi\right) = 0, \tag{21}$$
where $\xi = \xi(t, y) = x'(t, y)$ has been substituted for the derivative of x.

3. Translate the problem into an origin-preserving form by shifting to relative coordinates via the substitution $\zeta(t, y) = \xi(t, y) - x_0'(y)$. This defines the new function
$$\Psi(t, y, \zeta) = \Phi\bigl(t, y, \zeta(t, y) + x_0'(y)\bigr). \tag{22}$$

4. Since $\Psi(t_0, x_0, 0) = 0$, within the differential algebraic framework it is possible to write the first order truncation of Ψ without the constant part as
$$\Psi(t, y, \zeta) =_1 L_\zeta(\zeta) + L_R(t, y), \tag{23}$$
where L_ζ and L_R denote the linear parts in ζ and (t, y), respectively.

5. If L_ζ is regular, transform the previous expression into an equivalent fixed point formulation for ζ.
$$\zeta(t, y) = \mathcal{H}(\zeta) = -L_\zeta^{-1}\bigl(\Psi(t, y, \zeta) - L_\zeta(\zeta)\bigr). \tag{24}$$

6. Using the operator $\mathcal{H}$, define a sequence (a_ν) of DA vectors in ${}_nD_{1+v}$ by $a_0 = 0$ and
$$a_{\nu+1} = \mathcal{H}(a_\nu). \tag{25}$$
Then define the polynomial $P(t, y) = a_{n+1}$. This polynomial is the exact nth order expansion of the solution.

7. Construct a Taylor model T with the reference polynomial P over an appropriate domain $\boldsymbol{T} \times \boldsymbol{D} \subset \mathbb{R} \times \mathbb{R}^v$ containing the reference point (t_0, x_0) such that
$$\mathcal{H}(T) \subset T. \tag{26}$$

8. Compute a Taylor model X from T by using the relation
$$x(t, y) = y + \int_{t_0}^{t} \bigl(\zeta(\tau, y) + x_0'(y)\bigr)\, \mathrm{d}\tau. \tag{27}$$

Utilizing DA vectors and Taylor model methods for the verified integration of initial value problems allows the propagation of initial conditions by not only expanding the solution in time, but also in the transverse variables [10]. By representing the initial conditions as additional DA variables, their dependence can be propagated through the integration process, allowing Taylor model based integration schemes to reduce the wrapping effect to high orders [33]. Thus, the final Taylor model X is, in fact, a Taylor model for the flow $x(t, y)$ of the original ODE problem (18) for the particular choice of the initial derivative.

3.2. *Mathematical background*

The algorithm presented in the previous subsection rests on several nontrivial assertions that will be proven here. We provide its mathematical foundation and establish the basis for the discussion in the next subsection. The proofs can be split into two groups: the differential algebraic part of the algorithm and the Taylor model results. Implementation and user interface issues of the algorithm will be discussed in the next section.

3.2.1. *Differential algebraic results*

In this section we discuss the differential algebraic results needed to justify the presented algorithm for the verified integration of implicit ODEs. First, we show that the operator $\mathcal{H}$ introduced above is well defined and DA-contracting. We then show that up to an additive constant, its unique fixed point lies in the same equivalence class as the derivative of the flow of the original ODE problem.

Lemma 1. The operator $\mathcal{H}$ given by (24) is a contracting operator that maps the set $M = \{a \in {}_nD_{1+v} \mid \lambda(a) > 0\}$ into itself.

Proof. To first order regularity of L_ζ is equivalent to the assumed regularity of the Jacobian $\partial F(t, u, v)/\partial v$.

$$\frac{\partial \Psi}{\partial \zeta}(t_0, y_0, \zeta) = \frac{\partial \Phi}{\partial \zeta}(t_0, y_0, \zeta + x_0') =_1 \frac{\partial F}{\partial \zeta}(t_0, x_0, \zeta). \tag{28}$$

Thus by assumption, the linear map L_ζ is regular in a neighborhood of the initial conditions and the operator $\mathcal{H}$ is therefore well defined on all of ${}_nD_{1+v}$.

To show that $\mathcal{H}$ is contracting on the subset M of nilpotent DA vectors, let $a, b \in M$ be given and assume that a and b agree up to order k. Since L_ζ^{-1} is invertible, it suffices to show that

$$\lambda\Big(\big(\Psi(t_0, y_0, a) - L_\zeta(a)\big) - \big(\Psi(t_0, y_0, b) - L_\zeta(b)\big)\Big) > k. \tag{29}$$

Since $\Psi(t_0, x_0, 0) = 0$, the map $\Psi(t_0, x_0, \zeta)$ is origin-preserving and can be written as

$$\Psi(t_0, x_0, \zeta) = L_\zeta(\zeta) + L_R(t_0, x_0) + \mathcal{N}(t_0, x_0, \zeta), \tag{30}$$

where $\mathcal{N}$ is a purely nonlinear function. Thus, it suffices to show that $\mathcal{N}$ is contracting. However, if a and b agree up to order k, their images $\mathcal{N}(t_0, x_0, a)$ and $\mathcal{N}(t_0, x_0, b)$ trivially agree up to order $k + 1$. Finally, since Ψ is origin-preserving, it is indeed $\mathcal{H}(M) \subset M$, and therefore a contracting operator on M. □

While this lemma guarantees the existence of a unique fixed point of the operator $\mathcal{H}$, the next theorem summarizes the main result of the DA part of the presented algorithm.

Theorem 4. If we denote the flow of the implicit first order ODE initial value problem (18) by $x(t, y)$, then the fixed point of $\mathcal{H}$ is a representative for

$$\left[x'(t, y) - x_0'(y)\right]_n. \tag{31}$$

around the expansion point (t_0, x_0).

Proof. In principle, this assertion follows from the construction of the operator $\mathcal{H}$. However, in the following we summarize observations on existence, uniqueness, and smoothness of the solutions to the IVP (18) that are required for a full justification.

Since the original function F is of class $\mathcal{C}^{n+2}$ and its Jacobian matrix is assumed to be nonsingular over a suitable region containing (t_0, x_0, x_0'), at least one solution to the initial value problem exists. Moreover, once a consistent initial derivative x_0' has been fixed, the Inverse Function Theorem guarantees the existence of a unique solution $x(t, y)$ in a neighborhood of (t_0, x_0, x_0'). The solution is the unique solution to an explicit first order system, which can in principle be derived from the original implicit problem. Since the Inverse Function Theorem guarantees that the explicit system is also of class $\mathcal{C}^{n+2}$ in a neighborhood of the consistent initial conditions, smooth dependence on initial conditions ensures that the flow $x(t, y)$ is a $\mathcal{C}^{n+2}$ function of its variables. Thus, the equivalence class of its derivative is well defined in $\mathcal{C}^{n+1}$, and since the solution ζ is unique, the fixed point of $\mathcal{H}$ is indeed a representative for that class. □

3.2.2. *Taylor model results*

In this section we prove the main Taylor model result needed for the presented algorithm: the self-inclusion of the Taylor model in (26) is a sufficient condition to guarantee the enclosure of a fixed point of (24).

Definition 6. Let $T = (P, x_0, \boldsymbol{D}, R)$ be an nth order Taylor model and let $L > 0$ be a Lipschitz constant for P over $\boldsymbol{D}$. Then the L-Taylor model T_L is the set of all functions $f \in \mathcal{C}^0(\boldsymbol{D}, \mathbb{R}^w)$ such that

1. $f(x) - P(x) \in R$ for all $x \in \boldsymbol{D}$,
2. $f(x_0) = P(x_0)$,
3. $|f(x) - f(y)| \leqslant L|x - y|$ for all $x, y \in \boldsymbol{D}$.

Since $P \in T_L$, the set T_L is non-empty. Moreover, while there is an obvious connection to normal Taylor models, L-Taylor models contain a different set of functions:

1. Taylor models are “more restrictive” than the corresponding L-Taylor models, since the latter may contain nondifferentiable functions.
2. Taylor models are also “less restrictive” than the corresponding L-Taylor models, since they do not pose any limits on the Lipschitz constant of its members.

Most importantly though, unlike standard Taylor models, L-Taylor models are subsets of a Banach space, namely $\mathcal{C}^0(\boldsymbol{D}, \mathbb{R}^w)$.

Lemma 2. The L-Taylor model T_L defined as above is a convex subset of the Banach space $\mathcal{C}^0(\boldsymbol{D}, \mathbb{R}^w)$ [30].

Proof. Given two functions f_0 and f_1 in T_L and $t \in [0, 1]$, define $f_t = t \cdot f_1 + (1-t) \cdot f_0$. Since the sum of continuous functions is continuous, $f_t \in \mathcal{C}^0(\boldsymbol{D}, \mathbb{R}^w)$ for any $t \in [0, 1]$. Moreover, by convexity of R

$$f_t(x) - P(x) = \big(t \cdot f_1(x) + (1-t) \cdot f_0(x)\big) - P(x) \in R \tag{32}$$

for any $t \in [0, 1]$. Since

$$\begin{aligned} \big|f_t(x) - f_t(y)\big| &= \big|t \cdot f_1(x) + (1-t) \cdot f_2(x) - t \cdot f_1(y) - (1-t) \cdot f_2(y)\big| \\ &\leqslant t \cdot \big|f_1(x) - f_1(y)\big| + (1-t) \cdot \big|f_0(x) - f_0(y)\big| \\ &\leqslant t \cdot L \cdot |x-y| + (1-t) \cdot L \cdot |x-y| = L \cdot |x-y|, \end{aligned} \tag{33}$$

T_L is indeed a convex subset of the Banach space $\mathcal{C}^0(\boldsymbol{D}, \mathbb{R}^w)$. □

Lemma 3. The L-Taylor model T_L defined as above is a compact subset of the Banach space $\mathcal{C}^0(\boldsymbol{D}, \mathbb{R}^w)$ [30].

Proof. It suffices to show that every sequence (f_ν) in T_L has at least one limit point in T_L. According to the Arzela–Ascoli Theorem it suffices to show that (f_ν) is uniformly bounded and equicontinuous.

We will first show that the sequence is uniformly bounded. Since P is a polynomial and therefore continuous on $\boldsymbol{D}$, $|P|$ assumes its finite maximum M_1 over $\boldsymbol{D}$ at some point $\tilde{x}$:

$$M_1 = \big|P(\tilde{x})\big| = \max\big\{\big|P(x)\big| \mid x \in \boldsymbol{D}\big\}. \tag{34}$$

Since R is bounded, there is a constant $M_2 > 0$ such that

$$y \in R \;\Rightarrow\; |y| \leqslant M_2. \tag{35}$$

After defining $\delta_\nu = f_\nu - P$, $\delta_\nu(x) \in R$ for all $x \in \boldsymbol{D}$. Thus for any $x \in \boldsymbol{D}$ and any $\nu \in \mathbb{N}$

$$\big|f_\nu(x)\big| \leqslant \big|P(x)\big| + \big|\delta_\nu(x)\big| < M_1 + M_2 < \infty. \tag{36}$$

Hence, the sequence is indeed uniformly bounded. Moreover, since by definition the sequence is also Lipschitz with uniform Lipschitz constant L that is independent of ν, it follows that it is also equicontinuous.

Thus, the L-Taylor model T_L is a compact subset of the Banach space $\mathcal{C}^0(\boldsymbol{D}, \mathbb{R}^w)$. □

The main result of this paper is that the self-inclusion given in (26) is indeed sufficient to guarantee that the fixed point of $\mathcal{H}$ is contained in the Taylor model T. The proof of this assertion is based on the next theorem, which is an almost immediate consequence of the previous two lemmas.

Theorem 5. Let $\boldsymbol{T}$ and $\boldsymbol{D}$ be interval domains containing the points t_0 and x_0, respectively. Let $L > 0$ be a Lipschitz constant for P and define the L-Taylor model

$$T_L = \big(P, (t_0, x_0), \boldsymbol{T} \times \boldsymbol{D}, R\big). \tag{37}$$

If $\mathcal{H}(T_L) \subset T_L$, then the L-Taylor model T_L contains a fixed point of $\mathcal{H}$.

Proof. Since $\mathcal{H}$ is a continuous operator on $\mathcal{C}^0(\boldsymbol{T} \times \boldsymbol{D}, \mathbb{R}^v)$, the assertion follows from the last two lemmas and the Schauder Fixed Point Theorem. □

According to this theorem, a fixed point of $\mathcal{H}$ is contained in the L-Taylor model T_L. However, as indicated earlier, one of the fixed points is, in fact, equal to $x'(t, y) - x_0'(y)$ and is therefore of class $\mathcal{C}^{n+1}$. Moreover, according to theorem 4, the nth order Taylor expansion of the fixed point around (t_0, x_0) equals P. Thus, if $|\mathcal{H}| < 1$, the fixed point is also contained in the regular Taylor model

$$T = \big(P, (t_0, x_0), \boldsymbol{T} \times \boldsymbol{D}, R\big). \tag{38}$$

Since the reference polynomial P is already a very good approximation of the mathematically correct fixed point, in practice the self-inclusion of the image can almost always be achieved by choosing sufficiently small domains $\boldsymbol{T}$ and $\boldsymbol{D}$ around the reference points t_0 and x_0. While this can guarantee the inclusion in terms of the remainder bound R (i.e., the maximum norm), this leaves us with the task of proving the contractivity of $\mathcal{H}$. Obviously, for most operators this assertion would be difficult to satisfy. However, several peculiarities of $\mathcal{H}$ and its domain T help in ensuring that it meets the given requirements: $\mathcal{H}$ is purely nonlinear and it operates on T, which is a set of origin-preserving functions defined on small domains around zero. Thus, with the exception of the most extreme cases, contractivity of $\mathcal{H}$ over the Taylor model T is not a fundamental limitation of the method. Moreover, Taylor model methods can be used [16] to prove the contractivity of $\mathcal{H}$ over the Taylor model T on a computer, eliminating the need for a manual analysis of $\mathcal{H}$.

3.3. Higher order systems

Similar to the example presented in (14), the new method also allows the direct integration of higher order implicit ODE initial value problems without explicit order reduction. While this approach does not reduce the dimensionality of the Taylor models if dependence on initial conditions is desired, it can however improve the sharpness of the final remainder bounds. Since the magnitude of the time domain $\boldsymbol{T}$ is generally smaller than one, the definition of the Taylor model antiderivation ensures that the remainder bounds of the actual solution will often be smaller than the ones of the computed highest derivative.

To illustrate how the algorithm can be adapted to higher-order ODEs, consider the general second order implicit ODE initial value problem

$$G(t, x, x', x'') = 0, \quad x(t_0) = x_0,\ x'(t_0) = x'_0. \tag{39}$$

If we assume that the Jacobian matrix $\partial G(t, u, v, w)/\partial w$ is nonsingular in a suitable domain, this can be written as

$$\Phi(t, \xi) = G\left(x_0 + \int_{t_0}^{t}\left(x'_0 + \int_{t_0}^{\tau} \xi(\sigma)\,\mathrm{d}\sigma\right)\mathrm{d}\tau, x'_0 + \int_{t_0}^{t} \xi(\tau)\,\mathrm{d}\tau, \xi, t\right) = 0, \tag{40}$$

and the algorithm works with only minor adjustments. Similar arguments can be made for more general higher-order ODEs. The performance of this approach will be illustrated in the next section with the direct integration of a second order system.

3.3.1. A second order example

To illustrate how the new algorithm can be used for the direct integration of higher-order problems and to demonstrate how the method works in practice, consider the implicit second order ODE initial value problem

$$\mathrm{e}^{x''} + x'' + x = 0, \tag{41a}$$

$$x(0) = x_0 = 1, \tag{41b}$$

$$x'(0) = x'_0 = 0. \tag{41c}$$

While the demonstration in this section uses explicit algebraic transformations for illustrative purposes, it is important to note that the actual implementation uses the DA framework and does not rely on such explicit manipulations. We also mention that for purposes of keeping the exposition transparent, in this example we do not expand the solution in the transverse variables.

1. Compute a consistent initial value for $x''_0 = x''(0)$ such that $\mathrm{e}^{x''_0} + x''_0 + x_0 = 0$. A simple interval Newton method, with a starting value of 0, finds an enclosure of the unique solution $x''_0 = -1.278464542761074$ in just a few steps.

2. Rewrite the original ODE in a derivative-free form by substituting $\xi = x''$.

$$\Phi(\xi, t) = \mathrm{e}^{\xi(t)} + \xi(t) + \left(x_0 + \int_0^t\left(x'_0 + \int_0^{\tau} \xi(\sigma)\,\mathrm{d}\sigma\right)\mathrm{d}\tau\right) = 0. \tag{42}$$

3. Define the new dependent variable ζ as the relative distance of ξ to its consistent initial value and substitute $\zeta = \xi - x_0''$ in Φ to obtain the new function Ψ given by

$$\Psi(\zeta, t) = \zeta + x_0'' + \mathrm{e}^{x_0''}\mathrm{e}^{\zeta} + 1 + \frac{x_0''}{2}t^2 + \int_0^t \int_0^\tau \zeta(\sigma)\,\mathrm{d}\sigma\,\mathrm{d}\tau = 0. \tag{43}$$

4. The linear part $L_\zeta(\zeta)$ of Ψ is $1 + \mathrm{e}^{x_0''}$ where the 1 is the constant coefficient and $\mathrm{e}^{x_0''}$ results from the linear part of the exponential function e^{ζ}.

5. With L_ζ from the previous step, the solution ζ is a fixed point of the DA-contracting operator $\mathcal{H}$ defined by

$$\mathcal{H}(\zeta) = \frac{1}{1 + \mathrm{e}^{x_0''}}\left(\mathrm{e}^{x_0''}\left(\zeta - \mathrm{e}^{\zeta}\right) - x_0'' - 1 - \frac{x_0''}{2}t^2 - \int_0^t \int_0^\tau \zeta(\sigma)\,\mathrm{d}\sigma\,\mathrm{d}\tau\right). \tag{44}$$

   Since antiderivation raises the order by one and the linear parts have been subtracted from Ψ, $\mathcal{H}$ is a purely nonlinear operator in ζ. Thus, $\mathcal{H}$ is indeed DA-contracting as defined in definition 4.

6. Since $\mathcal{H}$ is DA-contracting, starting with an initial value of $\zeta^{(0)} = 0$, the nth order expansion P of ζ is obtained in exactly n steps.

$$\zeta^{(k+1)} = \mathcal{H}\left(\zeta^{(k)}\right). \tag{45}$$

7. The result is verified by constructing a Taylor model T with the computed reference polynomial P such that $\mathcal{H}(T) \subset T$. With the Taylor model

$$T = \left(P, 0, [0, 0.5], \left[-10^{-14}, 10^{-14}\right]\right), \tag{46}$$

$$\mathcal{H}(T) = \left(P, 0, [0, 0.5], \left[-0.659807722 \cdot 10^{-14}, 0.659857319 \cdot 10^{-14}\right]\right). \tag{47}$$

   Since P is a fixed point of $\mathcal{H}$, the inclusion $\mathcal{H}(T) \subset T$ can be checked by simply comparing the remainder bounds of T and $\mathcal{H}(T)$. By utilizing that $\mathcal{H}(T_{50}) \subset T_{50}$, we show that the inclusion requirement is indeed satisfied for the constructed T. Due to the nonlinear nature of $\mathcal{H}$, an inclusion can almost always be achieved by choosing T sufficiently small.

8. A Taylor model for x is obtained by using the antiderivation of Taylor models according to

$$x(t) = x_0 + \int_0^t \left(x_0' + \int_0^\tau \left(x_0'' + \zeta(\sigma)\right)\mathrm{d}\sigma\right)\mathrm{d}\tau. \tag{48}$$

   The listing in table 1 shows the resulting Taylor model of order 25 computed by COSY Infinity.

This example demonstrates how the new method can be used for the verified integration of implicit ODE initial value problems to high accuracy. In this context it is important to note that the width of the final enclosure of the solution is in the order of 10^{-14} for a relatively large time step of $h = 0.5$.

Table 1
Taylor model for the solution of the implicit second order ODE initial value problem given by (41c).

```
      RDA  VARIABLE:   NO= 25, NV= 1
           I  COEFFICIENT             ORDER
           1  1.000000000000000        0
           2  -.6392322713805370       2
           3  0.4166666666666668E-01   4
           4  -.1993921404777223E-02   6
           5  0.6314945441169959E-04   8
           6  0.2635524930464548E-05  10
           7  -.4411105791086625E-06  12
           8  -.1533094467519992E-07  14
           9  0.8104707776528831E-08  16
          10  -.3384116382961162E-09  18
          11  -.1389729003787960E-09  20
          12  0.1981078695604361E-10  22
          13  0.1549987273495670E-11  24

VAR   REFERENCE POINT         DOMAIN INTERVAL
  1 0.000000000000000      [0.00000, 0.50000]
               REMAINDER BOUND INTERVAL
  R     [-.2500253775762034E-014,0.2500000000000003E-014]
```

3.4. Validated and verified integration of DAEs

Within the context of the previously presented algorithm for the integration of ODEs, the regularity of L_ζ provides a sufficient criterion for the solvability of the derived ODEs. While the linear map L_ζ will generally be singular, if the DAE is solvable, an exhaustive search using repeated differentiation of the individual equations will eventually lead to a regular linear map L_ζ. Additionally, once the consistent initial derivative is computed, the minimum number of differentiations gives the differentiation index ν_d of the DAE.

Since the regularity of L_ζ in the neighborhood of a consistent point is a sufficient criterion for the existence and uniqueness of solutions and since since the regularity of L_ζ is equivalent to the regularity of the Jacobian (19), this approach can allow a rigorous determination of the differentiation index ν_d. In practical terms, this approach is simplified by the availability of the differentiation in the DA framework of COSY Infinity [2–4]. After determination of a consistent initial velocity, an exhaustive search, using automatic differentiation, can find the "right" combination of differentiated equations, thereby yielding the differentiation index of the system. This exhaustive search for a solvable system has the advantage that it is computationally efficient and guaranteed to terminate by either finding a solvable system (and the index) or determining that the index is larger than the current computation order. Once the system is analyzed, high-order Taylor model methods can then be used to rigorously guarantee regularity of the resulting system Jacobian and determine the existence of a consistent point for the initial conditions.

4. Constrained mechanical systems

Constrained mechanical systems can often be written as differential algebraic equations and are typically of the form

$$M(x) \cdot x'' = f(t, x, x') + G(x)\lambda, \tag{49a}$$

$$\Phi(x) = 0, \tag{49b}$$

where $x \in \mathbb{R}^v$ is the state vector, $\lambda \in \mathbb{R}^w$ is the Lagrange multiplier, and $\Phi' = \partial\Phi/\partial x = G^{\mathrm{T}}$. We will generally assume that M is positive-definite symmetric; so its diagonal is nonzero. Moreover we will also assume that Φ' has full row rank, implying that $v \geqslant w$.

Extensive theories have been developed that give *cookbook recipes* for solving these problems by using the constraint conditions (49b) to introduce new variables that reduce the problem's dimensionality [17]. However, these schemes usually rely, to a certain degree, on the user's intuition and often require substantial arithmetic to reorganize and simplify the resulting ODEs into explicit first order systems. Moreover, while utilizing the constraints (49b) lead to simplified ODEs, it generally does not reveal the hidden constraints of the system. On the other hand, within the context of DAEs, the regular and the hidden constraints often provide significant information to the experts. The availability of automated integration schemes for DAE initial value problems removes the need for intuition and physical insight into the problems and allows the automated integration of these important systems without user intervention [40,41].

4.1. Example. Double pendulum

As a prototypical example for constrained mechanical systems, consider a planar pair of connected pendulums in a frictionless environment. Assume that the pendulums are massless and inextensible, with point masses on the ends, as illustrated in figure 1.

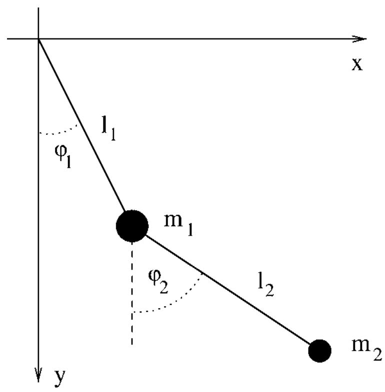

Figure 1. Illustration of a double pendulum with masses m_1 and m_2 connected by massless rods of lengths l_1 and l_2, respectively.

If we denote the tensions in the rods by λ_1 and λ_2, they take the roles of the Lagrange multipliers in this problem and the equations of motion, expressed in the Cartesian coordinates x_1, y_1, x_2, y_2, are given by

$$\begin{aligned} m_1 x_1'' + \frac{\lambda_1 x_1}{l_1} - \frac{\lambda_2(x_2 - x_1)}{l_2} &= 0, \\ m_2 y_1'' + \frac{\lambda_1 y_1}{l_1} - \frac{\lambda_2(y_2 - y_1)}{l_2} - m_1 g &= 0, \\ m_2 x_2'' + \frac{\lambda_2(x_2 - x_1)}{l_2} &= 0, \\ m_2 y_2'' + \frac{\lambda_2(y_2 - y_1)}{l_2} - m_2 g &= 0, \\ x_1^2 + y_1^2 - l_1^2 &= 0, \\ (x_2 - x_1)^2 + (y_2 - y_1)^2 - l_2^2 &= 0. \end{aligned} \tag{50}$$

While it is common practice in classical mechanics to reduce this problem to a two-dimensional ODE in the angles φ_1 and φ_2, here the focus is on showing how our method for the computation of self-validated DAEs can automatically treat this problem as an initial value problem in the six variables x_1, y_1, x_2, y_2, λ_1 and λ_2. By utilizing an exhaustive search and the L_ζ test, our software correctly identifies the DAE as an index-2 system and also determines that a solvable implicit ODE can be obtained by combining the first four equations of (50) with the second derivatives of the the algebraic constraints:

$$\begin{aligned} 2\bigl(x_1'^2 + x_1 x_1'' + y_1'^2 + y_1 y_1''\bigr) &= 0, \\ 2\bigl((x_2' - x_1')^2 + (y_2' - y_1')^2 + (x_2 - x_1)(x_2'' - x_1'') + (y_2 - y_1)(y_2'' - y_1'')\bigr) &= 0. \end{aligned} \tag{51}$$

Starting with Taylor models describing the initial conditions y for $\varphi_1(0) = \varphi_2(0) = 5°$, iterating the operator $\mathcal{H}$ finds the polynomial fixed point $\zeta(t, y)$:

$$\zeta(t, y) = (\zeta_1, \ldots, \zeta_6)(t, y) = \mathcal{H}(\zeta) = (x_1'', y_1'', x_2'', y_2'', \lambda_1, \lambda_2)(t, y). \tag{52}$$

We construct a Taylor model T from the computed polynomial fixed point ζ, domain intervals of ± 0.001 (cf. table 2) and the remainder bound intervals listed in (53).

$$\begin{aligned} &\bigl[-0.107588 \cdot 10^{-11}, 0.107563 \cdot 10^{-11}\bigr], &&\bigl[-0.325530 \cdot 10^{-12}, 0.324108 \cdot 10^{-12}\bigr], \\ &\bigl[-0.827930 \cdot 10^{-12}, 0.827411 \cdot 10^{-12}\bigr], &&\bigl[-0.135636 \cdot 10^{-11}, 0.135024 \cdot 10^{-11}\bigr], \\ &\bigl[-0.586048 \cdot 10^{-11}, 0.586802 \cdot 10^{-11}\bigr], &&\bigl[-0.296806 \cdot 10^{-11}, 0.297409 \cdot 10^{-11}\bigr]. \end{aligned} \tag{53}$$

Table 2
Taylor model describing the coordinate x_1 after the first integration step ($t = 0.001$). Shown are the reference describing the dependence on the eight initial conditions (570 coefficients omitted), reference points, domain intervals, and remainder bound.

```
RDA VARIABLE:   NO=    7, NV=      8
   I  COEFFICIENT           ORDER EXPONENTS
   1  0.8715569933562087E-01   0    0 0  0 0  0 0  0 0
   2  0.9999985208427887 1 1   0    0 0  0 0  0 0  0 0
   3  0.1294094804033084E-06   1    0 1  0 0  0 0  0 0
   4  0.4943135787532041E-06   1    0 0  1 0  0 0  0 0
   5  -.4324683436783440E-07   1    0 0  0 1  0 0  0 0
   6  0.9999995094698690E-03   1    0 0  0 0  1 0  0 0
   7  0.4291582558871565E-10   1    0 0  0 0  0 1  0 0
   8  0.1647712466901794E-09   1    0 0  0 0  0 0  1 0
   9  -.1441561616042412E-10   1    0 0  0 0  0 0  0 1
  10  0.3012457312051847E-06   2    2 0  0 0  0 0  0 0
  11  0.4396925540227502E-06   2    1 1  0 0  0 0  0 0
  12  -.4077392176614172E-07   2    0 2  0 0  0 0  0 0
  13  0.2503569837970065E-12   2    1 0  1 0  0 0  0 0
  14  0.4962017515784074E-06   2    0 1  1 0  0 0  0 0
 ...
 585  0.1432067646745945E-05   5    0 0  1 2  0 0  0 2
 586  -.1285873169976378E-06   5    0 0  0 3  0 0  0 2

VAR    REFERENCE POINT                     DOMAIN INTERVAL
  1 0.8715574274765817E-001 [0.8615574274765816E-001,0.8815574274765817E-001]
  2 0.9961946980917455      [0.9951946980917455,0.9971946980917455 ]
  3 0.1743114854953163      [0.1733114854953163,0.1753114854953163 ]
  4 1.992389396183491       [ 1.991389396183491, 1.993389396183491 ]
  5 0.000000000000000       [-.1000000000000000E-002,0.1000000000000000E-002]
  6 0.000000000000000       [-.1000000000000000E-002,0.1000000000000000E-002]
  7 0.000000000000000       [-.1000000000000000E-002,0.1000000000000000E-002]
  8 0.000000000000000       [-.1000000000000000E-002,0.1000000000000000E-002]
               REMAINDER BOUND INTERVAL
  R     [-.9116356403937544E-014,0.9118332138766990E-014]
```

The remainder bounds of $\mathcal{H}(T)$ are shown in (54), and since the polynomials are fixed under $\mathcal{H}$, we prove the required inclusion $\mathcal{H}(T) \subset T$ by comparing the remainder bounds of T (53) and $\mathcal{H}(T)$ (54).

$$
\begin{aligned}
&\left[-0.889002 \cdot 10^{-12}, 0.889003 \cdot 10^{-12}\right], && \left[-0.215533 \cdot 10^{-12}, 0.215533 \cdot 10^{-12}\right], \\
&\left[-0.550989 \cdot 10^{-12}, 0.550988 \cdot 10^{-12}\right], && \left[-0.792687 \cdot 10^{-12}, 0.792689 \cdot 10^{-12}\right], \\
&\left[-0.544366 \cdot 10^{-11}, 0.544366 \cdot 10^{-11}\right], && \left[-0.221796 \cdot 10^{-11}, 0.221795 \cdot 10^{-11}\right].
\end{aligned}
\tag{54}
$$

As a final illustration of the Taylor model approach and its ability to explicitly propagate guaranteed enclosures of the functional dependence on initial conditions, table 2 lists parts of the final Taylor model for the solution $x_1(y)$ at the end of the first integration step at $t = 0.001$. The width of the remainder bound is less than $2 \cdot 10^{-14}$,

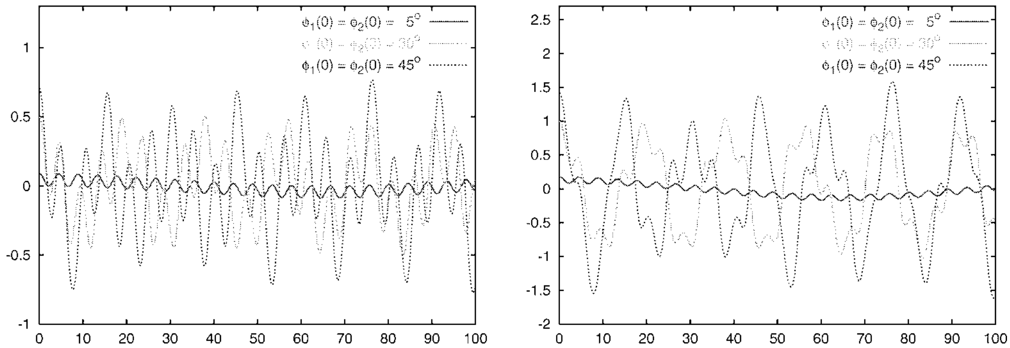

Figure 2. Coordinates x_1 and x_2 of the double pendulum for $\varphi_1(0) = \varphi_2(0) = 5°$, $30°$, $90°$, $x_1'(0) = x_2'(0) = y_1'(0) = y_2'(0) = 0$ and $0 \leqslant t \leqslant 100$ ($g = 1, l_1 = l_2 = 1, m_1 = m_2 = 1$).

providing an extremely tight enclosure of the solution and setting the stage for validated long-term integrations of differential algebraic equations.

Figure 2 illustrates the motion of the double pendulum by showing $x_1(t)$ and $x_2(t)$ for various initial conditions over the time interval [0, 100]. The graphs highlight the oscillations of the system with an energy transfer between the two masses m_1 and m_2. While the system follows an almost periodic path for small energies, it shows signs of chaotic motion for high energies.

Acknowledgements

We would like to thank John Pryce, Stephen Campbell, George Corliss, Ken Jackson, Ned Nedialkov, and Clifford Weil for fruitful discussions on differential algebraic equations.

References

[1] U.M. Ascher and L.R. Petzold, *Computer Methods for Ordinary Differential Equations and Differential–Algebraic Equations* (SIAM, Philadelphia, PA, 1998).

[2] M. Berz, Differential algebraic description of beam dynamics to very high orders, Particle Accelerators 24 (1989) 109.

[3] M. Berz, Forward algorithms for high orders and many variables, in: *Automatic Differentiation of Algorithms: Theory, Implementation and Application* (SIAM, Philadelphia, PA, 1991).

[4] M. Berz, *Modern Map Methods in Particle Beam Physics* (Academic Press, San Diego, 1999).

[5] M. Berz, Constructive generation and verification of Lyapunov functions around fixed points of nonlinear dynamical systems, Internat. J. Comput. Res. (2001).

[6] M. Berz, C. Bischof, G. Corliss and A. Griewank eds., *Computational Differentiation: Techniques, Applications, and Tools* (SIAM, Philadelphia, PA, 1996).

[7] M. Berz and J. Hoefkens, Verified high-order inversion of functional dependencies and superconvergent interval Newton methods, Reliable Computing 7(5) (2001) 379–398.

[8] M. Berz and G. Hoffstätter, Computation and application of Taylor polynomials with interval remainder bounds, Reliable Computing 4(1) (1998) 83–97.

[9] M. Berz, G. Hoffstätter, W. Wan, K. Shamseddine and K. Makino, COSY INFINITY and its applications to nonlinear dynamics, in: [6], pp. 363–365.

[10] M. Berz and K. Makino, Verified integration of ODEs and flows with differential algebraic methods on Taylor models, Reliable Computing 4(4) (1998) 361–369.

[11] M. Berz and K. Makino, COSY INFINITY version 8.1 reference manual, Technical Report MSUCL-1195, National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, MI 48824 (2001).

[12] M. Berz, K. Makino and J. Hoefkens, Verified integration of dynamics in the Solar system, Nonlinear Anal. 47 (2001) 179–190.

[13] K.E. Brenan, S.L. Campbell and L.R. Petzold, *Numerical Solution of Initial-Value Problems in Differential–Algebraic Equations* (North-Holland, Amsterdam, 1989).

[14] Y.F. Chang and G.F. Corliss, Solving ordinary differential equations using Taylor series, ACM Trans. Math. Software 8 (1982) 114–144.

[15] Y.F. Chang and G.F. Corliss, ATOMFT: Solving ODEs and DAEs using Taylor series, Comput. Math. Appl. 28 (1994) 209–233.

[16] B. Erdélyi, J. Hoefkens and M. Berz, Rigorous lower bounds for the domains of definition of extended generating functions, SIAM J. Appl. Dyn. Systems (2001) submitted.

[17] H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, MA, 1980).

[18] E. Griepentrog and R. März, *Differential–Algebraic Equations and Their Numerical Treatment* (Teubner, Stuttgart, 1986).

[19] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations II: Stiff and Differential–Algebraic Problems* (Springer, New York, 1991).

[20] E.R. Hansen, *Topics in Interval Analysis* (Oxford Univ. Press, London, 1969).

[21] J. Hoefkens, Rigorous numerical analysis with high-order Taylor models, Ph.D. thesis, Michigan State University, East Lansing, MI (2001); also MSUCL-1217.

[22] J. Hoefkens and M. Berz, Verification of invertibility of complicated functions over large domains, Reliable Computing 8(1) (2002) 1–16.

[23] J. Hoefkens, M. Berz and K. Makino, *Efficient High-Order Methods for ODEs and DAEs* (Springer, New York, 2001) pp. 343–350.

[24] J. Hoefkens, M. Berz and K. Makino, *Verified High-Order Integration of DAEs and Higher-Order ODEs* (Kluwer Academic, Dordrecht, 2001) pp. 281–292.

[25] R.B. Kearfott, *Rigorous Global Search: Continuous Problems* (Kluwer, Dordrecht, 1996).

[26] E.R. Kolchin, *Differential Algebraic Groups* (Academic Press, New York, 1985).

[27] R.J. Lohner, Enclosing the solutions of ordinary initial and boundary value problems, in: *Computer Arithmetic: Scientific Computation and Programming Languages*, eds. E.W. Kaucher, U.W. Kulisch and C. Ullrich, Wiley–Teubner Series in Computer Science (Teubner, Stuttgart, 1987) pp. 255–286.

[28] R. Lohner, Einschließung der Lösung gewöhnlicher Anfangs- und Randwertaufgaben und Anwendungen, Ph.D. thesis, Universität Karlsruhe (1988).

[29] K. Makino, Rigorous integration of maps and long-term stability, in: *1997 Particle Accelerator Conf.*, Vol. 2, 1997, pp. 1336–1340.

[30] K. Makino, Rigorous analysis of nonlinear motion in particle accelerators, Ph.D. thesis, Michigan State University, East Lansing, MI, USA (1998); also MSUCL-1093.

[31] K. Makino and M. Berz, Remainder differential algebras and their applications, in: [6], pp. 63–74.

[32] K. Makino and M. Berz, Efficient control of the dependency problem based on Taylor model methods, Reliable Computing 5(1) (1999) 3–12.

[33] K. Makino and M. Berz, Advances in verified integration of ODEs, in: *SCAN'2000*, 2000.

[34] K. Makino and M. Berz, Global optimzation with Taylor models, Internat. J. Comput. Res. (2001).

[35] K. Makino and M. Berz, Verified integration of transfer maps, Phys. Rev. ST-AB, submitted.

[36] R.E. Moore, *Interval Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1996).
[37] R.E. Moore, *Methods and Applications of Interval Analysis* (SIAM, Philadelphia, PA, 1979).
[38] N.S. Nedialkov, Computing rigorous bounds on the solution of an initial value problem for an ordinary differential equation, Ph.D. thesis, University of Toronto (1999).
[39] N.S. Nedialkov, K.R. Jackson and G.F. Corliss, Validated solutions of initial value problems for ordinary differential equations, Appl. Math. Comput. 105(1) (1999) 21–68.
[40] C.C. Pantelides, The consistent initialization of differential–algebraic systems, SIAM J. Sci. Statist. Comput. 9(2) (1988) 213–231.
[41] J.D. Pryce, A simple structural analysis method for DAEs, BIT 41(2) (2001) 364–394.
[42] J.F. Ritt, *Differential Equations from the Algebraic Viewpoint* (Amer. Math. Soc., Washington, 1932).
[43] J.F. Ritt, *Integration in Finite Terms – Liouville's Theory of Elementary Methods* (Columbia Univ. Press, New York, 1948).