Reliable Computing **7:** 379–398, 2001.

Verified High-Order Inversion of Functional Dependencies and Interval Newton Methods

MARTIN BERZ and JENS HOEFKENS
Department of Physics and Astronomy and National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, MI 48824, USA, e-mail: berz@msu.edu, hoefkens@msu.edu

(Received: 25 February 2000; accepted: 28 August 2000)

Abstract. A new method for computing verified enclosures of the inverses of given functions over large domains is presented. The approach is based on Taylor Model methods, and the sharpness of the enclosures scales with a high order of the domain. These methods have applications in the solution of implicit equations and the Taylor Model based integration of Differential Algebraic Equations (DAE) as well as other tasks where obtaining verified high-order models of inverse functions is required.

The accuracy of Taylor model methods has been shown to scale with the $(n + 1)$-st order of the underlying domain, and as a consequence, they are particularly well suited to model functions over relatively large domains. Moreover, since Taylor models can control the cancellation and dependency problems (see Makino, K. and Berz, M.: Efficient Control of the Dependency Problem Based on Taylor Model Methods, *Reliable Computing* **5** (1) (1999)) that often affect regular interval techniques, the new method can successfully deal with complicated multidimensional problems. As an application of these new methods, a high-order extension of the standard Interval Newton method that converges approximately with the $(n + 1)$-st order of the underlying domain is developed.

Several examples showing various aspects of the practical behavior of the methods are given.

1. Introduction

In the following we will develop methods for inclusion of inverses of general functions. As a prerequisite it is necessary to prove invertibility over the domain in question, and there are a variety of methods for this purpose. Recently, Taylor model methods based on high-order multivariate floating point polynomials with interval remainder bounds have been derived that allow the efficient and accurate determination of whether a given function can be guaranteed to be invertible [9]. The Taylor model methods allow verification of invertibility over domains that are often larger than the ones over which other methods succeed. Moreover, Taylor models are significantly less susceptible to blow up due to the complexity of the problem and the linear algebra in the inversion. In general, Taylor model methods provide sharp guaranteed inclusion of functional dependencies with an accuracy that scales with the $(n+1)$-st order of the domain over which the functional dependence is evaluated [4], [5], [13], and it has been shown [12] that Taylor models can often substantially alleviate the following problems inherent in naive interval arithmetic:

- Sharpness for large domain intervals,

- Cancellation and dependency problem,
- Dimensionality curse.

They have recently been used for a variety of applications, including verified bounding of highly complex functions [3], [6], solution of ODEs under avoidance of the wrapping effect [8], and high-dimensional verified quadrature [7].

This paper demonstrates how these techniques can be combined with the recently developed methods for the verification of invertibility [9] to derive a new method that allows the computation of Taylor models containing the inverses of given invertible functions, modeled by Taylor models. As an immediate application of this, a superconvergent high-order extension of the regular Interval Newton method will be presented.

2. Invertibility

Before inclusion of inverse functions is possible, invertibility of the function under consideration has to be guaranteed over the whole domain of interest. Invertibility, by virtue of the following theorem, is a local property.

THEOREM 2.1 Inverse Mapping Theorem. *Let $f : U \subset \mathbb{R}^v \to \mathbb{R}^w$ be of class $\mathcal{C}^n$, $n \geq 1$, $x_0 \in U$, and suppose that the derivative $(\mathbf{D}f)(x_0)$ of f at x_0 is a linear isomorphism. Then f is a $\mathcal{C}^n$ diffeomorphism of some neighborhood of x_0 onto some neighborhood of $f(x_0)$. Moreover, $(\mathbf{D}f^{-1})(y) = ((\mathbf{D}f)(f^{-1}(y)))^{-1}$ for y in this neighborhood of $f(x_0)$.*

It should be noted that this is a purely local theorem that does not give any estimates on the size of the neighborhood over which the function f is invertible, nor can such an estimate be deduced from it. The following, semi-local extension of Theorem 2.1 on the other hand can determine invertibility over a whole domain box $\boldsymbol{B}$.

THEOREM 2.2 Invertibility from First Derivatives. *Let $\boldsymbol{B} \subset \mathbb{R}^v$ be a box and $f : \boldsymbol{B} \to \mathbb{R}^v$ a $\mathcal{C}^1$ function. Assume that the matrix*

$$M = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(\chi_1) & \cdots & \dfrac{\partial f_1}{\partial x_v}(\chi_1) \\ \vdots & & \vdots \\ \dfrac{\partial f_v}{\partial x_1}(\chi_v) & \cdots & \dfrac{\partial f_v}{\partial x_v}(\chi_v) \end{pmatrix}$$

is invertible for every choice of $\chi_1, \ldots, \chi_v \in \boldsymbol{B}$. Then f has a $\mathcal{C}^1$-inverse defined on all of the range of f over the domain $\boldsymbol{B}$.

This theorem is proven in [9], and it has been shown how a practical application can benefit from the utilization of Taylor models. While intervals cannot adequately model the fact that individual rows of the matrix M can be evaluated at

the same point, Taylor models can model and use this inherent structural property of M. Moreover, the resulting Taylor model based algorithm scales better to higher dimensional problems, more complicated functional dependencies, and larger domain boxes than corresponding interval methods.

Finally, it should be noted that the computational complexity of a practical implementation of Theorem 2.2 can be significantly reduced, since the utilization of Taylor model methods allows all computations to be performed using a single set of v variables as opposed to the v^2 variables Theorem 2.2 seems to require [9].

3. Polynomial Inverses

As a first step to the computation of Taylor models of inverses, we develop methods to determine the multivariate polynomial part of the inverse. To automate this process, we utilize a particular fixed point theorem. For purposes of notation, we consider the set $\mathcal{C}^n(\mathbb{R}^v, \mathbb{R}^w)$ and equip it with following relation: for $f, g \in \mathcal{C}^n(\mathbb{R}^v, \mathbb{R}^w)$ we say $f =_n g$ (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. This gives an equivalence relation on $\mathcal{C}^n(\mathbb{R}^v, \mathbb{R}^w)$. The resulting equivalence classes are called DA vectors and the class containing $f \in \mathcal{C}^n(\mathbb{R}^v, \mathbb{R}^w)$ is denoted by $[f]$. The collection of these equivalence classes is called ${}_nD_v$ [2], [4]. It should be noted that $f \in \mathcal{C}^n(\mathbb{R}^v, \mathbb{R}^w)$ lies in the same equivalence class as its Taylor polynomial T_f of order n around the origin.

Elementary operations can be defined on ${}_nD_v$ such that it becomes an algebra and, with the appropriate definitions of the derivative operation, it can even be shown to be a differential algebra [4]. Efficient methods of using DA vectors have been implemented in the arbitrary order code COSY Infinity [1].

DEFINITION 3.1. For $[f] \in {}_nD_v$, the depth $\lambda([f])$ is defined by

$$\lambda([f]) = \begin{cases} \text{order of first non-vanishing derivative of } f, & \text{if } [f] \neq 0, \\ n+1, & \text{else.} \end{cases}$$

The function λ resembles a valuation on ${}_nD_v$. Using it, it becomes possible to define the notion of contracting operators on ${}_nD_v$.

DEFINITION 3.2. Let $\mathcal{O}$ be an operator on $M \subset {}_nD_v$. $\mathcal{O}$ is contracting on M, if for any $[f], [g] \in M$

$$\lambda\big(\mathcal{O}([f]) - \mathcal{O}([g])\big) \geq \lambda([f] - [g])$$

and equality iff $[f] = [g]$.

This definition has a striking similarity to the corresponding definitions on regular function spaces and as it turns out, a theorem that resembles the Banach Fixed Point Theorem can be established on ${}_nD_v$. But, unlike in the case of the Banach Fixed Point theorem, it will be guaranteed to converge in at most $n + 1$ steps, which makes it very well suited for practical applications.

THEOREM 3.1 DA Fixed Point Theorem. *Let $\mathcal{O}$ be a contracting operator and self-map on $M \subset {}_nD_v$. Then $\mathcal{O}$ has a unique fixed point $a \in M$. Moreover, for any $a_0 \in M$ the sequence*

$$a_k = \mathcal{O}(a_{k-1})$$

converges in at most $n+1$ steps to a.

Proof. We have

$$\begin{aligned} \lambda(a_{k+2} - a_{k+1}) &= \lambda\big(\mathcal{O}(a_{k+1}) - \mathcal{O}(a_k)\big) \\ &> \lambda(a_{k+1} - a_k). \end{aligned}$$

Because λ is integer-valued, one gets $\lambda(a_{k+2} - a_{k+1}) \geq \lambda(a_{k+1} - a_k) + 1$ and by induction it follows that

$$\begin{aligned} \lambda(a_{n+2} - a_{n+1}) &\geq \lambda(a_{n+1} - a_n) + 1 \\ &\geq \ldots \\ &\geq \lambda(a_1 - a_0) + n \geq n + 1 \end{aligned}$$

By definition of λ it follows that $a_{n+2} = a_{n+1}$ and hence

$$a_{n+1} = \mathcal{O}(a_{n+1}).$$

Thus, the fixed point is reached after at most $n+1$ steps. Moreover, it is unique, because assume that there were two distinct fixed points a and $\bar{a}$. Then $\lambda(a - \bar{a}) = \lambda(\mathcal{O}(a) - \mathcal{O}(\bar{a})) > \lambda(a - \bar{a})$ which would be a contradiction. □

We will now utilize the fixed point theorem to calculate the n-th order polynomial part of the inverse. Assume that $f \in \mathcal{C}^n(\mathbb{R}^v, \mathbb{R}^v)$ is an origin-preserving map (i.e. $f(0) = 0$). If the derivative M of f at 0 is a linear isomorphism, then Theorem 2.1 guarantees that there is a neighborhood V of 0, such that f^{-1} is also of class $\mathcal{C}^n$ on V. Once the linearization of f has been shown to be invertible, using the DA framework, it becomes possible to compute a representative of the equivalence class of $[f^{-1}]_n$ from one representative $\mathcal{M}$ of the equivalence class $[f]_n$ using the following steps:

Split the map $\mathcal{M}$ as

$$\mathcal{M} = M + \mathcal{N}_{\mathcal{M}}$$

where $\mathcal{N}_{\mathcal{M}}$ denotes the purely non-linear part of $\mathcal{M}$. Composing from the right with the (locally existing) inverse results in

$$\begin{aligned} \mathcal{I} &= \mathcal{M} \circ \mathcal{M}^{-1} = M \circ \mathcal{M}^{-1} + \mathcal{N}_{\mathcal{M}} \circ \mathcal{M}^{-1} \\ &\Rightarrow \mathcal{M}^{-1} = M^{-1} \circ (\mathcal{I} - \mathcal{N}_{\mathcal{M}} \circ \mathcal{M}^{-1}) = \mathcal{O}(\mathcal{M}^{-1}), \end{aligned}$$

where $\mathcal{I}$ is the identity map. The last equation is actually a fixed point relation for the map $\mathcal{M}^{-1}$. Viewing this equation as a relation on equivalence classes, it turns out that $\mathcal{O}$ is contracting because

$$\left(\mathcal{O}([\mathcal{A}]_n) - \mathcal{O}([\mathcal{B}]_n)\right) = M^{-1} \circ ([\mathcal{N}_{\mathcal{M}}]_n \circ [\mathcal{B}]_n - [\mathcal{N}_{\mathcal{M}}]_n \circ [\mathcal{A}]_n)$$

implies that with $[\mathcal{A}]_n$ and $[\mathcal{B}]_n$ agreeing up to order k, it follows that $[\mathcal{N}_{\mathcal{M}}]_n \circ [\mathcal{B}]_n$ and $[\mathcal{N}_{\mathcal{M}}]_n \circ [\mathcal{A}]_n$ agree up to order $k+1$.

Thus, according to Theorem 3.1, this assures the existence of a unique fixed point of the operator $\mathcal{O}$, which can be reached in at most $n+1$ iterations. Moreover, this method generates all derivatives of up to order n of f^{-1} at the origin in finitely many steps through mere iteration. Because it only requires the iteration of a rather simple operator, it is particularly useful for computational applications.

The presented method by itself allows only the computation of inverse polynomials for origin-preserving maps. But, as will be shown in Section 5, it can nevertheless be used for the computation of general inverse Taylor models.

4. Taylor Models

Taylor models are a combination of high-order Taylor polynomials and interval arithmetic [4], [5], [13]. Their fundamental properties are given by the following definition.

DEFINITION 4.1 *Taylor Model.* Let $\boldsymbol{D} \subset \mathbb{R}^v$ be a box with $x_0 \in \boldsymbol{D}$. Let $P_n : \boldsymbol{D} \to \mathbb{R}^w$ be a polynomial of order n ($n, v, w \in \mathbb{N}$) and $R \subset \mathbb{R}^w$ be an open non-empty set. Then $(P_n, x_0, \boldsymbol{D}, R)$ is called a Taylor model of order n with reference point x_0 over $\boldsymbol{D}$.

In general, we will view Taylor models as subsets of function spaces by virtue of the following definition.

DEFINITION 4.2 *Taylor Models as Sets of Functions.* Let $T = (P_n, x_0, \boldsymbol{D}, R)$ be a Taylor model of n-th order. Then, identify T with the set of functions $f \in \mathcal{C}^n(\boldsymbol{D}, \mathbb{R}^w)$ such that $P_n(x) - f(x) \in R$ for all $x \in \boldsymbol{D}$, and the n-th order Taylor series of f around x_0 equals P_n.

Furthermore, if a $\mathcal{C}^n$ function f is contained in a Taylor model T, we call T a Taylor model for f.

Methods have been developed that allow mathematical operations on Taylor models that preserve the inclusion relationships. I.e. for two given Taylor models T_1 and T_2 methods have been developed that compute Taylor models for the sum S, and product P, of T_1 and T_2, i.e.

$$f_1 \in T_1,\ f_2 \in T_2 \ \Rightarrow\ (f_1 + f_2) \in S,$$
$$f_1 \in T_1,\ f_2 \in T_2 \ \Rightarrow\ (f_1 \cdot f_2) \in P.$$

More information on arithmetic and intrinsic functions on Taylor models can be found in [4], [5], [12], [13].

The accuracy of the enclosures obtained by Taylor models scales with the $(n+1)$-st order of the domain size. This makes Taylor models particularly well suited for high-dimensional problems over large domains, since it reduces the number of domain splittings.

THEOREM 4.1. *Let f be a function that is represented by a finite number of elementary operations and intrinsic functions, and assume that B is an inclusion monotone polynomial bounder that scales at least linear with the size of the domain. If $T = (P_n, x_0, \boldsymbol{D}, R)$ is a Taylor model obtained by evaluating the code list representing f, then the remainder bound R scales with the $(n+1)$-st order of the size of $\boldsymbol{D}$.*

Proof. The proof follows by induction over the elementary operations and intrinsic functions that make up the code list for f. Firstly, the assertion is correct for the constant Taylor models and for the identity, since the respective reference polynomials can be bound with an arbitrary precision.

Since, for two Taylor models A and B, we have $R_{A+B} = R_A + R_B$, the assertion also holds for the sum of Taylor models. Denoting all terms of a polynomial P that are at least of order $n+1$ by $P^{(n+1)}$, the remainder bound of the product $A \cdot B$ is given by

$$R_{A\cdot B} = R_A \cdot R_B + R_A \cdot B(P_B) + B(P_A)R_B + B\big((P_A \cdot P_B)^{(n+1)}\big).$$

Since each of the terms scales with at least order $n+1$, so does the sum. Thus, the assertion holds for all elementary operations.

Finally, the computation of remainder bounds of intrinsic functions is exemplified by the exponential function. According to [11], we have

$$R_{\exp(A)} = \Big(B\big(P_A - P_A(0)\big) + R_A\Big)^{n+1} \cdot \exp\left([0,1] \cdot \Big(B\big(P_A - P_A(0)\big) + R_A\Big)\right).$$

By inclusion monotonicity, the second term never increases with a decreasing domain size and thus, the product scales with the $(n+1)$-st order of the domain size.

Using the complete definitions [11], [13], similar arguments can be made for all the intrinsic functions of Taylor models. Thus, the remainder bound of any finite Taylor model computation does indeed scale with the $(n+1)$-st order of the domain. □

5. Verified Inverses

The computation of verified inverses begins by modeling the functional dependence of interest by a Taylor model by evaluating its code in Taylor model arithmetic. Once this enclosure of the function has been obtained, the next step is the "in

principle" determination of invertibility, for example via recently developed Taylor model based techniques [9]. With the knowledge that the function contained in the Taylor model is actually invertible, the computation of an enclosure of the inverse will follow the steps outlined in this section.

First, it should be clarified what exactly constitutes an inverse Taylor model. In fact, due to compatibility requirements of the underlying domains, one has to consider two kinds of inverse Taylor models as given by the next definition. We denote by $\boldsymbol{f}^{\boldsymbol{u}}(\boldsymbol{D})$ the range of the the function f over the domain $\boldsymbol{D}$. Moreover, we denote by $B(T)$ an inclusion of the range of all functions in the Taylor model T and by I the Taylor model $(x, 0, \boldsymbol{D}, [0, 0])$ of the identity function over the domain $\boldsymbol{D}$. Finally, for a polynomial P and a Taylor model T, $P(T)$ denotes the result of the formal evaluation of the polynomial P with the argument T. This is well defined since it requires only additions and multiplications of Taylor models with Taylor models and real numbers.

DEFINITION 5.1. Let $T = (P_n, x_0, \boldsymbol{D}, R)$ and $S = (G_n, y_0, \boldsymbol{\Delta}, \Omega)$ be two Taylor models. S is called a *left-inverse Taylor model for* T if $G_n \circ P_n =_n \mathcal{I}$, $P_n(x_0) = y_0$, $f \in T \Rightarrow \boldsymbol{f}^{\boldsymbol{u}}(\boldsymbol{D}) \subset \boldsymbol{\Delta}$, and $B(G_n(T) - I) \subset \Omega$. Similarly, S is called a *right-inverse Taylor model for* T if T is a left-inverse Taylor model for S.

It turns out that the rather general definition of left-inverse Taylor models is already sufficient for the verified enclosure of left-inverses for all invertible functions contained in T. And, it should be noted that the computation of left-inverse Taylor models is often sufficient for practical applications because they do not require a modification of the original domains of definition and their domains are easily computed.

THEOREM 5.1. *Let* $T = (P_n, x_0, \boldsymbol{D}, R)$ *and* $S = (G_n, y_0, \boldsymbol{\Delta}, \Omega)$ *be given Taylor models such that* S *is a left-inverse Taylor model for* T. *Assume that* $f \in T$ *is invertible over* $\boldsymbol{D}$. *Then there is* $g \in S$ *such that* $g = f^{-1}$ *on* $\boldsymbol{f}^{\boldsymbol{u}}(\boldsymbol{D})$.

Proof. Define $\delta_f : \boldsymbol{D} \to \mathbb{R}^v$ by $\delta_f(x) := G_n(f(x)) - x$. Note that the Taylor expansion of δ_f around $x_0 \in \boldsymbol{D}$ vanishes up to order n since the expansion of f around x_0 agrees with P_n up to order n and by assumption $G_n(P_n(x)) =_n x$.

Since f is invertible over $\boldsymbol{D}$, $\forall\, y \in \boldsymbol{f}^{\boldsymbol{u}}(\boldsymbol{D})$ there is a unique $x_y \in \boldsymbol{D}$ such that $f(x_y) = y$. Then, for $y \in \boldsymbol{f}^{\boldsymbol{u}}(\boldsymbol{D})$ define the function

$$\tilde{g}(y) = G_n(y) - \delta_f(x_y) = G_n(y) - \delta_f\big(f^{-1}(y)\big).$$

Then, for any $x \in \boldsymbol{D}$ we have

$$\begin{aligned}\tilde{g}\big(f(x)\big) &= G_n\big(f(x)\big) - \delta_f\Big(f^{-1}\big(f(x)\big)\Big)\\ &= x + \delta_f(x) - \delta_f\Big(f^{-1}\big(f(x)\big)\Big) = x.\end{aligned}$$

Thus $\tilde{g}$ is the unique inverse to f on $\boldsymbol{f}^{\boldsymbol{u}}(\boldsymbol{D})$. Since $|\tilde{g}(y) - G_n(y)| = |\delta_f(f^{-1}(y))| \in \Omega$, it follows that $\tilde{g}$ is contained in the Taylor model $S_f := (G_n, y_0, \boldsymbol{f}^{\boldsymbol{u}}(\boldsymbol{D}), \Omega)$.

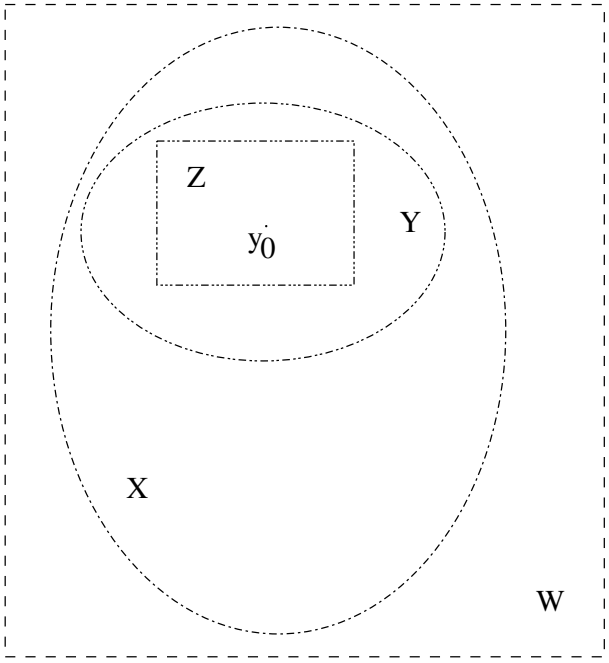

Figure 1. Domains in the definitions of inverse Taylor models.

As the continuous image of a compact set, $\boldsymbol{f}^{\boldsymbol{u}}(\boldsymbol{D})$ itself is compact and hence the function $\tilde{g}$ can be n-times continuously differentiably extended to a function g on the whole set $\boldsymbol{\Delta}$. Moreover, since Ω is open and $\tilde{\boldsymbol{g}}^{\boldsymbol{u}}(\boldsymbol{f}^{\boldsymbol{u}}(\boldsymbol{D})) \subset \Omega$, it is possible to satisfy $g - G_n \subset \Omega$ over $\boldsymbol{\Delta}$. Since the n-th order Taylor expansion of g around $y_0 = f(x_0) \in \boldsymbol{D}$ equals the n-th order expansion of $\tilde{g}$, which in turn equals G_n, it follows that $g \in S$. Finally, for $x \in \boldsymbol{D}$ it is $g(f(x)) = \tilde{g}(f(x)) = x$. □

The first step in translating the previous theorem into a practicable algorithm lies in the efficient computation of the inverse polynomial G_n. To this end, let $T = (P_n, x_0, \boldsymbol{D}, R)$ be a given Taylor model and assume that P_n is (locally) invertible around x_0. Now we consider the polynomial

$$\tilde{P}(x) = P(x + x_0) - P(x_0).$$

We apparently have $\tilde{P}(0) = 0$, and hence according to Section 3 it is possible to calculate $[\tilde{P}^{-1}]$ in $n + 1$ steps. With $y_0 = P(x_0)$ we then obtain $G_n(y)$ as

$$G_n(y) = x_0 + \tilde{P}^{-1}(y - y_0).$$

Once the reference polynomial of the inverse Taylor model has been computed, the next step is the determination of an appropriate domain $\boldsymbol{\Delta}$ such that either $f \in T \Rightarrow \boldsymbol{f}^{\boldsymbol{u}}(\boldsymbol{D}) \subset \boldsymbol{\Delta}$ or $g \in S \Rightarrow \boldsymbol{g}^{\boldsymbol{u}}(\boldsymbol{\Delta}) \subset \boldsymbol{D}$.

Figure 1 illustrates the various domains involved in this and their relationships; we have $X = \bigcup_{f \in T} \boldsymbol{f}^{\boldsymbol{u}}(\boldsymbol{D})$ and $Y = \bigcap_{f \in T} \boldsymbol{f}^{\boldsymbol{u}}(\boldsymbol{D})$, and hence X and Y are the smallest and largest sets that could be used as domains for left- and right-inverse Taylor models, respectively. Since a general computation and representation of X and Y is impossible, a practical application strives to find a good overestimation $\boldsymbol{W}$ for X and underestimation $\boldsymbol{Z}$ for Y such that $\boldsymbol{Z} \subset Y \subset X \subset \boldsymbol{W}$. An implementation of this method chooses $\boldsymbol{\Delta} = \boldsymbol{T}^{\boldsymbol{u}}(\boldsymbol{D})$ as the domain for the left-inverse Taylor model. Computation of the latter consists of bounding the range of the reference polynomial P_n over $\boldsymbol{D}$ and adding the remainder bound R. While the bound on the polynomial does not have to be accurate to obtain an inverse Taylor model, the practical application of left-inverse Taylor models benefits from a small overestimation to remain accurate.

6. A High-Order Newton Method

In this section we apply the methods derived in the previous section to a higher order version of the interval Newton method. We use this approach to illustrate that convergence does indeed happen with the order used for the Taylor model, and that the domain in which convergence occurs is sometimes enlarged. For practical root-finding problems, the new method offers a larger domain of convergence, and is more robust for complicated functions because of the decrease of the dependency problem. For standard cases, the conventional method may often be suffient, because different from the problem of charting of contstraint manifolds and the verified inclusion of the inverse over large domains necessary for DAE integrators, good approximations of the inverse are only needed in a very small domain near the previous guess for the root. Since the method converges with an order higher than two, it is also called superconvergent.

Denoting the centerpoint of an interval $\boldsymbol{X}$ by $m(\boldsymbol{X})$, the algorithm for the Taylor model based computation of zeros is elegant and straightforward. Moreover, it can be easily implemented using the COSY Infinity [1] language environment:

1. Set $k = 0$ and let $\boldsymbol{D}^{(0)}$ be a box that contains a zero of f and assume that f is invertible over $\boldsymbol{D}^{(0)}$.
2. Model the functional dependence of f over $\boldsymbol{D}^{(k)}$ by an n-th order Taylor model T with the reference point $m(\boldsymbol{D}^{(k)})$.
3. Compute a left-inverse Taylor model S for T.
4. Compute the intersection $\boldsymbol{D}^{(k+1)} = \boldsymbol{D}^{(k)} \cap (m(\boldsymbol{D}^{(k)}) + S(m(\boldsymbol{D}^{(k)})))$. If the enclosure $\boldsymbol{D}^{(k+1)}$ of the zero is not accurate enough, replace k by $k + 1$ and go back to 2.

Obviously, for $n = 1$ this method is equivalent to the traditional Interval Newton method [14], [15].

This algorithm allows for fast and accurate determination of guaranteed enclosures of the zeros of a given function. Since the accuracy of Taylor models scales

with the $(n+1)$-st order of the domain size, this method converges very rapidly once the domain size has become sufficiently small. This fact will be illustrated further in Section 7.2.4.

Using the notation of the previous algorithmic description, the following is an obvious consequence.

COROLLARY 6.1. *Given a Taylor model T for f as before. Assume that f is invertible over $\boldsymbol{f}^{\boldsymbol{u}}(\boldsymbol{D})$, has a zero $x^* \in \boldsymbol{D}$ and that $\boldsymbol{D} \cap (m(\boldsymbol{D}) + S(m(\boldsymbol{D}))) \subset \boldsymbol{D}$. Then, for $\boldsymbol{D}^{(0)} = \boldsymbol{D}$ we have $x^* \in \boldsymbol{D}^{(k)}$ for each $k \in \mathbb{N}_0$.*

It is important to note that it is generally not possible to give rigorous estimates on the speed of convergence of this method. But once the requirements of Corollary 6.1 have been satisfied, it will usually converge with the $(n+1)$-st order of the size of the inclusion $\boldsymbol{D}$ (see also Section 7.2.4). It should also be pointed out that unlike the extended Interval Newton method, the presented algorithm cannot handle local extrema of f within the domain. However, this can be circumvented by performing extended Interval Newton preconditioning steps that split the original domain until the Taylor models on the smaller boxes satisfy the requirements of Corollary 6.1.

7. Examples

In the following we demonstrate the presented methods in examples. All computations have been performed using the Taylor model object in the arbitrary order code COSY Infinity [1].

7.1. Inversion

The first two examples illustrate the verified inversion itself. Although not presented here, all computations are based on initially proving invertibility by the recently presented Taylor model methods in [9].

7.1.1. *Sine Function*

As a first example, consider the one-dimensional sine function over the domain $[-0.5, 0.5]$. The functional dependence is modeled by a 19-th order Taylor model over that domain, and the result is presented in Table 1. As can be seen, the 19-th order Taylor model encloses the sine function with an accuracy that is rather close to machine epsilon.

Invertibility of the sine function can be verified easily. The next step is the computation of a polynomial approximation of the inverse of the reference function. Not surprisingly, the methods introduced in Section 5 obtain the 19-th order Taylor expansion of the arcsine function up to machine epsilon. Proper computation of the remainder bounds and the domain gives the left-inverse Taylor model shown in Table 2.

Table 1. Example 7.1.1: Taylor model of the one-dimensional sine function as computed by COSY Infinity. Shown are Taylor coefficients, reference point, domain information and remainder bound.

```
RDA VARIABLE:   NO=  19, NV=     1
   I  COEFFICIENT            ORDER EXPONENTS
   1   1.000000000000000       1   1 0  0
   2  -.1666666666666667       3   3 0  0
   3  0.8333333333333333E-02   5   5 0  0
   4  -.1984126984126984E-03   7   7 0  0
   5  0.2755731922398589E-05   9   9 0  0
   6  -.2505210838544172E-07  11  11 0  0
   7  0.1605904383682162E-09  13  13 0  0
   8  -.7647163731819817E-12  15  15 0  0
   9  0.2811457254345521E-14  17  17 0  0
   ------------------------------------
VAR    REFERENCE POINT               DOMAIN INTERVAL
  1  0.000000000000000   [-.5000000000000000,0.5000000000000000]
              REMAINDER BOUND INTERVAL
  R     [-.1085432243394823E-014,0.1085432243394823E-014]
************************************************************
```

Table 2. Example 7.1.1: Left-inverse Taylor model for the Taylor model shown in Table 1. Shown are Taylor coefficients, reference point, domain information and remainder bound.

```
RDA VARIABLE:   NO=  19, NV=     1
   I  COEFFICIENT            ORDER EXPONENTS
   1  1.000000000000000       1   1 0  0
   2  0.1666666666666667      3   3 0  0
   3  0.7500000000000000E-01  5   5 0  0
   4  0.4464285714285714E-01  7   7 0  0
   5  0.3038194444444444E-01  9   9 0  0
   6  0.2237215909090909E-01 11  11 0  0
   7  0.1735276442307693E-01 13  13 0  0
   8  0.1396484375000000E-01 15  15 0  0
   9  0.1155180089613970E-01 17  17 0  0
  10  0.9761609529194068E-02 19  19 0  0
  ------------------------------------
VAR    REFERENCE POINT               DOMAIN INTERVAL
  1  0.000000000000000   [-.5210953054937487,0.5210953054937487]
              REMAINDER BOUND INTERVAL
  R     [-.7707363654262549E-008,0.7707363654262549E-008]
************************************************************
```

Figure 2 shows the difference between the arcsine function and the reference polynomial over the interval $\sin([-0.5, 0.5]) = [-0.479425\ldots, 0.479425\ldots]$. As

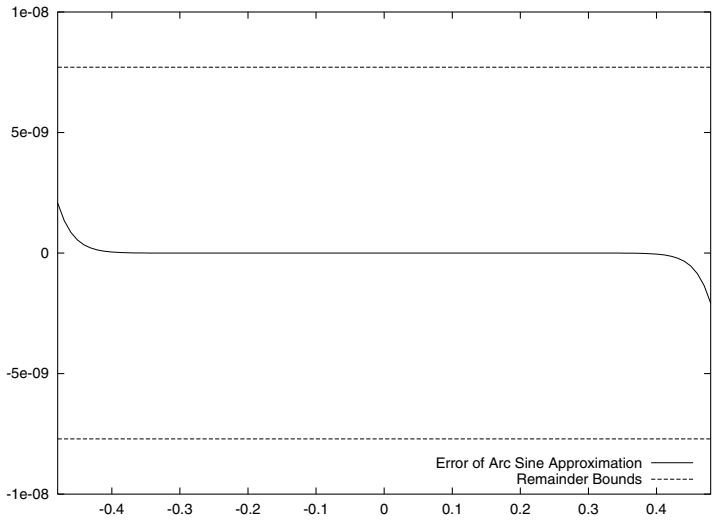

Figure 2. Example 7.1.1: Difference between arcsine and the reference polynomial shown in Table 2.

expected, the error stays within the remainder bounds (illustrated by the horizontal lines), and the remainder bounds represent an overestimation of only around a factor of 4. It should be noted, however, that the arcsine function as such is not contained in the resulting left-inverse Taylor model. But, as outlined earlier, the part of the arcsine function necessary to act as a left-inverse of the sine function over the domain $\sin([-0.5, 0.5])$ is contained and can be extended such that the resulting function is defined over the whole domain of the left-inverse Taylor model.

7.1.2. *Exponential Function*

The following example involves a 6 dimensional exponential function that has a dense 8-th order Taylor polynomial (each of the 6 reference polynomials consists of 3003 non-vanishing coefficients).

Let $A = (a_{ij})$ be an invertible 6×6 matrix and consider $f = (f_1, \ldots, f_6) : \mathbb{R}^6 \to \mathbb{R}^6$ defined by

$$f_i(x_1, \ldots, x_6) = \exp\left(\sum_{j=1}^{6} a_{ij} x_j\right) - 1.$$

If $B = (b_{ij})$ is the inverse matrix of A, the inverse function $g = (g_1, \ldots, g_6)$ of f is given by

$$g_i(y_1, \ldots, y_6) = \sum_{j=1}^{6} b_{ij} \log(y_j + 1).$$

Table 3. Example 7.1.2: Remainder Bounds of the left-inverse Taylor model for a 6 dimensional exponential function.

Component	Remainder Bound
1	[−0.4190638646976846E−011, 0.4184087823912867E−011]
2	[−0.2791908825275360E−011, 0.2791908821988238E−011]
3	[−0.2791908824574486E−011, 0.2791908821987869E−011]
4	[−0.1396454411975411E−011, 0.1396454410994258E−011]
5	[−0.1396454411909750E−011, 0.1396454410994186E−011]
6	[−0.1396454411225902E−011, 0.1396454410994267E−011]

Using the matrix A given below, the 8-th order Taylor model method for verified inversion has been applied to the function f over the domain box $D = [-0.01, 0.01]^6$:

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 & 1 & -1 \end{pmatrix}.$$

The resulting left-inverse Taylor model is defined over the domain $[-0.061686\ldots, 0.061686\ldots]^6$, and it has been verified that it contains the true inverse function g. The remainder bounds of the individual components of the left-inverse Taylor model are listed in Table 3. It should be noted that the range of the exponential function and its inverse are in the order of 10^{-1}. Thus, the relative accuracy of the remainder bounds is about 10^{-10}.

7.1.3. *Computational Complexity of Verified Inversion*

The computational complexity of individual Taylor model operations exceeds the cost of the corresponding standard interval arithmetic by a wide margin. For example, the size of an n-th order Taylor model in v variables scales approximately as

$$\binom{n+v}{v} = \frac{(n+v)!}{v!\,n!}.$$

However, since Taylor models can often avoid domain splitting, their real strength lies in high-dimensional problems with large domains. As an example, consider enclosing the inverse of a one-dimensional function as illustrated in Figure 3. Using only intervals, this requires domains of the size of the desired accuracy (i.e. the number of sub-intervals scales with the inverse of the desired accuracy).

As another example, consider the six-dimensional exponential function presented in Section 7.1.2. The computational cost of the eighth order Taylor models is

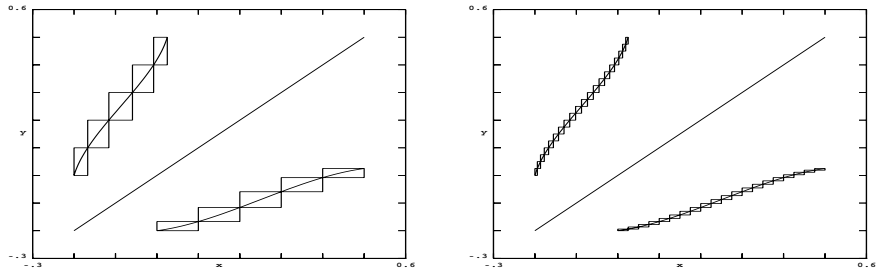

Figure 3. Example 7.1.3: Illustration of verified inversion with intervals. The accuracy scales with the size of the domain sub-intervals.

about 1500 times higher than that of intervals. But to achieve a uniform inclusion accuracy of 10^{-11} with intervals, as is needed for example for verified DAE integrators [10], one would need approximately 10^9 sub-intervals in each dimension, resulting in 10^{54} separate interval evaluations. Thus, the advantages of Taylor models outweigh their additional costs if enclosures of high precision are desired.

As a general rule of thumb, if simple interval methods and Taylor models compete over the same domain and box-splitting is not required to achieve the necessary sharpness, the simple interval methods will almost always win. But if using intervals requires splitting of the domains, the use of Taylor models will generally become favorable.

7.2. Newton Method

The following examples demonstrate the Taylor model based high-order extension of the Interval Newton method. In particular, they show that the method converges extremely fast over relatively large domains.

7.2.1. *One-Dimensional Function*

This example compares the presented DA and Taylor model high-order extensions of the Newton method and the floating point and interval versions of that algorithm in a one-dimensional setting. It shows how the high-order methods can indeed outperform the traditional algorithms by a wide margin.

The floating depth h of a cylindrical trunk of density $\rho = 0.66$ and radius r in water is given by

$$h = r\big(1 - \cos(\alpha_0 / 2)\big),$$

where $\alpha_0 = 3.655403079564624\ldots$ is the (unique) fixed point of

$$\alpha = \sin(\alpha) + 2\pi\rho.$$

This fixed-point relation for α_0 can readily be transformed into a root finding problem and the resulting function is shown in Figure 4.

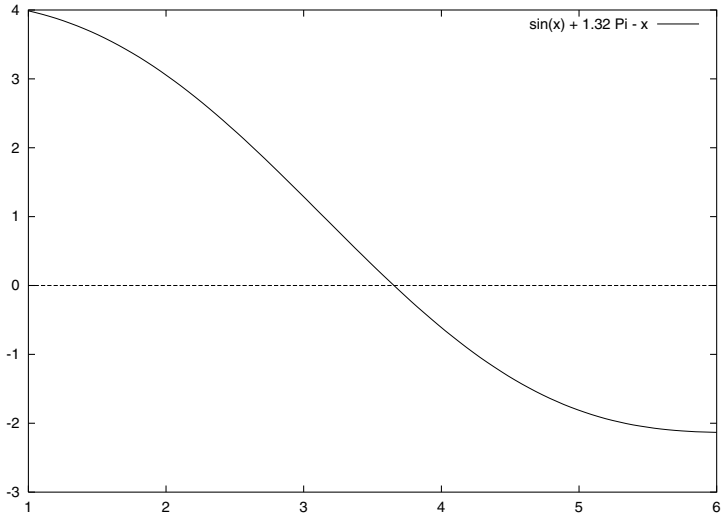

Figure 4. Example 7.2.1: Transformed Fixed-Point relation for the root finding of a one-dimensional problem. α_0 is the intersection point of the two graphs.

Table 4. Example 7.2.1: Comparison of different Newton methods for the determination of α_0.

Newton Method	Result	Steps	Time
Floating Point	3.655403079564624	4	0.0000166s
DA	3.655403079564624	1	0.0016249s
Interval	[3.655403079564622, 3.655403079564625]	4	0.0000390s
Taylor model	[3.655403079564619, 3.655403079564628]	1	0.0048408s

The various Newton methods have been used to determine the fixed point by solving the corresponding root-finding problem. The initial enclosure of the zero has been $[3.3, 4.3]$ (and the corresponding midpoint has been used as a starting value for the non-verified methods). The computations have been performed in 19-th order and the accuracy goal was 10^{-14}. The final enclosures of α_0 are shown in Table 4.

It should be noted that the high-order methods reach the desired accuracy after just one step. The traditional algorithms on the other hand need four iterations to achieve the desired precision for this rather simple function. However, since the computational complexity of Taylor models exceeds the one of intervals, and since all methods work without domain splitting, the traditional algorithms are still more efficient.

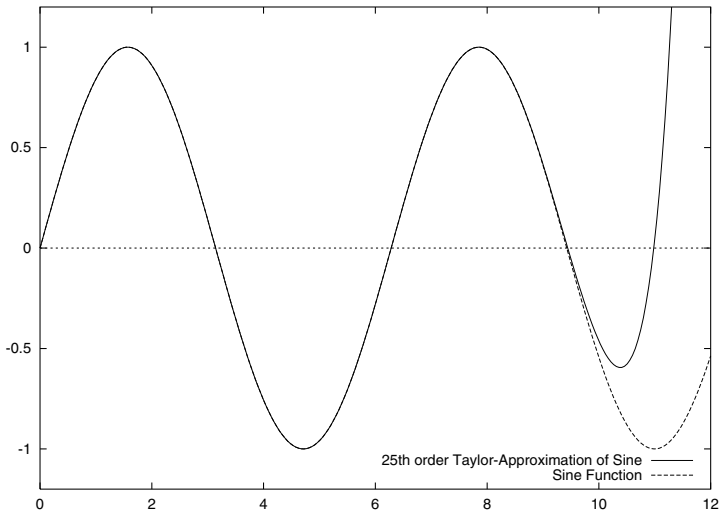

Figure 5. Example 7.2.2: 25-th order Taylor-Approximation of the sine function over the interval $[0, 12]$.

7.2.2. *Determining* π

The following one-dimensional example uses the verified Newton methods to enclose the value of π to almost machine epsilon—in fact the desired accuracy goal has been set to 10^{-12}. To that end, the sine function has been approximated by its 25-th order Taylor polynomial P, and the two Newton methods have been used to determine the unique zero of P in the interval $[1.8, 4]$. Note that the order of the approximation has been chosen in such a way that the value of that unique zero and π agree up to machine epsilon.

$$P(x) = \sum_{k=0}^{12} (-1)^k \frac{x^{2k+1}}{(2k+1)!}.$$

Due to the alternating sign of the terms, the naive implementation and evaluation of P exhibits strong cancellation. While it is clear that more efficient methods for the approximation of the sine function exist, this particular setup is typical for applications where the functional dependence is often given by complicated and redundant expressions that exhibit strong cancellation. While interval methods are known to have problems with this, the Taylor model approach can successfully avoid these problems.

Figure 5 shows the sine function and its 25-th order approximation between 0 and 12. In the region around 3 that is most interesting for this problem, the two functions agree to a very high degree.

This example illustrates how Taylor models can successfully control cancellation, and how the high-order methods converge much faster than the regular

Table 5. Example 7.2.2: Enclosures of π, as obtained by the interval Newton and Taylor model methods.

Method	Step	Enclosures of π
Interval Newton	0	[1.800000000000000, 4.000000000000000]
	1	[2.919077672711509, 4.000000000000001]
	2	[2.919077672711509, 3.427717355896160]
	3	[2.919077672711508, 3.165511666111006]
	4	[3.085798724266786, 3.165511666111007]
	5	[3.136629538119056, 3.154844930481172]
	6	[3.141102949424404, 3.141988434766514]
	7	[3.141592414707353, 3.141592894936869]
	8	[3.141592653589790, 3.141592653589801]
Taylor models	0	[1.800000000000000, 4.000000000000000]
	1	[3.141592653589765, 3.141592653589826]

first-order Newton methods. Over the same initial domain, it takes the regular interval Newton method 8 iterations to enclose the zero with the desired accuracy. The Taylor model method on the other hand, achieves a comparable accuracy in just a single step (Table 5). Moreover, the interval Newton method had to resort to extended interval divisions and subsequent domain splitting in the first four of the eight steps.

7.2.3. *Multidimensional Example*

This example illustrates the ability of the Taylor model based Newton method to easily extend to higher dimensional problems. The example shows the determination of the zeros of the previously introduced 6 dimensional exponential function (c.f. Section 7.1.2), with an initial enclosure of the zeros given by the domains $[-0.25, 0.25]^6$ (Taylor model) and $[-0.02, 0.02]^6$ (Gauss-Seidel). The computation has been performed with 8-th order Taylor models and the desired accuracy has been set to 10^{-12}. The enclosures returned after each step of the iteration are listed in Table 6.

First, it should be noted that the Taylor model based Newton method generalizes easily from the one-dimensional case to higher dimensional problems. Furthermore, the rapid convergence of the method over a relatively large domain demonstrates the power of the high-order approach. It converges in only two steps to an extremely accurate enclosure of the zero. But even more striking, the second step improves the sharpness by about 11 orders of magnitude. Once again, this is a vivid example of how the accuracy of Taylor models scales with the $(n+1)$-st order of the domain size.

Table 6. Example 7.2.3: Enclosures of the zero for the six-dimensional exponential function, as given by a Gauss-Seidel and the Taylor model method.

Method	Step	Magnitude of Enclosures of Zero
Interval Newton	0	[−0.20000000000000E−001, 0.20000000000000E−001]
	1	[−0.20000000000000E−001, 0.20000000000000E−001]
	2	[−0.20000000000000E−001, 0.20000000000000E−001]
	3	[−0.20000000000000E−001, 0.20000000000000E−001]
	4	[−0.87410175378563E−002, 0.87410175378563E−002]
	5	[−0.96852367808498E−003, 0.96852367808498E−003]
	6	[−0.54933148362165E−005, 0.54933148362165E−005]
	7	[−0.82475281503554E−010, 0.82475281503554E−010]
	8	[−0.94355649697386E−014, 0.94355649697386E−014]
Taylor models	0	[−0.25000000000000E−000, 0.25000000000000E−000]
	1	[−0.47478831445046E−003, 0.47478831445046E−003]
	2	[−0.60171167482408E−014, 0.60171167482408E−014]

The interval Gauss-Seidel, on the other hand, reached the desired accuracy in eight steps; and since the computational complexity of the Taylor models is about three orders higher than that of intervals, this improvement does not yet justify the use of Taylor models. However, the interval Gauss-Seidel required a much smaller initial enclosure of the zero to converge. Combining this with the dimensionality of the problem and assuming that the domain has to be split uniformly in each coordinate direction, the Taylor model method turns out to be two orders of magnitude more efficient than the Gauss-Seidel.

7.2.4. *Order-Dependence of the Newton Method*

This final example illustrates how the accuracy of Taylor model approach scales with the $(n+1)$-st order of the domain. For the 6-dimensional exponential function from Section 7.1.2, Figure 6 shows the magnitude of $\boldsymbol{D}^{(1)}$ after just one iteration of the 6-th order Taylor model Newton method as a function of $\boldsymbol{D}^{(0)}$.

The data show how that accuracy of the method does indeed scale with the 7-th order of the initial domain $\boldsymbol{D}^{(0)}$. For purposes of illustration, a line of slope 7 has been fitted to the data points. Deviations from that theoretical line for small domains are due to limits imposed on the verified computations by the machine epsilon. For large initial domains on the other hand, $\boldsymbol{D}^{(1)}$ is bigger than the theoretical line would imply. This is caused by a decreasing accuracy of the 6-th order Taylor polynomial approximations over the initial domain. Because of that, the Taylor models operate in a region where the rules of $(n+1)$-st order convergence fail to apply.

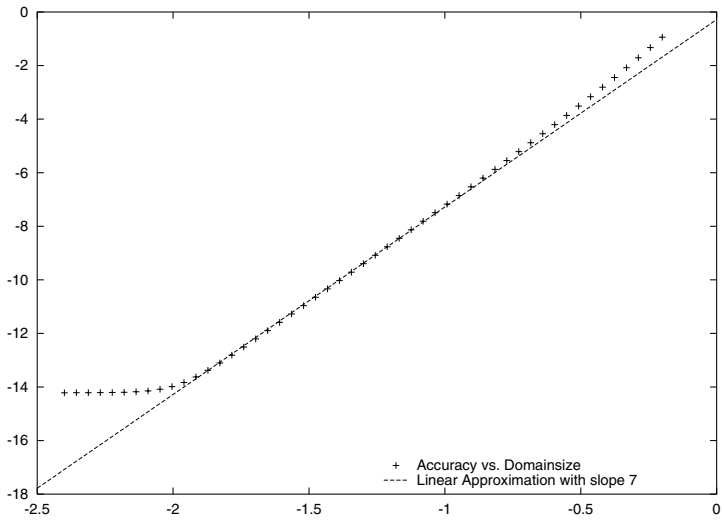

Figure 6. Example 7.2.3: Double-logarithmic plot of the accuracy of the first Newton step as a function of initial domain size; results for 6-th order Taylor models.

8. Conclusion

A new method for computing verified inverses of arbitrary complicated multidimensional functions has been presented. It has been demonstrated in the computation of verified enclosures of the inverse function for given invertible functions. The method combines Taylor models, a recently developed method for proving invertibility [9], and efficient differential algebra (DA) algorithms that allow the determination of exact n-th order polynomial inverses in finitely many steps.

The accuracy of the resulting enclosures scales with the $(n + 1)$-st order of the original domain and the resulting models of the inverse functions are extremely accurate. Moreover, the method can handle complicated functions by using Taylor models. These have been shown to control the dependency [12] and cancellation problems inherent in regular interval arithmetic. The latter cancellation is especially important in the establishment of invertibility that precedes the computation of the inverse Taylor model [9].

As an immediate application of the new method, an extension of the Interval Newton method has been presented. It has been demonstrated to converge much faster than the regular interval based version of that algorithm. Future applications of the new method will be in the computation of charts of constraint manifolds, used for verified Taylor model based DAE integration. Other applications are in the computation of verified solutions to non-linear equations. These will be used in a Taylor model based integrator for implicit Differential Equations, where the verified inversion is used to transform the given implicit equation to an explicit differential equation that can be solved using existing verified integration techniques [8].

References

1. Berz, M.: *COSY INFINITY Version 8 Reference Manual*, National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, MI 48824, MSUCL-1088, 1997, see also `http://www.cosy.nscl.msu.edu`.
2. Berz, M.: Differential Algebraic Description of Beam Dynamics to Very High Orders, *Particle Accelerators* **24** (1989), p. 109.
3. Berz, M.: Differential Algebras with Remainder and Rigorous Proofs of Long-Term Stability, in: *Fourth Computational Accelerator Physics Conference, AIP Conference Proceedings*, **391**, 1996, p. 221.
4. Berz, M.: *Modern Map Methods in Particle Beam Physics*, Academic Press, San Diego, 1999.
5. Berz, M. and Hoffstätter, G.: Computation and Application of Taylor Polynomials with Interval Remainder Bounds, *Reliable Computing* **4** (1) (1998), pp. 83–97.
6. Berz, M. and Hoffstätter, G.: Exact Bounds of the Long Term Stability of Weakly Nonlinear Systems Applied to the Design of Large Storage Rings, *Interval Computations* 2 (1994), pp. 68–89.
7. Berz, M. and Makino, K.: New Methods for High-Dimensional Verified Quadrature, *Reliable Computing* **5** (1) (1999), pp. 13–22.
8. Berz, M. and Makino, K.: Verified Integration of ODEs and Flows with Differential Algebraic Methods on Taylor Models, *Reliable Computing* **4** (4) (1998), pp. 361–369.
9. Hoefkens, J. and Berz, M.: Verification of Invertibility of Complicated Functions over Large Domains, *Reliable Computing* **8** (2002), in print.
10. Hoefkens, J., Berz, M., and Makino, K.: Efficient High-Order Methods for ODEs and DAEs, in: *Proceedings, Third International Workshop on Automatic Differentiation*, 2000.
11. Makino, K.: *Rigorous Analysis of Nonlinear Motion in Particle Accelerators*, Michigan State University, East Lansing, Michigan, USA, 1998, also MSUCL-1093 and `http://bt.nscl.msu.edu/pub`.
12. Makino, K. and Berz, M.: Efficient Control of the Dependency Problem Based on Taylor Model Methods, *Reliable Computing* **5** (1) (1999), pp. 3–12.
13. Makino, K. and Berz, M.: Remainder Differential Algebras and Their Applications, in: Berz, M., Bischof, C., Corliss, G., and Griewank, A.: *Computational Differentiation: Techniques, Applications, and Tools*, SIAM, Philadelphia, 1996, pp. 63–74.
14. Moore, R. E.: *Interval Analysis*, Prentice-Hall, Englewood Cliffs, 1966.
15. Moore, R. E.: *Methods and Applications of Interval Analysis*, SIAM, Philadelphia, 1979.