Reliable Computing **8:** 67–82, 2002.

Verification of Invertibility of Complicated Functions over Large Domains

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

(Received: 8 September 2000; accepted: 30 January 2001)

Abstract. A new method to decide the invertibility of a given high-dimensional function over a domain is presented. The problem arises in the field of verified solution of differential algebraic equations (DAEs) related to the need to perform projections of certain constraint manifolds over large domains. The question of invertibility is reduced to a verified linear algebra problem involving first partials of the function under consideration. Different from conventional approaches, the elements of the resulting matrices are Taylor models for the derivatives of the functions.

The linear algebra problem is solved based on Taylor model methods, and it will be shown the method is able to decide invertibility with a conciseness that often goes substantially beyond what can be obtained with other interval methods. The theory of the approach is presented. Comparisons with three other interval-based methods are performed for practical examples, illustrating the applicability of the new method.

1. Introduction

In [3], [4], a method involving high order Taylor polynomials with remainder bound has been presented that allows verified computations while avoiding some difficulties inherent in normal interval arithmetic. This Taylor model approach guarantees inclusion of functional dependencies with an accuracy that scales with the $(n + 1)$-st order of the domain over which the functions are evaluated.

In particular, as shown in [10], this method can often substantially alleviate the following problems inherent in naive interval arithmetic:

- Sharpness of the Result,
- Dependency Problem,
- Dimensionality Curse.

The method has recently been used for a variety of applications, including verified bounding of highly complex functions [2], [5], solution of ODEs under avoidance of the wrapping effect for practical purposes [8], and high-dimensional verified quadrature [7].

In this paper we will combine these techniques with a fresh look at the mathematics of invertibility to

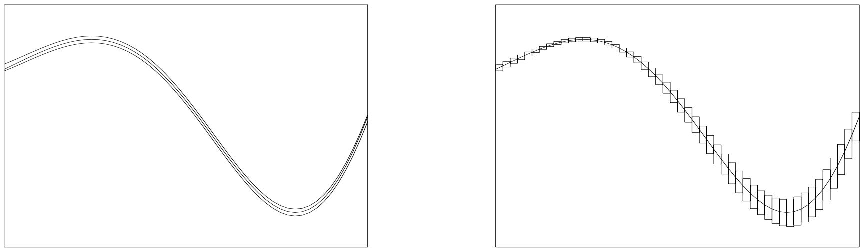

Figure 1. Left: Enclosing a function by a Taylor model of order eight. Right: Interval bounding of the same function (figures courtesy of Kyoko Makino).

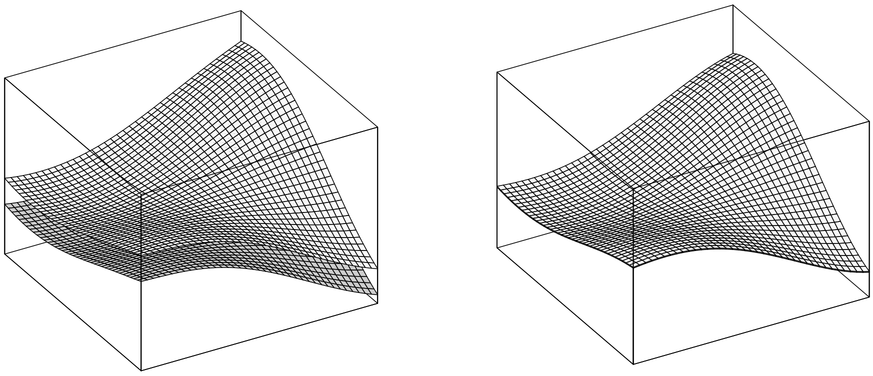

Figure 2. Interval bounding of a function by a Taylor model; orders 7 (left) and 10 (right).

- derive linear algebra tools that allow utilization of ability of Taylor models to provide sharpness;
- obtain very sharp bounds for first derivatives.

For purposes of illustration, Figure 1 compares the enclosure of a function by Taylor models and regular intervals. In this case, the virtue of the method lies in the fact that one Taylor model over a relatively large domain box guarantees a sharpness that interval bounding cannot even achieve with multiple smaller domains. This helps significantly in fighting the dimensionality curse inherent in interval bounding where it is of prime importance to avoid subdivisions of the domain boxes.

Figure 2 illustrates how the fact that the sharpness of Taylor models scales with the $(n + 1)$-st order of the domain allows to obtain sharp bounds quickly, even in the multidimensional case.

Recently we have begun work on a verified solver for Differential Algebraic Equations that uses a Taylor model based verified ODE solver [8] and verified inversion of implicit description of the constraint manifolds obtain verified enclosures for the solutions of the equations. Loosely the approach can be phrased as follows:

Given a function $f : \mathbb{R}^v \to \mathbb{R}^v$ (known only up to some accuracy) defined over a box $\boldsymbol{D}$, is the function invertible over its range $\boldsymbol{f^u}(\boldsymbol{D})$? And if so, find a representation of the inverse as accurately as possible.

The first step in this process of finding local coordinate representations of the constraint manifolds is the determination of invertibility of the constraints over the domain under consideration. In this paper we will derive new methods that give guaranteed answers to the question of whether a function is invertible over a given domain. We will first derive an interval arithmetic based method that requires only knowledge about first derivatives to determine invertibility. Capitalizing on some special structure of the Jacobians appearing in the approach, we then will extend this method to use high order Taylor models and show that it scales much better to high dimensional problems over larger domains than normal interval based algorithms. We will also utilize the fact that the Taylor models can be used to model the function and the derivatives much more accurately over larger domains than regular interval arithmetical methods.

2. Verified Invertibility from First Derivatives

There are a variety of ways to decide invertibility for a given vector function f, some of which rely on second- and higher order derivatives. These methods are computationally expensive since they require bounding complicated functions like the operator norm of the second derivative map [1] over the domain under consideration. We will present a method that requires only bounds on first partial derivatives to decide the question of invertibility. Furthermore and perhaps more importantly, the method exhibits some important structure regarding the points at which the derivatives actually have to be evaluated, which we will later capitalize on to significantly reduce cancellation problems in the necessary verified linear algebra. It should be noted that this method is not local in nature, but can indeed guarantee global invertibility everywhere in the given domain.

The following theorem enables us to decide whether a given function is invertible over a domain by just evaluating first derivatives and hence greatly reduces the computational overhead necessary. It seems to originate in works by K. Kovalevsky from the beginning of the twentieth century but has been “rediscovered” by many others (e.g. [13, E 5.3-4]).

THEOREM 2.1 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* $\boldsymbol{f^u}(\boldsymbol{B})$, *where* $\boldsymbol{f^u}(\boldsymbol{B})$ *denotes the range of* f *over* $\boldsymbol{B}$.

The proof of Theorem 2.1 rests on auxiliary functions $h_i(t) = f_i(x_1 \cdot (1-t) + x_2 \cdot t)$ that satisfy $h_i(0) = h_i(1)$. Applying the Mean Value Theorem to the h_i and using the regularity of M yields the result.

While regularity of the Jacobian at each point in the domain is a necessary condition for invertibility, it is in general not a sufficient condition. Thus, it may perhaps be worthwhile to note that Theorem 2.1 does not try to deduce invertibility by using the fact that the Jacobian is non-singular at each point within the domain. The theorem rather allows to guarantee invertibility over the whole domain by utilizing that the Jacobian as a function of v^2 variables is regular over $\boldsymbol{B}^v$.

An immediate corollary to Theorem 2.1 is the following interval formulation of necessary and sufficient conditions for the existence of the inverse function, which in this version is very well known (e.g. Theorem 5.1.6 in [12] and [15]).

COROLLARY 2.1. *Let* f *and* $\boldsymbol{B}$ *be as in Theorem 2.1. For* $i, j = 1, \ldots, v$ *let* $\boldsymbol{p}_{i,j} \subset \mathbb{R}$ *be compact intervals such that*

$$\frac{\partial f_i}{\partial x_j}(x) \in \boldsymbol{p}_{i,j} \qquad \forall\, x \in \boldsymbol{B}.$$

If the interval matrix $\boldsymbol{P} = (\boldsymbol{p}_{i,j})$ *is regular, then* f *has a* $\mathcal{C}^1$*-inverse defined on* $\boldsymbol{f^u}(\boldsymbol{B})$.

An algorithm based on Theorem 2.1 and Corollary 2.1 requires efficient and accurate methods to model partial derivatives and depends on the availability of efficient methods for the determination of regularity of an interval matrix with potentially large entries. Both these problems may become challenging for naive interval methods which have difficulties to model complicated functions (i.e. one usually has significant overestimation in bounding the derivatives of the functions) and suffer from the well known problems that come with an increasing number of variables and an increase in domain sizes.

Moreover, interval methods are not particularly well suited for practical applications of Theorem 2.1 via Corollary 2.1 because of difficulties in handling the fact that the individual rows of the Jacobian are all evaluated at one point. A new Taylor model based method that solves this will be presented in Section 4. But nevertheless, once the interval matrix $\boldsymbol{P}$ of Corollary 2.1 has been determined as accurately as possible, the next question is how to establish regularity of it. While in lower dimensions it may often be sufficient to compute the interval determinant of the matrix, more sophisticated methods are needed for higher dimensional problems.

It has been shown that regularity of a matrix is equivalent to the componentwise distance to the next singular matrix being greater than one [14]. In the following paragraphs we present two advanced interval based methods to prove the regularity of the interval Jacobian. We will use them in the following to compare interval based methods for the determination of invertibility with new approaches based on Taylor models.

The first method of verifying the regularity of an interval matrix $\boldsymbol{A} = \hat{A} + [-\Delta, \Delta]$ can be deduced from the singular value decompositions of both the midpoint matrix $\hat{A}$ and the Δ-matrix. If $\sigma(M, i)$ denotes the i-th singular value of a matrix M (sorted in non-increasing order), then the following assertion holds [16], [17]:

THEOREM 2.2. *Given* $\boldsymbol{A} = \hat{A} + [-\Delta, \Delta]$. *Then* $\sigma(\Delta, 1) < \sigma(\hat{A}, n)$ *implies that* $\boldsymbol{A}$ *is non-singular.*

The success of this method depends mostly on the sharpness of the models of the partial derivatives, and as such this method may have difficulties scaling with dimensionality and complexity of the functions of interest.

The next theorem has been shown in [12], and it often provides rather good estimates on the regularity of the given interval matrix A.

THEOREM 2.3. *Given* $\boldsymbol{A} = \hat{A} + [-\Delta, \Delta]$. *If* $\hat{A}$ *is regular let* S *be an approximate inverse of* $\hat{A}$. *If the spectral radius* $\rho(I - S \cdot \boldsymbol{A})$ *is less than* 1 *then any matrix in* $\boldsymbol{A}$ *is non-singular.*

This method uses a preconditioning of the matrix $\boldsymbol{A}$ to establish invertibility of a better conditioned matrix $S \cdot \boldsymbol{A}$. This allows the method to work very well for medium sized problems, but as will become clearer below, in higher dimensions this method still suffers from the fact that the entries of the product matrix $S \cdot \boldsymbol{A}$ are computed from addition and subtraction of intervals and as such are subject to significant overestimations.

The most practical method to establish the contraction of an interval matrix $\boldsymbol{A}$ is to start out with an arbitrary non-empty interval vector $\boldsymbol{x}_0$ and iterate $\boldsymbol{x}_{k+1} = \boldsymbol{B} \cdot \boldsymbol{x}_k$ with $\boldsymbol{B} = I - S \cdot \boldsymbol{A}$. If for any $k \in \mathbb{N}$ we can show that $\boldsymbol{x}_{k+1} \subset \boldsymbol{x}_k$ we have shown that $\boldsymbol{B}$ is indeed contracting and hence $\rho(\boldsymbol{B}) < 1$.

3. Taylor Models

In the following we will develop criteria asserting invertibility utilizing properties of the Taylor model approach introduced in [3], [4], [11]. Specifically, we define

DEFINITION 3.1 *Taylor Model.* 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 $(n, v, w \in \mathbb{N})$ and $R \subset \mathbb{R}^w$ be an open non-empty set. Then $(P, x_0, \boldsymbol{D}, R)$ is called a Taylor model of order n with expansion 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 3.2 *Taylor Models as Sets of Functions.* Let $T = (P, x_0, \boldsymbol{D}, R)$ 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 $f(x) - P(x) \in R$ for all $x \in \boldsymbol{D}$, and the n-th order Taylor series of f around x_0 equals P.

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 arithmetic operations on Taylor models that preserve the defining inclusion relationships and hence to obtain Taylor models for any smooth computer function. The next theorem shows how this can be done for the basic elementary operations. In the following, let P and $\boldsymbol{D}$ be as above and denote by $B(P, \boldsymbol{D})$ a guaranteed enclosure of the range of P over the box $\boldsymbol{D}$. Moreover, for $k \in \mathbb{N}$, $P_{(\leq k)}$ and $P_{(> k)}$ shall be the parts of P of orders up to k and greater than k, respectively.

THEOREM 3.1. *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 as above and define*

$$R_P = R_1 \cdot R_2 + R_1 \cdot B(P_2, \boldsymbol{D}) + B(P_1, \boldsymbol{D}) \cdot R_2 + B\big((P_1 \cdot P_2)_{(> n)}, \boldsymbol{D}\big).$$

Obtain new Taylor models T_S *and* T_P *by*

$$\begin{aligned} T_S &= (P_1 + P_2,\ x_0,\ \boldsymbol{D},\ R_1 + R_2), \\ T_P &= \big((P_1 \cdot P_2)_{(\leq n)},\ x_0,\ \boldsymbol{D},\ R_P\big). \end{aligned}$$

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* $f_2 \in T_2$*, we get*

$$(f_1 + f_2) \in T_S \quad \textit{and} \quad (f_1 \cdot f_2) \in T_P.$$

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

$$\big((f_1 + f_2) - (P_1 + P_2)\big)(x) = \delta_1(x) + \delta_2(x) \in R_1 + R_2 = R_S.$$

Since the n-th 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 $\boldsymbol{D}$it is

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

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

More information on arithmetic and intrinsic functions on Taylor models can be found in [4], [9].

It should also be noted that with availability of methods to obtain Taylor models of functions we also have means to compute tight enclosures for the derivatives of those functions by using either hand coded derivatives, automatic differentiation or new methods still under development that propagate not only $\mathcal{C}^0$ remainder terms in Taylor models but also take bounds on derivatives into account by demanding that the functions contained in such an extended Taylor model are also $\mathcal{C}^1$-close to the reference polynomial.

4. Invertibility via Taylor Models

We will now come back to the question whether a given function f, defined over some domain box $\boldsymbol{D}$, is invertible over its image. We will combine the results of Theorem 2.1 with the regularity criterion Theorem 2.3 and Taylor model based techniques to derive a new algorithm to decide whether a given function f is invertible. Compared to the corresponding interval version as presented in Corollary 2.1 we will be able to avoid an overly pessimistic behavior.

The proof of Theorem 2.1 is based on proving regularity of the matrix

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

with points $\chi_1, \ldots, \chi_v \in \boldsymbol{D}$. Thus, any combination of entries from the same row of M can be evaluated using the same set of domain variables. While interval arithmetic cannot capitalize on this intrinsic structure, as we shall see in the following, Taylor models are particularly well suited for this task. Instead of modeling each partial derivative of the functions f_i by intervals we now model each of the gradients by one single Taylor model and we are left with the task of proving regularity of a matrix valued Taylor model, which we will be doing using a method derived from Theorem 2.3.

1. For $i, j = 1, \ldots, v$ let $T_{i,j}$ be Taylor models for the partial derivatives of f, i.e.

$$\frac{\partial f_i}{\partial x_j} \in T_{i,j}.$$

In practice, the Taylor models $T_{i,j}$ can be obtained for example by evaluating the first order AD code belonging to f in Taylor model arithmetic. On the side we mention that currently, extensions of the Taylor model approach are being developed that allow the determination of the $T_{i,j}$ directly from the extended Taylor model of f.

As outlined earlier all Taylor models that have the same index i (belong to the same row of the matrix) can be evaluated using the same set of domain variables. The next step now is to compute a suitable preconditioning matrix for the transpose of T. So instead of showing that the matrix T is regular, we will show the equivalent statement that the transpose of T is regular.

2. Let $T_{\mathbb{R}}$ be an interval enclosure of the range of the Taylor models $T_{i,j}$, i.e. $g \in T_{i,j} \Rightarrow g(x) \in T_{\mathbb{R},i,j}$ for all $x \in \boldsymbol{D}$.
3. Let $\hat{T}_{\mathbb{R}}$ be the midpoint matrix of $T_{\mathbb{R}}$ and similarly for the transpose $\hat{T}^T_{\mathbb{R}}$.
4. Let S be an approximate floating point inverse of $\hat{T}^T_{\mathbb{R}}$.
5. Compute a new Taylor model N by $N_{i,j} = \delta_{i,j} - \sum_{k=1}^{v} S_{i,k} T^T_{k,j} = \delta_{i,j} - \sum_{k=1}^{v} S_{i,k} T_{j,k}$, where $\delta_{i,j}$ denotes the Kronecker Delta.

The advantage of working with the transpose T^T is that the resulting entries of the matrix N are all computed using linear combinations of Taylor models that can be evaluated over the same set of domain variables. Hence each entry of N is again a Taylor model in v variables, and the columns of N are now modeled over the same set of domain variables. As a consequence, the sum $\sum_{k=1}^{v} S_{i,k} T_{j,k}$ can be evaluated in Taylor model arithmetic, and in particular any blow-up is suppressed, as in other Taylor model computations.

The final step of the method is then to show that the spectral radius of the resulting matrix-valued Taylor model N is less than one. To establish this, the image of the unit ball (in the maximum norm) under the map N is considered. If it can be shown that the image is properly contained in the unit ball, invertibility is verified.

Similar to above, it proves advantageous to not work with the matrix N, but rather with its transpose N^T. In this way, mixing of Taylor models from different columns of N is again avoided, which allows suppression of blow up in the necessary linear algebra. This is possible since the transposed matrix has the same spectral radius as the original one. As a final application of Taylor models, we model the unit ball by Taylor models as follows: for any small $\varepsilon > 0$, the identity function over the interval $[-1, 1]$ is contained in $(x, 0, [-1, 1], (\varepsilon, \varepsilon))$. Thus, if we denote the $\boldsymbol{i}$-th such Taylor model by e_i, a bound on the range of e_i contains the closed unit interval in x_i. Then the algorithm proceeds as follows:

6. For $i = 1, \ldots, v$ compute Taylor models r_i with the Taylor models e_{v+j} that are defined over new domains, independent of those of the $N_{i,j}$'s:

$$r_i = \sum_{j=1}^{v} N^T_{i,j} \cdot e_{v+j} = \sum_{j=1}^{v} N_{j,i} \cdot e_{v+j}.$$

7. Find interval enclosures $\boldsymbol{I}_i$ for the Taylor models r_i.

It should be noted that due to the special nature of the Taylor models r_i, the bounding is actually quite simple: each of the v terms of the sum that constitutes the Taylor models r_i has a different set of variables. Thus, there will not be any cancellation in the sum, and hence the sharpest possible bound of the sum equals the sum of the bounds of the individual terms. Moreover, the latter are just Taylor models multiplied by the interval $[-1, 1]$, and hence their ranges are just the ranges of the Taylor model multiplied by the interval $[-1, 1]$.

8. If $\boldsymbol{I}_i \subset [-1, 1]$ for all $i = 1, \ldots, v$ the original function f is invertible.

To summarize, this method utilizes Taylor models at two crucial points. Firstly Taylor models are used to model the partial derivatives of the function f. As will be seen later this helps tremendously in fighting the issues of complexity and cancellation in modeling the derivatives. Moreover the Taylor models are used to compute an enclosure for the derivative matrices under consideration, and they are used for all arithmetic operations on these matrices which minimizes the overestimation due to cancellation, since the bulk of the functional dependence is propagated in the reference polynomial and therefore not subject to cancellation [10].

Finally we have been able to modify the computation in such a way that all arithmetic operations on Taylor models can be performed using only v variables, which allows for a very favorable computational complexity and allows the method to scale very well to high dimensional problems.

5. Examples

As outlined before we are mainly interested in proving invertibility for a given function over domains as large as possible. But not only the sheer size of the domains of invertibility matters, but an equally important question is how good we are able to fill out the maximum region of invertibility from within, i.e. how close to a critical point can we still prove invertibility. Moreover in the context of Differential Algebraic Equations we are likely to face high dimensional problems. The following examples will demonstrate how the new Taylor model based method performs in each of these areas as compared to interval based methods.

In each of the following examples we have tested a certain number of functions for invertibility over certain domains. All tests have been performed with each of the following methods and the results are represented in graphs and discussed. Note that the determinant test has only been used for problems of up to 6 variables.

(1) Interval tests based on Corollary 2.1 and an interval determinant;

(2) Interval tests based on Corollary 2.1 and Theorem 2.2;

(3) Interval tests based on Corollary 2.1 and Theorem 2.3;

(4) Taylor model based tests as presented in Section 4.

For the interpretation of the following results it is important to note that in all cases the necessary bounds on range enclosures have been computed using regular interval arithmetic. No dedicated range bounders utilizing domain decomposition etc. have been used either in the interval or Taylor model approaches to keep the computational overhead within reasonable limits.

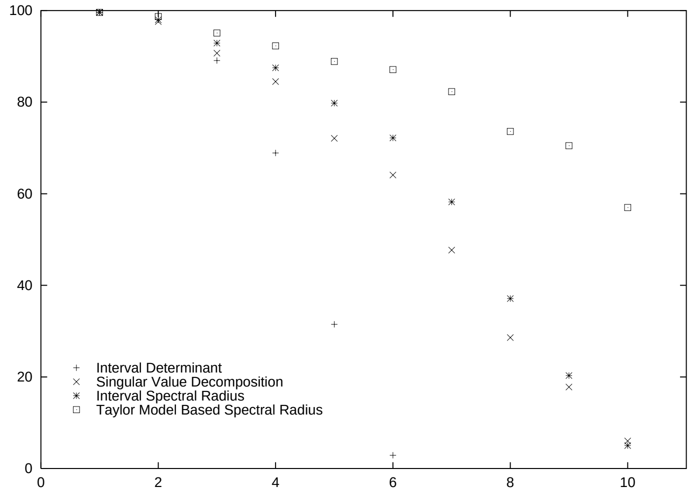

Figure 3. Example 5.1: Percentage of random functions that can be shown to be invertible as a function of dimensionality.

5.1. INVERTIBILITY AS A FUNCTION OF DIMENSIONALITY

The first example illustrates how the performance of the presented methods behaves as a function of dimensionality. Invertibility of one thousand 15-th order polynomials with uniformly distributed random coefficients in $[-1, 1]$ has been tested with the methods 1 to 4. The domain of the random polynomials has been $[-0.005, 0.005]^v$ for $v = 1, \ldots, 10$. Figure 3 shows the percentage of these random polynomials that could be verified to be invertible as a function of dimensionality. The Taylor models in this simulation have been of 10-th order and the resulting remainder bounds are in the order of 10^{-12}.

As it is to be expected for increasing dimensionalities the number of successfully established invertible polynomials is decreasing. But it is important to note that the Taylor model based method performs much better in establishing invertibility and suffers much less from an increase in the number of variables than the interval based methods do. Unfortunately there is no good way to assess what fraction of the original functions are truly invertible; but it is to be expected that this fraction decreases with the dimension, accounting for the drop of predicted invertibility in all approaches.

5.2. INVERTIBILITY AS A FUNCTION OF DOMAIN SIZE

For the next example we have investigated how the presented methods of establishing invertibility behave as functions of the domain size. We have restricted ourselves to the medium sized problem of 10-th order polynomials in 6 variables. For each polynomial we have applied the methods 1 to 4 over domains of increasing size. All

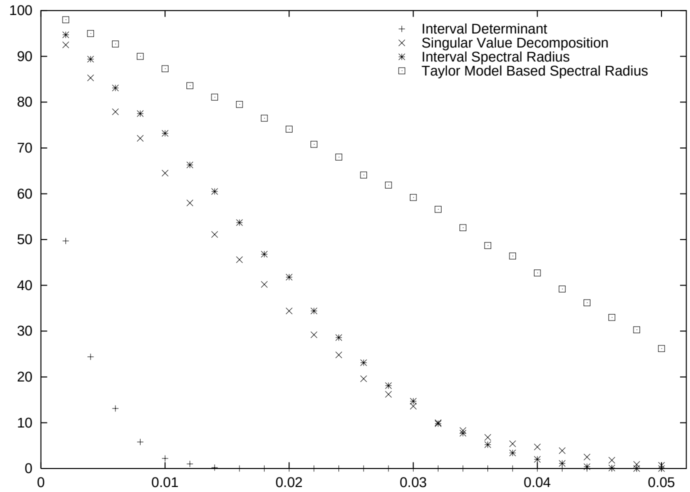

Figure 4. Example 5.2: Percentage of random functions that can be shown to be invertible as a function of domain size.

domain boxes have been placed symmetrically around the origin and the abscissa shows the magnitude of the domains. For each of the different domain sizes we have plotted the percentage of polynomials for which invertibility could be proven using the four different schemes. Like in the previous example the computation has been performed with one thousand random polynomials and the Taylor models have been extended with artificial remainder bounds in the order of 10^{-12}. The results of this simulation are shown in Figure 4.

It has been shown that the accuracy of Taylor models scales with the $(n+1)$-st order of the domain size [6], while interval arithmetic scales approximately linearly with the domain sizes. That explains why the Taylor model based method can assert invertibility over domain that are much larger than the ones interval based methods can handle. It should also be noted that the number of invertible polynomials decreases almost linear with increasing domain size.

5.3. Invertibility as a Function of Non-Linearity

This example demonstrates how the performance of the discussed methods changes with the non-linearity of the functions under consideration. To that end we have considered functions $f = (f_1, \dots, f_v) : [-1, 1]^v \subset \mathbb{R}^v \to \mathbb{R}^v$ given by

$$f_i(x_1, \dots, x_v) = \frac{\sum_{k=1}^{v} a_{i,k} x_k}{\left(1 - \sum_{k=1}^{v} b_{i,k} x_k\right)^2 + \varepsilon^2}$$

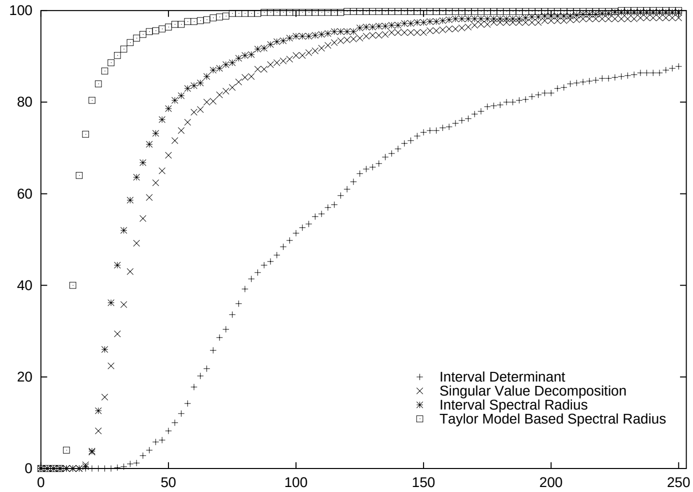

Figure 5. Example 5.3: Percentage of functions that can be shown to be invertible as a function of the non-linearity ε.

with $\varepsilon \in \mathbb{R}$ and $v \times v$ matrices $A = (a_{i,k})$ and $B = (b_{i,k})$ with coefficients in $[-1, 1]$. For small values of ε and appropriate coefficients $a_{i,k}$ and $b_{i,k}$, this may give ill-defined functions because of vanishing denominators. However, in these cases we count the functions as not invertible.

We have generated 500 functions by choosing the coefficients of the matrices A and B randomly in the interval $[-1, 1]$, and the presented methods have been used to prove invertibility over $[-1, 1]^6$. Figure 5 shows the percentage of functions that can be shown to be invertible depending on the non-linearity ε. All computations have been performed in 6-th order and the value of ε has been varied from 0 to 250 in steps of 2.5.

The non-linearity of these functions is mostly determined by the quantity ε, such that for $\varepsilon \gg 1$, the functions are linearly dominated and their invertibility depends mostly on the invertibility of A. Since almost all computer generated random matrices are invertible, it is to be expected that all methods succeed in proving invertibility for sufficiently large ε. Figure 5 shows that this is indeed the case.

For $\varepsilon \sim 1$ the resulting functions show a large non-linearity and all methods fail in proving invertibility. It is likely that due to the size of the domain box and the high degree of non-linearity, these functions are actually not invertible.

With increasing ε, the different methods become more successful in establishing invertibility. However, the Taylor model based method starts proving invertibility very suddenly for $\varepsilon \simeq 20$, while the success of the other methods sets in only for larger ε, and increases at a much slower rate.

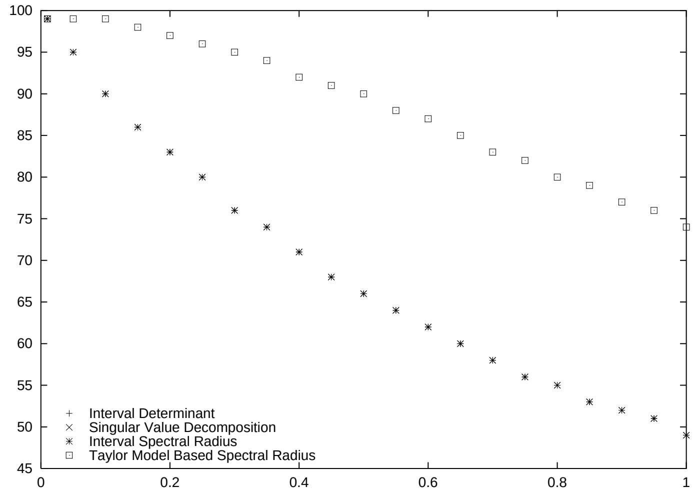

Figure 6. Example 5.4: Percentage of maximal domain size as a function of complexity and non-linearity in the partial derivatives.

5.4. Invertibility in the Vicinity of a Critical Point

As another example of how the presented methods scale to larger domain sizes and how they behave in the neighborhood of a singular point, consider the function $f = (f_1, \ldots, f_6) : \mathbb{R}^6 \to \mathbb{R}^6$ with components f_i for $i = 1, \ldots, 6$ given by

$$\begin{aligned} f_i(x_1, \ldots, x_6) \;=\; & (x_i - x_0)^2 + \lambda \cos\big((x_i - x_0)\cdot(x_{\pi(i)} - x_0)\big) \\ & - \lambda (x_{\pi(i)} - x_0)\sin(x_i - x_0), \end{aligned}$$

where $\pi(i) = i + 1$ for $i = 1, \ldots, 5$ and $\pi(6) = 1$.

At the point $(x_0, \ldots, x_0)$ the Jacobian of f vanishes and hence for symmetric domain boxes centered at the origin, $2x_0$ is an upper bound for the magnitude of the domain of invertibility. Figure 6 shows the percentage of this maximal magnitude $2x_0$ over which the different methods can prove invertibility as a function of the parameter λ, which is gradually increased from 0 to 1. For $x_0 = 0.1$, the invertibility tests have been performed for the domains $[-p\cdot 0.001, p\cdot 0.001]^6$ with $p = 1, \ldots, 100$. The plot shows the maximal value of p for which the various methods can determine invertibility as a function of λ. The results indicate that the Taylor model based test can guarantee invertibility over a much larger region and is less sensitive to the perturbation λ.

It is important to note that all interval based methods behave equally bad for increasing λ. This indicates that the real challenge of establishing invertibility in this example lies with the accurate modeling of the derivatives and not so much with the different methods to establish regularity of the interval matrix of derivatives. This example illustrates how Taylor models can model complicated functions extremely

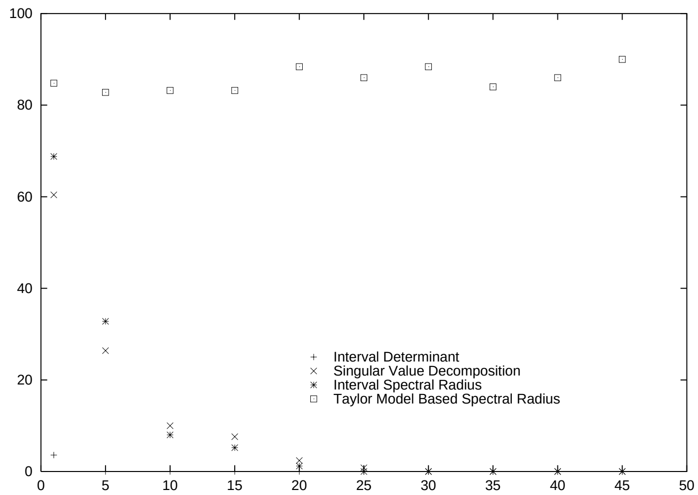

Figure 7. Example 5.5: Percentage of random functions that can be shown to be invertible as a function of functional complexity.

accurate—the remainder bounds of the Taylor models for the partial derivatives are all in the order of 10^{-12}. This aspect will be revisited in the next example.

5.5. Invertibility as a Function of Functional Complexity

Finally we would like to present an example that demonstrates how the presented methods behave as a function of computational complexity of the original function f under consideration. We have emulated functional complexity using the following method:

For the previously introduced random polynomials of order 10 in 6 variables with coefficients between -1 and 1 we have modeled computational complexity c by

$$f(x) = \frac{1}{\sqrt{c}} \sum_{i=1}^{c} f_i(x),$$

where each of the f_i is a random polynomial as before and the scaling factor has been introduced to maintain the standard deviation of the coefficients of the resulting polynomials. All tests were performed over the $[-0.005, 0.005]^6$ domain box and the results are shown in Figure 7.

It is important to note that the function f has been evaluated by adding the results of the individual polynomial evaluations. This is a realistic model of practical applications where the supplied functions often come as black boxes that do not permit any further simplifications to control cancellations.

The first observation worth noting here is that the Taylor model based method does not suffer from complexity, but rather even seems to improve with complexity. This caused by the way we simulate computational complexity. While the coefficients in the original polynomials are uniformly distributed in $[-1, 1]$ the coefficients of the resulting “complex” one are not uniform anymore, and this distributional change leads to a simpler behavior with improved chances for invertibility.

As outlined earlier Taylor models are particularly well suited to deal with the cancellation problems that arise in regular interval arithmetic, since the bigger part of that cancellation takes place in the polynomial coefficient real number arithmetic and only a small fraction of it contributes to the remainder bounds (this is opposed to normal interval arithmetic, where all the functional dependence is propagated in the “remainder” bound. As such Taylor models are expectedly much better in modeling the derivatives themselves and hence the Taylor model based methods can succeed in proving invertibility for computationally complex functions that cannot be properly modeled by intervals.

6. Conclusions

We have presented a new Taylor model based method to give a verified answer to the question of whether a given function is invertible over a certain domain of interest. This method uses only first derivatives and as such it can be implemented with little computational expense.

Using Taylor models we have been able to overcome some limitations of interval arithmetic based methods to determine invertibility. Namely we have been able to show that Taylor model based methods can model computationally complex functions much more accurately than interval methods. Moreover, since the accuracy of Taylor models scales with the $(n + 1)$-st order of the domain size, we have shown that the new methods work over significantly larger domains and scale better to high dimensional problems than interval based methods.

An important application of this method lies in the combination with recently presented methods for the computation of Taylor models of inverse functions [6]. In that case, it is necessary to verify invertibility before actually computing the inverse Taylor models.

Acknowledgements

We would like to thank several individuals for stimulating discussions and valuable suggestions, in particular the referees for providing us with helpful references, Jens von Bergmann for critical comments, and Kyoko Makino for her groundbreaking work on Taylor models. This work was supported in part by the U.S. Department of Energy under contract DE-FG02-95ER40931 and the Alfred P. Sloan Foundation.

References

1. Abraham, R., Marsden, J. E., and Ratiu, T.: *Manifolds, Tensor Analysis, and Applications*, *Applied Mathematical Sciences* **75**, Springer Verlag, second edition, 1988.
2. Berz, M.: Differential Algebras with Remainder and Rigorous Proofs of Long-Term Stability, in: *Fourth Computational Accelerator Physics Conference*, *AIP Conference Proceedings*, vol. 391, 1996, p. 221.
3. Berz, M.: *Modern Map Methods in Particle Beam Physics*, Academic Press, San Diego, 1999.
4. Berz, M. and Hoffstätter, G.: Computation and Application of Taylor Polynomials with Interval Remainder Bounds, *Reliable Computing* **4** (1998), pp. 83–97.
5. 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.
6. Berz, M. and Hoefkens, J.: Verified High-Order Inversion of Functional Dependencies and Interval Newton Methods, *Reliable Computing* **7** (5) (2001), pp. 379–398.
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. Makino, K.: *Rigorous Analysis of Nonlinear Motion in Particle Accelerators*, PhD thesis, Michigan State University, East Lansing, Michigan, 1998, also MSUCL-1093.
10. Makino, K. and Berz, M.: Efficient Control of the Dependency Problem Based on Taylor Model Methods, *Reliable Computing* **5** (1) (1999), pp. 3–12.
11. Makino, K. and Berz, M.: Remainder Differential Algebras and Their Applications, in: Berz, M., Bischof, C., Corliss, G., and Griewank, A. (eds), *Computational Differentiation: Techniques, Applications, and Tools*, SIAM, Philadelphia, 1996, pp. 63–74.
12. Neumaier, A.: *Interval Methods for Systems of Equations*, Cambridge University Press, 1990.
13. Ortega, J. M. and Rheinboldt, W. C.: *Iterative Solution of Nonlinear Equations in Several Variables*, Computer Science and Applied Mathematics, Academic Press, New York and London, 1970.
14. Rump, S. M.: Ill-Conditioned Matrices Are Componentwise Near to Singularity, *SIAM Review* **41** (1) (1999), pp. 102–112.
15. Rump, S. M.: Private Communication, 2000.
16. Rump, S. M.: Validated Solution of Large Linear Systems, in: Albrecht, R., Alefeld, G., and Stetter, H. J. (eds), *Computing Supplement* **9**, Springer-Verlag, Wien, 1993, pp. 191–212.
17. Rump, S. M.: Verification Methods for Dense and Sparse Systems of Equations, in: Herzberger, J. (ed.), *Topics in Validated Computations*, North-Holland, Amsterdam, 1994, pp. 63–135.