Reliable Computing **4:** 83–97, 1998.

Computation and Application of Taylor Polynomials with Interval Remainder Bounds

MARTIN BERZ and GEORG HOFFSTÄTTER
Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA, e-mail: berz@pilot.msu.edu

(Received: 20 February 1996; accepted: 12 May 1997)

Abstract. The expansion of complicated functions of many variables in Taylor polynomials is an important problem for many applications, and in practice can be performed rather conveniently (even to high orders) using polynomial algebras. An important application of these methods is the field of beam physics, where often expansions in about six variables to orders between five and ten are used.

However, often it is necessary to also know bounds for the remainder term of the Taylor formula if the arguments lie within certain intervals. In principle such bounds can be obtained by interval bounding of the $(n+1)$-st derivative, which in turn can be obtained with polynomial algebra; but in practice the method is rather inefficient and susceptible to blow-up because of the need of repeated interval evaluations of the derivative. Here we present a new method that allows the computation of sharp remainder intervals in parallel with the accumulation derivatives up to order n.

The method is useful for a variety of numerical problems, including the interval inclusion of very complicated functions prone to blow-up. To this end, the function is represented by a Taylor polynomial with remainder using the above method. Since at least for high orders, the remainder terms have a tendency to be very small, the problem is reduced to an interval evaluation of the Taylor polynomial. The method is used for guaranteed global optimization of blow-up prone functions and compared with some interval-based global optimization schemes.

1. Introduction

The idea of verified computation is based on the rigorous estimation of the influences of uncertainties on the calculation. Such uncertainties arise mainly from two sources.

- On one hand, there are computational inaccuracies based on the finite accuracy of computational environments.
- On the other hand, there are uncertainties in the variables of the model to be analyzed.

These two sources of uncertainty differ in that:

- in the first case, at least initially, the inaccuracies are small and comparable to the machine accuracy;
- in the second case, however, inaccuracies can be large even from the beginning.

The most extreme case is perhaps the problem of verified global optimization where often large areas of variable space have to be covered by intervals.

Interval methods allow a rigorous estimation of the influence of either form of uncertainty; however, it is well known that this is often at the expense of possibly significant overestimation of the actual ranges of the functions considered. This phenomenon of *blow-up* is usually particularly noticeable in the case of numerical algorithms of extended size as well as situations where the initial uncertainties are large to begin with. To control blow-up in its various forms is one of the interesting problems in the theory of interval algorithms, and a large variety of methods are employed for this purpose.

- On one hand, one can try to control blow-up for certain sub-algorithms; examples are the case of the exact scalar product, and also the evaluation of intrinsic functions by carefully using knowledge about local monotonicity and maxima instead of direct application of conventional numerical algorithms.
- On the other hand, there are various methods of retroactive error correction, typically based on iterative contraction methods using fixed point arguments.

If these approaches fail because of complexity or size of the initial intervals, as in the case of global optimization problems, frequently the starting intervals are split into smaller pieces and the algorithm is executed repeatedly. In the simplest form, the intervals are split evenly into m pieces, which however if the number v of input variables is large can quickly exhaust computational resources that scale with m^v. So in all but the most trivial cases, it is usually beneficial to develop strategies for uneven splits based on knowledge gained in previous evaluations; these methods are highly developed in the field of guaranteed global optimization [14], [17]–[19], [23], [25].

Early in the history of scientific computation it was realized that a given computational algorithm contains not only the information necessary to compute *interval bounds*, but also the information necessary to compute *partial derivatives* or *Taylor expansions* [12], [15]. In the forward mode of automatic differentiation, this only requires matrix methods for first order, and Taylor series methods [8] for higher orders. In a mathematical generalization of analysis on nonarchimedean fields [1], [3], [26], it allows the rigorous formulation of the intuitive concept of derivatives as differential quotients, and in practice can compute derivatives at points where Taylor methods may fail because of the lack of roots or inverses of Taylor series.

Besides the forward mode of automatic differentiation based on converting code to perform Taylor arithmetic, the main thrust of current algorithmic research probably lies in the so-called *reverse mode* which can even outperform numerical differentiation in speed if the number of initial variables is large, albeit at the expense of an often enormous growth of code size. In practice, these methods are used almost exclusively for the computation of first order derivatives.

- On one hand, this readily allows conventional sensitivity analysis describing the linearized relationship between output and input variables;

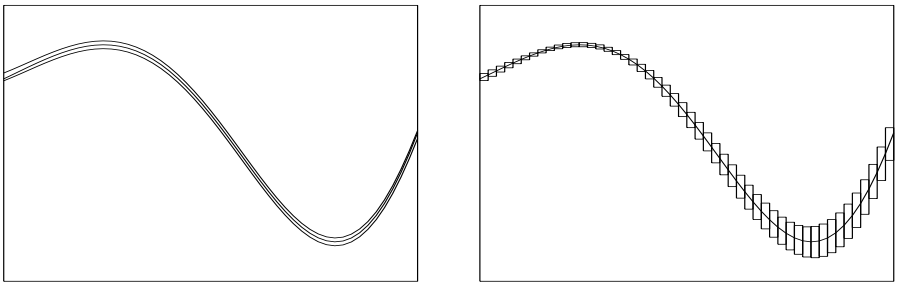

Figure 1. Left: enclosing a function by a Taylor polynomial of order eight with remainder bound; right: Interval bounding of the same function.

- on the other hand, it can be used for conventional numerical algorithms that are usually designed to try to avoid higher derivatives as much as possible because of the well-known instabilities arising in their numerical calculation.

Recently, efficient algorithms for high order multivariate Taylor methods were developed [8] and used in a differential algebraic framework [2] in the fields of optics and weakly nonlinear systems [10]. Different from other computational disciplines, they play a unique role because they represent the only known method to compute so-called *high-order aberrations*. Knowing these aberrations is of great importance for the understanding of high-precision *optics* or for the long term behavior of the dynamics in large *storage rings*.

The derivatives computed with Taylor methods are accurate up to machine errors, which can in turn be controlled by interval methods if necessary. This fact leads to the observation that even complicated functional evaluations can be done with limited blow-up by first computing the Taylor polynomial of the function, and then evaluating the polynomial on the interval of interest using interval methods. To be rigorous, this method critically depends on the ability to bound the remainder term of the Taylor expansion. This remainder could be estimated by bounding of the $(n+1)$-st derivatives, which in turn could be obtained by Taylor arithmetic. This, however, would require to generate code for the calculation of interval evaluations of derivatives and to repeatedly evaluate them using interval methods of global optimization. Here we present a new method that determines a bound for the Taylor remainder in an efficient way in parallel to the determination of the Taylor expansion.

As the result, the function under consideration can be enclosed in a bound around the Taylor polynomial over a wide range of arguments, as shown in the left part of Figure 1. On the other hand, obtaining a similarly tight inclusion using interval evaluations requires a large number of interval evaluations. As we will show later, the method will be particularly useful if in the interval case, the number of required intervals becomes large, which can happen in the case of many variables or of functions producing significant blow-up.

In the following section, after a quick review of calculus relating to extrema and Taylor formulas subsequently needed, we will introduce the concept of Taylor models. In Section 3, we show how Taylor expansions with remainder bounds for complicated functions can be obtained computationally. Section 4 gives various applications of the method, including very challenging global optimization problems related to predictions of long term stability in dynamical systems.

2. Taylor Models of Functions

The method of bounding of functions by Taylor polynomials with remainder, which is discussed in this paper, rather directly exploits very fundamental facts of calculus. In order to provide the proper background and introduction of the proper terminology, we begin with a brief summary of the required calculus tools.

2.1. Extrema and Taylor Formulas

Let $D \subset R$, and $f : D \to R$ be a continuous function, and let $[a, b] \subset D$, $a \neq b$. Then f assumes a maximum and a minimum on $[a, b]$, i.e., there are $\xi_l, \xi_u \in [a, b]$ such that $f(x) \geq f(\xi_l)$ and $f(x) \leq f(\xi_u)$ for all $x \in [a, b]$, which is known as the *Extremum Theorem*. Moreover, if $f(a) = f(b)$, at least one of ξ_l and ξ_u can be found in the interior of the interval. If f is differentiable in (a, b) and satisfies $f(a) = f(b)$, such interior point ξ (yielding an extremum) satisfies $f'(\xi) = 0$, and thus there is a zero of the derivative in (a, b), which is known as *Rolle's Theorem*.

Let now f and g be differentiable on D, and let $g'(x) \neq 0$ on (a, b). Then $g(a) \neq g(b)$, and the function h defined by

$$h(x) = f(x) - f(a) - (g(x) - g(a)) \cdot \frac{f(b) - f(a)}{g(b) - g(a)}$$

satisfies $h(a) = h(b) = 0$, is differentiable on D, and its derivative has the value $h'(x) = f'(x) - g'(x) \cdot (f(b) - f(a))/(g(b) - g(a))$. Application of Rolle's theorem yields the existence of $\xi \in (a, b)$ with $h'(\xi) = 0$, entailing

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)},$$

which is known as the *Mean Value Theorem*.

2.2. Taylor Formulas

By pure arithmetic, the mean value theorem entails *Taylor's Theorem*. Indeed, assume that the function f is $(n + 1)$ times differentiable on D, and let $x_0, x \in D$, $x \neq x_0$. Consider the function F defined by the formula

$$F(\overline{x}) = \sum_{\nu=0}^{n} \frac{f^{(\nu)}(\overline{x})}{\nu!} \cdot (x - \overline{x})^{\nu}.$$

Apparently we have $F(x) = f(x)$, and

$$F(x_0) = \sum_{\nu=0}^{n} f^{(\nu)}(x_0)/\nu!\ (x - x_0)^{\nu}.$$

Furthermore F is differentiable, with

$$F'(\overline{x}) = \frac{f^{(n+1)}(\overline{x})}{n!} \cdot (x - \overline{x})^n.$$

Applying the mean value theorem to F and an arbitrary function g with nonvanishing derivative strictly between x_0 and x then yields

$$\begin{aligned} f(x) \ &= \ F(x) \ = \ F(x_0) + \frac{g(x) - g(x_0)}{g'(\xi)} \cdot F'(\xi), \qquad \text{and thus} \\ f(x) \ &= \ \sum_{\nu=0}^{n} \frac{f^{(\nu)}(x_0)}{\nu!} \cdot (x - x_0)^{\nu} + \frac{g(x) - g(x_0)}{g'(\xi)} \cdot \frac{f^{(n+1)}(\xi)}{n!} \cdot (x - \xi)^n, \end{aligned}$$

which is *Taylor's formula with remainder.*

Specific choices for the function g yield various forms of the remainder term. For example, choosing $g(\overline{x}) = (x - \overline{x})^{(n+1)}$, observing that $g'(\overline{x}) = -(n+1) \cdot (x - \overline{x})^n$ is nonzero strictly between x_0 and x, yields *Lagrange's Remainder*

$$L_n = \frac{f^{(n+1)}(\xi)}{(n+1)!} \cdot (x - x_0)^{n+1},$$

where ξ lies strictly between x_0 and x. If f is $C^{(n+1)}$ on D, then, according to the maximum theorem, the derivatives are bounded on any interval in D, and thus so is the remainder. Furthermore, as $x \to x_0$, the remainder decreases like $(x - x_0)^{(n+1)}$.

It is easily possible to generalize the result to *Taylor's Formula for Multivariate Functions* by introducing a suitable one dimensional function. Let $D \subset R^d$ be a convex domain, let $f : D \to R$ be $(n+1)$ times continuously partially differentiable, and let $\vec{x_1}, \vec{x_0}$ be two different points from D. Then introduce a function f_R on $[0,1]$ by the formula $f_R(s) = f(\vec{x_0} + s \cdot (\vec{x_1} - \vec{x_0}))$, and note that according to the chain rule $f_R^{\nu}(s) = ([(\vec{x_1} - \vec{x_0}) \cdot \vec{\nabla}]^{(\nu)}\ f)\{\vec{x_0} + s \cdot (\vec{x_1} - \vec{x_0})\}$. Apply the Taylor formula for f_R and evaluate at $s = 1$ to obtain

$$f(\vec{x_1}) = \sum_{\nu=0}^{n} \frac{\left(\left[(\vec{x_1} - \vec{x_0}) \cdot \vec{\nabla}\right]^{\nu} f\right)(\vec{x_0})}{\nu!} + \frac{\left(\left[(\vec{x_1} - \vec{x_0}) \cdot \vec{\nabla}\right]^{n+1} f\right)(\vec{x_0} + \xi \cdot (\vec{x_1} - \vec{x_0}))}{(n+1)!}$$

where $\xi \in [0,1]$. Note that each term in the remainder term contains $(n+1)$ factors of the form $(x_1^{(j)} - x_0^{(j)})$ and so all of its derivatives up to order n at $\vec{x}_0$ vanish. This also entails that the n-th order Taylor polynomial $T^{(n)}_{f,\vec{x}}$ of f at the expansion point $\vec{x}_0$ can be written more conveniently as

$$T^{(n)}_{f,\vec{x}_0}(x) = \sum_{0 \le i_1 + \cdots + i_d \le n} \frac{\partial^{i_1+\cdots+i_d} f(\vec{x}_0)}{\partial x_1^{i_1} \ldots \partial x_d^{i_d}} \cdot \frac{(x^{(1)} - x_0^{(1)})^{i_1} \times \cdots \times (x^{(d)} - x_0^{(d)})^{i_d}}{i_1! \times \cdots \times i_d!}.$$

Using the sum and product rules of differentiation, this representation of the Taylor polynomial also rather directly reveals that the Taylor polynomial of the sum of two functions is the sum of the Taylor polynomials, and that the one of the product of two functions can be obtained by multiplying the respective Taylor polynomials and ignoring all terms of order higher than n.

2.3. Taylor Models

Let f be $C^{(n+1)}$ on $D_f \subset R^v$, and $\vec{B} = [a_1, b_1] \times \cdots \times [a_v, b_v] \subset D_f$ be an interval box containing the point $\vec{x}_0$. Let T be the Taylor polynomial of f around the point $\vec{x}_0$. We call the interval I an n-th order *Remainder Bound* of f on $\vec{B}$ if

$$f(\vec{x}) - T(\vec{x}) \in I \quad \text{for all} \quad \vec{x} \in \vec{B}.$$

We call the pair (T, I) an n-th order *Taylor Model* of f. The set of all remainder bounds is called the *Remainder Family*. Since all partial derivatives of f are continuous on the compact set $\vec{B}$, they are bounded there, and hence so is the $(n+1)$-st application of the directional derivative $(\vec{x} - \vec{x}_0) \cdot \vec{\nabla}$. This entails that for all $\vec{x} \in \vec{B}$, the Lagrange remainder is bounded, and hence a finite remainder bound exists. Furthermore, since $f(\vec{x}) - T(\vec{x})$ is continuous, it assumes extrema at x_l, x_u. So $I = [f(\vec{x}_l) - T(\vec{x}_l),\ f(\vec{x}_u) - T(\vec{x}_u)]$ is a remainder bound; all other remainder bounds must contain I, and thus I is called the *Optimal Remainder Bound*.

Since every polynomial of order not exceeding n agrees with its n-th order Taylor polynomial, the optimal remainder bound in this case is the interval $[0, 0]$.

For practical purposes, it is important to remark that if the original interval box $\vec{B}$ *decreases* in size, then, according to the Lagrange remainder formula, the optimal remainder bound will *decrease* in size as $(n+1)$-st power of the size of the box $\vec{B}$, and hence will become small quickly. In particular, this entails that the knowledge of a good Taylor model of a function on an interval box $\vec{B}$ allows a rather accurate estimate of the range of the function. We also note that for the case of $n = 0$, the method reduces to the conventional interval arithmetic; for the case $n = 1$, the method is somewhat similar to the the generalized interval arithmetic introduced by Hansen [16], except that in this case, the coefficients of the (linear) polynomials are also intervals.

As we will show in the next section, it is actually possible to compute remainder bounds to any order in a convenient way in parallel with a Taylor expansion of the function. Section 4 will then illustrate the practical use of the method.

3. Operations with Taylor Models

In this section we will show how from the knowledge of Taylor models of functions f and g, we can infer Taylor models of the sums and products as well as of elementary functions applied to one of the functions. This will allow the systematic computation of Taylor models for any function that can be represented in a computer environment. Throughout, we will employ standard notation for operations on sets, for example $A + B = \{a + b \mid a \in A,\ b \in B\}$ or $x + A = \{x + a \mid a \in A\}$ or $f(A) = \{f(a) \mid a \in A\}$, as well as standard rules like $A \cdot (B + C) \subset A \cdot B + A \cdot C$.

3.1. Addition of Taylor Models

Let (T_f, I_f) and (T_g, I_g) be n-th order Taylor models of the functions f and g on the interval box $\vec{B}$. As pointed out above, the Taylor polynomial of $(f + g)$ is simply $T_f + T_g$; on the other hand, we know that on $\vec{B}$, $f(\vec{x}) \in T_f(\vec{x}) + I_f$ and $g(\vec{x}) \in T_g(\vec{x}) + I_g$. Thus obviously,

$$(f + g)(\vec{x}) \in (T_f + T_g)(\vec{x}) + (I_f + I_g) \quad \text{for all} \quad \vec{x} \in \vec{B},$$

and so $(T_f + T_g,\ I_f + I_g)$ is a Taylor model for $(f + g)$ on $\vec{B}$. Furthermore, the interval $I_{f+g} = I_f + I_g$ is the sharpest remainder bound that can be inferred from the knowledge of the Taylor models for f and g. For practical purposes, it is also important to note that if I_f, I_g are "fine of order $\vec{B}^{n+1}$," i.e., their size scales with the size of $\vec{B}$ to the $(n + 1)$-st power, so is I_{f+g}. In the same way we see that $(T_f - T_g,\ I_f - I_g)$ is a Taylor model for $(f - g)$.

3.2. Multiplication of Taylor Models

Let (T_f, I_f) and (T_g, I_g) be n-th order Taylor models of the functions f and g on the interval box $\vec{B}$. As pointed out before, the Taylor polynomial $T_{f \cdot g}$ of $f \cdot g$ can then be obtained by multiplication of T_f and T_g and subtracting the polynomial $\overline{T}_{f \cdot g}$ consisting of the terms whose order exceeds n. For any $\vec{x} \in \vec{B}$, there are values $e_f \in I_f$ and $e_g \in I_g$ such that $f(\vec{x}) = T_f(\vec{x}) + e_f$ and $g(\vec{x}) = T_g(\vec{x}) + e_g$. So we obtain

$$\begin{aligned}(f \cdot g)(\vec{x}) &= (T_f(\vec{x}) + e_f) \cdot (T_g(\vec{x}) + e_g) \\ &= T_f(\vec{x}) \cdot T_g(\vec{x}) + T_f(\vec{x}) \cdot e_g + T_g(\vec{x}) \cdot e_f + e_f \cdot e_g \\ &= T_{f \cdot g}(\vec{x}) + \{\overline{T}_{f \cdot g}(\vec{x}) + T_f(\vec{x}) \cdot e_g + T_g(\vec{x}) \cdot e_f + e_f \cdot e_g\}.\end{aligned}$$

The first term is the Taylor polynomial of $f \cdot g$. The term in curly brackets describes the behavior of the error; it is a polynomial in the $v + 2$ variables $(\vec{x}, e_f, e_g) \in \vec{B} \times I_f \times I_g$ and is denoted by $E(\vec{x}, e_f, e_g)$. Since no knowledge about the correlation between $\vec{x}$ and e_f, e_g is contained in the Taylor models of f and g besides the fact that $e_f \in I_f$, $e_g \in I_g$, the sharpest possible remainder bound of $f \cdot g$ that can be inferred is given by the extrema of $E(\vec{x}, e_f, e_g)$ on the interval box $\vec{B} \times I_f \times I_g$.

Thus the determination of $I_{f \cdot g}$ is a polynomial bounding problem, and the quality of the polynomial bounds translates directly into the sharpness of $I_{f \cdot g}$.

For practical purposes, the necessary polynomial bounding can be performed in various ways. The most accurate but probably also most expensive way is the direct bounding of E using dedicated polynomial bounding techniques, like those based on evaluation at Tschebycheff points.

The effort can be decreased somewhat by first bounding $T_f(\vec{B})$, $T_g(\vec{B})$ and $\overline{T}_{f \cdot g}(\vec{B})$ as v dimensional polynomials and then computing a remainder bound as $I_{f \cdot g} = \overline{T}_{f \cdot g}(\vec{B}) + T_f(\vec{B}) \cdot I_g + T_g(\vec{B}) \cdot I_f + I_f \cdot I_g$. In this case the three interval additions can cause some slight blow-up due to cancellation; but this bounding method still has the property that if I_f and I_g are "fine of order $\vec{B}^{(n+1)}$," so is $I_{f \cdot g}$. The actual polynomial bounding can again be performed using a dedicated polynomial technique.

At the expense of loss of sharpness, the polynomial evaluations can also be performed by evaluations of Horner factorizations; even in this case the blow-up will not be substantial because of the multiplications with the very narrow intervals I_f and I_g and the fact that $\overline{T}_{f \cdot g}$ is of order $(n+1)$ and higher and hence consists only of small intervals raised to high powers.

3.3. Elementary Functions of Taylor Models

In this subsection we will show how it is possible to determine Taylor models of elementary functions composed with functions with known Taylor models. We will reduce the problem to the computation of interval evaluations of elementary functions (as well as the addition and multiplication of Taylor models). Let (T_f, I_f) be an n-th order Taylor model for f on $\vec{B}$, and let g be an elementary function appearing in a computer environment, e.g., exp or log. The goal is to find a Taylor model for $g \circ f$. While there appears to be no fully universal strategy that is applicable for any function g, in most cases it is possible to adhere to the following approach:

1. Introduce $c = T_f(\vec{x_o})$, $\overline{T}_f(\vec{x}) = T_f(\vec{x}) - c$; then $\overline{T}$ does not have a constant part, and when evaluated on $\vec{B}$, it is "fine of order $\vec{B}$."
2. Use an addition theorem for g (or a similar method) to separate the total computation into:
   - computations for c and
   - computations for $\overline{T}_f(\vec{x}) + I_f$, denoted by $h_i(\overline{T}_f(\vec{x}) + I_f)$.
3. Perform the computations for c *exactly* (or with verified methods yielding a tight inclusion).
4. Perform the computations for $\overline{T}(\vec{x})$ using an n-th order Taylor polynomial with remainder for h. Because $\overline{T}(\vec{x})$ does not have a constant part, powers of $(\overline{T}_f(\vec{x}), I_f)$ of order higher than n generate a vanishing contribution to the

Taylor polynomial, and it is sufficient to find the Taylor polynomial of the composition $g \circ f$.

5. The remainder term of the n-th order Taylor expansion for h which is determined by differentiation of h and has to be evaluated with ξ and $(x - x_0)$ in $\overline{T}_f(\vec{B}) + I$, is computed using standard interval arithmetic. Since it contains an $(n+1)$-st power of $\overline{T}_f(\vec{B}) + I$, which is "small of order $\vec{B}$," this yields an interval "small of order $\vec{B}^{n+1}$."

Because of the generality of the description and the associated element of vagueness, we want to illustrate the method on two examples:

- First, consider the case $g(x) = \exp(x)$. Writing $T_f(\vec{x}) = c + \overline{T}_f(\vec{x})$, we obtain

  $$\exp\left(T_f(\vec{x}) + I_f\right) = \exp(c) \cdot \exp\left(\overline{T}_f(\vec{x}) + I_f\right).$$

  For the second exponential, we use the Taylor formula of exp around the origin with Lagrange remainder, i.e.

  $$\exp(x) = \sum_{\nu=0}^{n} \frac{x^\nu}{\nu!} + x^{n+1} \exp(\xi)/(n+1)!,$$

  where ξ is between 0 and x. In our case, $\xi \in \overline{T}_f(\vec{B}) + I_f$, and so we obtain the following inclusion:

  $$\exp\left(\overline{T}_f(\vec{x}) + I_f\right) \subset \sum_{\nu=0}^{n} \frac{\left(\overline{T}_f(\vec{x}) + I_f\right)^\nu}{\nu!} + \left(\overline{T}_f(\vec{B}) + I_f\right)^{(n+1)} \cdot \frac{\exp\left(\overline{T}_f(\vec{B}) + I_f\right)}{(n+1)!}.$$

  The polynomial in the resulting Taylor model on the right is the Taylor polynomial of $\exp(f)$. Apparently, sharper inclusions are possible if there are dedicated methods for the computation of polynomials of Taylor models.

- In a very similar way, we can proceed for the cases $g = \sin$ and $g = \cos$ (using the respective addition theorems).

- As another example, consider the determination of a Taylor model for $\log(f)$ from a Taylor model for f. Let (T_f, I_f) be an n-th order Taylor model for f on $\vec{B}$, and let there be an evaluation of $T_f(\vec{B})$ such that $0 < T_f(\vec{B}) + I_f$. This implies in particular that the constant part c of T_f does not vanish. Write $T_f(\vec{x}) + I_f = c \cdot [1 + (\overline{T}(\vec{x}) + I_f)/c]$ and use $\log(ab) = \log(a) + \log(b)$ to separate the constant part from the rest. Consider $h(x) = \log(1+x)$; then for $|x| < 1$, we have $\log(1+x) = \sum_{\nu=1}^{n} (-1)^{\nu+1} x^\nu/\nu + L_n$, where the n-th order Lagrange remainder can be written as $L_n = (-1)^n x^{n+1}/\{(n+1)(1+\xi)^{n+1}\}$, with $0 \le \xi \le x$. So we obtain

  $$\begin{aligned} \log\left(T_f(\vec{x}) + I_f\right) &= \log(c) + \log\left(\frac{T_f(\vec{x}) + I_f}{c}\right) \\ &\subset \log(c) + \sum_{\nu=1}^{n} \frac{(-1)^{\nu+1}}{\nu} \left(\frac{\overline{T}(\vec{x}) + I_f}{c}\right)^\nu + \frac{(-1)^n}{n+1} \left(\frac{\overline{T}(\vec{B}) + I_f}{T(\vec{B}) + I_f}\right)^{(n+1)}. \end{aligned}$$

Again, if the original interval I_f is "fine of order $\vec{B}^{n+1}$," so are the intervals that are accumulated in the evaluation of the Taylor polynomial, and so is the interval generated by the Lagrange remainder.

3.4. Division of Taylor Models

The division of Taylor models will be achieved by application of the elementary function $1/x$. So let (T_f, I_f) be an n-th order Taylor model for f on $\vec{B}$, and let there be an evaluation of $T_f(\vec{B})$ such that $0 \notin T_f(\vec{B}) + I_f$. This implies in particular that the constant part c of T_f does not vanish. Write $T_f(\vec{x}) + I_f = c \cdot [1 + (\overline{T}(\vec{x}) + I_f)/c]$ and use $1/(ab) = (1/a)(1/b)$ to separate the constant part from the rest. Consider $h(x) = 1/(1+x)$; then for $|x| < 1$, we have

$$\frac{1}{1+x} = \sum_{\nu=0}^{n} (-x)^\nu + L_n,$$

where the n-th order Lagrange remainder can be written as $L_n = (-1)^{n+1} x^{n+1} / (1+\xi)^{n+2}$. So we obtain

$$\begin{aligned} \frac{1}{T_f(\vec{x}) + I_f} &= \frac{1}{c} \cdot \frac{1}{1 + (\overline{T}(\vec{x}) + I_f)/c} \\ &\subset \sum_{\nu=0}^{n} \frac{(-1)^\nu}{c} \cdot \left(\frac{\overline{T}(\vec{x}) + I_f}{c} \right)^\nu + \frac{(-1)^{n+1} \cdot (\overline{T}(\vec{B}) + I_f)^{n+1}}{(T(\vec{B}) + I_f)^{n+2}}. \end{aligned}$$

And as before, if the original interval I_f is "fine of order $\vec{B}^{n+1}$," so are the interval remainders that are accumulated in the evaluation in the Taylor polynomial, and so is the interval remainder generated by the Lagrange remainder.

3.5. Calculation of Taylor Models of Complicated Functions

In the previous subsections we showed how Taylor models for sums, products, and elementary functions can be computed as soon as the Taylor model of the original function is known. Using these rules repeatedly, it is thus possible to determine Taylor models of any function that can be expressed in terms of repeated applications of these elementary operations; in particular, this includes all functions that can be represented in a computer environment.

Apparently it is necessary to begin the calculation with one known Taylor model, and this is most readily achieved by writing the function f under consideration as $f \circ \vec{I}$, where $\vec{I}$ is the identity in R^v. In this case, all that is needed are Taylor models for the components I_v. Since $\vec{I}$ agrees with its first order Taylor polynomial, these are readily given as $(x_v, [0,0])$.

4. Applications

4.1. Implementation

In the last sections it became apparent that the computation of remainder bounds for Taylor series can be performed in parallel to the computation of Taylor expansions of functions and thus represents a generalization of Taylor methods. In case the order of the Taylor calculations is chosen to be zero, it reduces to *conventional interval arithmetic*, and hence also *generalizes interval methods.*

The computational expense needed to determine remainder bounds while computing Taylor expansions is limited. For the use of interval chains, it is usually almost negligible, while for the use of very sharp polynomial methods to estimate the bound for the product, the expense depends on the sophistication of the methods.

The practical use of the Taylor models method benefits from object oriented language environments like those offered in C++ or Object Oriented Pascal, or in many of the more dedicated packages used in scientific computation like Pascal XSC, or the various codes used for the automatic differentiation of FORTRAN or C. The actual implementation requires the availability of libraries for interval operations as well as Taylor multiplications and benefits from the existence of fast algorithms to bound the range of polynomials.

We implemented [21], [22] the method in the object oriented Pascal-like language environment of COSY INFINITY [4], [5], [13], [20]. The code is a large scale design and analysis tool with extensive macros to study the dynamics in weakly nonlinear systems; it has about 200 registered users in the accelerator community.

4.2. Bounding of Functions

A Taylor model defines an *approximation* of a multivariate function as well as a "*band*" around this approximation containing the values of the actual function. Figure 2 shows examples for a function of two variables. The band is obtained by arithmetic with Taylor models as implemented in COSY, where $\vec{B}$ described the region of the displayed coordinate space. With increasing order, the width of the band decreases quickly. Bounding the function with conventional interval arithmetic requires its evaluation at many different intervals, because the width of the bound is largely determined by the size of the interval to which the function is applied. The graphs in Figure 3 show the improved bound for a finer interval covering of the relevant region.

The potential of the Taylor model arithmetic becomes apparent when one considers that:

- the algorithm describing the function had to be evaluated at only one Taylor model to get the tightest bound shown, whereas

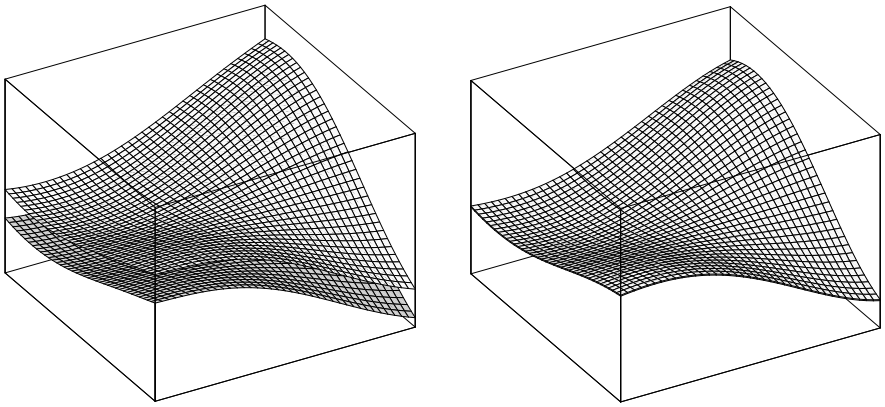

Figure 2. Interval bounding of a function by a Taylor model; orders 7 (left) and 10 (right).

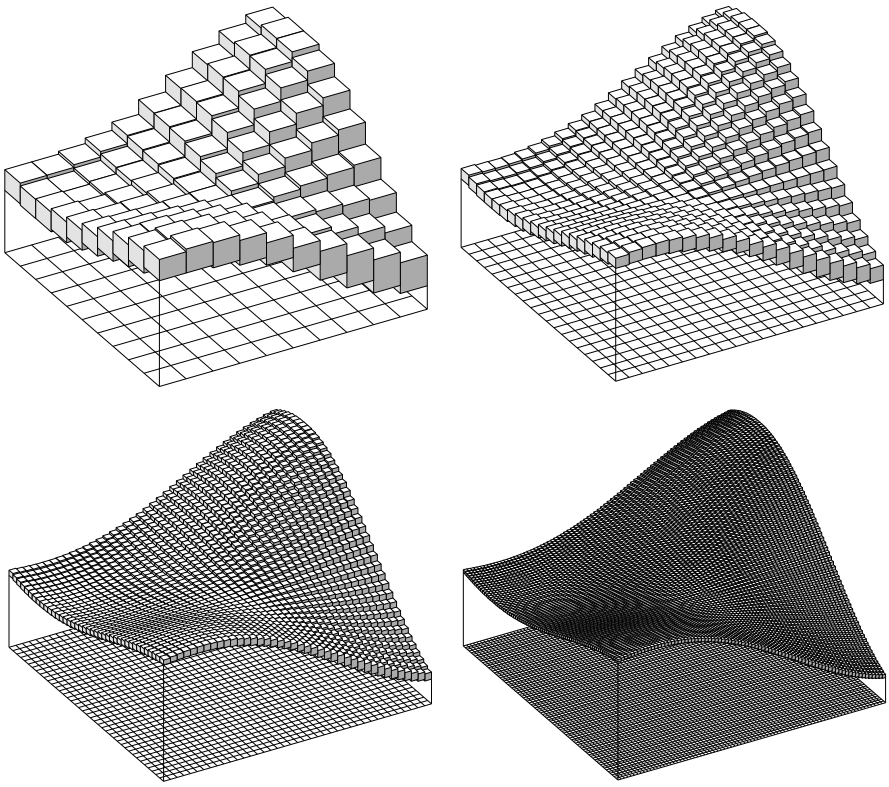

Figure 3. Conventional interval bounding of a function; from top left to bottom right the number of interval evaluations is 100, 400, 1600, and 6400.

- 6400 interval evaluations were needed to achieve the tightest bound computed by conventional interval arithmetic.

4.3. Applications for Guaranteed Stability Estimates

Estimating the time of stable motion for *planetary systems* has first started the interest in the stability of weakly nonlinear mechanical systems.

In *accelerator physics*, this question became important with the introduction of *storage rings*. In large storage rings particles often have to be kept in the accelerator for up to a billion turns.

In the past, the question of long term stability in storage rings has been analyzed by various methods [7], [9], [27], [28], [33]. Although some of the established methods are useful analysis tools, they all fail to give mathematically rigorous lower bounds on the time during which particles stay inside the accelerator (the motion is described by a nonlinear map).

The principle underlying the proof of the Nekhoroshev estimate [24] was used numerically to obtain lower bounds for the time which particles can orbit the accelerator [30]. For this approach, one first finds a function f which defines surfaces by $f = c$ with $c \in \mathbb{R}$ that enclose a volume around the origin and additionally do not change much during one application of the transfer map $\vec{M}$ of the accelerator. The deviation function $d_f = f \circ \vec{M} - f$ is therefore very small and f is called a *pseudo invariant* of $\vec{M}$. The pseudo invariant needed for this method is computed via nonlinear normal form theory [6], [10].

If a rigorous upper bound δ on d_f can be found in the volume with $f < c$, then particles starting in the region $f < c - N\delta$ will certainly not leave the region $f < c$ for at least N turns. To make refined rigorous estimates of the survival time, several tricks are needed, but these few lines already contain the basic thoughts behind rigorous lower bounds on the survival time in weakly nonlinear dynamics. The method was first used in [29], [31], [32], where the maximum of d_f was approximated by evaluating the function on a grid. First rigorous bounds based on interval methods were reported in [11].

A typical example of the deviation function is shown in Figure 4, where a nonlinear transfer map of order 6 was used.

- The rigorous maximization of the deviation function calls for *interval* arithmetic. However, since the displayed deviation function is a polynomial of order 72 in four variables, which in expanded form has 1,282,975 coefficients, enormous blow-up builds up during their evaluation. To bound the maximum by conventional interval arithmetic, in the order of 10^{24} interval evaluations would be needed. Evaluating such a complicated function in interval arithmetic 10^{24} times is certainly out of the question with today's (and probably tomorrow's) computer speeds.

- For this problem, *Taylor model arithmetic* can be applied very beneficially, resulting in a speed up of about 10^{16} as compared to conventional interval arithmetic.

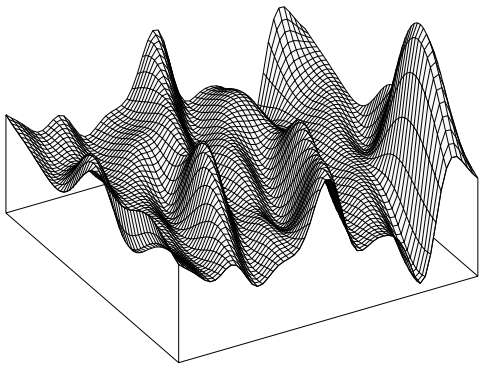

Figure 4. Two-dimensional projection of a typical deviation function for a six dimensional transfer map.

Acknowledgement

We would like to thank Kyoko Makino for performing the calculations in this paper, for valuable comments, and a careful reading of the manuscript.

References

1. Berz, M.: *Analyis on a Nonarchimedean Extension of the Real Numbers*, Technical Report MSUCL-753, National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, MI 48824, 1990.
2. Berz, M.: Arbitrary Order Description of Arbitrary Particle Optical Systems, *Nuclear Instruments and Methods* **A298** (1990), p. 426.
3. Berz, M.: Automatic Differentiation as Nonarchimedean Analysis, in: *Computer Arithmetic and Enclosure Methods*, Elsevier Science Publishers, Amsterdam, 1992, p. 439.
4. Berz, M.: *COSY INFINITY Version 7 Reference Manual*, Technical Report MSUCL-977, National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, MI 48824, 1996.
5. Berz, M.: COSY INFINITY Version 6, in: Berz, M., Martin, S., and Ziegler, K. (eds), *Proc. Nonlinear Effects in Accelerators*, IOP Publishing, 1992, p. 125.
6. Berz, M.: Differential Algebraic Formulation of Normal Form Theory, in: Berz, M., Martin, S., and Ziegler, K. (eds), *Proc. Nonlinear Effects in Accelerators*, IOP Publishing, 1992, p. 77.
7. Berz, M.: Differential Algebraic Treatment of Beam Dynamics to Very High Orders Including Applications to Spacecharge, *AIP Conference Proceedings* **177** (1988), p. 275.
8. Berz, M.: Forward Algorithms for High Orders and Many Variables, *Automatic Differentiation of Algorithms: Theory, Implementation and Application*, SIAM, 1991.
9. Berz, M.: High-Order Description of Accelerators Using Differential Algebra and First Applications to the SSC, in: *Proceedings, Snowmass Summer Meeting*, Snowmass, Co, 1988.
10. Berz, M.: High-Order Computation and Normal Form Analysis of Repetitive Systems, in: Month, M. (ed.), *Physics of Particle Accelerators*, volume AIP 249, American Institute of Physics, 1991, p. 456.
11. 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.
12. Berz, M., Bischof, C., Griewank, A., and Corliss, G. (eds), *Computational Differentiation: Techniques, Applications, and Tools*, SIAM, Philadelphia, 1996.

13. Berz, M., Hoffstätter, G., Wan, W., Shamseddine, K., and Makino, K.: COSY INFINITY and Its Applications to Nonlinear Dynamics, in: *Computational Differentiation: Techniques, Applications, and Tools*, SIAM, 1996.
14. Csenes, T.: Test Results of Interval Methods for Global Optimization, *Computer Arithmetic, Scientific Computation and Mathematical Modelling* **12** (1991).
15. Griewank, A. and Corliss, G. F. (eds), *Automatic Differentiation of Algorithms*, SIAM, Philadelphia, 1991.
16. Hansen, E. R. A Generalized Interval Arithmetic, in: Nickel, K. (ed.), *Interval Mathematics*, 1975.
17. Hansen, E. R.: Global Optimization Using Interval Analysis—the One-Dimensional Case, *J. Optim. Theor. and Appl.* **29** (1979), pp. 331–334.
18. Hansen, E. R.: Global Optimization Using Interval Analysis—the Multidimensional Case, *Numerische Mathematik* **34** (1980), pp. 247–270.
19. Jansson, C.: A Global Optimization Method Using Interval Arithmetic, in: *Computer Arithmetic and Scientific Computation, Proceedings of the SCAN 91*, Amsterdam, North-Holland, Elsevier, 1992.
20. Makino, K. and Berz, M.: COSY INFINITY Version 7, in: *Fourth Computational Accelerator Physics Conference*, AIP Conference Proceedings, 1996.
21. Makino, K. and Berz, M.: *Implementation and Applications of Taylor Model Methods.*
22. Makino, K. and Berz, M.: Remainder Differential Algebras and Their Applications, in: *Computational Differentiation: Techniques, Applications, and Tools*, SIAM, 1996.
23. Moore, R. E. and Ratschek, H.: Inclusion Functions and Global Optimization II. *Mathematical Programming* **41** (1988), pp. 341–356.
24. Nekhoroschev, N. N.: An Exponential Estimate of the Time of Stability of Nearly Integrable Hamiltonian Systems, *Uspekhi Mat. Nauk, English translation Russ. Math. Surv.* **32** (6) (1977).
25. Rokne, J. and Ratschek, H.: *New Computer Methods for Global Optimization*, Ellis Horwood Limited, Chichester, England, 1988.
26. Shamseddine, K. and Berz, M.: Exception Handling in Derivative Computation with Nonarchimedean Calculus, in: *Computational Differentiation: Techniques, Applications, and Tools*, SIAM, 1996.
27. Talman, R.: Long Term Prediction and the SSC, in: *Nonlinear Problems in Future Accelerators*, World Scientific, New York, 1991, pp. 215–231.
28. Turchetti, G.: Nekhoroshev Stability Estimates for Symplectic Maps and Physical Applications, in: *Number Theory and Physics, Springer Proceedings in Physics* **47**, Springer-Verlag, Berlin, Heidelberg, 1990.
29. Warnock, R. L.: Close Approximation to Invariant Tori in Nonlinear Mechanics, *Physical Review Letters* **66** (14) (1991), pp. 1803–1806.
30. Warnock, R. L. and Ruth, R. D.: Long-Term Bounds on Nonlinear Hamiltonian Motion, *Physica D* **56** (14) (1992), pp. 188–215, also SLAC-PUB-5267.
31. Warnock, R. L. and Ruth, R. D.: Stability of Orbits in Nonlinear Mechanics for Finite but Very Long Times, in: *Nonlinear Problems in Future Accelerators*, World Scientific, New York, 1991, pp. 67–76, also SLAC–PUB–5304.
32. Warnock, R. L., Ruth, R. D., Gabella, W., and Ecklund, K.: Methods of Stability Analysis in Nonlinear Mechanics, in: *1987 Accelerator Physics Summer School* AIP Conference Proceedings, 1988, also SLAC Pub 4846, 1989.
33. Yan, Y.: *Applications of Differential Algebra to Single-Particle Dynamics in Storage Rings*, Technical Report SSCL-500, Superconducting Super Collider Laboratory, 1991.