

Nonlinear Analysis 47 (2001) 3503-3514



www.elsevier.nl/locate/na

# Higher Order Verified Inclusions of Multidimensional Systems by Taylor Models

K. Makino and M. Berz

Department of Physics and Astronomy, and National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, MI 48824, USA

#### Abstract

Different from floating point computations, interval methods provide rigorous enclosures of functions, however the limitation of the methods is the overestimation mostly caused by the lack of information on functional dependency. The first cure to the problem is to use a smaller domain, but when a function is complicated, as it often is for practical problems, the number of subdivisions becomes quite large. In case of multidimensional systems, the computational expense by simple interval methods increases astronomically. A new approach, the Taylor model method, models a function by a higher order polynomial which keeps the majority of the functional dependency, and an interval which contains the small remaining error. The method naturally suppresses the dependency problem, and proves particularly effective for the treatment of complicated multidimensional systems.

Key words: Taylor model, high order multivariate polynomial, interval method, verified method, dependency problem 1991 MSC: 41A80, 65G20

### 1 Introduction

Interval methods [1,2] offer a simple mechanism to evaluate an enclosure of a function on computers, so the field of the applications is potentially large [1,3–5]. The practicality depends on the precision of the enclosure, and the methods often suffer by the overestimation problem. The conventional simple interval methods do not carry the information on functional dependency, and

Email address: makino@msu.edu, berz@msu.edu (K. Makino and M. Berz). URL: bt.nscl.msu.edu (K. Makino and M. Berz).

0362-546X/01/\$ - see front matter © 2001 Published by Elsevier Science Ltd.

PII: \$0362-546X(01)00467-9

simple interval arithmetic below shows a typical overestimation problem. Let I = [a, b], then by simple interval arithmetic,

$$I - I = [a, b] - [a, b] = [a, b] + [-b, -a] = [a - b, b - a],$$

results in a width that is not zero as it should, but rather twice as large as before. Such a situation cannot be avoided in conventional interval methods in practice if the canceling terms appear as a result of preceding complicated computations.

The first mean to suppress the overestimation is to have the domain smaller so that the overestimation also is proportionally small. Apparently the size of the divided sub-domain has to be quite small for a complicated function. When the system is multidimensional with the dimensionality v and assuming the complication is uniform on different dimensions, the computational expense is proportional to  $m^v$  with m the number of subdivision per dimension, and it is exceedingly large for larger m and higher v. One of examples which we will study later has v = 6 and  $m = 10^4$  to cover an interested area to get a reasonably precise enclosure by simple interval arithmetic. That means the problem requires to scan  $(10^4)^6 = 10^{24}$  sub-domains. Even though the cost to compute one sub-domain is low, the total cost is astonishingly high. One of the important applications of interval methods is in the field of global optimization problems [6,7], and interval arithmetic is used for the elimination process of regions. Practical optimization problems often involve many local extrema indicating many cancellations and often the problems are multidimensional, thus conventional interval arithmetic suffers a severe dependency problem.

A new method, the Taylor model method, has been developed and implemented in the code COSY Infinity [8–10]. The method models a sufficiently smooth function  $f(\vec{x})$  in a domain  $[\vec{a}, \vec{b}]$  by a polynomial P and bounds the error in an interval I. There are several advantages to using Taylor polynomials; so let P be the nth order Taylor polynomial, which requires f to be (n+1)times continuously partially differentiable in the domain, then the scaling of the size of the remainder interval I with domain width can be easily obtained. According to Taylor's theorem, the Taylor remainder error depends on  $\vec{x}$  via  $(\vec{x} - \vec{x}_0)^{n+1}$ , where  $\vec{x}_0$  is the expansion point. Thus, the size of the remainder interval I is proportional to  $|\vec{x} - \vec{x}_0|^{n+1}$ , which decreases very quickly with the decrease of the size of the domain, and it is in contrast with the linear decrease in the conventional interval methods. For example, if the size of the sub-domain is half the original size and a tenth order computation is performed, then roughly speaking, the size of the remainder interval I becomes more exact by the factor of  $(1/2)^{11}$  which is about  $5 \times 10^{-4}$ . Raising the order n also contributes to a decrease in the size of the remainder interval I quickly. We denote a Taylor model of f by

$$T_f = (P_f, I_f).$$

To estimate the computational expense for Taylor models, the most dominating factor is the total possible number of monomials in a Taylor polynomial, and it is given by [11]

$$N(n,v) = \binom{n+v}{v} = \frac{(n+v)!}{n! \cdot v!},\tag{1}$$

where v is the dimensionality. The formula (1) is symmetric in n and v, and N(n,v) increases moderately with the increase of n or v. For example, N(5,6), ..., N(10,6) are 462, 924, 1716, 3003, 5005, 8008 each, and either increasing the order by one or increasing the dimensionality by one results in having N less than doubled, in contrast to the increasing effort when scanning the whole area by subdivided intervals.

In Taylor models, the higher order functional dependency is kept in the polynomial part, and the cancellation happens automatically in the polynomial part, and the size of the remainder interval is naturally kept small. Thus, the overestimation problem is optimally controlled [12,13]. In the next sections, we first introduce the Taylor model arithmetic, then discuss the function inclusion problems with examples. For illustrative purposes, we will discuss one dimensional systems first, and both an easy example and a challenging one will be presented. Then, we will study a small three dimensional system and finally will study a six dimensional practical problem. In those more complicated problems, the simple interval method is used for the comparison purpose, but it will present the obvious limitation.

#### 2 Taylor model arithmetic

Any computer representable function  $f(\vec{x})$  can be modeled by nth order Taylor models if the function f is  $C^{(n+1)}$  in the domain  $[\vec{a}, \vec{b}]$ . When the order n and the expansion point  $\vec{x}_0$  are specified, the Taylor polynomial part  $P_f(\vec{x} - \vec{x}_0)$  is determined uniquely with coefficients described by floating point numbers on a computer. The remainder interval part  $I_f$  further depends on the domain  $[\vec{a}, \vec{b}]$  and the algorithm to compute it.

Taylor model arithmetic starts from preparing the variables of the function,  $\vec{x}$ , represented by Taylor models. The variables  $\vec{x}$  can be expressed in the

polynomial form at the reference point  $\vec{x}_0$  as  $\vec{x}_0 + (\vec{x} - \vec{x}_0)$ , which are linear, and by doing so, there is no error involved. Thus, a Taylor model of the identity function of the *i*th variable  $x_i$  is

$$T_{x_i} = (x_{i0} + (x_i - x_{i0}), [0, 0]).$$

Then, Taylor model arithmetic is carried through binary operations and intrinsic functions, which compose the function f sequentially. Suppose we have Taylor models for g and h as

$$T_{n,q} = (P_{n,q}, I_{n,q})$$
 and  $T_{n,h} = (P_{n,h}, I_{n,h}).$ 

Taylor models of the sum and difference of g and h can be obtained as

$$T_{n,g} \pm T_{n,h} = (P_{n,g} \pm P_{n,h}, I_{n,g} \pm I_{n,h}).$$

The Taylor model for the product of g and h can be obtained as

$$T_{n,g} \cdot T_{n,h} = (P_{n,g \cdot h}, I_{n,g \cdot h}), \text{ where } P_{n,g} \cdot P_{n,h} = P_{n,g \cdot h} + P_e$$

with  $P_{n,g\cdot h}$ , the *n*th order polynomial of the result of the left hand side, and  $P_e$ , the part of the product polynomial with order from n+1 to 2n, and

$$I_{n,g \cdot h} = B(P_e) + B(P_{n,g}) \cdot I_{n,h} + B(P_{n,h}) \cdot I_{n,g} + I_{n,g} \cdot I_{n,h},$$

where B(P) is an enclosure of P.

$$\begin{split} \exp(g) &= \exp{(c + \bar{g})} = \exp(c) \cdot \exp{(\bar{g})} \\ &= \exp(c) \cdot \left\{ 1 + \bar{g} + \frac{1}{2!} (\bar{g})^2 + \dots + \frac{1}{k!} (\bar{g})^k + \frac{1}{(k+1)!} (\bar{g})^{k+1} \exp{(\theta \cdot \bar{g})} \right\}, \end{split}$$

where  $0 < \theta < 1$ . Taking  $k \ge n$ , a Taylor model of the part

$$\exp(c) \cdot \left\{ 1 + \bar{g} + \frac{1}{2!} (\bar{g})^2 + \dots + \frac{1}{n!} (\bar{g})^n \right\}$$

can be obtained as a sequence of additions and multiplications, and it contributes to the polynomial part  $P_{n,\exp(g)}$  and to the remainder interval part  $I_{n,\exp(g)}$ . Since  $\bar{g}$  does not have a constant part,  $(\bar{g})^m$  starts from mth order. Thus, the remaining part in  $\exp(g)$  has a vanishing polynomial part, and contributes only to the remainder interval part  $I_{n,\exp(g)}$ . Refer to [8,9] for details, also on the treatment of other intrinsic functions.

#### 3 One dimensional examples

For illustrative purposes for verified inclusions by the conventional simple interval method and the Taylor model method, let us start our study from one dimensional systems.

#### 3.1 A simple function

The first example is the following simple one dimensional function

$$f(x) = x(x-1.1)(x+2)(x+2.2)(x+2.5)(x+3) \cdot \sin(1.7x+0.5)$$

and we try to find verified inclusions of the function in the domain [-0.5, 1.0]. All the pictures in Fig.1 show the function itself as the middle curve in the domain, and the function ranges in [-25.77, 1.16].

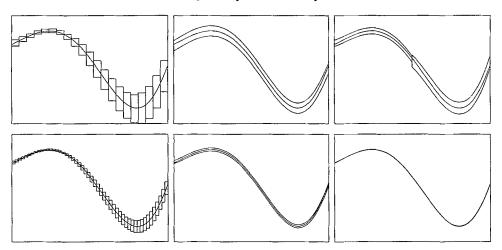


Fig. 1. Verified enclosures of a simple one dimensional function by 21 (left top) and 50 (left bottom) subintervals, and by one Taylor model with order seven (middle top) and eight (middle bottom), as well as two fourth order Taylor models (right top), and four fourth order Taylor models (right bottom).

With the interval method, compute the bounds of the function successively in smaller and smaller sub-domains, excluding sub-domains whose upper functional bounds fall below an already known function value. In this example, to keep the computational result fit inside the common sized frame, the whole domain has to be subdivided at least to 21 equal sized pieces (Fig.1, left top). In this case, the overall verified bounds of the function are [-31.03, 1.92]. The left bottom picture in Fig.1 shows the case of 50 equal sized sub-domains, resulting in the overall verified bounds [-27.94, 1.26].

The Taylor model approach provides the approximate polynomial  $P_{n,f}$ , and a band of uniform width  $I_{n,f}$  around  $P_{n,f}$  that encloses the function f. Not only the size of the sub-domain, but also the order n contribute to the size of the band. The two pictures in the middle in Fig.1 show one Taylor model covering the whole domain; the upper picture is to seventh order, and the lower to eighth order, having remainder intervals  $I_{7,f} = [-1.81, 1.79]$  and  $I_{8,f} = [-0.53, 0.54]$ , respectively. The pictures show that the band size of the seventh order non-divided Taylor model is already comparable to the local enclosure size of the 50 subdivided intervals. To compare the Taylor model method with the interval method, the subdivisions to two (right top) and to four (right bottom) equal sized sub-domains are shown in Fig.1 for Taylor models of the order four. With two subdivided Taylor models, the band size is comparable to the accuracy of the 50 subdivided intervals, and with four subdivided Taylor models, the band size almost cannot be identified in the picture.

#### 3.2 Gritton's function

The example here is the so-called Gritton's second problem from chemical engineering, addressed by Kearfott [14]. The system exhibits a typical cancellation problem, hence it presents a difficulty for conventional interval methods, even though it is a one dimensional system. The function is an 18th order polynomial

```
\begin{split} f(x) &= -371.9362500 - 791.2465656 \cdot x + 4044.944143 \cdot x^2 \\ &+ 978.1375167 \cdot x^3 - 16547.89280 \cdot x^4 + 22140.72827 \cdot x^5 \\ &- 9326.549359 \cdot x^6 - 3518.536872 \cdot x^7 + 4782.532296 \cdot x^8 \\ &- 1281.479440 \cdot x^9 - 283.4435875 \cdot x^{10} + 202.6270915 \cdot x^{11} \\ &- 16.17913459 \cdot x^{12} - 8.883039020 \cdot x^{13} + 1.575580173 \cdot x^{14} \\ &+ 0.1245990848 \cdot x^{15} - 0.03589148622 \cdot x^{16} \\ &- 0.0001951095576 \cdot x^{17} + 0.0002274682229 \cdot x^{18}, \end{split}
```

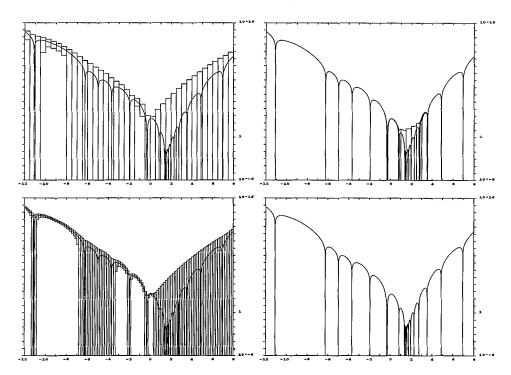


Fig. 2. Gritton's function in [-12, 8] evaluated by 40 (left top) and 120 (left bottom) subdivided intervals, and by 40 subdivided Taylor models with fourth order (right top) and eighth order (right bottom). The pictures show the absolute value of the function in logarithmic scale.

and has 18 roots in the range [-12,8]. The function varies roughly from  $-4 \times 10^{13}$  to  $6.03 \times 10^{14}$ , and all the local maxima and minima have different magnitudes. As an illustration of its complicated structure, we note that there are four local extrema in the range [1.4, 1.9] with function values varying only between around -0.1 and 0.1, while outside this region, function values are exceedingly large. Simple interval computation shows a severe overestimation due to cancellation.

To show the complicated behavior of the function, we choose a logarithmic scale. The pictures in Fig.2 over the whole domain [-12,8] show the absolute value of the function. The eighteen visible dips to the bottom frame correspond to zeros of the function. The left two pictures show enclosures by 40 (top) and 120 (bottom) subdivided intervals; when the local bound interval contains zero, the lower end of the interval range box reaches the bottom frame. As seen in the pictures, there is no advantage of smaller subdivisions in x > 0. The right two pictures are with 40 subdivided Taylor models. At order four (right top), there still remains a visible band width of the Taylor models in the range  $0.5 \le x \le 3$ , but at order eight (right bottom), the verified inclusion of the original function reaches printer resolution.

Table 1 Widths of the local bounds of Gritton's function around  $x_0 = 1.5$  by non-verified rastering, the simple interval method, and the 5th order Taylor model (TM) method. Also listed are the width of the remainder intervals of the Taylor models.

Sub-domain	Width of Local Bounds of the Function		Width of TM 5th	
Width	Rastering	Interval	TM 5th	Remainder Intervals
0.4	0.2323	144507.	1.455	0.7185
0.2	$2.353\times10^{-2}$	55675.	0.1274	$9.119 \times 10^{-3}$
0.1	$1.854\times10^{-2}$	24555.	$3.361\times10^{-2}$	$1.284 \times 10^{-4}$
0.05	$1.057\times10^{-2}$	11766.	$1.322\times10^{-2}$	$1.904 \times 10^{-6}$
0.025	$5.478\times10^{-3}$	5788.	$6.000\times10^{-3}$	$2.898 \times 10^{-8}$
0.0125	$2.764\times10^{-3}$	2872.	$2.877\times10^{-3}$	$4.469 \times 10^{-10}$

The performance in local range enclosures shows a dramatic difference between the simple interval method and the Taylor model method. For the purpose of a quantitative understanding, Table 1 lists the resulting width of the bounds in the sub-domains with various sizes at the reference point  $x_0 = 1.5$ . Listed are the simple interval evaluation, the fifth order Taylor model computation, and the non-verified rastering for comparison. The bounds in Taylor models used a very simple polynomial bound evaluator. For intervals, to achieve the comparable result to the Taylor models in the 0.1 width sub-domain, the domain would have to be cut to a practically hard to achieve size of  $10^{-7}$ .

## 4 Multidimensional systems

Multidimensional systems require more sub-domains with simple interval method as discussed in section 1, and even a simple function may show some challenge as will be shown below. Practical problems often have severe cancellation problems as in Gritton's case, and the last example in this paper discusses such a difficult case to obtain verified enclosures of the system.

#### 4.1 A small three dimensional function

We study a randomly chosen three dimensional function

$$f(x_1, x_2, x_3) = \frac{4\tan(3x_2)}{3x_1 + x_1\sqrt{\frac{6x_1}{77(x_1 - 8)}}} - 120 - 2x_1 - 7x_3(1 + 2x_2)$$

$$+\frac{5x_1 \tanh(0.9x_3)}{\sqrt{5x_2}} + \frac{(3x_2+13)^2}{3x_3} - \sinh\left(0.5 + \frac{6x_2}{8x_2+7}\right) -20x_3(2x_3-5) - 20x_2\sin(3x_3).$$

There are nine terms contributing to the result, each of which consists of not completely trivial arithmetic. Since each variable appears several times, terms depend on one another. Hence, a certain amount of overestimation due to the dependency problem is to be expected in conventional interval arithmetic.

To study the effect, we ask for the range enclosure of function over a domain  $[1.95, 2.05] \times [0.95, 1.05] \times [0.95, 1.05]$ . The non-verified range of the function is obtained by scanning at  $11^3$  equidistant points as [-2.31166, 1.78169]. On the

Range enclosures of the three dimensional function by non-verified rastering, the simple interval method, and the Taylor model method.

Non-Verified Real Number Rastering

Taylor Model Method

10

Table 2

Sampling Points	Non-Verified Range
113	[-2.31166, 1.78169]
Interval Method	
Subdivision	Total Range Enclosure
13	[-16.36393,16.09747]
$4^3$	[ -5.73777, 5.43391]
$16^{3}$	[ -3.16171, 2.69842]
$64^{3}$	[ -2.52387, 2.01107]

Order	Remainder Interval	Total Range Enclosure
1	[-0.39140 , 0.72524 ]	[-2.80268, 2.35080]
2	[-0.33950E-1 , 0.33940E-1 ]	[-2.48316, 1.84826]
3	[-0.10202E-2 , 0.16096E-2 ]	[-2.47884, 1.84454]
4	[-0.84132E-4 , 0.84028E-4 ]	[-2.47871, 1.84429]
5	[-0.24107E-5 , 0.43833E-5 ]	[-2.47866, 1.84424]
6	[-0.33555E-6 , 0.33431E-6 ]	[-2.47866, 1.84424]
7	[-0.16319E-7 , 0.20518E-7 ]	[-2.47866, 1.84424]
8	[-0.24246E-8 , 0.24107E-8 ]	[-2.47866, 1.84424]
9	[-0.17219E-9 , 0.17367E-9 ]	[-2.47866, 1.84424]

[-2.47866, 1.84424]

[-0.23138E-10, 0.22986E-10]

other hand, the simple interval arithmetic without subdivision gives a verified enclosure as [-16.36393, 16.09747], which is almost ten times wider than the range evaluated by scanning. To obtain a resulting precision closer to the one obtained by scanning, the domain has to be subdivided to 64<sup>3</sup> sub-domains, as listed in Table 2.

On the other hand, the Taylor model evaluation gives quite accurate enclosures without any subdivisions, and even a first order Taylor model gives a comparable result to the one obtained by scanning. The Taylor model arithmetic is performed around the center point (2,1,1), and Table 2 lists Taylor model computation as a function of the order, and again a very simple polynomial bound evaluator was used to get the total range enclosures. The size of remainder interval drops down with order quickly as expected, reaching ten digits of accuracy with order ten.

#### 4.2 A normal form deviation function

The last example is the pseudo-Lyapunov function of a weakly nonlinear dynamical system, characterizing a deviation from a normal form invariance. The function is a six dimensional polynomial up to roughly 200th order which involves a large number of local minima and maxima, but the function value itself is almost zero. So there is a large amount of cancellation, thus the problem represents a substantial challenge for interval methods.

As an example calculation, we ask for the range enclosure in a domain box [0.04, 0.06] in each of the six coordinate variables. The value of the function at the center point  $\vec{x}_0 = \overline{0.05}$  is

$$f(\vec{x}_0) = 6.97670 \times 10^{-6}$$
.

To give a rough measure of the actual size, a real number rastering at  $3^6$  equidistant points and 1000 random points is performed, resulting in the non-verified range as

$$[-3.12119 \times 10^{-6}, \ 4.21243 \times 10^{-5}] \qquad \text{(by non-verified rastering)}.$$

The simple interval computation without subdivision gives the bounds

$$[-4.47134, 4.80775]$$
 (by non-divided interval),

which dramatically shows the overestimation. Table 3 shows local enclosures in successively smaller sub-domain at the center. Only the smallest sub-domain

Table 3
Range enclosures of a normal form deviation function by non-verified rastering, the simple interval method, and the Taylor model method.

Non-Verified Real Number Rastering				
Sampling Points Non-Verified Range				
3 <sup>6</sup> and 1000 [-3.12119E-6, 4.21243E-5]				
Interval Method				
Local Domain Local Range Enclosure				
[0.040000, 0.060000] <sup>6</sup> [-4.47134 , 4.80774 ]				
[0.049000, 0.051000] <sup>6</sup> [-5.44365E-3, 1.00332E-2]				
$[0.049900, 0.050100]^6$ [-6.97752E-4, 7.62484E-4]				
[0.049990, 0.050010] <sup>6</sup> [-6.57704E-5, 8.02314E-5]				
[0.049999, 0.050001] <sup>6</sup> [-3.20844E-7, 1.42793E-5]				
Taylor Model Method				
Order Remainder Interval Total Range Enclosure				
6 [-5.3585E-6, 5.3588E-6] [-3.466E-5, 5.358E-5]				
7 [-8.3873E-7, 8.3884E-7] [-3.016E-5, 4.902E-5]				
8 [-1.2321E-7, 1.2321E-7] [-2.945E-5, 4.831E-5]				

yields the local enclosure of a size comparable to the scanned estimate. However, to cover the entire domain in this fashion requires  $10^{24}$  sub-domains, showing the practical limitations of the interval approach for this problem.

On the other hand, the Taylor model computation gives very small remainder intervals even without subdivision as shown in Table 3, and the total enclosures even using a very simple polynomial bound evaluator, are comparable to the scanned estimate. As discussed in section 1, the total possible number of monomials increases very moderately with order compared to the division number necessary for a comparable interval evaluation.

# Acknowledgements

We thank Baker Kearfott, George Corliss and William Walster for various stimulating discussions. The 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] R. E. Moore. Methods and Applications of Interval Analysis. SIAM, 1979.
- [2] G. Alefeld and J. Herzberger. Introduction to Interval Computations. Academic Press, 1983.
- [3] R. B. Kearfott and V. Kreinovich, editors. *Applications of Interval Computations*, volume 3. Kluwer, 1996.
- [4] G. W. Walster et al. FORTE<sup>TM</sup> FORTRAN/HPC, Interval Arithmetic, http://www.sun.com/forte/fortran/interval/index.html.
- [5] G. W. Walster. Philosophy and practicalities of interval arithmetic. In R. E. Moore, editor, Reliability in Computing, pages 309–323. Academic Press, 1988.
- [6] R. B. Kearfott. Rigorous Global Search: Continuous Problems. Kluwer, 1996.
- [7] G. F. Corliss and R. B. Kearfott. Rigorous global search: Industrial applications. In T. Csendes, editor, *Developments in Reliable Computing*, pages 1–16. Kluwer, 2000.
- [8] K. Makino and M. Berz. Remainder differential algebras and their applications. In M. Berz, C. Bischof, G. Corliss, and A. Griewank, editors, Computational Differentiation: Techniques, Applications, and Tools, pages 63-74, Philadelphia, 1996. SIAM.
- [9] K. Makino. Rigorous Analysis of Nonlinear Motion in Particle Accelerators. PhD thesis, Michigan State University, East Lansing, Michigan, USA, 1998. Also MSUCL-1093.
- [10] M. Berz. COSY INFINITY Version 8 reference manual. Technical Report MSUCL-1088, National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, MI 48824, 1997. See also http://cosy.nscl.msu.edu.
- [11] M. Berz. Modern Map Methods in Particle Beam Physics. Academic Press, San Diego, 1999.
- [12] K. Makino and M. Berz. Efficient control of the dependency problem based on Taylor model methods. *Reliable Computing*, 5:3–12, 1999.
- [13] M. Berz, K. Makino, and J. Hoefkens. Verified integration of dynamics in the solar system. *Nonlinear Analysis*, in print.
- [14] R. B. Kearfott. Private communication.