RIGOROUS INTEGRATION OF MAPS AND LONG-TERM STABIITY

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

Abstract

In spite of its importance, it has been very difficult to estimate long-term stability of particles in repetitive systems in a fully rigorous way. One of the main causes of the difficulty is the inaccuracy of the maps of the system; while to any fixed order they can be computed easily using Differential Algebraic (DA) techniques, it has so far not been possible to determine bounds for the remainders. Another difficulty is that most methods to rigorously formulate the problem lead to the need for global optimization of highly complex multi-dimensional objective functions. The Remainder-enhanced Differential Algebraic (RDA) method, an extension of the DA method that simultaneously provides rigorous bounds for the remainders, can solve both problems. The Taylor maps are evaluated rigorously with interval remainders, using the verified integral method within the framework of RDA. And rigorous RDA global optimization allows to efficiently get bounds on long-term stability.

1 INTRODUCTION

The problem to estimate the long-term stability of weakly nonlinear systems finds its origin in the detailed study of the solar system. Many perturbative methods for repetitive motion have been developed from this question. The question of long-term stability of particles in repetitive systems like circular accelerators and storage rings is one of new applications in this category.

High order Taylor transfer maps of the action on phase space describe the motion of particles very well. The differential algebraic (DA) techniques [1] [2] [3] [4] have offered a very elegant and accurate way to obtain such transfer maps. Typically derivatives of up to order ten in six variables are needed, so other methods are far from providing a robust way to study the weakly nonlinear behavior of beams.

Recently, ideas of Lyapunov, Nekhoroshev and others triggered an analysis of stability in particle accelerator based on approximate invariants[5][6][7], and the question of long-term stability can be re-cast into a highly complicated optimization problem [6] [8]. Figure 1 shows tracking pictures of a repetitive system described by a six dimensional Taylor map in the sixth order in actual coordinates and in normal form coordinates, where in the latter case the motion is seen as approximately stable. The deviations from invariance directly relate to the time for particles to be lost. The deviation functions, which are multidimensional polynomials up to roughly 500th order, consist of about 10^5 floating-point operations. The irregularity of the functions as well as the high dimensionality makes the question very troublesome for conventional optimization methods. The sharpness of the bounds of the functions is important in order to guarantee a large number of stable turns, but in reality the functions have a very large number of local maxima. To be useful, the maxima have to be sharp to about 10^{-6}, and for some applications to 10^{-12}. Figure 2 shows the deviations from a normal form invariant circle in Figure 1 as a function of two angles. Interval methods give a mathematically rigorous estimate, but complicated functions like ours cannot avoid a severe blow-up problem, the control of which is the key to get a practical estimate.

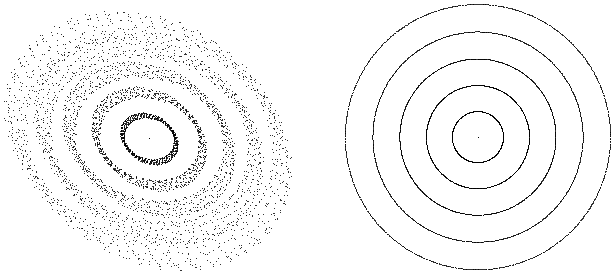

Figure 1: Tracking pictures of a repetitive system. The left is in particle optical coordinates x and $a = p_x/p_0$, and the right is in normal form coordinates.

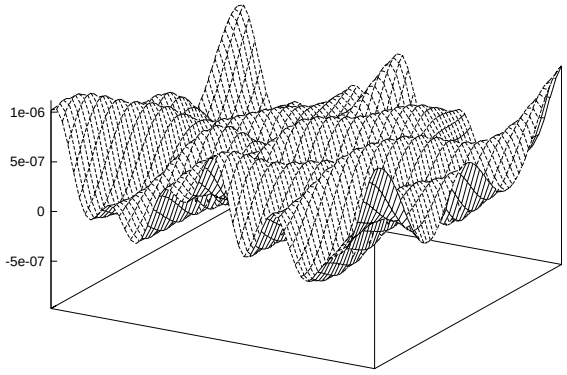

Figure 2: Deviation from normal form invariance of a repetitive system.

A new technique, the method of Remainder-enhanced Differential Algebras (RDA) [9][10][11][12], combines the DA technique to express the model function by a Taylor polynomial, and interval computation to evaluate the interval bound of the Taylor remainder. The resulting error bounds are usually rather sharp, in particular at higher or-

*Work in collaboration with Martin Berz, Department of Physics and Astronomy and National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, MI 48824.

0-7803-4376-X/98/$10.00 © 1998 IEEE

ders. In normal form theory, it is known that if the transfer map has a converging Taylor series in normal form coordinates, the system is stable. So the sharpness of the remainder term of the deviation function is the key to the problem. Thus, the RDA method now can give a practical answer to the question.

To complete the question, the other important aspect is to have rigorous Taylor transfer maps to describe the weakly nonlinear systems. This paper discusses the theory for rigorous integration of ordinary differential equations within the framework of RDA, which enables to obtain rigorous Taylor transfer maps[13]. The method has been implemented in the code COSY INFINITY[14][15], and the paper covers some example calculations.

2 RIGOROUS COMPUTATION BY INTERVAL METHODS

An interval represents a number on computers in a rigorous way, by rounding the number downwards to the lower bound and upwards to the upper bound. It can also represent an extended domain of numbers. These two features enable interval calculations on computers to make reliable statements. However, in practice, interval methods have some severe disadvantages, which make the mere interval method useless for calculation of complicated functions: The width of resulting intervals scales with the width of original intervals; and a blow-up occurs in extended calculations, which can be seen in the simple example, $[a,b]-[a,b] = [a-b, b-a] \neq [0,0]$. Furthermore, a difficulty exists in case of multiple dimensions d with n sampling points, because computational expense scales with n^d.

A new approach, the Remainder-enhanced Differential Algebraic (RDA) method, provides remedy to these disadvantages of interval computations[9].

3 REMAINDER-ENHANCED DIFFERENTIAL ALGEBRAS

3.1 Taylor Models

A C^∞ function $f : [\vec{a}, \vec{b}] \subset R^v \to R$ can be expressed by the n-th order Taylor polynomial P_n expanded around the reference point $\vec{x}_0$ and a remainder ε_n as $f(\vec{x}) = P_n(\vec{x} - \vec{x}_0) + \varepsilon_n(\vec{x} - \vec{x}_0)$. Let the interval I_n be such that $\forall \vec{x} \in [\vec{a}, \vec{b}], \varepsilon_n(\vec{x} - \vec{x}_0) \in I_n$. Then

$$\forall \vec{x} \in [\vec{a}, \vec{b}], \quad f(\vec{x}) \in P_n(\vec{x} - \vec{x}_0) + I_n.$$

Because of the special form of the Taylor remainder term ε_n, in practice usually the remainder decreases as $|\vec{x} - \vec{x}_0|^{n+1}$. Hence, if $|\vec{x} - \vec{x}_0|$ is chosen to be small, the interval remainder bound I_n can become very small. We say a pair (P_n, I_n) is an n-th order Taylor model of f.

3.2 Taylor Models for addition, multiplication and intrinsic functions

The question now is how to efficiently determine Taylor models for any arbitrary functional dependency that can be expressed on a computer[9]. The key is to begin with the Taylor model for the identity function, which is trivial, and then successively build up Taylor models for the total function from its pieces. This requires methods to determine Taylor models for sums and products from those of the summands or factors.

Let the functions $f, g : [\vec{a}, \vec{b}] \subset R^v \to R$ have Taylor models $(P_{n,f}, I_{n,f})$, $(P_{n,g}, I_{n,g})$. Then an n-th order Taylor model for $f + g$ is obviously obtained as

$$(P_{n,f} + P_{n,g}, I_{n,f} + I_{n,g}).$$

An n-th order Taylor model for $f \cdot g$ is obtained as

$$\begin{aligned} & (P_{n,f\cdot g}, I_{n,f\cdot g}), \quad \text{where} \\ I_{n,f\cdot g} \;=\; & B(P_{n,f} \cdot P_{n,g} - P_{n,f\cdot g}) + B(P_{n,f}) \cdot I_{n,g} \\ & + B(P_{n,g}) \cdot I_{n,f} + I_{n,f} \cdot I_{n,g}, \end{aligned}$$

with $B(P)$ denoting a bound of the polynomial P. The key idea of computing Taylor models for intrinsic functions is to employ Taylor's theorem of the function under consideration.

3.3 Taylor Models for Derivations and Anti-derivations

For many practical problems, in particular the efficient solution of differential equations, it is actually important to complement the set of operations by a derivation ∂, as well as its inverse ∂^{-1}, similar as in other differential algebraic approaches.

Given an n-th order Taylor model (P_n, I_n) of a function f, we can determine a Taylor model for the indefinite integral $\partial_i^{-1} f = \int f \, dx_i'$ with respect to variable i. The operator ∂_i^{-1}on the space of Taylor models is defined as

$$\begin{aligned} & \partial_i^{-1}(P_n, I_n) \\ = \; & \left(\int_0^{x_i} P_{n-1} dx_i', \; (B(P_n - P_{n-1}) + I_n) \cdot B(x_i) \right). \end{aligned}$$

Similar to the case of the Differential Algebra on the set of Truncated Power Series, and following one of the main thrusts of the theory of Differential Algebras, we will use these for the solution of the initial value problem

$$\frac{d}{dt}\vec{r}(t) = \vec{F}(\vec{r}(t), t), \tag{1}$$

where $\vec{F}$ is continuous and bounded. We are interested in both the case of a specific initial condition $\vec{r}_0$, as well as the case in which the initial condition $\vec{r}_0$ is a variable, in which case our interest is in the flow of the differential equation

$$\vec{r}(t) = \mathcal{M}(\vec{r}_0, t).$$

3.4 Example computation

We made a comparison computation to show how the new method works to obtain a rigorous bound. We used a deviation function from a normal form invariance used in Figures 1 and 2 to get a bound within the domain intervals $[.04, .06]$ in each six coordinate variables. To get the rough idea of the actual size, we made a real number scan at 1000 random points and at 3^6 points at edges and center in each dimension. The bound of values by the scan is

```
[-.3121185581961283E-05,0.4212429306152572E-04].
```

The mere interval computation gave the bound

```
[-4.471335284762441    , 4.807741733133240    ],
```

which is rigorous but useless because of a severe blow-up. Now, the remainder bound carried by the sixth order Taylor model computation is

```
[-.5358533718862318E-05,0.5358814729171932E-05]
```

and it added up to a total bound of

```
[-.3466186723563667E-04,0.5352931790602934E-04].
```

The comparison with the estimate by the scan shows the practical strength of the RDA method.

4 VERIFIED INTEGRAL WITH TAYLOR MODELS

Our goal is now to determine a Taylor model for the flow $\mathcal{M}(\vec{r}_0, t)$ of the differential equation (1). The remainder bound should be fully rigorous for all initial conditions $\vec{r}_0$ and times t that satisfy

$$\vec{r}_0 \in [\vec{r}_{01}, \vec{r}_{02}] = \vec{B}, \quad t \in [t_0, t_1].$$

In particular, $\vec{r}_0$ itself may be a Taylor model, as long as its range is known to lie in $\vec{B}$.

Since conventional numerical integrators don't provide rigorous estimates for the integration error but only approximate estimates, we have to start from scratch from the foundations of the theory of differential equations[13].

4.1 Schauder's Fixed Point Theorem

We re-write the differential equation as an integral equation

$$\vec{r}(t) = \vec{r}_0 + \int_{t_0}^{t} \vec{F}(\vec{r}(t'), t')\, dt',$$

noting that the initial value problem has a (unique) solution if and only if the corresponding integral equation has a (unique) solution. Now we introduce the operator $A : \vec{C}^0[t_0, t_1] \to \vec{C}^0[t_0, t_1]$ on the space of continuous functions from $[t_0, t_1]$ to R^v via

$$A\left(\vec{f}\right)(t) = \vec{r}_0 + \int_{t_0}^{t} \vec{F}\left(\vec{f}\,(t'), t'\right)\, dt'.$$

Then the problem of finding a solution to the differential equation is transformed to a fixed-point problem on the space of continuous functions

$$\vec{r} = A(\vec{r}).$$

We will now apply Schauder's fixed point theorem to rigorously obtain a Taylor model for the flow.

Theorem (Schauder): *Let A be a continuous operator on the Banach Space X. Let $M \subset X$ be compact and convex, and let $A(M) \subset M$. Then A has a fixed point in M, i.e. there is an $\vec{r} \in M$ such that $A(\vec{r}) = \vec{r}$.*

4.2 Strategy to Satisfy the Requirements of Schauder's Theorem

In our specific case, $X = \vec{C}^0[t_0, t_1]$, the Banach space of continuous functions on $[t_0, t_1]$, equipped with the maximum norm, and the integral operator A is continuous on X. The process to apply Schauder's theorem consists of the following steps:

- Determine a family Y of subsets of X, the Schauder Candidate Sets. Each set in Y should be compact and convex, it should be contained in a suitable Taylor model, and its image under A should be in Y.
- Using differential algebraic methods on Taylor models, determine an initial set $M_0 \in Y$ that satisfies the inclusion property $A(M_0) \subset M_0$. Then all requirements of Schauder's theorem are satisfied, and M_0 contains a solution.
- Iteratively generate the sequence $M_i = A(M_{i-1})$ for $i = 1, 2, 3, ...$ Each M_i also satisfies $A(M_i) \subset M_i$, and we have $M_1 \supset M_2 \supset ....$ We continue the iteration until the size stabilizes sufficiently.

4.3 Schauder Candidate Sets

For the first step, it is necessary to establish a family of sets Y from which to draw candidates for M_0. Let $(\vec{P} + \vec{I})$ be a Taylor model depending on time as well as the initial condition $\vec{r}_0$. Then we define the associated set $M_{\vec{P}+\vec{I}}$ as follows:

$$\begin{aligned}
M_{\vec{P}+\vec{I}} &\subset \vec{C}^0[t_0, t_1]; \text{ and for } \vec{r} \in M_{\vec{P}+\vec{I}}: \\
\vec{r}(t_0) &= \vec{r}_0 \\
\vec{r}(t) &\in \vec{P} + \vec{I}\, \forall t \in [t_0, t_1]\ \forall \vec{r}_0 \\
|\vec{r}(t') - \vec{r}(t'')| &\leq k|t' - t''|\ \forall t', t'' \in [t_0, t_1]\ \forall \vec{r}_0.
\end{aligned}$$

In the last condition, k is a bound for $\vec{F}$, which exists because $\vec{F}$ is continuous and the solutions can cover only a finite range over the interval $[t_0, t_1]$. The last condition means that all $\vec{r} \in M_{\vec{P}+\vec{I}}$ are uniformly Lipschitz with constant k. Define the family of candidate sets Y as

$$Y = \bigcup_{\vec{P}+\vec{I}} M_{\vec{P}+\vec{I}}.$$

4.4 *Convexity, Compactness, and Invariance of Schauder Candidate Sets*

Let $M \in Y$. Then M is convex, because $\vec{x}_1, \vec{x}_2 \in M \Rightarrow \alpha\vec{x}_1 + (1-\alpha)\vec{x}_2 \in M \; \forall \alpha \in [0,1]$.

Furthermore, M is compact, i.e. any sequence in M has a clusterpoint in M. To see this, let $(\vec{x}_n)$ be a sequence of functions in M. Then by definition of M, $(\vec{x}_n)$ is uniformly Lipschitz, and thus uniformly equicontinuous. $(\vec{x}_n)$ is also uniformly bounded, and hence according to the Ascoli-Arzela Theorem, has a uniformly convergent subsequence. Since the $\vec{x}_n$ are continuous, so is the limit $\vec{x}^*$ of this subsequence, and since M is closed, the limit $\vec{x}^*$ is in M.

Finally, A maps Y into itself, and the uniform Lipschitzness follows because $\vec{F}$ is bounded by k.

4.5 *Satisfying the Schauder Inclusion Requirement with Differential Algebraic Methods*

The only remaining requirements for Schauder's theorem is to find a Taylor model $\vec{P} + \vec{I}$ such that

$$A(\vec{P} + \vec{I}) \subset \vec{P} + \vec{I}.$$

This condition can be checked computationally using the differential algebraic operations on the set of Taylor models. To succeed with the inclusion requirement depends on finding suitable choice for $\vec{P}$ and $\vec{I}$. Furthermore, it is desirable to have $\vec{I}$ tight. Both benefit from the choice of a polynomial $\vec{P}$ that is already "close" to the true solution of the ODE.

Attempt sets M^* of the form

$$M^* = M_{\vec{P}^* + \vec{I}^*}, \quad \text{where} \quad \vec{P}^* = \mathcal{M}_n(\vec{r}_0, t),$$

the n-th order Taylor expansion of the flow of the ODE. It is to be expected that $\vec{I}^*$ can be chosen smaller and smaller as the order n of $\vec{P}^*$ increases.

This requires the knowledge of the n-th order flow $\mathcal{M}_n(\vec{r}_0, t)$, including time dependence. It can be obtained by iterating in conventional DA. To this end, one chooses an initial function $\mathcal{M}_n^{(0)}(\vec{r}, t) = \mathcal{I}$, where $\mathcal{I}$ is the identity function, and then iteratively determines

$$\mathcal{M}_n^{(k+1)} =_n A(\mathcal{M}_n^{(k)}).$$

In case $\vec{F}$ is origin preserving, this process converges to the exact DA result $\mathcal{M}_n$ in exactly n steps.

Now try to find $\vec{I}^*$ such that

$$\mathcal{M}_n + \vec{I}^* \subset A(\mathcal{M}_n + \vec{I}^*),$$

the Schauder inclusion requirement. The suitable choice for $\vec{I}^*$ requires experimenting, but is greatly simplified by the observation

$$\begin{aligned} \vec{I}^* &\supset \vec{I}^{(0)}, \quad \text{where} \\ \mathcal{M}_n(\vec{r}, t) + \vec{I}^{(0)} &= A(\mathcal{M}_n(\vec{r}, t) + [\vec{0}, \vec{0}]). \end{aligned}$$

Evaluating the right hand side in the RDA yields a lower bound for $\vec{I}^*$, and a benchmark for the size to be expected. Now iteratively try

$$\vec{I}^{(k)} = 2^k \cdot \vec{I}^{(0)},$$

until a computational inclusion is found, i.e.

$$A(\mathcal{M}_n(\vec{r}, t) + \vec{I}^{(k)}) \subset \mathcal{M}_n(\vec{r}, t) + \vec{I}^{(k)}.$$

4.6 *Iterative Refinement of the Inclusion*

Once a computational inclusion has been determined, the solution of the ODE is known to be contained in the Taylor model $\mathcal{M}_n(\vec{r}, t) + \vec{I}^{(k)}$. Set $\vec{I}_{(1)} = \vec{I}^{(k)}$; since the solution is a fixed point of A, it is even contained in

$$A^k(\mathcal{M}_n(\vec{r}, t) + \vec{I}_{(1)}) \quad \text{for all } k.$$

Furthermore, the iterates of A are shrinking in size, i.e.

$$A^k(\mathcal{M}_n(\vec{r}, t) + \vec{I}_{(1)}) \subset A^{k-1}(\mathcal{M}_n(\vec{r}, t) + \vec{I}_{(1)}) \; \forall k.$$

So the width of the remainder bound of the flow can be decreased by iteratively determining

$$\mathcal{M}_n(\vec{r}, t) + \vec{I}_{(k)} = A(\mathcal{M}_n(\vec{r}, t) + \vec{I}_{(k-1)}),$$

until no further significant decrease in size is achieved. As a result,

$$\mathcal{M}_n(\vec{r}, t) + \vec{I}_{(k)}$$

is the desired sharp inclusion of the flow of the original ODE.

5 EXAMPLES OF INTEGRATION

In this section, we will provide two examples for the practical use and performance of the method discussed above. The first example is to test the integration algorithm; it is the motion on a circle defined by the differential equations and initial conditions

$$\dot{x} = -y, \quad \dot{y} = x, \quad x(0) = 1, \; y(0) = 0.$$

The integration from 0 to 2π was performed using tenth order Taylor models with a fixed step size of $\pi/36$. The resulting interval inclusions based on double precision interval arithmetic are

```
+1.00000000E+00+[-.43837892E-13,+0.43837892E-13]
-0.63043563E-14+[-.43587934E-13,+0.43587934E-13].
```

Another example is to analyze the motion of a charged particle in a homogeneous dipole magnet. The flow of the differential equation over a region of initial conditions is determined. The integration was carried out through the dipole with the deflection radius of 1m over a deflection angle of 36 degrees with a fixed step size of 4 degrees. The initial conditions of four phase space variables, x,

$a = p_x/p_0$, y and $b = p_y/p_0$, are within the domain intervals

$$[-.02, .02] \times [-.02, .02] \times [-.02, .02] \times [-.02, .02],$$

and the Taylor polynomial describing the dependence of the four final coordinate values on the four initial coordinate values was determined. The order in time and initial conditions was chosen to be 12, and the step size was estimated so as to ascertain an overall accuracy below 10^{-9}; since no automatic step size control was utilized, the estimate proved conservative and the actual resulting error was somewhat lower:

```
[-0.4496880372277553E-09,+0.3888593417126594E-09]
[-0.1301070602141642E-09,+0.1337099965985420E-09]
[-0.3417079805637740E-10,+0.3417079805637740E-10]
[-0.0000000000000000E+00,+0.0000000000000000E+00]
```

The resulting Taylor polynomials describing the dependence of final on initial coordinates were compared with those obtained by our particle optics code COSY INFINITY[14][15], and agreement was found. A further check was to compare a large collection of rays through the dipole obtained by COSY INFINITY with ones through the results of the flow calculated by the verified integrator. For all rays studied, the difference between the final coordinates determined geometrically by the dipole element in COSY INFINITY and those predicted by the twelfth order Taylor polynomial were within the calculated remainder bounds.

6 OUTLOOK

Automatic step size and order control is expected to allow the integrator to automatically assure pre-specified accuracy goals at nearly optimal speed.

For the study of rigorous long-term stability, the development of other invariant classes based on wavelets, or possibly Fourier-splines is important, as well as the utilization of multi-turn maps and the development of more sophisticated understanding of actual escape mechanisms. Finally, the method discussed here will facilitate a modification of existing accelerator lattices to minimize occurring invariant defects.

7 ACKNOWLEDGEMENTS

We are grateful for the financial support of this research to the Department of Energy, Grant No. DE-FG02-95ER40931, and the Alfred P. Sloan Foundation.

8 REFERENCES

[1] M. Berz. Differential algebraic description of beam dynamics to very high orders. *Particle Accelerators*, 24:109, 1989.

[2] M. Berz. Arbitrary order description of arbitrary particle optical systems. *Nuclear Instruments and Methods*, A298:426, 1990.

[3] M. Berz. *High-Order Computation and Normal Form Analysis of Repetitive Systems, in: M. Month (Ed), Physics of Particle Accelerators*, volume AIP 249, page 456. American Institute of Physics, 1991.

[4] M. Berz. Modern map methods for charged particle optics. *Nuclear Instruments and Methods*, 363:100, 1995.

[5] R. L. Warnock and R. D. Ruth. Long–term bounds on nonlinear Hamiltonian motion. *Physica D*, 56(14):188–215, 1992. also SLAC–PUB–5267.

[6] M. Berz and G. Hoffstätter. Exact bounds of the long term stability of weakly nonlinear systems applied to the design of large storage rings. *Interval Computations*, 2:68–89, 1994.

[7] G. H. Hoffstätter. *Rigorous bounds on survival times in circular accelerators and efficient computation of fringe–field transfer maps*. PhD thesis, Michigan State University, East Lansing, Michigan, USA, 1994. also DESY 94-242.

[8] M. Berz, G. Hoffstätter, W. Wan, K. Shamseddine, and K. Makino. COSY INFINITY and its applications to nonlinear dynamics. *in: Computational Differentiation: Techniques, Applications, and Tools*, SIAM, 1996.

[9] K. Makino and M. Berz. Remainder differential algebras and their applications. *in: Computational Differentiation: Techniques, Applications, and Tools*, SIAM, 1996.

[10] M. Berz. Differential algebras with remainder and rigorous proofs of long-term stability. In *Fourth Computational Accelerator Physics Conference*. AIP Conference Proceedings, 1996.

[11] M. Berz. From Taylor series to Taylor models. In *Proceedings, First Accelerator Theory Institute*. AIP, 1997.

[12] K. Makino and M. Berz. Implementation and applications of Taylor model methods. *Reliable Computing*, submitted, 1997.

[13] M. Berz and K. Makino. Integration of ODEs and flows with differential algebraic methods on Taylor models. *Reliable Computing*, submitted, 1997.

[14] M. Berz. COSY INFINITY Version 7 reference manual. Technical Report MSUCL-977revised, National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, MI 48824, 1996.

[15] M. Berz. COSY INFINITY version 7. In *1997 Particle Accelerator Conference*. these proceedings, 1997.