Arbitrary Order Maps, Remainder Terms, and Long Term Stability in Particle Accelerators

Martin Berz and Kyoko Makino

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

ABSTRACT

It is reviewed how transfer maps of arbitrary order can be calculated in a convenient way using Differential Algebraic (DA) Techniques. Such high-order maps are of use for high-precision optical systems, as well as for the case of repetitive particle accelerators, where nonlinear effects often have a tendency to build up over many turns.

It is shown how the methods can be extended to also allow a rigorous treatment of the Lagrange remainder term of the Taylor expansion. The methods can be applied for the estimation of the effects going beyond the aberration order; it is also for the study of the repetitive motion where, leaning on ideas of Lyapunov, Nekhoroshev and others, it for the first time allows a rigorous determination of guaranteed particle survival times.

Keywords: Transfer map, High order aberrations, Differential Algebras, Truncated Power Series Algebra, Remainder-enhanced Differential Algebras, Taylor model, COSY INFINITY, Verified integration

1. MAP COMPUTATION

The effect of an optical system is described by the transfer map, which relates final conditions $\vec{z}_f$ to initial conditions $\vec{z}_i$ and possibly system parameters $\vec{\delta}$ via

$$\vec{z}_f = \mathcal{M}(\vec{z}_i, \vec{\delta}).$$

Motion in optical systems is dominated by linear effects, and because of this the map is often Taylor-expanded; the expansion coefficients have a practical meaning, as they describe the aberrations of the system under consideration.

The treatment of the nonlinear effects could be unified to all orders by the differential algebraic method.[1,2] It is based on the observation that for any functional dependence on the computer, it is possible to determine the Taylor expansion of the function in a very convenient way by building up the final Taylor expansion step by step via the respective arithmetic operations. The knowledge of the Taylor expansions of summands or factors allows the determination of the Taylor expansion of the sum or product; hence letting T denote the process of Taylor expansion to a given order n, it is possible to close the diagrams

$$\begin{tikzcd} f,g \arrow[r, "T"] \arrow[d, "{+,-}"'] & F,G \arrow[d, "{\oplus,\ominus}"] & f,g \arrow[r, "T"] \arrow[d, "{\cdot,/}"'] & F,G \arrow[d, "{\odot,\oslash}"] \\ f \pm g \arrow[r, "T"'] & F \genfrac{}{}{0pt}{}{\oplus}{\ominus} G & f \genfrac{}{}{0pt}{}{\cdot}{/} g \arrow[r, "T"'] & F \genfrac{}{}{0pt}{}{\odot}{\oslash} G \end{tikzcd} \tag{1}$$

with the introduction of suitable arithmetic operations $\oplus, \ominus$ as well as $\odot, \oslash$. The resulting structure, together with the multiplication with reals, forms an algebra, the so-called Truncated Power Series Algebra TPSA.[3,4] It allows the

Other author information:
M. B.: Email: berz@pa.msu.edu; WWW: http://www.beamtheory.nscl.msu.edu/; Telephone: 517-333-6313; Fax: 517-353-5967.
K. M.: Email: makino@nscl.msu.edu; WWW: http://www.beamtheory.nscl.msu.edu/; Telephone: 517-333-6441; Fax: 517-353-5967.

convenient computation of the Taylor expansion to any order of a functional dependence that can be represented in a computer environment.

In this vein, the computation of transfer maps can be performed by taking a conventional integrator that integrates the equations of motion, and perform all of its arithmetic in TPSA. Practically straightforward, care must be taken only to adjust the step size in the properly, since a step size that was sufficiently small to integrate particles is usually not sufficiently small to obtain higher aberrations accurately.

While robust and easy to use, the main drawback of the method is its efficiency. Significant improvements of several orders of magnitude in speed can be obtained by custom-made integrators.[2] These are based on augmenting TPSA by the derivation operation ∂, which naturally exists on function spaces. So we have an additional commuting diagram

$$\begin{tikzcd} f \arrow[r, "T"] \arrow[d, "{\partial, \partial^{-1}}"'] & F \arrow[d, "{\partial_{\bigcirc}, \partial_{\bigcirc}^{-1}}"] \\ \partial f, \partial^{-1} f \arrow[r, "T"'] & \partial_{\bigcirc} F, \partial_{\bigcirc}^{-1} F \end{tikzcd} \tag{2}$$

and the resulting structure becomes a differential algebra.[5,1] The integrators based on the differential algebraic derivation are of arbitrary order, and to obtain high precision, in practice[6,7] usually one utilizes large steps and orders in the range of 25 through 35.

2. TREATMENT OF REMAINDERS OF MAP BY RDA METHODS

Recently it was realized[5,8,9] that not only is it possible to obtain Taylor maps of any functional relationship, but that one can even obtain rigorous bounds on the remainder terms. This is convenient for many practical applications, as it provides an estimate about the accuracy of the map of the given order; it can also be used for the estimation of the effect of measurement uncertainties.[10] Furthermore, for applications demanding any kind of rigor, it is even essential. One important case in the field of optics is the estimation of long-term stability described below.

Interval methods have a long history of providing rigorous statements based on computations. The idea is to represent all numbers on a computer by including them in an interval. By rounding down for lower bounds and up for upper bounds, rigorous bounds for sums and products can be found, and thus the diagrams in eq. (1) can be made to commute. Similarly, intervals can be used to represent an extended domain of numbers.

However, in practice, interval methods have some severe disadvantages, which in certain cases even 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.

The new approach presented here, the Remainder-enhanced Differential Algebraic (RDA) method, provides a remedy to these disadvantages of interval computations.[8]

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.

Having introduced Taylor models, the question is now how to make the diagrams in eq. (1) commute. This requires methods to determine Taylor models for sums and products from those of the summands or factors.

Let the functions $f, g : [\vec{a}, \vec{b}] \subset R^v \rightarrow R$ 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

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

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; but for reasons of space, we will restrict ourselves from discussing details here.

For questions of integration of ODEs, it is also necessary to study the differential algebraic operators ∂ and ∂^{-1}. 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

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

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

$$\frac{d}{dt}\vec{r}(t) = \vec{F}(\vec{r}(t), t), \tag{3}$$

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. VERIFIED INTEGRATION 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 (3). 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 do not provide rigorous estimates for the integration error but only approximate estimates, the brute force method of utilizing a numerical integrator and viewing it as a functional dependence cannot be applied. Rather, we have to start from scratch from the foundations of the theory of differential equations.[11] To this end, 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 $A : \vec{C}^0[t_0, t_1] \rightarrow \vec{C}^0[t_0, t_1]$ 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 $A(\vec{r}) = \vec{r}$.*

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.

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

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

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.

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.

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

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.

As an example for the practical use of the method, we 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,[7,6] 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.

4. LONG-TERM STABILITY ESTIMATES

The RDA methods described in this paper can be used for the rigorous estimate of long-term stability of motion in repetitive systems. The method is based on the search of families of functions that are almost invariants of the motion, and the rigorous determination of the defect from invariance. It is commonly used in theoretical arguments about stability (see for example,[12] but in practice it presents a significant challenge to make the necessary estimates. Within a computer environment, using the RDA methods these estimates can be made sharply and economically.

As an example, we made a comparison computation to show how the new method works to obtain a rigorous bound. The family of nearly-invariant functions were obtained with normal form techniques,[2] and we used RDA techniques to obtain a bound within the domain intervals $[.04, .06]$ in each of the 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 quite 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.

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.

REFERENCES

1. M. Berz, "Differential algebraic description of beam dynamics to very high orders," *Particle Accelerators* **24**, p. 109, 1989.
2. M. Berz, *High-Order Computation and Normal Form Analysis of Repetitive Systems, in: M. Month (Ed), Physics of Particle Accelerators*, vol. AIP 249, p. 456. American Institute of Physics, 1991.
3. M. Berz, "The new method of TPSA algebra for the description of beam dynamics to high orders," Tech. Rep. AT-6:ATN-86-16, Los Alamos National Laboratory, 1986.
4. M. Berz, "The method of power series tracking for the mathematical description of beam dynamics," *Nuclear Instruments and Methods* **A258**, p. 431, 1987.
5. M. Berz, "From Taylor series to Taylor models," in *Proceedings, First Accelerator Theory Institute*, vol. in print, AIP, 1997.
6. K. Makino and M. Berz, "COSY INFINITY Version 7," in *Fourth Computational Accelerator Physics Conference*, vol. 391, p. 253, AIP Conference Proceedings, 1996.
7. M. Berz, "COSY INFINITY Version 7 reference manual," Tech. Rep. MSUCL-977, National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, MI 48824, 1996. see also http://www.beamtheory.nscl.msu.edu/cosy.
8. K. Makino and M. Berz, "Remainder differential algebras and their applications," *in: Computational Differentiation: Techniques, Applications, and Tools, M. Berz, C. Bischof, G. Corliss, A. Griewank (Eds.)* **SIAM**, 1996.
9. M. Berz and G. Hoffstätter, "Computation and application of Taylor polynomials with interval remainder bounds," *Interval Computations, in print* , 1997.
10. K. Makino and M. Berz, "Arbitrary order aberrations for elements characterized by measured fields," in *these proceedings, Optical Science, Engineering and Instrumentation '97*, SPIE, 1997.
11. M. Berz and K. Makino, "Integration of ODEs and flows with differential algebraic methods on Taylor models," *Reliable Computing* **submitted**, 1997.
12. N. N. Nekhoroschev, "An exponential estimate of the time of stability of nearly integrable hamiltonian systems," *Uspekhi Mat. Nauk 32:6, English translation Russ. Math. Surv.* **32:6,5**, p. 1, 1977.