2.3.5 Differential Algebraic Techniques

M. Berz, MSU

The study of the behavior of motion in the vicinity of a chosen reference solution is a central problem arising in many subfields of dynamical systems, including beam dynamics. The Taylor expansions of these solutions can be obtained by solving the so-called variational equations, which in beam physics has been carried out to orders two and three in the code Transport [1], to orders three for example in the codes TRIO [2], GIOS [3] and MaryLie [4], and to order five in COSY 5.0 [5]. This approach is laborious in practice, and the development of the DA techniques has greatly simplified this endeavor in beam physics and other fields. In their latest versions [6]–[9], the unprecedented accuracy these methods afford for the solution of differential equations has been awarded the Moore prize for rigorous computing.

A review of DA as used in our field can be found in [10]. The DA techniques allow the convenient computation of high-order Taylor expansions of the transfer map $\mathcal{M}$ which relates final particle coordinates $\vec{z}_f$ to initial coordinates $\vec{z}_i$ and parameters $\vec{\delta}$,

$$\vec{z}_f = \mathcal{M}(\vec{z}_i, \vec{\delta}) \qquad (1)$$

relative to one (in most beamlines, microscopes, or synchrotrons) or several (in FFAs, cyclotrons and some spectrographs) reference orbits. The map $\mathcal{M}$ is the *flow* (solution depending on initial conditions) of ODEs

$$\vec{z}' = \vec{f}(\vec{z}, s), \quad \text{independent variable } s \qquad (2)$$

Since their introduction [11, 12], DA techniques have been utilized in most newly developed codes [13]–[18] and the DA engines forming the core of COSY INFINITY [13] also form the backbone of various other codes. DA methods have their origin in the algebraic study of problems involving differentiation and integration for the purpose of solving complicated integrals, ODEs, and PDEs. Pioneered by Liouville [19] in connection with the search of elementary integrals for elementary functions, it was put on a solid foundation and significantly enhanced by Ritt [20] and Kolchin [21]. Now the methods have gained prominence in the field of formula manipulation, where they provide the backbone of the theory of analytic quadrature and integration of ODEs [22].

For a given function f of v variables, we form a vector that contains all Taylor expansion coefficients at $\vec{x} = \vec{0}$ up to a certain order n. Knowing this vector for two functions f and g allows to compute the respective vector for $f + g$ and $f \cdot g$, since the derivatives of the sum and product function are uniquely defined from those of f and g. The resulting operations of addition and multiplication lead to an algebra, the Truncated Power Series Algebra (TPSA) [12, 23]. One can also introduce elementary functions like exp, sin etc. on TPSA. TPSA allows the convenient computation of derivatives of any functional dependency on a computer [24].

For the solution of ODEs and PDEs it is necessary to introduce another operation. For any fixed function g with $g(0) = 0$, it is possible to determine the Taylor coefficients of $g \cdot \partial f/\partial x_i = \partial_{g,i} f$ from those of f. Including the operation $\partial_{g,i}$ and its inverse with addition and multiplication leads to a differential algebra (DA). Details on this particular DA can be found in [10, 24]. In passing it is worth mentioning that the DA structures also provide a novel way to a theory of differentials as infinitely small numbers [10]. For practical work with DA, care has to be taken to provide elementary operations $+, \cdot, \partial_{g,i}$ that work to any order and any number of variables. Since usually many derivatives vanish due to symmetry, the active support of sparsity is essential [23].

Solutions of ODEs and PDEs To determine the map Eq. (1), it is necessary to solve the ODEs Eq. (2). Since usually fields are known only in the midplane, it is often also necessary to solve their PDEs in the process. The crudest approach to solve Eq. (2) is to replace all operations in a tracking code by the corresponding ones in DA [12, 11]. This replacement is similar in flavor to the modification of existing code through pre-processors performing "automatic differentiation" [24], although in that field derivatives are typically only obtained to first or second order and the challenge lies in the efficient handling of large numbers of independent variables. An important practical problem of this map integration approach is to obtain higher order terms accurately, which can be done by reducing the step size. This approach has been used to retrofit several existing tracking codes, including recently MAD-X for the extraction of high-order DA maps.

However, using DA techniques it is possible to entirely avoid the approximate time stepping of conventional integrators, but rather develop new integrators of arbitrary order in time. Strikingly, they require only a single evaluation of the r.h.s. per step, resulting in greatly increased efficiency and robustness [10, 6, 7]. One way is based on the common rewriting of the ODE as a fixed point problem

$$\vec{z}_f = \vec{z}_i + \int_{s_i}^{s_f} \vec{f}(\vec{z}, \tilde{s}) d\tilde{s} \tag{3}$$

Utilizing the operation ∂^{-1} for the integral, the problem can be iterated in DA with $\tilde{s}$ as an additional variable. It can be shown that iteration converges to the exact result in n steps, where n is the order of the DA operations; moreover, this is not affected if in the ith iteration step the overall order is reduced to i. The result is an nth order integrator; for a given accuracy demand, the integrator typically results in a speed-up of about an order of magnitude. Moreover, the integrator also affords a rigorous and sharp estimate of all integration errors [6, 7, 8, 9].

Similarly, it is also possible to solve PDEs in finitely many steps. For this purpose, one eliminates differentiation with respect to one variable by integration. For example, the PDE

$$\begin{aligned} a_1 \tfrac{\partial}{\partial x}\left(a_2 \tfrac{\partial V}{\partial x}\right) + b_1 \tfrac{\partial}{\partial y}\left(b_2 \tfrac{\partial V}{\partial y}\right) \\ + c_1 \tfrac{\partial}{\partial z}\left(c_2 \tfrac{\partial V}{\partial z}\right) = 0 \end{aligned} \tag{4}$$

which describes the scalar potential in particle optical relative coordinates, is rewritten as

$$\begin{aligned} V = V\Big|_{y=0} + \frac{1}{b_2}\int_y \left\{ \frac{\partial V}{\partial y}\Big|_{y=0} \right. \\ \left. - \int_y \frac{a_1}{b_1}\frac{\partial}{\partial x}\left(a_2 \frac{\partial V}{\partial x}\right) + \frac{c_1}{b_1}\frac{\partial}{\partial z}\left(c_2 \frac{\partial V}{\partial z}\right) \right\} \end{aligned} \tag{5}$$

and again, iteration converges to the exact result in finitely many steps [10]. In addition to producing general PDE solvers, this approach allows the execution of the commonly performed out-of-plane expansion to arbitrary order [25].

It is important to observe that the accuracy of the out-of-plane expansion rests on the quality of field derivatives in the midplane. If the field arises from models, DA methods can directly be used to obtain these derivatives accurately to any order of interest. If on the other hand, the field is based on measurements, it is highly advantageous to represent the field in space through integrals over surface fields based on the Helmholtz theorem which resembles the Cauchy formula in complex analysis. DA methods can be used elegantly to expand fields simultaneously in the coordinates in the midplane, as well in the surface coordinates, and the integration is carried out directly utilizing the DA operation ∂^{-1} [25].

For ODEs that are time independent and for which $\vec{z} = \vec{0}$ is a solution, which is the case in particle optical relative coordinates, another method can be applied. For a given function on phase space $g(\vec{z}, s)$, it is possible to obtain its derivative along the true solution via

$$\frac{d}{ds} g(\vec{z}, s) = \vec{f} \cdot \vec{\nabla} g + \frac{\partial}{\partial t} g = L_{\vec{f}} g \tag{6}$$

Apparently we also have $d^n/ds^n \; g = L_{\vec{f}}^n g$. If g is not explicitly time dependent, the operator $L_{\vec{f}}^n$ can be evaluated directly within DA for any n; using it for $g = z_\nu$, the components of the vector $\vec{z}$, we obtain an integrator of adjustable order. This method is utilized for the main fields of elements in COSY; an element is typically traversed in one step, and orders of 25–30 are usually chosen to obtain integration to nearly machine precision. Note that for certain particle optical systems, including the notoriously improperly treated fringe fields, there are other efficient and fast perturbative methods to obtain approximate solutions of high accuracy, like the method of symplectic scaling [26].

Advanced DA operations and manipulation of maps Given the nth order representations of two maps $\mathcal{M}_n$ and $\mathcal{N}_n$, it is possible to determine the nth order representation of their composition $\mathcal{M}_n \circ \mathcal{N}_n$ as long as $\mathcal{N}_n$ has no constant parts. To this end, one simply inserts the components of $\mathcal{N}_n$ into the polynomial defined by the coefficients of $\mathcal{M}_n$. This allows to determine the map of a combined system from the maps of the subsystems.

It is also possible to obtain the nth order representation $\mathcal{N}_n$ of the inverse of $\mathcal{M}_n$ as long as its linear part is invertible. To this end, one writes $\mathcal{M}_n = \mathcal{M}_1 + \mathcal{M}_n^*$, where $\mathcal{M}_1$ is the linear part of $\mathcal{M}_n$. Then we have

$$\begin{aligned} \mathcal{I}_n &= (\mathcal{M}_1 + \mathcal{M}_n^*) \circ \mathcal{N}_n \\ &= \mathcal{M}_1 \circ \mathcal{N}_n + \mathcal{M}_n^* \circ \mathcal{N}_n \quad \Rightarrow \\ \mathcal{N}_n &= \mathcal{M}_1^{-1}(\mathcal{I} - \mathcal{M}_n^* \circ \mathcal{N}_n) \end{aligned} \tag{7}$$

a fixed point problem for $\mathcal{N}_n$. Beginning iteration with $\mathcal{N}_n = \mathcal{I}_n$ yields convergence to the exact result in n steps because $\mathcal{M}_n^*$ is purely nonlinear. The map of the reversion [27, 10] of a system, i.e.

the system traversed backwards, is related to the inverse of the map; but since angles change directions when going backwards, the reversed map $\mathcal{M}_r$ is $\mathcal{M}_r = R \circ \mathcal{M}^{-1} \circ R$, where R is the linear map that changes the signs of angles.

Symplectic tracking with maps Inverse maps can also be used to establish mixed-variable generating functions that represent the map, which are useful for symplectic tracking. Beginning from the n-th order map $\mathcal{M}_n$, we create $\mathcal{N}_{n,1}$, consisting of the position part of $\mathcal{M}_n$, and an identity for the momenta; thus $(\vec{q}_f, \vec{p}_i) = \mathcal{N}_{n,1}(\vec{q}_i, \vec{p}_i)$. Similarly we construct $\mathcal{N}_{n,2}$ via $(\vec{q}_i, \vec{p}_f) = \mathcal{N}_{n,2}(\vec{q}_i, \vec{p}_i)$. Inversion of $\mathcal{N}_{n,1}^{-1}$ (if possible) yields $(\vec{q}_i, \vec{p}_i) = \mathcal{N}_{n,1}^{-1}(\vec{q}_f, \vec{p}_i)$, and composition with $\mathcal{N}_{n,2}$ from the right yields

$$(\vec{q}_i, \vec{p}_f) = \mathcal{N}_{n,2} \circ \mathcal{N}_{n,1}^{-1}(\vec{q}_f, \vec{p}_i) \tag{8}$$

From this mixed-variable representation, the underlying generating function can be obtained by a mere integration along a suitable path, again using the DA operation ∂^{-1}; in a similar way as the case shown here, also all the other three common generating functions can be determined [10].

However, in practice it is frequently observed that different symplectification schemes result in different tracking results; so it is highly desirable to achieve symplectification with minimal modification of the prediction of Taylor transfer map. Utilizing Hofer's metric on spaces of Hamiltonians, this can actually be achieved in a unified way using DA techniques, resulting in the so-called EXPO symplectification scheme [28].

Normal forms Another important manipulation of maps is the transformation to normal form [10]. The first step is the transformation to the parameter-dependent fixed point $\vec{z}(\vec{\delta})$ which satisfies $(\vec{z}(\vec{\delta}), \vec{\delta}) = \mathcal{M}_n(\vec{z}(\vec{\delta}), \vec{\delta})$. Subtracting the non-parameter identity map $\mathcal{I}_n^{\vec{z}}$ on both sides we have $(\vec{0}, \vec{\delta}) = (\mathcal{M}_n - \mathcal{I}_n^{\vec{z}})(\vec{z}(\vec{\delta}), \vec{\delta})$, and thus

$$(\vec{z}(\vec{\delta}), \vec{\delta}) = (\mathcal{M}_n - \mathcal{I}_n^{\vec{z}})^{-1}(0, \vec{\delta}) \tag{9}$$

from which we read off $\vec{z}(\vec{\delta})$ in the non-parameter lines.

The linear part of the normal form algorithm itself is based on a study of the eigenvalue spectrum of the map; if they are distinct and complex, then it amounts to a diagonalization to the form $\mathcal{R}$, where the phases of the diagonal terms correspond to the tunes. The nonlinear part of the normal form algorithm consists of a sequence of transformations $\mathcal{S}_n$ that to order n have the form $\mathcal{S}_n = \mathcal{I} + S_n$. The orders higher than n in $\mathcal{S}_n$ can in principle be picked freely, and frequently they are chosen so that all transformations $\mathcal{S}_n$ belong to the same symmetry group as the original map.

Up to order n, we have $\mathcal{S}_n^{-1} =_n \mathcal{I} - S_n$; letting $\mathcal{N}_n = \mathcal{R} + \mathcal{O}_n$ denote the map that has been normalized to order n, we have

$$\begin{aligned}\mathcal{N}_n &= \mathcal{S}_n \circ \mathcal{N}_{n-1} \circ \mathcal{S}_n =_n (\mathcal{I}+S_n) \circ \mathcal{N}_{n-1} \circ (\mathcal{I}-S_n) \\ &=_n \mathcal{R} + \mathcal{O}_n + \{S_n \circ \mathcal{R} - \mathcal{R} \circ S_n\}\end{aligned} \tag{10}$$

Apparently, the S_n in the commutator $\{S_n \circ \mathcal{R} - \mathcal{R} \circ S_n\}$ can now be chosen to remove nth order terms in $\mathcal{O}_n$. In the symplectic case, it turns out that the remaining terms are just those that describe motion on circles, with a frequency that depends on the radius and possible parameters $\vec{\delta}$, allowing direct computation of amplitude dependent tune shifts [10].

The radius vectors in each phase space pair represent invariants, the quality of which is directly determined by the magnitude of the coefficients of S_n, the so-called resonance strengths. Trying to minimize their size is an efficient way to enhance the overall nonlinear behavior of the system [10].

Other DA-based tools It is also possible to obtain Lie factorizations,

$$\begin{aligned}\mathcal{M}_n &=_n \{\mathcal{M}_1 \exp(:f_3:) \exp(:f_4:) ...\} \mathcal{I} \\ \text{or} \quad \mathcal{M}_n &=_n \{... \exp(:f_4:) \exp(:f_3:) \mathcal{M}_1\} \mathcal{I}\end{aligned} \tag{11}$$

of symplectic maps $\mathcal{M}_n$ [10], where $=_n$ means two functions are equal up to order n; in fact, the DA approach currently represents the only method to obtain them to arbitrary order. The procedure is order-by-order; in the ith step, all orders less then i have already been taken care of, and the problem is reduced to finding f_{i+1} such that $\exp(:f_{i+1}:)\mathcal{I} = \mathcal{I} + \vec{\nabla} f_{i+1} \cdot \hat{J}$ agrees with the given map to order i. This is an integration problem similar to the case of the generating functions above, which is resolved once more with the DA operator ∂^{-1} [10].

Some other DA-based methods not discussed above shall be mentioned briefly. For the analysis of spin dynamics, there are methods that allow the computation of the spin map, spin tracking, as well as invariant spin axis $\hat{n}$. There are also extensions of the DA methods to allow for a mathematically rigorous treatment of the remainder terms of Taylor's formula [6], which allow for the

development of rigorous error bounds for integrators [7, 8] and also a rigorous bounding of stability times of nonlinear motion [29, 6] when combined with methods of verified global optimization [6]. There are methods for the solution of algebraically constrained differential equations, so-called differential algebraic equations [30], and methods to develop high order versions of the Fast Multipole Method for 3-D space charge computation [31].

Applications The differential algebraic methods form the core tools for the computation of maps and aberrations for a large number of codes. First developed and used in the code COSY INFINITY [13] which currently has about 2,500 registered users, the methods also form the core engines in many other codes [14]–[18].

Maps are used for symplectic tracking of dynamics in synchrotrons resulting in speed increases for the LHC [32] and the Tevatron; at the latter, they have also been used for purposes of linear decoupling [33] and a rigorous estimate of long-term stability [31]. Another application is the tracking and correction based on more sophisticated models for fringe fields or other nonlinearities in large acceptance rings [34], symplectic tracking in light sources under consideration of wigglers and undulators, all the way to various applications for novel accelerator types like the FFA [36].

The methods have also been used for the design and correction of fragment separators and particle spectrographs. In addition to computation and hardware correction, a particularly important application is the reconstructive correction of aberrations in high resolution spectrographs [35]. This technique uses combinations of the map and its inverse to eliminate aberrations by measurement data in two planes, i.e. knowledge of (x_f, y_f, a_f, b_f). Since this is uniquely determined as a function of $(y_i, \delta_i, a_i, b_i,)$, utilizing map inversion techniques discussed above yields reaction energy and angles to precisions that are sometimes two orders of magnitude higher than without correction.

Finally, over the last years, use is also being made of the methods in electron microscopy, where they afford the ability to compute and correct high-order aberrations when the fields of the devices are known.

Muon $g-2$ ring As an illustration, consider the case of a muon $g-2$ model ring that requires magnets with exceedingly uniform field. Using the fact that muons decay with different rates depending on the orientation of their spin relative to their direction of motion, and very large statistics, it is possible to measure the precession of the spin and thus an accurate measure of muon $g-2$. (See also Sec. 1.6.16.)

From the perspective of beam dynamics, particles moving strictly in the midplane would experience simple circular motion, and the transfer map would be unity. However, in the interim before the particles return to their initial condition, the motion is actually quite nonlinear. This becomes important because in order to provide vertical focusing, it is necessary to introduce electrostatic quadrupoles which nominally do not disturb the spin motion. However, in practice, the quads are necessarily not perfect. After shaping their electrodes to suppress the first allowed harmonic, the twelve pole, the next allowed harmonic, the twenty pole, still introduces nonlinearities of order 9 into the dynamics.

In order to analyze the motion to the precision required for the experiment, DA-based normal form methods are used to compute tune shifts and show the results for an illustrative ensemble of particles on a grid within the acceptance of the ring in Fig. 1 at four different contemplated settings of the voltage of the electrostatic quads. Analysis to order 9 shows significant horizontal and vertical tune shifts, while analysis to order 7 (superimposed in green) misses most of these and shows mostly the effect of the symmetry breaking of the circular motion.

The careful study of losses in the system is of prime importance since lost particles can potentially bias the statistics of the muon decay study which has to be taken into account. So a careful preparatory beam treatment after injection called scraping that removes high-amplitude

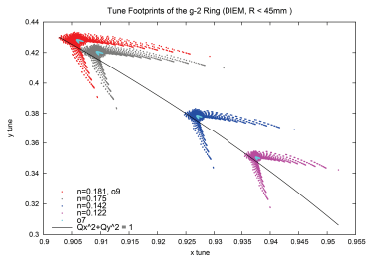

Figure 1: Tune footprint of muon $g-2$ ring with four contemplated operation settings.

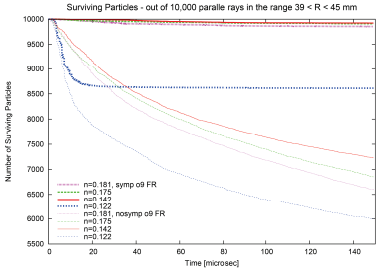

Figure 2: particle loss predictions using different simulation models.

particles by various deliberate time-dependent mispowering of the quads is carried out. For this purpose, it is necessary to treat full nonlinear motion without approximation in the Hamiltonian and include all field effects including the fringe fields of the quadrupoles, which precludes commonly used split-operator-based symplectic tracking, which is avoided using the map-based EXPO symplectic tracking scheme. Figure 2 shows muon losses for the four voltage configurations using the EXPO method (upper curves) and non-symplectic tracking (lower curves) for a selection of large-amplitude particles before scraping, and the results show that avoiding symplectification in tracking leads to a qualitatively different result that incorrectly predicts much larger losses than what appears in reality.

References

[1] K.L. Brown, SLAC Tech. Rep. 91 (1979)
[2] T. Matsuo et al., Mass Spectrometry 24
[3] H. Wollnik et al., AIP CP 177 (1988) 74
[4] A.J. Dragt et al., IEEE Tran. Nuc. Sci. NS-3, 5, 2311 (1985)
[5] M. Berz et al., NIM A258 (1987) 402
[6] K. Makino, PhD thesis, Michigan State U., MSUCL-1093 (1998)
[7] M. Berz, K. Makino, Reliable Computing 4 (4) (1998) 361
[8] K. Makino, M. Berz, Int. J. Diff. Eq. & Appl. 10, 4 (2005) 353
[9] M. Neher, K. R. Jackson, N. S. Nedialkov, On Taylor Model Based Integration of ODEs, SIAM Journal of Numerical Analysis, 45 (1) 236-262 (2007)
[10] M. Berz, Modern Map Methods in Particle Beam Physics, Academic Press, San Diego, 1999; M. Berz, K. Makino, W. Wan, An Introduction to Beam Physics, Taylor & Francis, 2014, ISBN 978-0-7503-0263-0
[11] M. Berz, PA 24, 109 (1989)
[12] M. Berz, NIM A258, 431 (1987)
[13] M. Berz et al., http://cosyinfinity.org
[14] Y. Yan, AIP Conf. 297 (1993) p. 279
[15] J. van Zeijts, AIP Conf. 297 (1993) p. 285
[16] L. Michelotti, Fermilab Tech. rep. (1990)
[17] W. Davis et al., Proc. Workshop on Nonlinear Effects in Accelerators, 1993
[18] F.C. Iselin, AIP Conf. 297 (1993) p. 325
[19] J.F. Ritt, Integration in Finite Terms - Liouville's Theory of Elementary Methods, Columbia U. Press, 1948
[20] J.F. Ritt, Differential Algebra, APS, Washington, DC, 1950
[21] E.R. Kolchin, Differential Algebraic Groups, Academic, NY, 1985
[22] R.H. Risch, APS Bulletin 76, 605 (1970)
[23] M. Berz, Forward algorithms for high orders and many variables, Automatic Differentiation of Algorithms: Theory, Implementation and Application SIAM
[24] M. Berz et al., SIAM 1996
[25] S. Manikonda, PhD thesis, Michigan State U., 2006
[26] G. Hoffstaetter, M. Berz, PR E54, 4
[27] W. Wan, M. Berz, PR E54, 3, 2870 (1996)
[28] B. Erdélyi, M. Berz, PRL 87, 11, 114302 (2001)
[29] K. Makino et al., ECMI Newsletter 39
[30] J. Hoefkens, PhD thesis, Michigan State U. (2001)
[31] H. Zhang, M. Berz, Fast Multipole Method in the Differential Algebra Framework, Nuclear Instruments and Methods A, 645, 338-344 (2011) doi:10.1016/j.nima.2011.01.053
[32] M.L. Shashikant, M. Berz, B. Erdélyi, IOP CP 175, 299 (2002)
[33] P.V. Snopok et al., NIM 558, 142 (2005)
[34] M. Berz, B. Erdélyi, K. Makino, PRST-AB 3, 124001 (2000)
[35] M. Berz et al., PR C 47, 2, 537 (1993)
[36] Carol Johnstone, Martin Berz, Pavel Snopok (Eds.), Proceedings, FFAG 2009 International Conference, International Journal Modern Physics A, 26, 10-11, World Scientific (2011) DOI:10.1142/S0217751X11053079