Symplectic scaling of transfer maps including fringe fields

Georg Heinz Hoffstätter* and Martin Berz
Department of Physics and Astronomy and National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, Michigan 48824
(Received 5 January 1996)

A method is introduced that provides an accurate and fast approximation of high-order maps of fringe fields and other fields that change along the reference trajectory. While the effects of main fields of optical elements can be determined very efficiently with differential algebraic (DA) methods via exponentiation of the respective propagator, the computation of high-order maps of nonstationary fields in general requires time-consuming DA integration. The method of symplectic scaling presented in this paper provides a very fast approximation of such maps by relating an arbitrary map to a specific previously computed map. This is achieved by a combination of geometric scaling and scaling with rigidity performed in a canonically perturbative treatment of a strength parameter. The method is useful for detailed analysis of nonlinear motion in particle optics, which in many cases is strongly influenced or even dominated by the presence of fringe fields. The use of the symplectic scaling method typically speeds up the computation of fringe-field effects by around two orders of magnitude and thus approaches speeds similar to that of the main-field calculation. The method has been implemented in the code COSY INFINITY; several examples from various subfields of beam physics are given to illustrate the accuracy and speed of the method. [S1063-651X(96)04911-2]

PACS number(s): 41.85.−p, 02.70.−c

I. INTRODUCTION

Beam physical systems can be described by a map that relates coordinates $\vec{z}_i$ in an initial plane to coordinates $\vec{z}_f$ in a final plane, and all information about the optical properties is contained in the map. This map $\vec{M}$ for a system of d degrees of freedom depends on a collection $\vec{\delta}$ of p parameters of the system and thus maps R^{2d+p} into R^{2d} via

$$\vec{z}_f=\vec{M}(\vec{z}_i,\vec{\delta}). \tag{1}$$

Since the relationship between initial and final variables is weakly nonlinear, often the map is represented by its Taylor expansion, a concept that goes back all the way to Hamilton [1], who utilized what he called the characteristic function to compute expansions for rotationally symmetric optical systems to third order. Since then the method has been used beneficially in many cases and even now forms the basis of various important beam physics codes [2–5].

The motion of the particles is described by a set of ordinary differential equations

$$\frac{d}{ds}\vec{z}=\vec{f}(\vec{z},\vec{\delta},s), \tag{2}$$

where s is the coordinate along the central trajectory and $\vec{z}$ describes the usual particle optical coordinates. Under mild conditions for $\vec{f}$, the equation has a unique solution through any given initial point $\vec{z}_i$ and thus defines the map $\vec{M}$.

Because of the simplifications involved in the regions where the field does not depend on s, traditionally the effect of a particle optical device is represented by a field-free drift and a main-field region. If fringe-field effects are to be considered, they are represented by the perturbative so-called fringe-field map that is sandwiched between the drift and the main-field map at the position s_0. Hence the fringe-field map consists of a negative drift to the region where the field vanishes, the map through the varying field, and the application of an inverse main-field map back to s_0, as shown in Fig. (1). So the fringe-field map represents the necessary corrections to the simple step function field model described by

$$\vec{B}_{\mathrm{MF}}(x,y,s)=\begin{cases}\vec{B}(x,y,s_m) & \text{for } s \text{ in the main field}\\ \vec{0} & \text{for } s \text{ outside the main field,}\end{cases} \tag{3}$$

where s_m describes the center of the main field. The abrupt change of the field in this model of course violates Laplace's equation and therefore cannot represent a physical system. To describe a realistic system, the main-field map has to be composed with fringe-field maps, which describe the connection of the main-field to the field-free region outside the element.

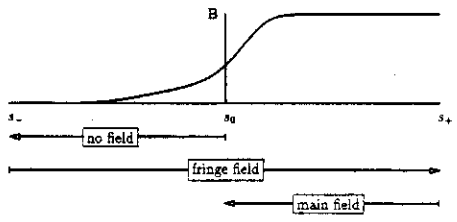

FIG. 1. Fringe-field region of an optical element with effective edge at s_0. For computational purposes, all fringe-field effects are usually concentrated in a map of length zero sandwiched between a pure drift map up to s_0 and a main-field map beginning at s_0.

*Present address: DESY, 22603 Hamburg, Germany.

The main-field map of optical elements can usually be computed more easily than that of the fringe fields. This holds both for the case of low-order codes based on analytic formulas for the nonlinear effects [2–6] and also for high-order codes based on differential algebraic (DA) methods [7–10]. In the latter case, the main-field map can be evaluated very quickly by applying the propagation operator of the autonomous equation of motion. Let L_f denote the directional derivative, i.e., the operator $\vec{f}^T\vec{\partial}+\partial_s$. Then the equation of motion $d\vec{z}/ds=\vec{f}(\vec{z},\vec{\delta},s)$ allows one to compute the transfer map via

$$\vec{M}=\sum_{i=1}^{\infty}\frac{1}{i!}L_f^i\vec{z} \tag{4}$$

provided mild smoothness conditions are satisfied. Unfortunately, for nonautonomous systems the propagator cannot be evaluated in DA [11] due to the derivative ∂_s, which reduces the order in s of the DA vector on which it acts.

As mentioned above, the fringe-field map $\vec{M}_{\mathrm{FF}}$ is defined as a correction to be inserted at the edge of the element and is formalized as follows. Let s_0 denote the effective field boundary at the entrance of the element, s_- a position so far before the optical device that the field can be neglected, and s_+ a position so far inside the element that $\vec{B}(x,y,s)$ already changes very little with s. Let the map $\vec{M}_{\mathrm{MF},s_0\to s_+}$ describe particle motion through the main field given in Eq. (3) from the effective field boundary to s_+. The fringe-field map is constructed in such a way that a drift $\vec{D}_{s_-\to s_0}$ from s_- to the effective field boundary composed with $\vec{M}_{\mathrm{FF}}$ and then with $\vec{M}_{\mathrm{MF},s_0\to s_+}$ yields the transfer map $\vec{M}_{s_-\to s_+}$ from s_- to s_+:

$$\vec{M}_{s_-\to s_+}=\vec{M}_{\mathrm{MF},s_0\to s_+}\circ\vec{M}_{\mathrm{FF}}\circ\vec{D}_{s_-\to s_0}. \tag{5}$$

Here, $\circ$ indicates the composition of maps. Hence the fringe-field map has the form

$$\vec{M}_{\mathrm{FF}}=\vec{M}^{-1}_{\mathrm{MF},s_0\to s_+}\circ\vec{M}_{s_-\to s_+}\circ\vec{D}^{-1}_{s_-\to s_0}. \tag{6}$$

Computing fringe-field maps requires the computation of the map $\vec{M}_{s_-\to s_+}$ of a system where $\vec{B}(x,y,s)$ and hence the differential equation (2) depends on s. For this, methods to find general solutions of nonautonomous differential equations are needed. The method of numerical integration in DA yields the Taylor map to such an equation of motion to arbitrary expansion order [12,13,11,14–16]. However, numerical integration in DA is very time consuming in comparison to evaluating the propagator. Figure (2) shows the ratio of computational expense when calculating fringe-field effects and main-field effects of dipoles for various orders in the code COSY INFINITY. For quadrupoles and higher-order multipoles the relative speed of main-field computations is similarly high.

This phenomenon manifests itself dramatically in typical calculations. As an example, we consider an extensive optimization problem of a total of 20 first- and second-order conditions for the case of the multipole fields of the S800 beam line and the S800 spectrograph under construction at the National Superconducting Cyclotron Laboratory (NSCL). Listed are the computational expense for the hard-edge model, for a full fringe-field integration, as well as the SYSCA method discussed in this paper.

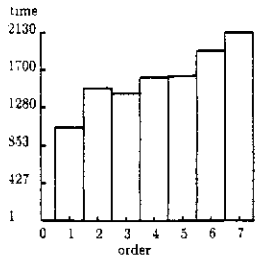

FIG. 2. Computational expense of DA map computation for fringe fields relative to that of main fields for a dipole magnet as a function of expansion order.

Only main fields with propagator	51 sec
Fringe fields with DA integration	7 h, 10 min, 50 sec
Fringe fields with SYSCA	6 min, 38 sec

Unfortunately, merely ignoring the fringe-field effects leads to quite substantial errors in the nonlinearities of the element. To stress this point, we compute the maps of a dipole both with and without fringe-field effects and compare the results. The dipole has simple wedge shape with a radius of 2 m, an angle of 30°, and an aperture of 2.54 cm. The chosen ion is $^{16}O^{3+}$ with an energy of 25 MeV per nucleon. The fringe fields used are those of the Enge model [7,17,18].

In order to analyze whether the transfer map can be represented well by the main-field map, we define the normalized average deviation Δ_m between the Taylor coefficients of the transfer map and those of the main-field map in order m for a typical charged particle optics devices. The quantity Δ_m approaches zero for high similarity between two maps and one for the case of purely randomly chosen coefficients. Figure 3 shows examples for the values of Δ_m for a dipole and various calculation orders. It can be seen that to first order, the errors are very slight, while for the cases in which high-order terms do not vanish due to symmetry many high-order aberrations are so wrong when fringe fields are ignored that Δ_m is close to unity, the value corresponding to a totally

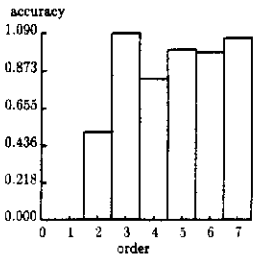

FIG. 3. Accuracy of the "sharp cutoff" (SCOFF) approximation for different orders of a dipole transfer map. The errors are substantial and often even approach unity, which would also occur for a plainly random choice of coefficients.

random selection of map coefficients. For quadrupoles and higher-order multipoles the errors are similarly unacceptable.

The importance of fringe fields becomes apparent when realizing that many nonlinear properties of an electric or magnetic field are due to the dependence of the field on s. For quadrupoles and higher multipoles, this can most easily be seen from the scalar potential V_B, which defines the field by $\vec{B} = -\vec{\partial} V_B$ and can be expanded as

$$V_B(r,\phi) = \sum_{\nu=2}^{\infty} \sum_{\lambda=0}^{\infty} r^{\nu+2\lambda} \frac{(-1/4)^\lambda \nu!}{\lambda!(\lambda+\nu)!} \times \mathrm{R}\{[\Psi_\nu(s) e^{(i\theta_\nu)}]^{[2\lambda]} e^{(-i\nu\phi)}\}, \quad (7)$$

a well-known relationship found frequently in the literature, for example, in [19]. Many nonlinear contributions only arise because of the nonvanishing derivatives of $\Psi_\nu(s)$. But the functions involved and also their derivatives can be quite complicated; one frequently used model is the so-called Enge function, which has the form

$$\Psi_\nu(s) = \frac{\Psi_\nu(s_+)}{1+\exp[b_{k,0}+b_{k,1}(s/a)+b_{k,2}(s/a)^2+\cdots]}, \quad (8)$$

where a is the aperture of the device under consideration and $b_{k,l}$ are real coefficients modeling the details of the field's falloff.

Altogether, a careful treatment of fringe-field effects seems imperative for a detailed study of nonlinear effects, and the advantages of very fast DA evaluation of propagators can only be taken advantage of if an efficient fringe-field approximation can be found.

Our search for ways out of this dilemma leads to the method of symplectic scaling. The particle motion through and therefore the transfer map of an optical device depends on the field excitation in that device. There are, however, some scaling properties of particle motion in electric and in magnetic fields that allow one to relate the motion of particles through an optical device at a certain field excitation to the motion through a similar but bigger device with a different field excitation. We call this property "geometric scaling." Similarly, the motion of a particle with different mass, energy, or charge can be related to the motion through the same optical device excited to a different field. We call this well-known property "scaling with magnetic rigidity." These two scaling laws illustrate that it is sufficient to know the transfer map of an optical device as a function of excitation field strength in order to compute the map of that device for the transport of all different charged particles and for similar devices of all sizes. It is in general not possible to compute and store such a field-dependent transfer map; however, via DA methods we can approximate this function by the Taylor expansion of such maps in respect to the field excitation.

However, approximating transfer maps in this way leads to maps that only approximately satisfy the symplectic condition. In order not to compromise on the very important symplectic symmetry, we do not store the field-dependent transfer map but field-dependent symplectic representations of it. Approximating these representations by Taylor expansions with respect to the field excitation leads to approximate Taylor transfer maps that satisfy the symplectic condition in every evaluation order exactly.

The symplectic scaling process for obtaining a transfer map including fringe-field effects therefore starts with reading from a file a previously computed and stored symplectic representation that depends on the field excitation. Now one figures which field excitation one has to choose to scale this symplectic representation to the special particle (i.e., mass, charge, and energy) and to the special size of the element of interest. Inserting this field excitation leads to the symplectic representation from which the symplectic transfer map can be extracted.

II. DESIRABLE PROPERTIES OF ARBITRARY ORDER FRINGE-FIELD APPROXIMATIONS

In the preceding section it became apparent that, on the one hand, a detailed study of nonlinear motion requires careful consideration of fringe-field effects, while, on the other hand, the expense of DA integration makes the use of efficient models very desirable. Any approximation for fringe-field effects should satisfy various requirements. First and foremost, it should be accurate, but besides, it should have certain properties important in typical problems encountered in beam physics simulations. Altogether, it is desirable that an approximation should (i) lead to order-n symplectic maps, (ii) represent the fringe effect well for a wide range of apertures, and (iii) be usable for arbitrary orders.

The simplest approximation, as already described, is SCOFF, where fringe fields are simply ignored. As illustrated in Sec. IV, this method strongly violates the accuracy requirement and point (ii). The impulse approximation [20] used in the code TRANSPORT [21] violates points (ii) and (iii) and the method of fringe-field integrals [22–24] used in the computer code GIOS [25,3] violates points (i) and (iii). In the following, we will illustrate the particular requirements and briefly discuss their importance.

A. Order-n symplecticity

The necessity for order-n symplecticity is especially apparent when long-term behavior in storage rings is to be analyzed. If the Taylor map does not approximate a symplectic map very well, then the phase-space volume is not necessarily conserved, which can have detrimental effects on simulation. In [26], for instance, it is illustrated that in the case of the Superconducting Super Collider lattice, expansion to 12th order is needed to have a sufficient degree of symplecticity of the total map; but if the fringe-field approximation does not preserve symplecticity, even very high evaluation orders cannot lead to trustworthy predictions of long-term behavior.

There are a variety of methods to perform symplectic tracking with generating functions computed from the Taylor map [27–30]. These methods rely on the order-n symplecticity of the underlying map and therefore only a fringe-field approximation that leads to order-n symplectic maps can be used for symplectic tracking.

Yet another important reason for order-n symplectic fringe-field maps is the fact that symplecticity imposes interrelations between different Taylor coefficients. Some systems rely on these interrelations in order to correct aberra-

tions. Examples are the high-order achromats described in Refs. [31,32] and the opening aberration correction for electron microscopes by means of hexapoles described in [33,34], to mention just a few. All these systems cannot be studied properly if fringe-field maps are used that are not order-n symplectic.

B. Accuracy for a wide range of apertures and fields

Designing spectrographs, simulating electron microscopes, and analyzing beam lines and high-energy storage rings are all topics that rely on the same principle of calculating transfer maps from the equation of motion. It therefore seems highly desirable to formulate calculation methods that work equally well for all these subfields of particle optics. The effect of the fringe fields in these different areas is often of different importance and frequently this is connected to the range of apertures of the elements. In the case of high-energy accelerators, fringe fields in many cases have limited influence because apertures are small and so is the amount of bending per element.

On the other hand, nuclear spectrographs are often characterized by exceedingly large apertures and very extensive fringe fields. Thus a good fringe-field approximation has to work over an extended range of apertures.

C. Usability for arbitrary orders

With the use of DA methods, it is easily possible to compute Taylor maps to very high order, and the limit is only given by computer memory and computation time. To unleash the potentials of the DA approach under the presence of fringe fields, the computation of fringe-field maps should be possible for arbitrary evaluation orders. Especially since an efficient way of computing fringe-field maps was sought for the arbitrary order code COSY INFINITY, this requirement was important for the development of symplectic scaling (SYSCA).

III. THE SYSCA METHOD

The general problem is to determine the transfer map of interest for a beam with reference particle of energy E, mass m, and charge q in a magnetic or electric particle optical device. The device is characterized by a size parameter A, a reference field strength F, as well as possibly additional parameters $\vec{\delta}$. Thus the task is to find the transfer map

$$\vec{M}^{E,m,q,A,F}(\vec{z},\vec{\delta}), \tag{9}$$

which will be achieved by relating it through a sequence of transformations to a suitably chosen previously computed and stored reference map. The SYSCA method relies on two different scaling mechanisms, one being based on geometric observations and the other one relying on the linear relation between rigidity of a particle and the field.

A. Geometric scaling

The first scaling method used in the determination of the map of a general element is based on a simple geometric observation. Assume that a certain space-dependent field and an orbit through it are given. If the geometry of the field is scaled up by a factor α and simultaneously the field strength is scaled down by the factor α, then the original orbit can be scaled up by the factor α along with the geometry of the device (Fig. 4). The reason for this phenomenon is that the momentary radius of curvature of the orbit, which is geometric in nature, is inversely proportional to the field. This fact can easily be deduced from the Lorentz equation. If a trajectory described by $x(t)$ and $p(t)$ satisfies

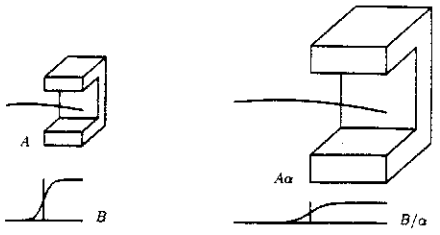

FIG. 4. If the size of an element is scaled by a factor α and simultaneously the strength of the field is scaled by a factor $1/\alpha$, then coordinates of particle trajectories scale with the factor α.

$$\frac{d\vec{p}}{dt}=q\left(\frac{d\vec{x}}{dt}\times\vec{B}(\vec{x})+\vec{E}(\vec{x})\right), \tag{10}$$

then a scaled trajectory $\vec{X}(t)=\alpha\vec{x}(t/\alpha)$ and $\vec{P}(t)=\vec{p}(t/\alpha)$ satisfies

$$\frac{d\vec{P}}{dt}=q\left(\frac{d\vec{X}}{dt}\times\frac{1}{\alpha}\vec{B}(\vec{X}/\alpha)+\frac{\vec{1}}{\alpha}E(\vec{X}/\alpha)\right). \tag{11}$$

There is one limitation to this approach connected to the fact that the shape of the field is not only influenced by the geometry generating it, but also by possible saturation effects that depend on the strength of the field. This limits the range of the factor α to a size where changes of saturation can be ignored.

It is also important to note that once the ratio of length to aperture of a device exceeds a certain minimum, the exact shape of the fringe field is almost unaffected by the actual length of the device. So for the treatment of fringe fields, in a good approximation, the aperture of the element can be used as the size parameter A.

For the actual use of the method, it is important to observe that only geometric quantities associated with the particle, such as positions, slopes, and lengths of trajectories, scale with the geometry, while the canonical momenta used in many codes, including COSY INFINITY [7], do not scale properly. Thus before scaling it is necessary to transform the map to purely geometric coordinates such as those used in the code TRANSPORT [21]. The transformation between these two sets of coordinates depends on the energy E_0 and the mass m_0 of the reference particle and is denoted by $\vec{T}(E_0,m_0)$. The coordinates are denoted by $\vec{z}_{E_0}=(x,a,y,b,\delta_E,\tau)$ and $\vec{z}_{p_0}=(x,x',y,y',l,\delta_p)$ and the transformation is given by

$$x'=\frac{a}{\sqrt{\left(\frac{p}{p_0}\right)^2-a^2-b^2}},\quad y'=\frac{b}{\sqrt{\left(\frac{p}{p_0}\right)^2-a^2-b^2}},$$

$$l=l_i-\frac{v}{v_0}\frac{2+\eta_0}{1+\eta_0}(\tau-\tau_i)+\left(\frac{v}{v_0}-1\right)s,$$

$$\delta_p=\frac{\sqrt{[E_0(1+\delta_E)]^2+2E_0mc^2(1+\delta_E)}}{p_0c}-1, \qquad (12)$$

which has the inverse

$$a=\left(\frac{p}{p_0}\right)\frac{x'}{\sqrt{1+x'^2+y'^2}},\quad b=\left(\frac{p}{p_0}\right)\frac{y'}{\sqrt{1+x'^2+y'^2}},$$

$$\tau=\tau_i-\frac{v_0}{v}\frac{1+\eta_0}{2+\eta_0}\left[l-l_i-\left(\frac{v}{v_0}-1\right)s\right],\delta_E$$

$$=\frac{\sqrt{[p_0c(1+\delta_p)]^2+(mc^2)^2}-mc^2}{E_0}-1. \qquad (13)$$

The maps $\vec{T}$ and $\vec{T}^{-1}$ can be evaluated conveniently in DA to arbitrary order. It is worthwhile to point out that in order to perform the transformation, the knowledge of the total arc length s of the system under consideration as well as the mass and energy of the reference particle is required; thus these quantities have to be stored along with the reference map. In geometric coordinates, the map $\vec{S}_\alpha$ performing the scaling is characterized simply by

$$x_2=x_1\alpha,\quad y_2=y_1\alpha,\quad l_2=l_1\alpha,$$
$$x_2'=x_1',\quad y_2'=y_1',\quad \delta_{p2}=\delta_{p1}; \qquad (14)$$

in the case that there are any parameters $\vec{\delta}$, their scaling behavior also has to be considered.

Altogether, the geometric scaling thus allows one to express a map that is associated with the size scale A in terms of a stored reference map whose size scale is $A^*=\alpha A$. This is accomplished by the transformations

$$\vec{M}^{E,m,q,F,A}(\vec{z}_E,\vec{\delta})=\vec{T}^{-1}\circ\vec{S}_\alpha^{-1}\circ\vec{T}\circ\vec{M}^{E,q,m,A^*,F/\alpha}$$
$$(\vec{z}_E,\vec{\delta})\circ\vec{T}^{-1}\circ\vec{S}_\alpha\circ\vec{T}. \qquad (15)$$

Since the stored reference map has to be evaluated at F/α, it is clear that the reference map has to be known as a function of the field strength.

B. Rigidity scaling

The next step in the transformation is to make adjustments for the fact that the properties of the reference particle differ in the stored map and in the map that is to be computed. To this end, the rigidity χ of the reference particle of interest is computed and compared to that of the stored map χ^*; since electric and magnetic rigidities depend on the quantities E,q,m in different ways, this requires that electric and magnetic fields not be present simultaneously.

In the following we will discuss the practically more important case of the magnetic field in detail and only briefly mention the electric case. Let $\beta=\chi^*/\chi$ be the ratio of the rigidity associated with the stored map and the map under consideration. Because of $v\vec{p}'=q\vec{v}\times\vec{B}$, a simultaneous scaling of magnetic rigidity and magnetic field has no influence on the orbit and we have $\vec{M}^{E,m,q,A^*,F/\alpha}(\vec{z}_p,\vec{\delta})=\vec{M}^{E^*,m^*,q^*,A^*,F\cdot\beta/\alpha}(\vec{z}_p,\vec{\delta})$. The change of a trajectory induced by a relative energy deviation δ_E depends on the energy E of the reference particle. It is essential that this is not true for δ_p. Due to the scaling law for magnetic rigidity, a relative momentum deviation δ_p creates the same changes in a trajectory no matter which reference momentum p is used. Thus the full transformation to the reference map is obtained as

$$\vec{M}^{E,m,q,A,F}(\vec{z}_E,\vec{\delta})=\vec{T}^{-1}(E,m)\circ\vec{S}_\alpha^{-1}\circ$$
$$\times\{\vec{T}(E^*,m^*)\circ\vec{M}^{E^*,q^*,m^*,A^*}(\vec{z}_E,\vec{\delta},\bar{F})$$
$$\circ\vec{T}^{-1}(E^*,m^*)\}|_{\bar{F}=F\beta/\alpha}$$
$$\circ\vec{S}_\alpha\circ\vec{T}(E,m). \qquad (16)$$

Again it is apparent that besides the dependence of the original map on the parameters $\vec{\delta}$, its dependence on the field F has to be known. While the exact dependence of the map on the field is usually difficult to obtain, the DA method very conveniently allows one to determine the expansion in terms of the field strength.

For an electric field of strength F the trajectory does not change if the quantity Fq/vp does not change because $v\vec{p}'=q\vec{E}$. Therefore TRANSPORT coordinates $\vec{z}_p$ with δ_p are not appropriate, but it is necessary to use the relative deviation δ_{vp} of the quantity vp from the corresponding value for the reference particle.

Even though this expansion can be obtained to rather high order, it still represents a source of errors and hence a study of the influence of these errors is warranted. The first question is that of the accuracy of the expansion. For this purpose it is useful to analyze the range of values the quantity

$$\bar{F}=F\frac{\beta}{\alpha} \qquad (17)$$

can assume; this establishes how far away from a chosen expansion point the extrapolation will have to be made. However, as it turns out, the quantity $\bar{F}$ has a rather simple geometric meaning. Note that both in the process of size scaling and in the process of rigidity scaling, the total deflection of a particle transversing the field at the aperture is not changed. Thus the quantity $\bar{F}$ plays the role of an universal strength parameter.

Clearly apertures and especially fields of elements vary widely in the different subdisciplines of charged particle optics encompassing electron microscopy, mass spectrometry, momentum spectrometry, and high-energy accelerators; but the amount of deflection a particle at the aperture experiences in any one given element is always limited.

This is particularly true for the fringe fields, which usually only represent a small contribution to the overall deflection effect of the device. Thus, in this case, a very good performance is to be expected by just choosing an expansion point in a neighborhood of typical deflection values.

But even the treatment of entire elements by the scaling method is expected to work reasonably well as long as the overall focusing power of the devices to be considered does not vary considerably. While this is not of importance for quadrupoles since their main fields can be determined very quickly, it proves very beneficial for the treatment of solenoids in electron microscopy and other applications. In this case, it is useful to estimate an approximate range for the strength parameter of the solenoid of interest and then choose the expansion point in the respective range when creating the reference representation.

Another reason why the SYSCA method can be expected to yield rather accurate results lies in the simple fact that it is based on the same type of polynomial expansion that forms the basis of all commonly used map perturbation methods, which are usually deemed accurate enough for most purposes.

The other important consideration in the process of approximating the field dependence in terms of an expansion is the question of symplecticity. Even if the errors produced by insertion into the truncated polynomial are minor, without additional considerations they violate the symplectic symmetry of the map.

This problem can be avoided by storing a symplectic representation of the map. For the nonlinear part of the map, it appears advantageous to choose a one-operator Dragt-Finn representation [35] using the pseudo-Hamiltonian, which can be readily determined from the map

$$\vec{M}(\vec{z}_E,\vec{\delta},F)=L(\vec{\delta},F)e^{:P(|Aw)z_E,\vec{\delta},F):}\vec{I}. \tag{18}$$

On the other hand, in the linear part the Dragt-Finn representation contains just the linear matrix in ordinary nonsymplectic form. In order to preserve the symplecticity of this linear part, it is advantageous to represent it in terms of a generating function in mixed variables [36], which again can be conveniently calculated from the map [37].

C. Extensions

There are various ways to extend the method of symplectic scaling. First of all, it is worthwhile to note that the required knowledge of the dependence of the map on a relative field change δ_F can be indirectly obtained from the conventional map itself. To this end, it is merely necessary to express the chromatic dependence of the map not in terms of the conventional energy deviation δ_E, but in terms of the momentum deviation δ_p and substitute

$$\vec{M}^{E,m,q,A,F}(\vec{z}_E,\delta_F)=\vec{T}(E,m)\circ\vec{M}^{E,m,q,A,F}\times(x,x',y,y',l,\delta_p/(1+\delta_F)). \tag{19}$$

This observation is particularly useful for the practical use of the method since it does not require any special setup to determine the stored map.

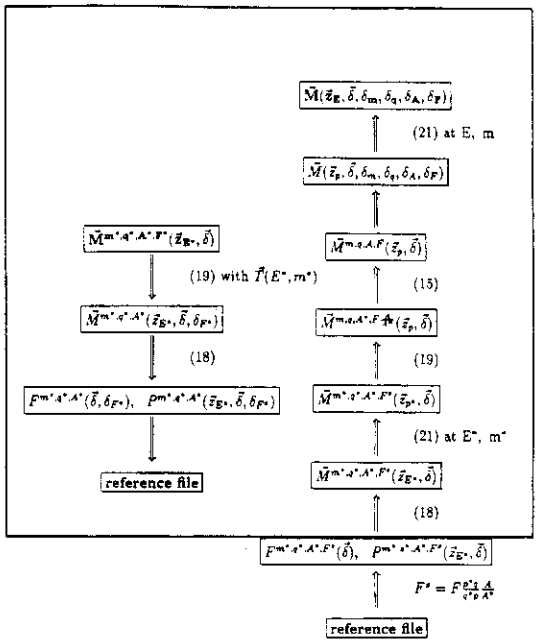

FIG. 5. Schematic outline of the procedures leading to the reference representation and from the reference representation to the general transfer map.

Similarly, the two scaling properties can be used to compute the map's dependence on a relative change of the aperture δ_A, the mass δ_m, and the charge δ_q. Therefore, no special parameters $\vec{\delta}$ have to be introduced to compute these dependences if they are desired and we write the most general parameter-dependent map as $\vec{M}^{E,m,q,A,F}(\vec{z}_E,\vec{\delta},\delta_m,\delta_q,\delta_A,\delta_F)$.

Another possible extension is the use of the parameters $\vec{\delta}$ for the description of other degrees of freedom of the field under consideration. These could include details about the falloff or about the geometric form of the effective field boundary. Especially the last topic is of great practical importance and can also conveniently be treated by rotating the fringe-field map of a wedge dipole appropriately as discussed in [38].

Finally, for practical considerations it is worthwhile to mention that it is sufficient to store only maps of entrance fringe fields. Given the entrance fringe field map $\vec{L}\circ[\exp(:P:)\vec{z}]$, the exit fringe-field map is described by the reversed map

$$R[\exp(:-P:)\vec{z}]\circ(L^{-1}R\vec{z}),$$

where the matrix R exchanges the sign of a, b, and τ. In addition, if a representation approximates maps of 2ν poles with fields close to B, then a rotation by an angle of $180°/\nu$ allows approximations close to the field $-B$.

D. Summary: Algorithm of SYSCA method

The left-hand side of Fig. 5 describes the procedure of computing the symplectic reference representation, which is

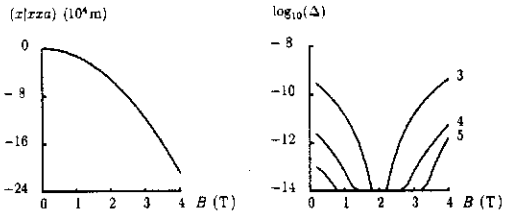

FIG. 6. Left: $(x|xxa)$ for a quadrupole as a function of the field at the pole tip. Right: error Δ of the approximation of $(x|xxa)$ with different expansion orders for the reference representation at $B=2$ T.

stored in a reference file. First the Taylor map $\vec{M}^{m^*,q^*,A^*,F^*}(\vec{z}_{E^*},\vec{\delta})$ for a certain size of the optical element and for a certain particle is computed in canonical variables. In our implementation this is done with the code COSY INFINITY. Scaling with rigidity yields the field dependent map $\vec{M}^{m^*,q^*,A^*}(\vec{z}_{E^*},\vec{\delta},\delta_{F^*})$. From this map the parameter-dependent symplectic representations $F(\vec{\delta},\delta_{F^*})$ and $P(\vec{z}_{E^*},\vec{\delta},\delta_{F^*})$ are computed and stored in a file [39].

The Taylor expansion of the general parameter-dependent map is written as $\vec{M}(\vec{z}_E,\vec{\delta},\delta_m,\delta_q,\delta_A,\delta_F)$. The goal is to compute this Taylor map from the reference representation. The right-hand side of Fig. 5 illustrates the procedure that will achieve this goal. After reading the reference file, the first step is to insert a value δ_{F^*} into the field dependence of the symplectic representation. δ_{F^*} is chosen in such a way that the symplectic representation describes particle motion that differs only by a scale in geometrical size from the trajectories of the desired map $\vec{M}^{m,q,A,F}(\vec{z}_E)$. For magnetic elements, the appropriate field is $B_s=B(p^*qA/pq^*A^*)$ and therefore

$$\delta_{F^*}=\left(F\frac{p^*qA}{pq^*A^*}-F^*\right)\Big/F^* \tag{20}$$

has to be inserted. For an electric element the required field is given by $F(v^*p^*qA/vpq^*A^*)$.

From the so-scaled symplectic representation one computes the order-n symplectic Taylor map, which is then transformed to TRANSPORT coordinates. The transformation

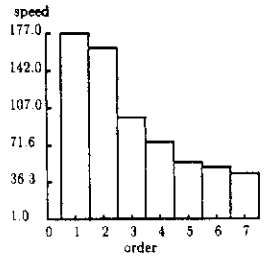

FIG. 7. Relative savings of computational expense of SYSCA to numerical integration for a dipole.

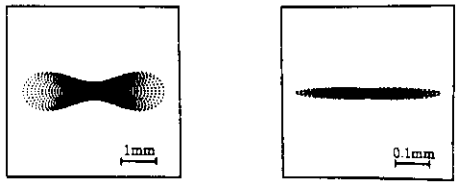

FIG. 8. Beam spots predicted by SYSCA (left) and SCOFF (right) approximations. The plot of the spot predicted by using the exact fringe fields cannot be distinguished from the one produced with SYSCA. Note the difference in scale.

map $\vec{T}(E,m)$ from canonical to transport coordinates depends on the properties of the particle which's motion the map describes:

$$\vec{M}(\vec{z}_p,\vec{\delta})=\vec{T}(E,m)\circ\vec{M}(\vec{z}_E,\vec{\delta})\circ\vec{T}^{-1}(E,m). \tag{21}$$

The map in TRANSPORT coordinates $\vec{z}_p$ can then be scaled to the desired field by scaling with rigidity and to the correct size of the element by geometric scaling. As mentioned before, we can obtain the general parameter-dependent map from the specific map by scaling. With DA this procedure automatically leads to the Taylor expansion with respect to all the parameters. Finally, one obtains the required general parameter-dependent order-n symplectic Taylor map $\vec{M}(\vec{z}_E,\vec{\delta}\delta_m,\delta_q,\delta_A,\delta_F)$ by transforming back to canonical coordinates.

Symplectic scaling can be applied to any map, but it is especially useful for fringe-field maps. The following examples show the accuracy and speed of obtaining fringe-field maps by symplectic scaling of a stored symplectic reference representation.

IV. EXAMPLES

In this section, we will illustrate the profitable use of the method with several examples. In order to evaluate the speed and accuracy of the proposed approximation, we analyze a certain aberration coefficient of a quadrupole. Figure 6 shows the dependence of the expansion coefficient $(x|xxa)$ on the field B at the pole tip. Because functions like this can be closely approximated by polynomials, it can be expected that symplectic scaling (SYSCA) is very accurate. The quadrupole used here has a length of 41.9 cm and an aperture of

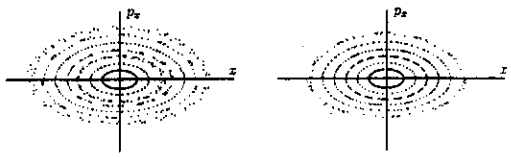

FIG. 9. 500 turn tracking without (left) and with (right) fringe fields trough a 1-GeV proton storage ring.

TABLE I. Amplitude- and energy-dependent tune shifts with and without fringe fields.

SCOFF	SYSCA
$\nu_x = 0.34$	$\nu_x = 0.34$
$+0.00\delta_E$	$+0.00\delta_E$
$+30.21\epsilon_x$	$+23.82\epsilon_x$
$+156.69\epsilon_y$	$+743.50\epsilon_y$
$-19.33\delta_E^2$	$+28.34\delta_E^2$
$+534.29\epsilon_x\delta_E$	$+271.39\epsilon_x\delta_E$
$-1792.76\epsilon_y\delta_E$	$-9685.27\epsilon_y\delta_E$
$+1260.32\delta_E^3$	$+1072.70\delta_E^3$
$+25599.56\epsilon_x^2$	$+13240.99\epsilon_x^2$
$+68007.73\epsilon_x\epsilon_y$	$-625777.29\epsilon_x\epsilon_y$
$-89715.35\epsilon_y^2$	$-1377087.43\epsilon_y^2$
$+32939.09\epsilon_x^2\delta_E^2$	$+593.75\epsilon_x^2\delta_E^2$
$+51138.52\epsilon_y^2\delta_E^2$	$-72396.79\epsilon_y^2\delta_E^2$
$+9918.05\delta_E^4$	$-448338.65\delta_E^4$

2.54 cm. The chosen ion is $^{16}O^{3+}$ with an energy of 25 MeV per nucleon.

Even at the border of the range in Fig. 6, the method presented is within the default tolerance allowed in the automatic step size control of the eight order Runge-Kutta integrator in COSY INFINITY. Close to the value with which the reference file was produced, the accuracy increases drastically. The results in Fig. 6 were obtained by evaluating the symplectic reference representation to third, fourth, and fifth order. The accuracy can be further improved by increasing this order, which of course increases the computation time that has to be invested for creating the reference map in advance. This investment can be very rewarding, especially when beam lines or spectrometers are being fitted or when system errors are analyzed so that maps of similar fringe fields are needed over and over again with only slightly different parameters.

The SYSCA approximation is especially helpful in the design of realistic systems after approximate parameters of the elements have been obtained by neglecting fringe fields. These values can be used to create a reference file for symplectic scaling. In this way, a very high accuracy almost equivalent to accurate but time intensive numerical integration can be obtained. The time savings obtained by this method is illustrated in Fig. 7.

A very important example is the computation of beam spots under the presence of nonlinear effects. For this purpose, a cone of particles was sent through the seventh-order map of the A1200 isotope separator at the NSCL. The images with the SCOFF and SYSCA approximations are shown in Fig. 8. The maximum angle used is 15 mrad. Note the difference in scale; the beam spot predicted with SCOFF is only one-tenth as big. Hence trusting SCOFF would lead to a loss of most of the beam.

Also for long-term tracking in storage rings, fringe fields have influences. In Fig. 9, typical tracking pictures are displayed. An example storage ring for 1-GeV protons was optimized for a big dynamical aperture. Eight particles were tracked with phase-space coordinates that had x and y components in order not to avoid x-y coupling. The left-hand picture was computed using the SCOFF approximation, whereas the right-hand picture was computed with SYSCA. Note the difference in the short-term dynamic aperture. Also the amplitude- and energy-dependent tune shifts of this system indicate that fringe fields can have very significant influences in repetitive systems. In Table I, the nonlinear dependence of the tune is given; in order to make the comparison as fair as possible, all fitting procedures were performed both without fringe fields and then with fringe fields. In particular, the linear tune and the chromaticity were adjusted to the same values.

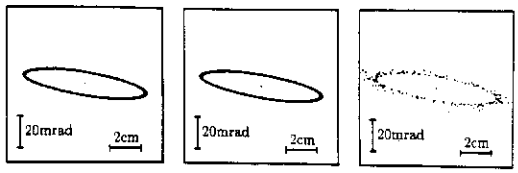

FIG. 10. 5000-turn tracking with fringe fields obtained by numerical integration (left), SYSCA (middle), and a nonsymplectic fringe-field approximation (right). The initial position of the particle is $(x,y) = (3\text{ cm}, 3\text{ cm})$ with no initial inclinations x' and y'.

The consequences of using symplectic approximations for the fringe-field maps are important for the case of repetitive tracking. As an example of this point, we use the lattice of the proposed Proton Storage Ring II of the Los Alamos National Laboratory . The ninth-order 5000-turn tracking pictures are displayed in Fig. 10.

The tracking was performed with the described standard numerical integration, with SYSCA, and with a nonsymplectic fringe-field approximation obtained by low-accuracy numerical integration. Nonsymplectic tracking strongly violates the conservation of phase-space volumes.

However, due to limited accuracy, sometimes even standard numerical integration produces maps that slightly violate symplecticity. To eliminate this possibility, the first tracking picture in Fig. 10 was obtained by symplectification via generating functions. For SYSCA this additional time investment is not required, since maps obtained by SYSCA are completely order-n symplectic. The corresponding ninth-order maps were produced with the SYSCA mode in COSY INFINITY in 30 min, whereas the standard numerical integration took 15 h, and the nonsymplectic approximation took 44 min on a VAX 4000 90 computer.

In this paper we applied symplectic scaling mostly to fringe-field maps. In conclusion, we would like to emphasize again that symplectic scaling is a method that can be used to speed up the computation of symplectic transfer maps in many other cases where nonautonomous differential equations have to be solved.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation, Grant No. PHY 89-13815, and the Alfred P. Sloan Foundation.

[1] W. R. Hamilton, in *Abhandlungen zur Strahlenoptik*, edited by G. Prange (Akademische-Verlagsgesellschaft, Leipzig, 1933).
[2] T. Matsuo and H. Matsuda, Mass Spectrom. **24**, 19 (1975).
[3] H. Wollnik, GIOS user's manual, Justus Liebig University Gießen, 1992 (unpublished).
[4] K. L. Brown, The ion optical program TRANSPORT, Stanford Linear Accelerator Center Technical Report No. 91, 1979 (unpublished).
[5] A. J. Dragt, L. M. Healy, F. Neri, and R. Ryne, IEEE Trans. Nucl. Sci. **5**, 2311 (1985).
[6] M. Berz, H. C. Hofmann, and H. Wollnik, Nucl. Instrum. Methods A **258**, 402 (1987).
[7] M. Berz, COSY INFINTY version 6 reference manual, Michigan State University Technical Report No. MSUCL-869, 1992 (unpublished).
[8] Y. T. Yan, in *Computational Accelerator Physics*, edited by Robert Ryne, AIP Conf. Proc. No. 297 (AIP, New York, 1994), pp. 279–284.
[9] L. Michelotti, in *Computational Accelerator Physics* (Ref. [8]), pp. 264-266.
[10] J. van Zeijts, in *Computational Accelerator Physics* (Ref. [8]), pp. 285–290.
[11] M. Berz, Nucl. Instrum. Methods A **298**, 426 (1990).
[12] M. Berz, *Computer Arithmetic and Enclosure Methods* (Elsevier Science, Amsterdam, 1992), pp. 439–450.
[13] M. Berz, *Automatic Differentiation of Algorithms* (SIAM, Philadelphia, 1991).
[14] M. Berz, Part. Accel. **24**, 109, (1989).
[15] M. Berz, Nucl. Instrum. Methods A **528**, 431 (1987).
[16] L. B. Rall, *Automatic Differentiation: Techniques and Applications* (Springer, New York, 1981).
[17] S. Kowalski and H. A. Enge, RAYTRACE user's manual, Massachusetts Institute of Technology Technical Report, 1987 (unpublished).
[18] K. L. Brown and J. E. Spencer, IEEE Trans. Nucl. Sci., **3**, 2568 (1981).
[19] H. Rose, Nucl. Instrum. Methods A **258**, 374 (1987).
[20] R. H. Helm, Stanford Linear Accelerator Center Technical Report No. 24, 1963 (unpublished).
[21] K. L. Brown, F. Rothacker, D. C. Carey, and Ch. Iselin, TRANSPORT user's manual, Stanford Linear Accelerator Center Technical Report No. SLAC-91, 1977.
[22] B. Hartmann, M. Berz, and H. Wollnik, Nucl. Instrum. Methods A **297**, 343 (1990).
[23] B. Hartmann, H. Irnich, and H. Wollnik, in *Nonlinear Problems in Accelerator Physics*, edited by M. Berz, S. Martin, and K. Ziegler, IOP Conf. Proc. No. 131 (Institute of Physics and Physical Society, London, 1993), pp. 87–96.
[24] H. Wollnik, Nucl. Instrum. Methods **38**, 56 (1965).
[25] H. Wollnik, B. Hartmann, and M. Berz, in *Linear Accelerator and Beam Optics Codes*, edited by Charles R. Eminhizer, AIP Conf. Proc. No. **177** (AIP, New York, 1988), p. 74.
[26] Y. Yan, in *The Physics of Particle Accelerators*, edited by M. Month and M. Dienes, AIP Conf. Proc. No. 249 (AIP, New York, 1992), p. 378.
[27] M. Berz, in *Linear Accelerator and Beam Optics Codes* (Ref. [25]), p. 275.
[28] M. Berz, *Nonlinear Problems in Future Accelerators* (World Scientific, Singapore, 1991), pp. 288–296.
[29] Y. Yan, Superconducting Super Collider Laboratory Technical Report No. 157, 1993 (unpublished).
[30] I. M. Gjaja, In *Nonlinear Problems in Accelerator Physics* (Ref. [23]), pp. 185–192.
[31] W. Wan, E. Goldmann, and M. Berz, in *Nonlinear Problems in Accelerator Physics* (Ref. [23]), pp. 201–207.
[32] W. Wan, E. Goldmann, and M. Berz, in *Computation Accelerator Physics* (Ref. [8]), pp. 143–149.
[33] H. Rose, Optic **85**, 19 (1990).
[34] G. H. Hoffstätter, Master's thesis, Technische Hochschule Darmstadt, 1991 (unpublished).
[35] A. J. Dragt and J. M. Finn, J. Math. Phys. **17**, 2215 (1976).
[36] H. Goldstein, *Classical Mechanics*, 2nd ed. (Addison-Wesley, Reading, MA, 1980).
[37] M. Berz, *Symplectic Tracking in Circular Acceleratos with High Order Maps* (World Scientific, Singapore, 1991), p. 288.
[38] G. H. Hoffstätter and M. Berz, Nucl. Instrum. Methods A **363**, 124 (1995).
[39] G. H. Hoffstätter, Ph.D. thesis, Michigan State University, 1994 (unpublished).