A Hamiltonianfree description of single particle dynamics for hopelessly complex periodic systems

Étienne Forest

Citation: Journal of Mathematical Physics **31**, 1133 (1990); doi: 10.1063/1.528795
View online: http://dx.doi.org/10.1063/1.528795
View Table of Contents: http://scitation.aip.org/content/aip/journal/jmp/31/5?ver=pdfcov
Published by the AIP Publishing

A Hamiltonian-free description of single particle dynamics for hopelessly complex periodic systems

Étienne Forest
Exploratory Studies Group, Accelerator and Fusion Research Division, Lawrence Berkeley Laboratory, 1 Cyclotron Road, MS 47-112, Berkeley, California 94720

(Received 27 July 1989; accepted for publication 10 January 1990)

A picture of periodic systems that does not rely on the Hamiltonian of the system, but on maps between a finite number of time locations, is developed. Moser or Deprit-like normalizations are done directly on the maps, thereby avoiding the complex time-dependent theory. Linear and nonlinear Floquet variables are redefined entirely in terms of maps. This approach relies heavily on the Lie representation of maps introduced by Dragt and Finn [J. Math. Phys. **20**, 2649 (1979); J. Geophys. Res. **81**, 13 (1976)]. One might say that although the Hamiltonian is not used in the normalization transformation, Lie operators are used, which are themselves, in some sense, pseudo-Hamiltonians for the maps they represent. The techniques find application in accelerator dynamics or in any field where the Hamiltonian is periodic, but hopelessly complex, such as magnetic field design in stellarators.

I. INTRODUCTION

The study of dynamical systems in traditional branches of classical mechanics uses the Hamiltonian as a starting point, which means that the numerical, as well as the analytical work, is done directly on the Hamiltonian of the system. In certain applications, such as dynamics of charged particle beams in accelerators, this approach may lead to a dead end. To illustrate this point, one needs only to compare celestial mechanics to accelerator physics. In accelerator physics, we try to understand the potential behavior of circular machines whose Hamiltonian is a very complex position-dependent function (the timelike variable is actually a length denoted by s). For example, even a small synchrotron radiation ring can consist of several dozens of bending magnets, quadrupoles, sextupoles, orbit correction dipoles, and rf cavities. The necessary inclusion of random errors in the simulation of such a device implies that the Hamiltonian will be a horrible periodic s-dependent function with the number of parameters ranging in the thousands. On the other hand, the problems of celestial mechanics, which might have a higher phase space dimensionality, are parametrized by a relatively small number of variables. In addition, accelerator Hamiltonians are very discontinuous in the timelike variable, which adds to the complexity of using a formalism devised for smooth time dependence.

Analytical computations (such as normalization procedures) have emphasized the "flow" (i.e., the Hamiltonian) instead of the study of a one-period "map" (i.e., a turn around a circular storage ring). Accelerator theorists have tried to adapt these tools to the study of circular machines, but have always had to restrict themselves to less than realistic problems. For this reason, a new approach for understanding our systems has been developed in recent years: It emphasizes the computation and analysis of large time maps. We believe that the tools that have been developed (software and theory) can be of use in other fields.

A. A bit of history

This new approach did not develop overnight. In fact, the approach has its roots deep in the field of accelerator physics. To orient the reader, we will present a subjective (not exhaustive) historical perspective of the use of flow and maps in accelerator theory and simulation. Since pioneering work in accelerator physics has often been obscure, a historical perspective may be viewed as an attempt at the proper recognition of such work.

Originally, the use of maps entered in the design of circular accelerators as paraxial (linear) representations of the ray propagation. The theory was derived from the light optics equivalent and consequently, the periodic structure of the systems was not properly exploited. Eventually, Courant and Snyder,[1] in their seminal paper on strong focusing, parametrized the motion around a linear ring in terms of invariant quantities, taking full advantage of the pseudoharmonic oscillator structure of the motion. In fact, as we will see in this paper, the Lie operator associated to the so-called Courant–Snyder invariant is proportional to the Lie operator of the linear one-turn map.

On another front, the inclusion of sextupoles in a ring worried a few people in the late 1950's. Because computers were not very powerful, Meier and Symon and Laslett *et al.*[2] used simple one-dimensional maps to guess at the potential harm caused by nonlinearities in a circular machine. As computers improved, maps disappeared from simulation and were replaced by the so-called kick codes, which are in fact second-order explicit symplectic integrators. In these codes, each time step of the integrator is derivable from a Hamiltonian. These codes are still the main ingredients of brute force simulations.[3]

On the nonlinear theoretical front, accelerator theorists tried to compute various relevant quantities such as frequency shifts (shear terms known as tune shifts in accelerator physics) and distortions of the invariants with the help of canonical perturbation theory. The theorists adapted to ac-

 0022-2488/90/051133-12$03.00

celerator problems the algorithms of Poincaré (-Von Zeippel)[4]; Moser–Birkhoff[2]; and later, Deprit.[5] The inherent complexity of our Hamiltonians leads to great technical difficulties in the application of such algorithms to realistic problems.

Meanwhile, the use of maps kept creeping into the linear theory. This probably culminated in a series of papers by Chao[6] in the late 1970's, where he computed the equilibrium emittances and spin polarization in a circular electron ring with the help of a map-based theory instead of a flow or Hamiltonian approach. One might ask why accelerator theorists continued to write papers on the evaluation of nonlinear quantities, relying entirely on the usual canonical perturbation theory, while linear calculations often used the more suitable map approach: We venture to suggest that the answer is the Lie representation of the map. Although it is obvious after a little thought that the quantities obtained by standard canonical perturbation theory are present in a power series expansion of the one-turn map,[7] the expansion of the final position and momentum (+ spin if you care about it) in terms of their initial components is very different in form from the central object of canonical perturbation theory: the Hamiltonian. Consequently, without at least an awareness of the Lie representation, accelerator theorists could not have been expected to rephrase the nonlinear theory in a way suitable for the circular machine, such as they did for the linear case.

At this point enters Dragt. In the 1970's Dragt and Finn worked on some version of the Deprit algorithm[8] and applied it to various problems of plasma physics.[9,8] Dragt became involved with accelerator theory and with the help of Douglas, they wrote the first version of a code for beam optics (MARYLIE)[10] which parametrizes the Taylor series maps in terms of their Lie generators. In collaboration with the present author, a normal form algorithm was first introduced in the code MARYLIE by Dragt *et al.*[8] Finally, Dragt, in an obscure report, introduced for the first time the concept of phase advance from a map point of view[11]: His ideas were not complete, but they planted the right seed in the present author's mind. At this point it became clear that the extraction of maps and their subsequent analysis (normalization) provided a powerful approach to numerical and analytic computation in accelerator theory. Remarkably, in 1959, Meier and Symon (Ref. 2) used a Lie representation of the map without knowing it. Meier and Symon were studying a map consisting of a rotation followed by a sextupolar kick. By writing a time-dependent pseudo-Hamiltonian that generates the exact same map, Meier and Symon were able to compute a canonical transformation in order to simplify it: Their pseudo-Hamiltonian was the factorized Lie representation of the map proposed by Dragt and Finn.[8] Meier and Symon's only error was in not trying to extirpate from the theory and normalization the bogus time dependence.

Most recently, software development has increased the numerical power of the map approach enormously. Indeed, the extraction of Taylor series representation of maps from simulation codes can be very tedious. Chao and, later, the present author restricted themselves to codes where the individual magnets had a simple representation (thin lenses): At most we could extract fifth-degree polynomial maps in six-dimensional phase space.[12] Fortunately, Berz has created a software package, the Differential Algebra Package, which permits exact automatic differentiation of any quantities integrated on the computer, in particular, the position and momentum vector that is evaluated in our simulation codes.[13] The same tools of Berz permit a user to manipulate the resulting power series map into any type of representation and in particular the factored Lie representation suggested by Dragt and Finn.[8] It suffices to say that the theoretical concepts discussed in this paper have *all* been implemented for the power series representation thanks to Berz's package.[14]

To set the tone, we will review a few concepts concerning symplectic maps.

B. A few words about symplectic maps[15–17]

A symplectic map $\mathbf{M}$ transforms a differentiable function $f(\mathbf{z}_0)$ of the initial phase space into another function $(\mathbf{M}f)(\mathbf{z}_0)$, where $\mathbf{z}_0 = (q_{01},p_{01},\ldots,q_{0N},p_{0N})$. We say that $\mathbf{M}$ is symplectic if it preserves the Poisson bracket of two functions f and g:

$$\mathbf{M}[f,g] = [\mathbf{M}f,\mathbf{M}g], \tag{1a}$$

$$\mathbf{z}_0 = (q_{01},p_{01},\ldots,q_{0N},p_{0N}). \tag{1b}$$

Now consider the motion generated by a Hamiltonian K from location s_0 to s. We know that any function $f(\mathbf{z}_0)$ will be transformed at location s into a new function $f_s(\mathbf{z}_0)$. The two functions are connected by a symplectic map $\mathbf{M}(s_0,s)$. Using the properties of Hamilton's equations, one can show that $\mathbf{M}(s_0,s)$ obeys[17]

$$\frac{d}{ds}\mathbf{M}(s_0,s) = \mathbf{M}(s_0,s):-K(\mathbf{z}_0;s):, \tag{2a}$$

$$:f:g = [f,g], \tag{2b}$$

$$\mathbf{M}(s_0,s_0) = \mathbf{E} = \text{identity map}. \tag{2c}$$

Here we follow Dragt's notation for the Lie operator [Eq. (2b)]. From Eqs. (2) we deduce that any map of the form $\mathbf{M}(\mathbf{z}_0) = \exp(:f(\mathbf{z}_0):)$ is a symplectic map. Equation (2a) is very similar to Schrödinger's equation for the unitary transformation in quantum mechanics. However, here the resulting map will be symplectic. Notice that the generator $:-K(\mathbf{z}_0;s):$ depends only on the initial phase space variable $\mathbf{z}_0$. It is also easy to show that the differential properties of any Lie operator such as $:-K:$ imply that

$$f_s(\mathbf{z}_0) = (\mathbf{M}(s_0,s)f)(\mathbf{z}_0) = f(\mathbf{M}(s_0,s)\mathbf{z}_0)$$
$$\equiv f(\mathbf{z}_s(\mathbf{z}_0)). \tag{3}$$

Finally, it is worth remembering that symplectic maps act in the reverse order from matrix multiplication when expressed in terms of the initial phase space variables. To show this property, we imagine a two-step process:

$$0 \overset{M}{\rightarrow} s' \overset{N}{\rightarrow} s,$$

$$\mathbf{M}(\mathbf{z}_0)\mathbf{z}_0 = \mathbf{z}_{s'}, \tag{4}$$
$$\mathbf{N}(\mathbf{z}_{s'})\mathbf{z}_{s'} = \mathbf{z}_s.$$

Here the map $\mathbf{M}$ transforms functions of the initial phase space at $s=0$, while $\mathbf{N}$ acts on functions of the phase space at s' [throughout this paper, the notation $\mathbf{M}(\mathbf{z}_0)$ indicates that the Lie operators of $\mathbf{M}$ are expressed in terms of $\mathbf{z}_0$].

Clearly, we can propagate a function $f(\mathbf{z}_0)$ to a location s by the composition of functions rule:

$$f_s(\mathbf{z}_0) = f(\mathbf{z}_s), \tag{5a}$$

but

$$f(\mathbf{z}_s) = \mathbf{N}(\mathbf{z}_{s'})f(\mathbf{z}_{s'}) = \mathbf{N}(\mathbf{z}_{s'})\mathbf{M}(\mathbf{z}_0)f(\mathbf{z}_0), \tag{5b}$$
$$= \mathbf{N}(\mathbf{M}(\mathbf{z}_0)\mathbf{z}_0)\mathbf{M}(\mathbf{z}_0)f(\mathbf{z}_0), \tag{5c}$$
$$= \mathbf{M}(\mathbf{z}_0)\mathbf{N}(\mathbf{z}_0)\mathbf{M}^{-1}(\mathbf{z}_0)\mathbf{M}(\mathbf{z}_0)f(\mathbf{z}_0), \tag{5d}$$
$$= \mathbf{M}(\mathbf{z}_0)\mathbf{N}(\mathbf{z}_0)f(\mathbf{z}_0). \tag{5e}$$

Equation (5a) is the result of simple composition. In (5b) we apply the definition of the two maps to (5a). Finally, (5c) and (5d) are the results of the differential properties of the Lie operators associated to the maps.

Indeed, it can be shown using properties (1) and (3) that

$$\exp(:f(\mathbf{M}\mathbf{z}_0):) = \exp(:\mathbf{M}f(\mathbf{z}_0):)$$
$$= \exp(\mathbf{M}:f(\mathbf{z}_0):\mathbf{M}^{-1})$$
$$= \mathbf{M}\exp(:f(\mathbf{z}_0):)\mathbf{M}^{-1}. \tag{6}$$

The reverse ordering is also present in Eq. (2a), as seen by integrating it from s to $s+ds$.

We are now ready to introduce our map description of complex periodic systems. In Sec. II and Appendix A we review in very general terms the basic Hamiltonian and canonical transformations used in circular machine theory and simulations. In Sec. III A and B we present the equivalent map description of our system by viewing the ring as an ordered set of maps. In Sec. III C we sketch the one-turn map normalization. In Sec. IV we use the one-turn map normalization to define the Floquet ring. In Sec. V these concepts are applied to the second-order normalization of a perturbed Floquet ring. In Sec. VI we apply the map approach to linear systems in the continuous limit (i.e., in the Hamiltonian limit): A set of well-known results follows explicitly. In Appendix B we sketch the proof of a few theorems.

II. THE PROBLEM OF CIRCULAR MACHINES

As mentioned in Sec. I, the need for a map-based theory is most apparent in accelerator physics. Therefore, we will describe in very general terms the central problem of accelerator design.

Consider a Hamiltonian $H(\mathbf{x},\Delta;s)$, where $\mathbf{x}$ is a phase space vector of dimension $2N$ and Δ is the set of N_p parameters describing the departure of our system from its design value (i.e., by definition, the ideal machine is described by the case $\Delta=0$). We also assume that H is periodic in s with period $s=1$:

$$H(\mathbf{x},\Delta;s+1) = H(\mathbf{x},\Delta;s). \tag{7}$$

In addition, we can select a new set of canonical variables $\mathbf{z}(\epsilon)$ that is generated by a periodic Lie operator associated to the function $w(\mathbf{z},\Delta;s;\epsilon)$:

$$\frac{d\mathbf{z}}{d\epsilon} = [\mathbf{z},w] = :-w:\mathbf{z}, \quad \mathbf{z}(\epsilon=0) = \mathbf{x}. \tag{8}$$

Ultimately the parameter ϵ is set to 1. It has been shown that the variable $\mathbf{z}(\epsilon=1)$ is propagated by a Hamiltonian $K(s)$ obeying:[18,19]

$$K(\mathbf{z},\Delta;s) = \mathbf{A}(\mathbf{z},\Delta;s)\Big(H(\mathbf{z},\Delta;s) + \int_0^{\epsilon=1} d\epsilon$$
$$\times \mathbf{A}^{-1}(\mathbf{z},\Delta;s)\frac{\partial}{\partial s}w(\mathbf{z},\Delta;s;\epsilon)\Big),$$
$$\frac{d\mathbf{z}}{ds} = [\mathbf{z},K]. \tag{9}$$

Here $\mathbf{A}^{-1}$ is the periodic canonical map that transforms $\mathbf{x}$ into $\mathbf{z}$; it is generated by w. Equations (8) and (9) can be viewed as the fundamental equations of an accelerator in the absence of collective or dissipative effects. The *fundamental problem* of accelerator dynamics is to study the stability of the motion generated by K or H as one iterates n turns around the machine from $s=s_0$ to $s=s_0+n$ $(n\rightarrow\infty)$.

Often the theorist attempts to select the generator w in a way that will simplify the structure of K. We refer to this kind of process as a normalization process. In general, the computation of K is extremely difficult because it requires a knowledge of $\mathbf{A}$ for every value of s! [For completeness, see Appendix A for the map equivalent of Eqs. (7)–(9) and a derivation of $K(\mathbf{z},\Delta;s)$.]

Typically, no attempt is made to simplify the Hamiltonian and one integrates the motion generated by H with the help of a symplectic integrator.[20] The phase space data are then examined at a finite number of surface of sections (often only one!) and all hope of analytical understanding is abandoned.

The map description of dynamics described in this paper was developed as a direct consequence of the impossibility of normalizing a realistic accelerator Hamiltonian by blindly applying a Deprit-type algorithm.

III. THE HAMILTONIAN-FREE OR MAP DESCRIPTION

A. The motivation

The map approach is based on a redefinition of the system in terms of a finite number of maps. In accelerator theory, we are motivated in redefining the problem by the following facts.

(i) Most simulations are performed by symplectic integrators.[21] Only a finite set of location $\{s_i\}$ are examined during this process. The maximum number of locations ever to be examined is the number of integration steps around the ring.

(ii) The form of the equations of motion for a computer simulation may not and will not in general be suitable for a normal form analysis.

(iii) Although symplectic maps act on the infinite-dimensional space of functions, the property displayed in Eq.

(3) permits us to restrict ourselves to the coordinate representation of the map $\mathbf{z}_s(\mathbf{z}_0) = \mathbf{M}\mathbf{z}_0$.

Statement (i) simply says that most simulations involve very discontinuous Hamiltonians in s. While it is cumbersome to fold a discontinuous Hamiltonian into a Deprit style algorithm, the production of maps and their subsequent analysis are ideally suited for discontinuous systems.

Statement (ii) points to the necessary discrepancies between the world of a computer and the world of theoretical analysis. The process we will describe allows for a total decoupling. One may extract maps using a noncanonical set of coordinates and later transform these maps into canonical variables quite independent of the system that produced them. For example, the calculation of the motion through some complex fringe field may be easiest in noncanonical variables, in some cases using even time as the Hamiltonian parameter.

Statement (iii) is extremely important: It implies that our efforts should be in the direction of extracting a representation of $\mathbf{z}_s$. For example, in the case of a Taylor series representation of $\mathbf{z}_s$, we mentioned that Berz has developed powerful software tools (the Differential Algebra Package) that perform automatic differentiation to arbitrary order on a computer,[13] making it possible to extract $\mathbf{z}_s$ as a power series around some trajectory in phase space (usually the periodic closed orbit). The same tools used by Berz allowed Irwin and the present author to write the necessary software for the normalization of the one-turn map.[14] Needless to say, a Taylor series representation may not be always suitable. Presently, nonpower series representations are being studied by Warnock *et al.* who have also developed methods to normalize the map. Warnock *et al.*'s representation can permit the study of very nonlinear processes and they succeed in many cases at finding numerically approximate invariants of the motion near chaotic regions.[22] Unfortunately, Warnock *et al.* do not have tools as flexible as the Differential Algebra Package of Berz. For this reason, thanks to Berz's tools, the power series representation of $\mathbf{z}_s$ is the only representation for which all the concepts presented in this paper are and have been implemented.

B. The redefinition of the system

1. Definition of the ring

A ring is an ordered m-tuple $\mathfrak{R} = (\mathbf{N}_{\mathbf{i}\,\mathbf{i}+\mathbf{1}})$ of m maps connecting m surfaces of sections or observation points (see Fig. 1). Here, the index $\mathbf{i}$ runs from 1 to m with the convention $\mathbf{i}+m=\mathbf{i}$. The maps in $\mathfrak{R} = (\mathbf{N}_{\mathbf{i}\,\mathbf{i}+\mathbf{1}})$ are symplectic. Without loss of generality, we assume that these maps transform the origin of phase space into itself:

$$\forall i \quad \mathbf{N}_{\mathbf{i}\,\mathbf{i}+\mathbf{1}}\mathbf{z}_0|_{\mathbf{z}_0=\mathbf{0}} = \mathbf{0}. \tag{10}$$

Clearly, the number of maps will depend on the particular aspect of the problem being studied: Again, we emphasize that it cannot exceed the number of steps in our symplectic integrator.

We are now in a position to define standard concepts on our ring. The reader must remember that the underlying assumption in this paper is our ability to extract and manipulate maps to any order in the perturbation [i.e., in the case of a power series this would be the degree in the vector $(\mathbf{z}_0,\delta)$, where δ is a subset of the total parameter set Δ of Eq. (7)].[13,14]

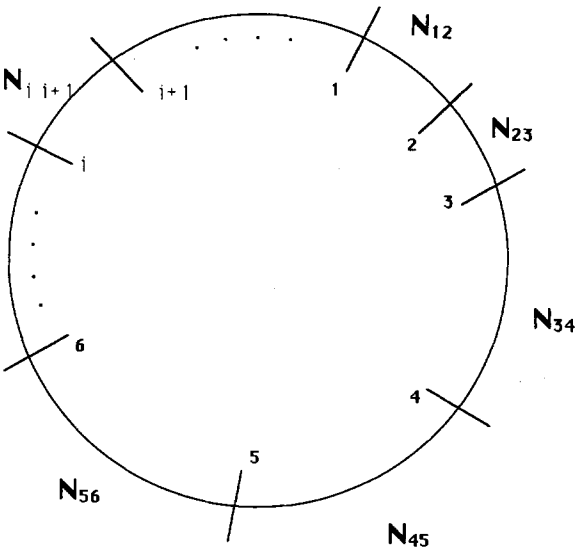

FIG. 1. Schematic view of the ring $\mathfrak{R}$.

C. The one-turn maps and their normalization[23]

In the following we sketch the steps of the normalization of a one-turn map. This formal procedure is explained in detail in Ref. 14. Given the ring $\mathfrak{R}$, we can easily compute the one-turn map from location s_i back to s_i. This map is simply given by a left to right product:

$$\mathbf{M}_i = \prod_{k=i}^{i+m-1} \mathbf{N}_{\mathbf{k}\,\mathbf{k}+\mathbf{1}} \tag{11a}$$

and

$$\mathfrak{I} = (\mathbf{M}_i). \tag{11b}$$

The new m-tuple $\mathfrak{I}$ that is created out of $\mathfrak{R}$ is not equivalent to $\mathfrak{R}$, as we will see in Sec. IV when we define the Floquet ring.

The first property we will assign to our map is linear stability. To define linear stability we expand the coordinate representation of $\mathbf{M}_i$ around the origin:

$$\mathbf{z}_{s_i}(\mathbf{z}_0) = L_i\mathbf{z}_0 + \cdots. \tag{12}$$

We assume that the $2N \times 2N$ matrix L_i has $2N$ *distinct* eigenvalues on the unit circle:

$$\lambda_j^{\pm} = \exp(\pm i2\pi\nu_j), \quad 0<\nu_j<1, \quad j=1,N. \tag{13}$$

Associated to the matrix L_i is a Lie map $\mathbf{L}_i$. Clearly, the eigenvalues of L_i do not depend on location since $\mathbf{L}_j$ is obtained from $\mathbf{L}_i$ by a similarity transformation (the same is true for the full nonlinear map):

$$\mathbf{L}_j = \mathbf{L}_{ij}{}^{-1}\mathbf{L}_i\mathbf{L}_{ij}, \text{ where } \mathbf{L}_{ij} = \prod_{k=i}^{j-1} \mathbf{L}_{\mathbf{k}\,\mathbf{k}+\mathbf{1}}. \tag{14}$$

For the linear part of the map, we claim that their exists a linear transformation $\mathbf{A}_{\mathbf{L}i}$ such that

$$\mathbf{A}_{\mathbf{L}i}\mathbf{L}_i\mathbf{A}_{\mathbf{L}i}{}^{-1} = \exp(:-\boldsymbol{\mu}\cdot\mathbf{J}:) = \mathbf{R}_{\mathbf{L}}, \tag{15a}$$

$$\mu_j = 2\pi\nu_j, \tag{15b}$$

$$J_j = ((z_{0;2j-1})^2 + (z_{0;2j})^2)/2. \tag{15c}$$

The transformation $\mathbf{A}_{\mathrm{Li}}$ depends on the location s_i. However, the map $\mathbf{R}_{\mathrm{L}}$ is universal once a tune ν_j has been assigned a given plane j. The transformation done in Eqs. (15) is always possible for a stable linear map with distinct eigenvalues. Let us go back to the nonlinear map $\mathbf{M}_i$. Following Dragt and Finn,[16] we express it in a factored form:

$$\mathbf{M}_i = \mathbf{L}_i(\mathbf{L}_i^{-1}\mathbf{M}_i) = \mathbf{L}_i \prod_{k=1}^{N_0} \exp(:f_k(\mathbf{z}_0):) \tag{16}$$

In the case of a power series representation of $\mathbf{M}_i$, $N_0 - 1$ is the degree of the polynomial approximation of the function $\mathbf{z}_{s_i}(\mathbf{z}_0)$ and $k+2$ is the degree in $\mathbf{z}_0$ of the Lie exponent. For a different type of approximation the reader can view N_0 and k as the degree of some smallness parameter. Our ultimate goal is to normalize $\mathbf{M}_i$ partially or totally. In a total normalization, we must find a transformation $\mathbf{A}_i$ (analytic around the origin) such that

$$\mathbf{A}_i\mathbf{M}_i\mathbf{A}_i^{-1} = \exp(:-\boldsymbol{\mu}\cdot\mathbf{J} + D(\mathbf{J}):) = \mathbf{R}. \tag{17}$$

For analytic $\mathbf{A}_i$ one can show that $\mathbf{R}$ is unique for a given ordering of the planes and independent of the location (see Appendix B).

Obviously, using Eqs. (15), we factor $\mathbf{A}_i$ into linear and nonlinear parts:

$$\mathbf{A}_i = \mathbf{A}_{\mathrm{Ni}}\mathbf{A}_{\mathrm{Li}}, \tag{18a}$$

$$\mathbf{A}_{\mathrm{Ni}} = \mathbf{A}_{\mathrm{N},i}\cdots\mathbf{A}_{1i}, \tag{18b}$$

$$\mathbf{A}_{ki} = \exp(:F_k:), \quad k = 1, N_0. \tag{18c}$$

Applying (18) on $\mathbf{M}_i$ gives us

$$\mathbf{A}_i\mathbf{M}_i\mathbf{A}_i^{-1} = \mathbf{A}_{\mathrm{Ni}}\mathbf{R}_{\mathrm{L}} \prod_{k=1}^{N_0} \exp(:g_{ki}(\mathbf{z}_0):)\mathbf{A}_{\mathrm{Ni}}^{-1}, \tag{19a}$$

where

$$g_{ki}(\mathbf{z}_0) = \mathbf{A}_{\mathrm{Li}}f_{ki}(\mathbf{z}_0) = f_{ki}(\mathbf{A}_{\mathrm{Li}}\mathbf{z}_0). \tag{19b}$$

The normalization of the nonlinear map starts with Eqs. (19). To see the type of operators involved, let us compute the effect of $\mathbf{A}_{1i}$ on the second-degree map $\mathbf{R}_{\mathrm{L}}\exp(:g_{1i}:)$:

$$\begin{aligned}
&\mathbf{A}_{1i}\mathbf{R}_{\mathrm{L}}\exp(:g_{1i}:)\mathbf{A}_{1i}^{-1} \\
&\quad = \exp(:F_{1i}:)\mathbf{R}_{\mathrm{L}}\exp(:g_{1i}:)\exp(:-F_{1i}:) \\
&\quad = \mathbf{R}_{\mathrm{L}}\mathbf{R}_{\mathrm{L}}^{-1}\exp(:F_{1i}:)\mathbf{R}_{\mathrm{L}}\exp(:g_{1i}:)\exp(:-F_{1i}:) \\
&\quad = \mathbf{R}_{\mathrm{L}}\exp(:\mathbf{R}_{\mathrm{L}}^{-1}F_{1i}:)\exp(:g_{1i}:)\exp(:-F_{1i}:).
\end{aligned} \tag{20}$$

To first order in the Lie exponents, we can rewrite (20) as

$$\begin{aligned}
&\mathbf{R}_{\mathrm{L}}\exp(:\mathbf{R}_{\mathrm{L}}^{-1}F_{1i}:)\exp(:g_{1i}:)\exp(:-F_{1i}:) \\
&\quad = \mathbf{R}_{\mathrm{L}}\exp(:-(\mathbf{E}-\mathbf{R}_{\mathrm{L}}^{-1})F_{1i} + g_{1i}:) + \mathrm{order}(g_2) \\
&\quad = \mathbf{R}_{\mathrm{L}}\exp(:-\mathbf{T}F_{1i} + g_{1i}:) + \mathrm{order}(g_2)\cdots,
\end{aligned} \tag{21a}$$

$$\mathbf{T} = \mathbf{E} - \mathbf{R}_{\mathrm{L}}^{-1}, \tag{21b}$$

where $\mathbf{E}$ is identity map.

From the Eqs. (21), we see that the operator $\mathbf{T}$ is central to the understanding of the effect of any similarity transformation. Since $\mathbf{T}$ is essentially $\mathbf{R}_{\mathrm{L}}$, we must study the Lie operator $:\boldsymbol{\mu}\cdot\mathbf{J}:$. As pointed out by Cushman *et al.*,[24] as well as Dragt and Finn,[8] the operator $:\boldsymbol{\mu}\cdot\mathbf{J}:$ is a semisimple endomorphism of the space P_k of homogeneous polynomials of degree $k \geqslant 1$ in $\mathbf{z}_0$; hence it is true that

$$P_k = \mathrm{Im}\,\mathbf{T} \oplus \mathrm{Ker}\,\mathbf{T}. \tag{22}$$

In fact, it is easy to derive Eq. (22) by simply constructing the linear eigenfunctions of $\mathbf{T}$ (or $:\boldsymbol{\mu}\cdot\mathbf{J}:$) (Ref. 8):

$$:\boldsymbol{\mu}\cdot\mathbf{J}:h_j^{\pm} = \pm i\mu_j h_j^{\pm}, \tag{23a}$$

$$h_j^{\pm} = z_{2j-1} \pm iz_{2j} = \sqrt{2J_j}\exp(\mp i\varphi_j),$$

where

$$[\varphi_j, J_j] = \delta_{ij} \tag{23b}$$

are the only nonzero brackets,

$$J_j = \tfrac{1}{2}h_j^{+}h_j^{-}, \quad j = 1, N. \tag{23c}$$

For completeness, in (23b) we displayed the connection between our eigenfunctions and the usual set of action-angle variables (J_j, φ_j).

Using these linear eigenvectors, we can construct an eigenbasis for P_k:

$$\begin{aligned}
&|\mathbf{m},\mathbf{n}\rangle = (h_1^{+})^{m_1}(h_1^{-})^{n_1}\cdots(h_N^{+})^{m_N}(h_N^{-})^{n_N} \\
&:\boldsymbol{\mu}\cdot\mathbf{J}:|\mathbf{m},\mathbf{n}\rangle = i(\mathbf{m}-\mathbf{n})\cdot\boldsymbol{\mu}|\mathbf{m},\mathbf{n}\rangle.
\end{aligned} \tag{24}$$

Since $\mathbf{T}$ is diagonal in the $|\mathbf{m},\mathbf{n}\rangle$ basis, it follows that P_k decomposes into the direct sum of its image and its kernel. In fact, if the μ's are irrational among each other, the kernel is given by

$$|\mathbf{m},\mathbf{n}\rangle \in \mathrm{Ker}\,\mathbf{T} \Rightarrow \mathbf{m}-\mathbf{n} = \mathbf{0}, \quad \text{i.e.,} \quad |\mathbf{m}-\mathbf{n}| = 0. \tag{25}$$

In a partial normalization, we decide to leave in the final map terms for which $\mathbf{m}-\mathbf{n} \neq \mathbf{0}$ [or often $(\mathbf{m}-\mathbf{n})\cdot\boldsymbol{\mu} \cong 0$]. This allows us to study islands produced by a resonance.

We say that $\mathbf{M}_i$ is partially normalized into the map $\mathbf{R}$ if

$$\mathbf{A}_i\mathbf{M}_i\mathbf{A}_i^{-1} = \exp\Big(:-\boldsymbol{\mu}\cdot\mathbf{J} + D_i(\mathbf{J}) + \sum_{\mathbf{m}-\mathbf{n}\in I_r} D_{\mathbf{m},\mathbf{n};i}|\mathbf{m},\mathbf{n}\rangle:\Big) = \mathbf{R}_i, \tag{26}$$

where $I_r = \{\mathbf{k}\in Z^N \mid \mathbf{k} = \text{selected resonances}\}$. As indicated in (26) by the index i, in a partial normalization the final map will depend on the location in the ring. According to Eqs. (21), we obtain $\mathbf{R}$ by inverting $\mathbf{T}$. In fact we can redefine $\mathbf{T}^{-1}$ using a projection operator:

$$\begin{aligned}
\mathbf{T}_{\mathrm{im}}^{-1}g &= \mathbf{T}^{-1}\mathbf{P}_{\mathrm{im}}g \\
&= \sum_{\mathbf{m}-\mathbf{n}\in I_r\cup\{0\}} \frac{\mathbf{A}_{\mathbf{m},\mathbf{n}}}{1-\exp(i(\mathbf{m}-\mathbf{n})\cdot\boldsymbol{\mu})}|\mathbf{m},\mathbf{n}\rangle,
\end{aligned} \tag{27a}$$

$$\mathbf{P}_{\mathrm{im}}g = \sum_{\mathbf{m}-\mathbf{n}\in I_r\cup\{0\}} \mathbf{A}_{\mathbf{m},\mathbf{n}}|\mathbf{m},\mathbf{n}\rangle, \tag{27b}$$

$$g = \sum_{\mathbf{m},\mathbf{n}} \mathbf{A}_{\mathbf{m},\mathbf{n}}|\mathbf{m},\mathbf{n}\rangle. \tag{27c}$$

Provided that one knows how to compose the maps involved in the normalization and extract their leading order Lie representation, the maps $\mathbf{A}_{ki}$ can be computed by iteration using $\mathbf{T}_{\mathrm{im}}^{-1}$ as defined in Eqs. (27). This procedure was first implemented to third order in the Taylor series by the present author and Dragt in the context of the code MAR-

YLIE.[10] Later, Neri and Dragt pushed the process to fifth order using the same code. Recently, the present author and Irwin, in close collaboration with Berz,[14] developed the algorithm and software necessary to extend the map normalization to an arbitrary order. The process is semianalytic since in practice only a small number of components of Δ [see Eqs. (7)–(9)] can be retained. This number can vary depending on the order of the normalization, the phase space dimension, and the power of the computer used.

For our purpose, it suffices to know that one can define (exactly for linear maps and formally in the nonlinear case) a normalized map $\mathbf{R}$. Although we concentrate in this article on a normal form algorithm based on the semisimple operator $:\boldsymbol{\mu}\cdot\mathbf{J}:$, it is possible and sometimes desirable to study systems that are not semisimple.[14,24] What can be done on the Hamiltonian can also be done on the map.

IV. THE FLOQUET RING

Let us assume that we have achieved a complete normalization of the ring. As described by Eq. (17), we have

$$\begin{aligned} &\rho(\mathbf{M}_i)=\mathbf{A}_i, \quad \rho:\mathfrak{I}\rightarrow\text{symplectic maps},\\ &\mathbf{A}_i\mathbf{M}_i\mathbf{A}_i^{-1}=\exp(:-\boldsymbol{\mu}\cdot\mathbf{J}+D(\mathbf{J}):)=\mathbf{R}. \end{aligned} \tag{28}$$

The transformation ρ introduced in (28) can be viewed as a map over the set of one-turn maps $\mathfrak{I}$ defined over $\mathfrak{R}$. Using ρ we can define a new ring: the Floquet ring. We first proceed by mapping the m-tuple $\mathfrak{I}$:

$$\mathfrak{I}_\rho=\rho(\mathfrak{I})=(\rho(\mathbf{M}_i)\mathbf{M}_i\rho(\mathbf{M}_i)^{-1})=(\mathbf{R},\mathbf{R},\ldots,\mathbf{R}). \tag{29}$$

The fact that $\mathfrak{I}_\rho$ contains only $\mathbf{R}$ is demonstrated in Appendix B by generalizing Eq. (14) to the maps $\mathbf{N}_{k\,k+1}$ and using the assumed analyticity of the maps that are involved.

More important, we must find out what happens to $\mathfrak{R}$. We first state the result.

(i) The new ring $\rho(\mathfrak{R})$ (or $\mathfrak{R}_\rho$) is made out of amplitude-dependent rotations (called phases). The angles of these rotations reduce to the so-called linear phase advance in the linear regime.

(ii) Two different normalizations ρ and ρ' can only differ by a phase for a given ordering of the tunes.

Corollary: The phase advance between two matched locations ($\mathbf{M}_i=\mathbf{M}_j$) is the same for any definition of ρ.

Proof: We now prove the above results.

Property (i): Using the normalization ρ, we conclude from Eq. (29) that

$$\rho(\mathbf{M}_i)\mathbf{M}_i\rho(\mathbf{M}_i)^{-1}=\rho(\mathbf{M}_j)\mathbf{M}_j\rho(\mathbf{M}_j)^{-1}=\mathbf{R}. \tag{30}$$

Using the definition of $\mathfrak{I}$, we may write

$$\mathbf{M}_j=\mathbf{N}_{ij}^{-1}\mathbf{M}_i\mathbf{N}_{ij}. \tag{31}$$

We can substitute (31) into (30):

$$\begin{aligned} &\rho(\mathbf{M}_i)\mathbf{M}_i\rho(\mathbf{M}_i)^{-1}\\ &\quad=\rho(\mathbf{M}_j)\mathbf{N}_{ij}^{-1}\mathbf{M}_i\mathbf{N}_{ij}\rho(\mathbf{M}_j)^{-1}\\ &\quad\Rightarrow\mathbf{M}_i=\rho(\mathbf{M}_i)^{-1}\rho(\mathbf{M}_j)\mathbf{N}_{ij}^{-1}\mathbf{M}_i\mathbf{N}_{ij}\rho(\mathbf{M}_j)^{-1}\rho(\mathbf{M}_i)\\ &\quad\Rightarrow\mathbf{R}=\rho(\mathbf{M}_j)\mathbf{N}_{ij}^{-1}\rho(\mathbf{M}_i)^{-1}\mathbf{R}\rho(\mathbf{M}_i)\mathbf{N}_{ij}\rho(\mathbf{M}_j)^{-1}\\ &\quad\Rightarrow\mathbf{R}=\mathbf{B}_{ij}^{-1}\mathbf{R}\mathbf{B}_{ij}, \end{aligned}$$

where

$$\mathbf{B}_{ij}=\rho(\mathbf{M}_i)\mathbf{N}_{ij}\rho(\mathbf{M}_j)^{-1}. \tag{32}$$

We now take advantage of the Lie algebraic representation of $\mathbf{R}$:

$$\begin{aligned} &\mathbf{R}=\mathbf{B}_{ij}^{-1}\exp(:-\boldsymbol{\mu}\cdot\mathbf{J}+D(\mathbf{J}):)\mathbf{B}_{ij}\\ &\quad\Rightarrow\exp(:-\boldsymbol{\mu}\cdot\mathbf{J}+D(\mathbf{J}):)\\ &\quad=\exp(:\mathbf{B}_{ij}^{-1}(-\boldsymbol{\mu}\cdot\mathbf{J}+D(\mathbf{J})):). \end{aligned} \tag{33}$$

Using the assumed analyticity of the various maps involved in Eq. (33), one can show that $\mathbf{B}_{ij}$ can depend only on $\mathbf{J}$ (see Appendix B); hence it can be written with a single Lie operator Φ_{ij}:

$$\mathbf{B}_{ij}=\exp(:-\Phi_{ij}(\mathbf{J}):). \tag{34}$$

The angle of the rotation produced by $\mathbf{B}_{ij}$ is simply

$$\Delta\Phi_{ij}=-\frac{\partial\Phi_{ij}}{\partial\mathbf{J}}. \tag{35}$$

Using Eqs. (32) and (35), we can define the Floquet ring $\mathfrak{R}_\rho$ associated to ρ to be the m-tuple

$$\begin{aligned} &\mathfrak{R}_\rho=(\mathbf{B}_{k\,k+1}) \text{ such that}\\ &\mathbf{B}_{k\,k+1}=\rho(\mathbf{M}_k)\mathbf{N}_{k\,k+1}\rho(\mathbf{M}_{k+1})^{-1}. \end{aligned} \tag{36}$$

Property (ii): Finally, from Eq. (30) and the uniqueness of $\mathbf{R}$, we obtain a relation identical to Eq. (33) in the presence of two different normalizations ρ and ρ':

$$\begin{aligned} \mathbf{R}&=\exp(:-\boldsymbol{\mu}\cdot\mathbf{J}+D(\mathbf{J}):)\\ &=\rho(\mathbf{M}_i)\rho'(\mathbf{M}_i)^{-1}\exp(:-\boldsymbol{\mu}\cdot\mathbf{J}+D(\mathbf{J}):)\\ &\quad\times\rho'(\mathbf{M}_i)^{-1}\rho(\mathbf{M}_i). \end{aligned} \tag{37}$$

Hence $\rho(\mathbf{M}_i)\rho'(\mathbf{M}_i)^{-1}$ is a rotation and equivalent normalizations can only differ by a phase. It is a simple exercise to prove the corollary on matched locations.

V. PERTURBATION OF THE RING $\mathfrak{R}$

Often one perturbs a Hamiltonian at several locations. One would like to know how the ring $\mathfrak{R}$ and its Floquet counterpart $\rho(\mathfrak{R})$ are affected by perturbations, in particular Hamiltonian perturbations.

Let us assume that the ring is perturbed at the ith location by a Lie operator $\mathbf{C}_i=\exp(:-V_i:)$. In accelerator physics, this kind of question is often asked. For example, $\mathbf{C}_i$ could represent a nonlinear multipole error or a beam–beam kick: The list is endless. Clearly, the perturbed ring $\mathfrak{R}^p$ is just the m-tuple

$$\mathfrak{R}^p=(\mathbf{C}_i\mathbf{N}_{i\,i+1}). \tag{38}$$

More interesting, we would like to examine the perturbed Floquet ring:

$$\begin{aligned} \mathfrak{R}_f^p&=(\mathbf{A}_i\mathbf{C}_i\mathbf{A}_i^{-1}\mathbf{B}_{i\,i+1})\\ &=(\exp(:-\mathbf{A}_iV_i:)\mathbf{B}_{i\,i+1})\\ &=(\exp(:-V_i(\mathbf{A}_i\mathbf{z}_0):)\mathbf{B}_{i\,i+1}). \end{aligned} \tag{39}$$

In Dragt's original paper on lattice functions,[11] he refers to $\mathbf{A}_i^{-1}$ as the "irritability": In some sense, it gives the true extent of the damage done on the Floquet ring.

As an example of the use of Eq. (39), let us completely normalize $\mathfrak{R}_f^p$ to second order in the perturbation V_i. This has practical application in the design of a large synchrotron ring, where one needs to keep the shear terms resulting from sextupoles under control. This process is schematically displayed in Fig. 2.

To proceed as before, we first compute the one-turn maps:

$$\mathbf{R}_i = \prod_{k=i}^{i+m-1} \exp(:-V_k(\mathbf{A}_k \mathbf{z}_0):)\mathbf{B}_{k\,k+1},$$
$$\mathfrak{I}_f^p = (\mathbf{R}_i). \tag{40}$$

We then isolate the perturbations on the rhs of the factored product of $\mathbf{R}_i$:

$$\mathbf{R}_i = \mathbf{R}\mathbf{R}^{-1}\left(\prod_{k=i}^{i+m-1} \exp(:-V_k(\mathbf{B}_{i\,k}\mathbf{A}_k \mathbf{z}_0):)\right)\mathbf{R},$$
$$\mathbf{R}_i = \mathbf{R}\left(\prod_{k=i}^{i+m-1} \exp(:-V_k(\mathbf{B}_{k\,i+m}{}^{-1}\mathbf{A}_k \mathbf{z}_0):)\right), \tag{41}$$
$$\mathbf{R}_i = \mathbf{R}\Pi_i.$$

To second order in the perturbation, we can factor Π_i:

$$\Pi_i = \exp(:W_{1i}:)\exp(:W_{2i}:)\cdots, \tag{42a}$$
$$W_{1i} = \sum_{k=i}^{i+m-1} -V_k(\mathbf{B}_{k\,i+m}{}^{-1}\mathbf{A}_k \mathbf{z}_0), \tag{42b}$$
$$W_{2i} = \frac{1}{2}\sum_{k=i}^{i+m-1}\sum_{k'=k+1}^{i+m-1}[-V_k(\mathbf{B}_{k\,i+m}{}^{-1}\mathbf{A}_k \mathbf{z}_0), - V_{k'}(\mathbf{B}_{k'i+m}{}^{-1}\mathbf{A}_{k'} \mathbf{z}_0)]. \tag{42c}$$

The expression for W_{2i} is the result of a simple application of the Campbell–Baker–Hausdorff formula.

The normalization starts with the application of $\mathbf{A}_{1i} = \exp(:F_{1i}:)$:

$$\mathbf{A}_{1i}\mathbf{R}_i\mathbf{A}_{1i}{}^{-1} = \mathbf{A}_{1i}\mathbf{R}\Pi_i\mathbf{A}_{1i}{}^{-1},$$
$$F_{1i} = \mathbf{T}_{\mathrm{im}}{}^{-1}W_{1i}, \quad D_{w1}(\mathbf{J}) = (\mathbf{E} - \mathbf{P}_{\mathrm{im}})W_{1i}. \tag{43}$$

In the case of a complete normalization, $D_{w1}(\mathbf{J})$ is often known as the average (or secular) term, which in the usual action-angle representation of (23b) has the form

$$D_{w1}(\mathbf{J}) = \langle W_1 \rangle = \frac{1}{(2\pi)^N}\int_0^{2\pi}\cdots\int_0^{2\pi} W_{1i}(\boldsymbol{\varphi},\mathbf{J})d^N\boldsymbol{\varphi}. \tag{44}$$

The resulting map is given by

$$\mathbf{A}_{1i}\mathbf{R}_i\mathbf{A}_{1i}{}^{-1} = \mathbf{R}\exp(:D_{w1}(\mathbf{J}):)\exp(:W_{2i}^1:),$$
$$W_{2i}^1 = \tfrac{1}{2}[D_{w1},\mathbf{P}_{\mathrm{im}}W_{1i} - 2F_{1i}] + \tfrac{1}{2}[F_{1i},\mathbf{P}_{\mathrm{im}}W_{1i}] + W_{2i}. \tag{45}$$

We now proceed with the second-order calculation:

$$F_{2i} = \mathbf{T}_{\mathrm{im}}{}^{-1}W_{2i}^1$$
$$= \mathbf{T}^{-1}\tfrac{1}{2}[D_{w1},\mathbf{P}_{\mathrm{im}}W_{1i} - 2F_{1i}] + \mathbf{T}_{\mathrm{im}}{}^{-1}(\tfrac{1}{2}[F_{1i},\mathbf{P}_{\mathrm{im}}W_{1i}] + W_{2i}), \tag{46a}$$
$$D_{w2}(\mathbf{J}) = (\mathbf{E} - \mathbf{P}_{\mathrm{im}})(\tfrac{1}{2}[F_{1i},\mathbf{P}_{\mathrm{im}}W_{1i}] + W_{2i}). \tag{46b}$$

The first term of W_{2i}^1 is entirely in the range of the operator $\mathbf{T}$. This completes the second-order normalization process. To second order in the $\mathbf{C}_i$'s, the Floquet ring is given by

$$\mathfrak{R}_f = (\exp(:F_{2k}:)\exp(:F_{1k}:)\mathbf{A}_k\mathbf{C}_k\mathbf{A}_k{}^{-1}\mathbf{B}_{k\,k+1} \times \exp(:-F_{1\,k+1}:)\exp(:-F_{2\,k+1}:)), \tag{47a}$$
$$\mathfrak{I}_f = (\mathbf{R}\exp(:D_{w1}(\mathbf{J}) + D_{w2}(\mathbf{J}):)) = (\exp(:-\boldsymbol{\mu}\cdot\mathbf{J} + D(\mathbf{J}) + D_{w1}(\mathbf{J}) + D_{w_2}(\mathbf{J}):)). \tag{47b}$$

VI. THE LINEAR PHASE ADVANCE

When new techniques are introduced, it is instructive to compare the approach with the old techniques whenever they exist. The difficulty in doing so is proportional to the enhanced power the new methods provide over the old ones. Therefore, while the mathematical equivalence is not in doubt, it is hard to work out a nontrivial and nonlinear example which explicitly displays the mathematical equivalence. Therefore, we will settle for a linear example. The reader with a knowledge of accelerator theory will see here an explicit connection between the two methods by allowing our ring to become an "∞-tuple," i.e., by reverting to the Hamiltonian. Clearly, from Eq. (2a), the Hamiltonian picture corresponds to the maximum ring $\mathfrak{R}_\infty$:

$$\mathfrak{R}_{\rho\infty} = \lim_{\substack{ds\to 0 \\ s\in[0,1]}} (\mathbf{N}_{s\,s+ds}) = \lim_{\substack{ds\to 0 \\ s\in[0,1]}} (\mathbf{E} + ds:-H(\mathbf{x}_0;s):) \tag{48}$$

First, let us state a few well-known results. In the one-dimensional case, where the Hamiltonian is given by

$$H = \tfrac{1}{2}(p^2 + k(s)q^2), \quad \mathbf{x} = (q,p), \tag{49}$$

a stable one-turn map can be parametrized by the so-called Twiss parameters[15]:

$$\mathbf{L}_s = \exp(:-\mu\mathbf{I}:) = \exp(:-\tfrac{1}{2}\mu c:), \tag{50a}$$
$$c = \gamma(s)q_0{}^2 + 2\alpha(s)q_0p_0 + \beta(s)p_0{}^2, \quad 1 + \alpha^2 = \beta\gamma. \tag{50b}$$

It is easy to verify that the matrix representation L_s of $\mathbf{L}_s$ is given by[1,15]

$$L_s = \begin{pmatrix} \cos\mu + \alpha\sin\mu & \beta\sin\mu \\ -\gamma\sin\mu & \cos\mu - \alpha\sin\mu \end{pmatrix}. \tag{51}$$

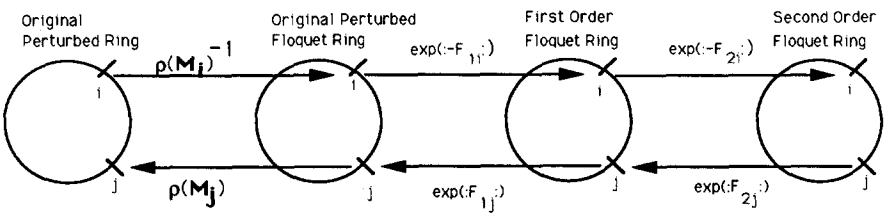

FIG. 2. Schematic view of the second-order normalization process.

The invariant c of Eqs. (50) is called the Courant–Snyder invariant in accelerator physics literature. From our previous discussion it is clear that the Twiss parameters (α,β,γ) are s dependent. Incidently, the maps $\mathbf{L}_s$ define our one-turn ∞-tuple, while H defines the ring

$$\mathfrak{R}_\infty = \lim_{\substack{ds\to 0 \\ s\in[0,1]}} \left(\mathbf{E} + ds{:}-\frac{1}{2}(p_0{}^2 + k(s)q_0{}^2){:}\right), \tag{52a}$$

$$\mathfrak{I}_\infty = (\mathbf{L}_s)_{s\in[0,1]}. \tag{52b}$$

To proceed further, we must define the map $\rho(\mathbf{L}_s)$. Following Courant–Snyder,[1] we define $\rho(\mathbf{L}_s)$ as

$$A_{L_s} = \rho(L_s) = \begin{pmatrix} \sqrt{\beta} & 0 \\ -\alpha/\sqrt{\beta} & 1/\sqrt{\beta} \end{pmatrix}. \tag{53}$$

Equation (53) uniquely defines $\mathbf{A}_{\mathbf{L}_s}$ given a linear map $\mathbf{L}_s$. Given (52) and (53), we will prove, using our concepts, that the Floquet ring and resulting phase advance are given by

$$\begin{aligned}\mathfrak{R}_{\rho\infty} &= \lim_{\substack{ds\to 0 \\ s\in[0,1]}} \left(\mathbf{E} + ds{:}-\frac{1}{2\beta}(p_0{}^2 + q_0{}^2){:}\right) \\ &= \lim_{\substack{ds\to 0 \\ s\in[0,1]}} \left(\mathbf{E} + ds{:}-\frac{1}{\beta}J{:}\right),\end{aligned} \tag{54a}$$

$$\Delta\Phi_{s_0,s_1} = \int_{s_0}^{s_1} \frac{ds}{\beta}. \tag{54b}$$

Before proving a generalization of this result, we point out that the choice of Courant–Snyder was dictated by the kind of perturbation expected in an accelerator. In our machine, we expect the perturbation $\mathbf{C}_i = \exp(:-V_i:)$ of Sec. V to depend mostly on the position vector $\mathbf{q}$ because the leading contribution to the perturbed Hamiltonian is proportional to the longitudinal component of a magnetic vector potential. Therefore, a choice of $\rho(L_s)$ that minimizes the change in the functional form of V_i is best. We can generalize the Courant–Snyder choice to a higher dimensionality. The resulting phase advance formula is given for the Hamiltonian

$$H(\mathbf{x};s) = \frac{1}{2}\sum_{\substack{i=1 \\ j=1}}^{2N} \mathbf{H}_{ij}(s) X_i X_j \tag{55a}$$

and for $A = \rho(L)$ such that

$$A_{2i-1\,2i-1} > 0, \quad A_{2i-1\,2i} = 0, \quad i = 1,N; \tag{55b}$$

by [(55a) and (55b)]

$$\Rightarrow \frac{d\Phi_i}{ds} = \sum_{j=1}^{N} \frac{\mathbf{H}_{2ij}\mathbf{A}_{j2i}}{\mathbf{A}_{2i-1\,2i-1}}. \tag{55c}$$

Proof: We rewrite Eq. (32) for an infinitesimal change in s:

$$\begin{aligned}\mathbf{B}_{\Delta\Phi} &= \exp(:-\Delta\Phi\cdot\mathbf{J}:) \\ &= \mathbf{A}_s(\mathbf{E} + ds{:}H(\mathbf{x}_0;s){:})\mathbf{A}_{s+ds}{}^{-1} \\ &\Rightarrow \mathbf{A}_{s+ds} \\ &= (\mathbf{E} + {:}\Delta\Phi\cdot\mathbf{J}{:})\mathbf{A}_s(\mathbf{E} + ds{:}-H(\mathbf{x}_0;s){:}) \\ &\quad + O(ds^2)\cdots.\end{aligned} \tag{56}$$

Next we assume that $\mathbf{A}_s$ obeys (55b) and we impose on $\mathbf{A}_{s+ds}$ the same condition (j and k are summed over):

$$\begin{aligned}\mathbf{A}_{s+ds}X_{0;2i-1} &= (\mathbf{E} + {:}\Delta\Phi\cdot\mathbf{J}{:})\mathbf{A}_s(X_{0;2i-1} + ds[-H(\mathbf{x}_0;s),X_{0;2i-1}]) \\ &= (\mathbf{E} + {:}\Delta\Phi\cdot\mathbf{J}{:})\mathbf{A}_s(X_{0;2i-1} + ds\,\mathbf{H}_{2ij}X_{0;j}) \\ &= (\mathbf{E} + {:}\Delta\Phi\cdot\mathbf{J}{:})(A_{2i-1\,k}X_{0;k} + ds\,\mathbf{H}_{2ij}A_{jk}X_{0;k}) \\ &= (A_{2i-1\,k}X_{0;k} + (-\Delta\Phi_k A_{2i-1\,2k-1}X_{0;2k} + \Delta\Phi_k A_{2i-1\,2k}X_{0;2k-1} + ds\,\mathbf{H}_{2ij}A_{jk}X_{0;k})).\end{aligned} \tag{57}$$

We extract from (57) the $(2i-1,2i)$ component of A_{s+ds} and set it to zero:

$$0 = ds\,\mathbf{H}_{2ij}A_{j2i} - \Delta\Phi_i A_{2i-1\,2i-1}$$

$$\Rightarrow \text{property (55c)}. \tag{58}$$

Q.E.D.

We can apply formula (58) to a problem already solved by Edwards and Teng[25], where they accidently chose the same definition for the transformation $\rho(L)$.

In Edwards and Teng's case, the Hamiltonian matrix $\mathbf{H}_{ij}$ was

$$\mathbf{H} = \begin{pmatrix} F & 0 & K & -L \\ 0 & 1 & L & 0 \\ K & L & G & 0 \\ -L & 0 & 0 & 1 \end{pmatrix}; \tag{59}$$

they parametrized the matrix A as

$$A = BC, \quad B = \begin{pmatrix} \mathbf{I}\cos\phi & \mathbf{D}^{-1}\sin\phi \\ -\mathbf{D}\sin\phi & \mathbf{I}\cos\phi \end{pmatrix}, \tag{60a}$$

$$\cos(\phi) > 0, \quad \mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad |\mathbf{D}| = 1, \tag{60b}$$

$$C = \begin{pmatrix} \mathbf{b}_1 & 0 & 0 \\ & 0 & 0 \\ 0 & 0 & \\ 0 & 0 & \mathbf{b}_2 \end{pmatrix}, \mathbf{b}_i = \begin{pmatrix} \sqrt{\beta_i} & 0 \\ -\alpha_i/\sqrt{\beta_i} & 1/\sqrt{\beta_i} \end{pmatrix}. \tag{60c}$$

Because the map A obeys (55b), we can compute the phase advance using (55c):

$$\frac{d\Phi_1}{ds} = \frac{(1 - Lb\tan(\phi))}{\beta_1}, \tag{61a}$$

$$\frac{d\Phi_2}{ds} = \frac{(1 + Lb\tan(\phi))}{\beta_2}. \tag{61b}$$

Results (61) are exactly those obtained by Edwards and Teng through a totally different method.[25]

In a final exercise, we would like to explore the relationship between two definitions of the phase advance. Going back to the original one-dimensional problem of Courant–Snyder,[1] we can examine the following definition:

$$A_{L_s} = \rho'(L_s) = \begin{pmatrix} 1/\sqrt{\gamma} & -\alpha/\sqrt{\gamma} \\ 0 & \sqrt{\gamma} \end{pmatrix}. \tag{62}$$

If the world was made out of primarily velocity-dependent potentials, Courant and Snyder would have selected ρ'. The phase advance is obtained by a symmetry argument [canonically exchanging q and p and applying (55c)]:

$$\Delta\Phi_{s_0 s_1} = \int_{s_0}^{s_1} \frac{k(s)ds}{\gamma}. \tag{63}$$

Using the ring given by (48) and the underlying Hamiltonian given by (49), we can derive a famous set of rules for the evolution of the Twiss parameters (α,β,γ). This will allow us to relate the phase advances of ρ and and ρ' explicitly.

$$\begin{aligned}
\mathbf{L}_{s+ds} &= \mathbf{N}^{-1}_{s\,s+ds}\mathbf{L}_s\mathbf{N}_{s\,s+ds} \\
&= \mathbf{N}^{-1}_{s\,s+ds}\exp(:-\tfrac{1}{2}\mu c:)\mathbf{N}_{s\,s+ds} = \exp(:-\tfrac{1}{2}\mu\mathbf{N}^{-1}_{s\,s+ds}c:) \\
&= \exp(:-\tfrac{1}{2}\mu(\mathbf{E}-ds:-\tfrac{1}{2}(p_0^2+k(s)q_0^2):)(\gamma(s)q_0^2+2\alpha(s)q_0p_0+\beta(s)\rho_0^2):) \\
&= \exp(:-\tfrac{1}{2}\mu(c-ds[-\tfrac{1}{2}(p_0^2+k(s)q_0^2),\gamma(s)q_0^2+2\alpha(s)q_0p_0+\beta(s)p_0^2]):) \\
&\Rightarrow \frac{d\alpha}{ds} = \beta k - \gamma, \quad \frac{d\beta}{ds} = -2\alpha, \quad \frac{d\gamma}{ds} = 2\alpha k.
\end{aligned} \tag{64}$$

Using the relations of (64) and the relation $1+\alpha^2=\beta\gamma$, we obtain

$$\begin{aligned}
\Delta\Phi_{s_0 s_1} &= \int_{s_0}^{s_1}\frac{k(s)ds}{\gamma} = \int_{s_0}^{s_1}\frac{\gamma+d\alpha/ds}{\beta\gamma}ds \\
&= \int_{s_0}^{s_1}\frac{1}{\beta}ds + \tan^{-1}(\alpha)\Big|_{s=s_0}^{s=s_1}.
\end{aligned} \tag{65}$$

In Eq. (65) we explicitly demonstrated that matched $(\alpha_1=\alpha_2)$ locations are separated by the same phase advance. If we then multiply ρ'^{-1} by ρ [Eqs. (53) and (62)], we obtain another advertised result:

$$\begin{aligned}
&\begin{pmatrix} \sqrt{\gamma} & \alpha/\sqrt{\gamma} \\ 0 & 1/\sqrt{\gamma} \end{pmatrix}\begin{pmatrix} \sqrt{\beta} & 0 \\ -\alpha/\sqrt{\beta} & 1/\sqrt{\beta} \end{pmatrix} \\
&\quad = \begin{pmatrix} 1/\sqrt{1+\alpha^2} & \alpha/\sqrt{1+\alpha^2} \\ -\alpha/\sqrt{1+\alpha^2} & 1/\sqrt{1+\alpha^2} \end{pmatrix}.
\end{aligned} \tag{66}$$

As we stated previously, the matrix in (66) is a rotation; two definitions of ρ can only differ by a phase. In fact, the angle of the rotation in (66) is $\tan^{-1}(\alpha)$, in perfect agreement with (65).

We close the present discussion with a remark on Dragt's original definition of the phase advance. In Ref. 11, Dragt defined the canonical transformation ρ in terms of maps. However, Dragt artificially introduced the timelike variable s of the original Hamiltonian in the definition of ρ. As a result, it was not true that

$$\mathbf{M}_i = \mathbf{M}_j \Rightarrow \rho(\mathbf{M}_i) = \rho(\mathbf{M}_j). \tag{67}$$

Although it is conceivable to imagine cases where (67) should be discarded on the basis of connecting two different types of perturbations, it is unacceptable to do so at random using the Hamiltonian parameter s. In that sense, Dragt's treatment was not totally Hamiltonian-free.

VII. CONCLUSION

We would like to summarize the actual achievements of the Hamiltonian-free theory. First, as we emphasized throughout this paper, our apporach goes directly to the quantities of interest; this greatly simplifies the theory for any representation of the map.

Second, our ability to generate and analyze Taylor series maps allows us to study arbitrarily complex systems, in particular, circular accelerators. One can define a Floquet ring and perturb it by Hamiltonian and/or stochastic effects. For example, one can easily implement the stochastic calculation of the final emittances proposed by Chao in *any* tracking code.[6] We are no longer restricted to simple models: This could become important in understanding the behavior of small light sources because of the nontrivial fringe fields they generate.

Finally, other areas of physics could benefit from such an approach. For example, in the design of toroidal stellarators, one can show that the magnetic field line pattern is (in some variables) a two-dimensional symplectic map. The computation of this map is extremely complex and tedious since one must integrate the Biot–Savart law around the stellarator. Hanson and Cary, in a paper on the stochastic nature of this map,[26] did exactly that: Had they known of the automatic differentiation of Berz,[13] they could have attempted to compute a one-turn map with some dependence on the current parameters they used to reduce the stochasticity. In fact, the stellarator problem seems to typify a proper use of a map-based theory: the map is simple (two-dimensional), but the Hamiltonian generating it is extremely complex (i.e., Maxwell's equations). In addition, the field lines are best integrated using non-Hamiltonian variables. One can convert the two-dimensional map into canonical variables at the end of the calculation, just before feeding it into some canonical perturbation theory algorithm.

By this example, we just wanted to point out the generality of certain concepts. Since not all problems are identical, we are convinced that the greater the selections of tools, the more efficiently a researcher or designer can attack a complex problem.

VIII. PROSPECT FOR THE FUTURE

We began this paper by pointing out that the essential problem of accelerator dynamics is to study the long-

term stability of the one-turn map. In fact, from a strict analysis of error propagation, accelerator simulation integrates the motion of particles far beyond a rigorously reasonable limit. Given this fact, can we then use a one-turn map in our simulation? Experience has shown that truncated Taylor series produced nonsymplectic maps with vastly different long-term behavior. Indeed, the motion can settle on a fixed point in phase space after a relatively short number of turns, despite a highly accurate Taylor series representation. However, with our ability to extract maps and manipulate them, we can reexpress the Taylor series representation into various exactly symplectic representations. This is being extensively studied at the moment, driven by projects such as third-generation synchrotron light sources, small "pocket" light sources and large hadron rings such as the contemplated Superconducting Super Collider (SSC).

We are also trying to understand quasisymplectic maps. For example, in light sources and so-called "beauty factories," electrons radiate a substantial amount of energy. In the classical regime, this leads to a nonlinear map with damping. This map can be easily extracted with automatic differentiation techniques, but its analysis in the nonlinear regime will require new developments beyond those advertised in this paper. In particular, it will not be possible to express the one-turn maps using symplectic Lie generators, but it is hoped that a new expanded set of Lie generators can be found.

These problems and others are now within reach thanks to the type of rethinking introduced in this paper.

ACKNOWLEDGMENTS

I would like to thank Professor Alex Dragt, who has been my mentor and has initiated this vastly expanded usage of maps in accelerator theory. More recently, I had the pleasure to collaborate with John Irwin of SLAC. In addition, I am extremely grateful to Martin Berz for developing the Differential Algebra Package and attaching to it extra subroutines to ease the development of map normalization algorithms. Berz's work on automatic differentiation reinforced the seriousness and generality of the map approach to accelerator design: The usefulness of his methods goes far beyond the domain of dynamical systems. Finally, I am also grateful to my co-workers and, in particular, to my supervisor Dr. Swapan Chattopadhyay for providing the necessary support.

This work was supported by the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

APPENDIX A: A MAP DERIVATION OF THE NEW HAMILTONIAN USING THE ADJOINT REPRESENTATION OF THE LIE OPERATOR ALGEBRA

Assuming that we are interested in all surfaces of sections, the ring becomes an "∞-tuple," as described in Eqs. (2), (8), and (9). In terms of maps, (2), (8), and (9) take the form

$$\frac{d}{d\epsilon}\mathbf{A}^{-1}(\mathbf{z}_0;s;\epsilon)=\mathbf{A}^{-1}(\mathbf{z}_0;s;\epsilon):-w(\mathbf{z}_0;s;\epsilon):, \tag{A1a}$$

$$\frac{d}{ds}\mathbf{M}(\mathbf{z}_0;s)=\mathbf{M}(\mathbf{z}_0;s):-K(\mathbf{z}_0;s):, \tag{A1b}$$

$$K(\mathbf{z}_0;s)=\mathbf{A}(\mathbf{z}_0;s)\Big(H(\mathbf{z}_0;s)+\int_0^{\epsilon=1}d\epsilon\,\mathbf{A}^{-1}(\mathbf{z}_0;s)\frac{\partial}{\partial s}w(z_0;s;\epsilon)\Big), \tag{A1c}$$

$$\mathbf{M}(\mathbf{z}_0;s=s_0)=\mathbf{E}=\text{identity}. \tag{A1d}$$

Equations (A1) must be integrated from s_0 to s_0+1 if one is to obtain a one-turn map. For a complete normalization, (A1) must be accompanied by the boundary conditions

$$K(\mathbf{z}_0;s)=\frac{\partial}{\partial s}\Phi_s(\mathbf{J}), \tag{A2a}$$

$$\mathbf{A}^{-1}(\mathbf{z}_0;s;\epsilon=1)=\mathbf{A}^{-1}(\mathbf{z}_0;s+1;\epsilon=1). \tag{A2b}$$

The periodicity of (A2b) and the normalized form imposed on K in (A2a) renders the direct solution of Eqs. (A1) and (A2) very difficult unless the original Hamiltonian H is simple. The process involves the computation of various Green's functions upon which the proper boundary conditions are imposed (see Ref. 5).

In this Appendix, we will concentrate on Eq. (A1c) (first obtained by Dewar[18] and then Cary[19]): It can be derived very elegantly using homomorphic Lie algebras. Let us denote by $\mathbf{N}$ the map generated by H:

$$\frac{d}{ds}\mathbf{N}(\mathbf{z}_0;s)=\mathbf{N}(\mathbf{z}_0;s):-H(\mathbf{z}_0;s):. \tag{A3}$$

Using Lie properties of maps, we can write $\mathbf{M}(\mathbf{z}_0;s)$ as

$$\mathbf{M}_s=\mathbf{A}_0\mathbf{N}_s\mathbf{A}_s^{-1}. \tag{A4}$$

In Eq. (A4) all the maps transform functions of the initial coordinates $\mathbf{z}_0$. First, $\mathbf{A}_0(\mathbf{z}_0)$ brings us to the original variables; these are propagated with the help of $\mathbf{N}_s(\mathbf{z}_0,s_0)$ and are finally taken back into the new variables by $\mathbf{A}_s(\mathbf{z}_0)^{-1}$. To obtain the new Hamiltonian K, we take the derivative of (A4) with respect to s:

$$\begin{aligned}\frac{d}{ds}\mathbf{M}_s&=\mathbf{A}_0\Big(\frac{d}{ds}\mathbf{N}_s\Big)\mathbf{A}_s^{-1}+\mathbf{A}_0\mathbf{N}_s\Big(\frac{d}{ds}\mathbf{A}_s^{-1}\Big)\\&=\mathbf{A}_0\mathbf{N}_s\mathbf{A}_s^{-1}\mathbf{A}_s\\&\quad\times:-H:\mathbf{A}_s^{-1}+\mathbf{A}_0\mathbf{N}_s\mathbf{A}_s^{-1}\mathbf{A}_s\Big(\frac{d}{ds}\mathbf{A}_s^{-1}\Big)\\&=\mathbf{M}_s\Big(:-\mathbf{A}_sH:+\mathbf{A}_s\Big(\frac{d}{ds}\mathbf{A}_s^{-1}\Big)\Big)\ .\end{aligned} \tag{A5}$$

Comparing (A5) with (A1b), we conclude that $\mathbf{A}_s((d/ds)\mathbf{A}_s^{-1})$ must be a Lie operator: To evaluate it, we use the equation of $\mathbf{A}_s$ and $\mathbf{A}_s^{-1}$:

$$\begin{aligned}&\frac{d}{d\epsilon}\mathbf{A}_s\Big(\frac{d}{ds}\mathbf{A}_s^{-1}\Big)\\&=:w:\mathbf{A}_s\Big(\frac{d}{ds}\mathbf{A}_s^{-1}\Big)+\mathbf{A}_s\Big(\frac{d}{ds}\mathbf{A}_s^{-1}:-w:\Big)\\&=\Big\{:w:,\mathbf{A}_s\Big(\frac{d}{ds}\mathbf{A}_s^{-1}\Big)\Big\}+:-\frac{\partial}{\partial s}w:,\end{aligned}$$

$$\frac{\partial}{\partial s} w = \frac{d}{ds} w. \tag{A6}$$

Here $\{\ ,\ \}$ denotes the commutator of two Lie operators. Denoting by $\mathbf{G}$ the operator $\mathbf{A}_s((d/ds)\mathbf{A}_s^{-1})$, we rewrite (A6) as

$$\frac{d}{d\epsilon}\mathbf{G} - \#w\#\mathbf{G} = :-\frac{\partial}{\partial s} w:. \tag{A7}$$

Here $\#w\#$ is a super operator that acts on the space of Lie transforms by taking a commutator. These commutators form a Lie algebra. Notice the homomorphism between the Lie algebra of super operators, Lie operators, and Poisson brackets[15]:

$$\#f\#:g: = \{:f:,:g:\} = :[f,g]:. \tag{A8}$$

To solve (A7), we make use of (A8) by writing $\mathbf{G}$ in terms of an ϵ-dependent super operator:

$$\mathbf{G}(\epsilon) = \mathbf{P}\mathbf{G}_0. \tag{A9}$$

We first solve the homogeneous equation

$$\frac{d}{d\epsilon}\mathbf{P} - \#w\#\mathbf{P} = 0 \Rightarrow \mathbf{P} = \mathbf{A}(\#w(\epsilon)\#). \tag{A10}$$

To obtain $\mathbf{P}$ in (A10), we notice that the formal functional dependence of $\mathbf{P}$ on the super operator $\#w\#$ must be the same as the dependence of $\mathbf{A}$ on Lie operator $:w:$; similarly, $\mathbf{P}^{-1}$ must have the same functional dependence on $\# - w\#$ as $\mathbf{A}^{-1}$ has on $: - w:$.

To solve the nonhomogeneous equation, we allow $\mathbf{G}_0$ to depend on ϵ (variation of parameters). For the particular solution $\mathbf{G}_p$ we obtain

$$\mathbf{G}_p = \mathbf{A}(\#w(\epsilon)\#)\int_0^\epsilon d\epsilon'\, \mathbf{A}^{-1}(\# - w(\epsilon')\#) :-\frac{\partial}{\partial s} w(\epsilon'):. \tag{A11}$$

The general solution is the sum of the homogeneous and particular solutions:

$$\mathbf{G} = \mathbf{A}(\#w(\epsilon)\#)\left(\mathbf{G}_0 + \int_0^\epsilon d\epsilon'\, \mathbf{A}^{-1}(\# - w(\epsilon')\#) :-\frac{\partial}{\partial s} w(\epsilon'):\right). \tag{A12}$$

We impose the boundary condition at $\epsilon = 0$:

$$(\mathbf{G}(\epsilon = 0) = \mathbf{0} \Rightarrow \mathbf{G}_0 = 0);$$

therefore,

$$\mathbf{G} = \mathbf{A}(\#w(\epsilon)\#)\int_0^\epsilon d\epsilon'\, \mathbf{A}^{-1}(\# - w(\epsilon')\#): -\frac{\partial}{\partial s} w(\epsilon'):. \tag{A13}$$

Finally, here the homomorphism enters between the three Lie algebras of (A8):

$$\mathbf{A}(\#w(\epsilon)\#)\int_0^\epsilon d\epsilon'\, \mathbf{A}^{-1}(\# - w(\epsilon')\#): -\frac{\partial}{\partial s} w(\epsilon'): = :-\mathbf{A}(:w(\epsilon):)\int_0^\epsilon d\epsilon'\, \mathbf{A}^{-1}(: - w(\epsilon'):)\frac{\partial}{\partial s} w(\epsilon'):. \tag{A14}$$

Substitution of (A14) into (A5) gives the advertised result.

APPENDIX B: THE UNIQUENESS OF R AND THE PHASE ADVANCE

We first prove that given a map $\mathbf{M}$, the assumed analyticity of the similarity transformation insures the uniqueness of $\mathbf{R}$.

We start by postulating the existence of two normalized rotations:

$$\mathbf{A}_1\mathbf{M}\mathbf{A}_1^{-1} = \mathbf{R}_1, \quad \mathbf{A}_2\mathbf{M}\mathbf{A}_2^{-1} = \mathbf{R}_2, \tag{B1}$$

which in turn imply that

$$\mathbf{A}_2\mathbf{A}_1^{-1}\mathbf{R}_1\mathbf{A}_1\mathbf{A}_2^{-1} = \mathbf{R}_2. \tag{B2}$$

Equation (B2) is a generalization of Eq. (33) for the phase advance:

$$\exp(:\mathbf{A}_2\mathbf{A}_1^{-1}(-\boldsymbol{\mu}_1\cdot\mathbf{J} + D_1(\mathbf{J})):) = \exp(:-\boldsymbol{\mu}_2\cdot\mathbf{J} + D_2(\mathbf{J}):). \tag{B3}$$

Using analyticity, we follow Dragt and Finn[16] by factorizing $\mathbf{A}_2\mathbf{A}_1^{-1}$:

$$\mathbf{A}_2\mathbf{A}_1^{-1} = \cdots\mathbf{r}_k\mathbf{r}_{k-1}\cdots\mathbf{r}_1\mathbf{r}_L, \tag{B4}$$

where $\cdots\mathbf{r}_k = \exp(:p_{k+2}:)$ and p_{k+2} is a homogeneous polynomial of degree $k+2$ in the phase space variables.

Except for a mere relabeling of the planes, let us assume that the uniqueness of $\mathbf{R}$ is true in the linear regime (the proof would be quite different for linear maps and amounts to the uniqueness of eigenvalues!). Then (B3) takes the form

$$\exp(:\cdots\mathbf{r}_k\mathbf{r}_{k-1}\cdots\mathbf{r}_1(-\boldsymbol{\mu}\cdot\mathbf{J} + D_1(\mathbf{J})):) = \exp(:-\boldsymbol{\mu}\cdot\mathbf{J} + D_2(\mathbf{J}):). \tag{B5}$$

To go further, we proceed by induction. Assuming that for $k < j-1$, the $\mathbf{r}_k$'s are rotations, we collect the terms of degree $j+2$ and obtain an equation for p_{j+2}:

$$:\boldsymbol{\mu}\cdot\mathbf{J}:p_{j+2} = D_2(\mathbf{J}) - D_1(\mathbf{J})|_{j+2\text{ component}}. \tag{B6}$$

Using the direct sum decomposition (or the eigenbasis) and the mutual irrationality of the tunes [Eqs. (22), (23), and (25)], we conclude from (B6) that p_{j+2} cannot contain anything from $\mathrm{Im}(:\boldsymbol{\mu}\cdot\mathbf{J}:)$. In addition, since the lhs of (B6) must be in $\mathrm{Ker}(:\boldsymbol{\mu}\cdot\mathbf{J}:)$, the only consistent solution to (B6) is

$$p_{j+2} \in \mathrm{Ker}(:\boldsymbol{\mu}\cdot\mathbf{J}:), \quad D_2(\mathbf{J}) - D_1(\mathbf{J})|_{j+2\text{ component}} = 0. \tag{B7}$$

Q.E.D.

The rest follows by induction, starting with $j = 1$.

Equation (B7) also proves the statement on the phase advance because it is a special case of (B1).

Finally, by generalizing Eq. (14) to the nonlinear maps, we can easily see that $\mathbf{R}$ does not depend on the location:

$$\mathbf{M}_j = \mathbf{N}_{ij}^{-1}\mathbf{M}_i\mathbf{N}_{ij}, \text{ where } \mathbf{N}_{ij} = \prod_{k=i}^{} \mathbf{N}_{k\,k+1}, \tag{B8a}$$

$$\mathbf{A}_i\mathbf{M}_i\mathbf{A}_i^{-1} = \mathbf{R}_i, \quad \mathbf{A}_j\mathbf{M}_j\mathbf{A}_j^{-1} = \mathbf{R}_j, \tag{B8b}$$

$$(\text{B8a}) \text{ and } (\text{B8b}) \Rightarrow \mathbf{A}_j\mathbf{N}_{ij}^{-1}\mathbf{M}_i\mathbf{N}_{ij}\mathbf{A}_j^{-1} = \mathbf{R}_j. \tag{B8c}$$

Equations (B8a) and (B8b) violate the uniqueness of $\mathbf{R}$ for a given map $\mathbf{M}$.

[1]E. D. Courant and H. S. Snyder, Ann. Phys. (NY) 3, 1 (1958).
[2]H. Meier and K. R. Symon, in *Proceedings of the International Conference*

on High Energy Accelerators and Instrumentation, Cern, edited by L. Kowarski (CERN, Geneva, 1959). See, also, L. J. Laslett, E. M. MacMillan, and J. Moser, *Long Term Stability for Particle Orbits* (Courant Institute of Mathematical Science, New York Univ. NYO-1480-101, 1968): In this reference, one finds some comments by Moser on the Birkhoff normalization of the Hamiltonian.

[3]The list of kick codes is endless. For example, R. Talman and L. Schachinger, Part. Accel. **22**, 35 (1987) or A. Wrulich, DESY Internal Rep. 84-026 (1984).

[4]A. Schoch, CERN 57-21 (1958).

[5]L. Michelotti, Part. Accel. **16**, 233 (1985).

[6]A. Chao, J. Appl. Phys. **50**, 595 (1979).

[7]In single-pass systems such as spectrometers, the extraction of Taylor series maps is a time honored practice. Indeed, the aberrations of interest are directly the coefficients of the Taylor series. Again, the list of power series codes is endless.

[8]A. J. Dragt and J. M. Finn, J. Math. Phys. **20**, 2649 (1979).

[9]A. J. Dragt and J. M. Finn, J. Geophys. Res. **81**, 13 (1976).

[10]A. J. Dragt *et al.*, Ann. Rev. Nucl. Sci. **38**, 455 (1988).

[11]A. J. Dragt, in *Proceedings of the 1984 Summer Study on the Design and Utilization of the Superconducting Super Collider, Snowmass*, edited by R. Donaldson and J. Morfin (Am. Phys. Soc., New York, 1985).

[12]E. Forest, Part. Accel. **22**, 15 (1987).

[13]M. Berz, Part. Accel. **24**, 109 (1989).

[14]E. Forest, J. Irwin, and M. Berz, Part. Accel. **24**, 91 (1989).

[15]A. J. Dragt, *Lectures on Nonlinear Orbit Dynamics*, American Institute of Physics Conference Proceedings No. 87, edited by R. A. Carrigan, F. R. Huson, and M. Month (AIP, New York, 1982).

[16]A. J. Dragt and J. M. Finn, J. Math. Phys. **17**, 2215 (1976).

[17]A. J. Dragt and E. Forest, J. Math. Phys. **24**, 2734 (1983).

[18]R. L. Dewar, J. Phys. A **9**, 2043 (1976).

[19]J. Cary, Phys. Rep. **79**, 129 (1981).

[20]R. D. Ruth, IEEE Trans. Nucl. Sci. **NS-30**, 2669 (1983). In addition, Ruth and the present author are preparing a paper on symplectic integration and its connection with the Lie representation of maps near the identity.

[21]Because we use symplectic integrators, the maps we extract are Hamiltonian in nature to computer accuracy. The theory presented here still applies for nonsymplectic integrators. Obviously, the violation of "symplecticity" is of the same order as the error due to the finite size of the integration steps.

[22]R. L. Warnock, R. D. Ruth, W. Gabella, and K. Ecklund, Internal Rep. SLAC-PUB -4846 (1989).

[23]The normalization of a one-turn Taylor series map using the Differential Algebra Package of Berz is explained in detail in Ref. 14. Here we summarize the results, deemphasizing the Taylor series representation. One must realize that even in other representations, normalization procedures are formal manipulations. Most problems are not integrable and, therefore, the expressions for the invariants are asymptotic series in some smallness parameter.

[24]R. Cushman, A. Deprit, and R. Mosak, J. Math. Phys. **24**, 2102 (1983).

[25]D. A. Edwards and L. C. Teng, IEEE Trans. Nucl. Sci. **NS-20**, 885 (1983).

[26]J. D. Hanson and J. R. Cary, Phys. Fluids **27**, 767 (1984).