Dispersive lattice functions in a six-dimensional pseudo-harmonic-oscillator

Étienne Forest
National Laboratory for High Energy Physics (KEK), 1-1 Oho, Tsukuba, Ibaraki 305, Japan
(Received 9 September 1997; revised manuscript received 11 February 1998)

We derive dispersivelike lattice functions in a way totally invariant under canonical transformation. This bridges the gap between invariant treatments that use only the coefficients of the coupled Courant-Snyder invariants as lattice functions and treatments that introduce dispersive lattices functions that depend on particular parametrizations. [S1063-651X(98)12907-0]

PACS number(s): 29.20.Dh, 42.15.Eq

I. INTRODUCTION

In this paper I would like to use certain symmetries present in a periodic system in an attempt to classify the types of lattice functions that can be defined in the case of a linear oscillatory map. The main result of this paper concerns the existence of ''dispersive'' lattice functions when all the planes are oscillating. Dispersion is a mathematically well defined concept when the energy is constant (no cavity and no radiation); however, it does not seem to exist in a three-dimensional pseudo-harmonic-oscillator. In this paper I define dispersive lattices functions that are invariant under the choice of canonical transformations. In the symplectic case the invariance is connected to ergodic averages, which can be defined ''experimentally'' and thus must be invariant under the theoretical technique used to compute them. I show, as it is well known, that ergodic averages of quadratic monomials are related to the usual lattices functions (Twiss parameters in one-dimension) while stroboscopic (or adiabatic) averages are related to dispersive quantities.

Finally, I express the one-turn matrix in terms of these lattice functions; the natural appearance of the dispersive lattice functions in such a parametrization explains why ''Courant-Snyder–like'' parametrizations [1] of the matrix in terms of lattice functions are not found in the literature (see the one-turn map of Ref. [2]) in more than one degree of freedom. Nevertheless, I succeed in expressing the present dispersive functions entirely in terms of the old Courant-Snyder parameters, even in the general case of the damped (nonsymplectic, radiative) pseudo-harmonic-oscillator relevant to electron rings.

II. DIAGONALIZATION AND INVARIANTS

In a periodic or a repetitive symplectic system such as a ring, it is normal to ask questions concerning the ''at infinity behavior.'' Are particles confined and if so on what trajectories do they sit? Therefore, one finds that many averages over distributions are closely related to ergodic averages over a single trajectory. This is true at least for the symplectic system. Indeed, a tracking code will display ellipses or Lissajous figures in phase space. A knowledge of the parametrization of the surfaces provides us with the ''infinite time'' behavior. Clearly, whatever at infinity property a trajectory has, it is invariant under initial conditions chosen on this trajectory. Any mathematical attempt to compute this trajectory will lead to invariant functions.

The symplectic or Hamiltonian case is easiest to understand and has this physical interpretation based on ergodic averages. Therefore, I will discuss it first. A more dry approach will be introduced later to prove the invariance of these lattice functions in the nonsymplectic case.

Let us assume that the one-turn matrix M for a ring is symplectic (derivable from a Hamiltonian). Then this implies that in a judicious choice of coordinates the matrix M and its transposed $\tilde{M}$ must obey

$$J = MJ\tilde{M} \tag{1}$$

where

$$J = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix}.$$

We then assume that the motion produced by M is pseudo-harmonic. This is a fancy way of saying that the matrix M can be diagonalized as

$$M = ARA^{-1}, \tag{2}$$

where A, it turns out, can be a symplectic matrix and R is a rotation:

$$R = \begin{pmatrix} r_1 & 0 & 0 \\ 0 & r_2 & 0 \\ 0 & 0 & r_3 \end{pmatrix},$$

$$r_i = \begin{pmatrix} \cos\mu_i & \sin\mu_i \\ -\sin\mu_i & \cos\mu_i \end{pmatrix}. \tag{3}$$

The angles of the rotation, known as the tunes, are certainly unique modulo 2π, but the matrix A is not unique. This can be seen by adding a rotation r to A:

$$\text{if } M=ARA^{-1}\Rightarrow M=\underbrace{Ar}_{B}\,R\,\underbrace{r^{-1}A^{-1}}_{B^{-1}}. \tag{4}$$

Thus we have a certain freedom in choosing A. The fact that A may vary at most by a rotation (provided A is restricted to symplectic, i.e., canonical, matrices) implies that the radii of the new trajectory are invariants as well. Let us compute one of these radii. If a particle has initial conditions $\vec{z}_0=(x,p_x,y,p_y,t,p_t)$, then in normalized variables it will have the initial conditions

$$A^{-1}\vec{z}_0=A^{-1}\begin{pmatrix} x\\ p_x\\ \vdots\end{pmatrix}=\begin{pmatrix} A_{11}^{-1}x+A_{12}^{-1}p_x+A_{13}^{-1}y+A_{14}^{-1}p_y+A_{15}^{-1}t+A_{16}^{-1}p_t\\ A_{21}^{-1}x+A_{22}^{-1}p_x+A_{23}^{-1}y+A_{24}^{-1}p_y+A_{25}^{-1}t+A_{26}^{-1}p_t\\ \vdots\end{pmatrix}, \tag{5}$$

which, we want to emphasize, are not unique. However, the radii are unique and characterize a trajectory. Denoting the square of the radius in the first plane by ε_1, in two degrees of freedom it is given by

$$\begin{aligned}\varepsilon_1(\vec{z})&=(A_{11}^{-1}x+A_{12}^{-1}p_x+A_{13}^{-1}y+A_{14}^{-1}p_y)^2+(A_{21}^{-1}x+A_{22}^{-1}p_x+A_{23}^{-1}y+A_{24}^{-1}p_y)^2\\ &=\{(A_{11}^{-1})^2+(A_{21}^{-1})^2\}x^2+\{(A_{12}^{-1})^2+(A_{22}^{-1})^2\}p_x^2+2\{A_{11}^{-1}A_{12}^{-1}+A_{21}^{-1}A_{22}^{-1}\}xp_x+\{(A_{13}^{-1})^2+(A_{23}^{-1})^2\}y^2\\ &\quad+\{(A_{14}^{-1})^2+(A_{24}^{-1})^2\}p_y^2+2\{A_{11}^{-1}A_{13}^{-1}+A_{21}^{-1}A_{23}^{-1}\}xy+2\{A_{11}^{-1}A_{14}^{-1}+A_{21}^{-1}A_{24}^{-1}\}xp_y+2\{A_{12}^{-1}A_{13}^{-1}+A_{22}^{-1}A_{23}^{-1}\}p_xy\\ &\quad+2\{A_{12}^{-1}A_{14}^{-1}+A_{22}^{-1}A_{24}^{-1}\}p_xp_y+2\{A_{13}^{-1}A_{14}^{-1}+A_{23}^{-1}A_{24}^{-1}\}yp_y.\end{aligned} \tag{6}$$

In one degree of freedom this reduces to the usual Courant-Snyder invariant

$$\varepsilon=\gamma x^2+\beta p^2+2\alpha xp, \tag{7}$$

where

$$\gamma=(A_{11}^{-1})^2+(A_{21}^{-1})^2,$$

$$\alpha=A_{11}^{-1}A_{12}^{-1}+A_{21}^{-1}A_{22}^{-1},$$

$$\beta=(A_{12}^{-1})^2+(A_{22}^{-1})^2.$$

The coefficients of this invariant as well as the multidimensional equivalents must themselves be invariant under the choice of A^{-1}. In other words, if a matrix $B^{-1}=r^{-1}A^{-1}$ as in Eq. (4) is used to define the functions ε_i, these ε_i's should be the same as the one defined using A^{-1}. Two polynomial functions are identical if the coefficients multiplying each monomial are the same (monomials form a basis in the vector space of functions). This implies that the coefficients denoted here as α, β, and γ, as well as all the others in Eq. (6), are invariant under a change of the matrix A^{-1}.

In summary, the radii in normalized variables are invariant along the trajectory. The invariance of these functions implies that the coefficients that define them are invariants of the diagonalization process. We should not forget the obvious: The tunes themselves are invariants of the diagonalization process.

We can even say more about these functions if we use a bit more physical intuition. Consider any quantity that is obviously time (or turn) invariant such as the average of a function or the extrema reached by a function. Such a quantity will depend only on the initial value of the invariants defined above. Why? If the averages or extrema exist, then they have to be the same for any point along the trajectory, i.e., they cannot depend on ''time.'' In normalized variables, time is just the action of the matrix R; thus it is not surprising that the invariants have to be made out of ''contractions'' of A or A^{-1} that are invariant under rotation.

For example, in one degree of freedom, it is easy to show that the ergodic averages of x^2, p^2, and xp are given by the formulas (here we assume that the tune is irrational)

$$\langle x^2\rangle=\frac{\beta\varepsilon}{2},\quad \langle p^2\rangle=\frac{\gamma\varepsilon}{2},\quad \langle xp\rangle=-\frac{\alpha\varepsilon}{2}. \tag{8}$$

In conclusion, the so-called lattice functions emerge naturally whenever we examine properties that are invariant under iteration of the map. We will see how it is possible to derive such formulas using the canonical transformation A and the symplectic condition.

III. ERGODIC AVERAGES

In this section I will derive two types of ergodic averages. One is a regular average over the trajectory and the other one is a stroboscopic average. Both will lead to invariants. We start, as we did before, by transforming $\vec{z}$ into normalized space, each subspace characterized by a tune μ_i:

$$\vec{w} = A^{-1}\vec{z} = \sum_i (\underbrace{A^{-1}_{1i}z_i, A^{-1}_{2i}z_i}_{\mu_1}, \underbrace{A^{-1}_{3i}z_i, A^{-1}_{4i}z_i}_{\mu_2}, \underbrace{A^{-1}_{5i}z_i, A^{-1}_{6i}z_i}_{\mu_3}). \tag{9}$$

In this space the trajectories are circles by assumption. Therefore, we can express a trajectory as

$$\vec{w}(n) = R^n A^{-1}\vec{z} = [\sqrt{\varepsilon_1}\cos(n\mu_1+\phi_1), -\sqrt{\varepsilon_1}\sin(n\mu_1+\phi_1), \ldots, \sqrt{\varepsilon_3}\cos(n\mu_3+\phi_3), -\sqrt{\varepsilon_3}\sin(n\mu_3+\phi_3)]. \tag{10}$$

The ray at $n=0$ must correspond to the initial ray of Eq. (9). Both quantities ϕ_i and ε_i can be chosen to satisfy this need. As we have seen, the quantity ε_i is invariant and, in fact, the canonical nature of the original variables implies that the Poisson bracket $[\phi_i, \varepsilon_i]$ is equal to 2. Thus one can identify $J_i = \varepsilon_i/2$ with the usual action variable canonically conjugate to ϕ_i. Let us return to ergodic averages.

A. Regular ergodic averages

We first assume that the three tunes μ_i are prime among each other, i.e., they are not on a resonance. We then reexpress the trajectory in real space $\vec{z}(n)$ in terms of the trajectory in normalized space

$$\vec{z}(n) = A\vec{w}(n). \tag{11}$$

Away from resonances, it is clear that the ergodic average over all three tunes of Eq. (11) will be zero because it amounts to an average of sines and cosines over their respective phases: Thus the linear moments $\langle z_a \rangle$ are null.

The next possibility is to consider the so-called beam envelope $\langle z_a z_b \rangle$ defined by an ergodic average. We can express this ergodic average as

$$\langle z_a z_b \rangle = \left\langle \sum_{i,\sigma} A_{a(2i-\sigma)} w_{2i-\sigma} \sum_{j,\eta} A_{b(2j-\eta)} w_{2j-\eta} \right\rangle, \tag{12}$$

where the latin letters i,j take the value 1, 2, or 3, while the greek letters are either 0 or 1. To proceed further we notice that

$$\langle w_{2i-\sigma} w_{2j-\eta} \rangle = \frac{1}{2}\varepsilon_i \delta_{ij} \delta_{\sigma\eta}, \tag{13}$$

where δ_{ij} and $\delta_{\sigma\eta}$ are Kronecker delta functions, and

$$\begin{aligned} \langle z_a z_b \rangle &= \frac{1}{2}\sum_{i=1,3}\left\{\sum_{\sigma=0,1} A_{a(2i-\sigma)} A_{b(2i-\sigma)}\right\}\varepsilon_i \\ &= \frac{1}{2}\sum_{i=1,3}\{A_{a(2i-1)}A_{b(2i-1)} + A_{a(2i)}A_{b(2i)}\}\varepsilon_i . \end{aligned} \tag{14}$$

Using the symplectic condition, we can rewrite all the above formulas in terms of the inverse of A and thus make a one-to-one connection between the coefficients that define the invariants ε_i and the coefficients of the beam envelope [see Eqs. (33) and (34)].

It should be said that the results of this section are well known. In the case of a distribution of particles they are still valid formulas when the distribution is static, i.e., the phase space dependence of the distribution is a function of the ε_i's: In that case one replaces $\varepsilon_i/2$ by the average of ε_i over the distribution.

We will call the lattice functions of this section betaoids because they appear naturally in the Hamiltonian theory of a pseudo-oscillator. The Lie operator for the one-turn linear map is none other than the invariants ε_i themselves; in fact, the function $\frac{1}{2}\{\mu_1\varepsilon_1 + \mu_2\varepsilon_2 + \mu_3\varepsilon_3\}$ is associated with the Lie operator of the one-turn map and can viewed as a pseudo-Hamiltonian for the matrix M.

B. Stroboscopic or adiabatic ergodic averages

There are other averages that can be built in terms of the matrix A. Their physical meaning is not so obvious. We will look at them in two different ways. First we will take the dispersion route. Our goal is to construct objects that are obviously invariant when the motion in one of the three harmonic plane freezes. The standard dispersion is defined in the absence of a cavity, that is to say, in the absence of longitudinal oscillations. The normal form associated with such a map is different from the pseudoharmonic normal form. In that case we have only two tunes and five distinct eigenvalues. This is because the motion in the longitudinal plane is ''driftlike'' in nature. The energy is a constant (like the momentum in a drift) while the time (or path length) grows proportionally with the energy. This is exactly true in a region of the ring with no dispersion, i.e., the ray $(0,0,0,0,z_5,z_6)$ remains (0,0,0,0) in the transverse planes for all values of the energy z_6. In a dispersive region it can still be true if the map is reexpressed around the energy-dependent fixed point; the derivative of this fixed point with respect to z_6 is the dispersion vector. We will not go into the details of this type of nonoscillatory normal form because it might confuse the reader needlessly. Suffice it to say that this is what happens if there is no longitudinal focusing in a ring: The energy is constant and the transverse closed orbit varies with energy (for example, the cyclotron). That variation is the dispersion.

Returning to our three-dimensional oscillator, we can ask the following question: Under what condition do we see the effect of dispersion in a system without energy conservation? Physically, one should slowly lower the voltage on the rf system until it is zero. As we do this, the main linear effect will be the lowering of the longitudinal tune μ_3 until it is zero. The transverse phase space will move slowly as the longitudinal phase space evolves. The slow sloshing back and forth of the transverse coordinates is closely related to the usual ''cavity-free'' dispersion. We will see that this quantity, which seems to be well defined as an adiabatic limit, is nevertheless an invariant of the diagonalization process for arbitrary tunes.

I will now compute this adiabatic average and argue that it is an invariant using a mathematical and physical argument. Let us start with a ray whose initial condition is

$$\vec{z}=(0,0,0,0,0,z_6) \tag{15}$$

and transform it into normalized space using Eq. (9):

$$\vec{w}=A^{-1}\vec{z}=z_6(\underbrace{A^{-1}_{16},A^{-1}_{26}}_{\mu_1},\underbrace{A^{-1}_{36},A^{-1}_{46}}_{\mu_2},\underbrace{A^{-1}_{56},A^{-1}_{66}}_{\mu_3}). \tag{16}$$

The next step consists in letting the ray of Eq. (16) evolve under the action of the rotation R as in Eq. (10). If we assume that the motion is adiabatic in the third plane, $\mu_1^{-1},\mu_2^{-1}\ll\mu_3^{-1}$, then the average of $\langle\vec{w}\rangle$ over the short time scale of min $(1/\mu_1,1/\mu_2)$ will be given by

$$\langle\vec{w}\rangle_{1,2}=z_6(0,0,0,0,A^{-1}_{56},A^{-1}_{66}). \tag{17}$$

Of course this simply says that in normalized variables the first two planes, on their respective circular trajectories, average to zero before the positions (w_5,w_6) have any time to move and thus are frozen at their initial values. These values are of course dependent on the normal form; however, if we project this ray back into the original physical space we should get the dispersion

$$\langle\vec{z}\rangle_{1,2}=z_6\vec{\eta}=A\langle\vec{w}\rangle_{1,2}=z_6\begin{pmatrix}A_{15}A^{-1}_{56}+A_{16}A^{-1}_{66}\\A_{25}A^{-1}_{56}+A_{26}A^{-1}_{66}\\A_{35}A^{-1}_{56}+A_{36}A^{-1}_{66}\\A_{45}A^{-1}_{56}+A_{46}A^{-1}_{66}\\A_{55}A^{-1}_{56}+A_{56}A^{-1}_{66}\\A_{65}A^{-1}_{56}+A_{66}A^{-1}_{66}\end{pmatrix}. \tag{18}$$

The first four entries must reduce to the cavity-free dispersion in the limit of vanishing μ_3; the fifth and sixth entries are, respectively, zero and one if the map is symplectic and the longitudinal motion is not very dependent on the transverse positions.

It is clear that, in the limit of μ_3 going to zero, the vector created in Eq. (18) cannot depend on the choice of canonical transformation. This is not *a priori* obvious if μ_3 is arbitrary. However, it is true. Before proving this explicitly in the general nonsymplectic case, I would like to argue this on the basis of a gedanken experiment.

First of all, it is clear that one can measure the three tunes μ_1, μ_2, and μ_3 using a turn-Fourier transform of some quantity such as the energy or position. From this one can extract μ_3 with any desired accuracy (theoretically). Second, one can slightly change the machine so that some multiple of μ_3 is a multiple of 2π. Theoretically, this can be done with infinitesimal changes in the machines because rational numbers are dense in the real numbers. Let us assume that indeed $k\mu_3=m2\pi$, where both k and m are integers. We then launch a particle with initial conditions given by Eq. (15) and we observe this ray every k turns and average over the turns. The result will be given by Eq. (18) as well. In this case all the quantities necessary for performing this measurement are measurable, unique, and do not depend on the choice of the transformation A. More importantly, there is nothing required concerning the relative sizes of the three tunes. We only need that the two remaining tunes must be irrational.

Mathematically, the argument is even simpler: One averages around the invariant tori of first and second tunes. While the actual phase of a ray is arbitrary and depends on A, the integral around each torus cannot depend on A but just on the radius, which we know is an invariant in canonical perturbation theory.

The above considerations imply that one could have selected any initial ray and any tune in lieu of $(0,0,0,0,0,z_6)$ and μ_3 and one would still have produced invariant quantities. Therefore, we define the stroboscopic invariants

$$\eta^i_{jk}=A^{-1}_{(2i-1)j}A_{k(2i-1)}+A^{-1}_{(2i)j}A_{k(2i)}. \tag{19}$$

The dispersion of Eq. (18) is a special case of Eq. (19). I call these functions etaoids because they are dispersive in nature in the adiabatic limit or stroboscopic interpretation. The regular lattice functions of Eq. (14)

$$\sum_{\sigma=0,1}A_{a(2i-\sigma)}A_{b(2i-\sigma)}=A_{a(2i-1)}A_{b(2i-1)}+A_{a(2i)}A_{b(2i)} \tag{20}$$

will be called betaoids since they are, like the usual Twiss parameters, related to the envelope $\langle z_a z_b\rangle$ and to the Hamiltonian (Lie) representation of the map.

We have seen physical justifications for the existence of the betaoid and etaoid invariants and they are based on the Hamiltonian nature of the flow of a pseudo-harmonic-oscillator. It is remarkable that the invariants have an extension to the nonsymplectic case most relevant to electron rings. The proof of this is simple but somewhat dry. It is presented in the next section.

IV. MATHEMATICAL POINT OF VIEW

We have seen how lattice functions emerge from asking questions about the properties at infinity, a very natural thing to do in the study of dynamical systems. There is a dry mathematical way to get the same answers and a little bit more. This way has the advantage of being extendable to damped systems. If a small amount of radiation is added to a ring, the closed orbit will move slightly and the eigenvalues will go off the unit circle by small amounts [3–5]. In this case we have six complex eigenvalues of the form

$$\lambda_{2a-\sigma}=\exp[(-1)^{\sigma}i\mu_a-\alpha_a]; \tag{21}$$

where $a=1,2,3$ and $\sigma=0,1$. The map can be put into a normal form analogous to that of the pseudo-harmonic-oscillator:

$$M=A\Lambda RA^{-1}. \tag{22}$$

None of the matrices involved in this normalization are symplectic except for R. The matrices R and Λ are, respectively, a phase space rotation and a diagonal damping matrix

$$R=\begin{pmatrix} R_1 & 0 & 0\\ 0 & R_2 & 0\\ 0 & 0 & R_3\end{pmatrix}, \tag{23}$$

where

$$R_i=\begin{pmatrix} \cos\mu_i & \sin\mu_i\\ -\sin\mu_i & \cos\mu_i\end{pmatrix}, \tag{24}$$

and

$$\Lambda=\begin{pmatrix} \Lambda_1 & 0 & 0\\ 0 & \Lambda_2 & 0\\ 0 & 0 & \Lambda_3\end{pmatrix}, \tag{25}$$

where

$$\Lambda_i=\begin{pmatrix} \exp(-\alpha_i) & 0\\ 0 & \exp(-\alpha_i)\end{pmatrix}. \tag{26}$$

This normal form is appropriate to electron rings in the presence of classical radiation. It is also useful when considering the stochastic maps on moments [5]. Here we will restrict our discussion to the deterministic damped map.

As in the symplectic case we know that the eigenvalues of M are unique and thus the matrices R and Λ are unique provided we associate each eigenvalue with a definite plane. The map A, however, is not unique. This is because the matrix ΛR commutes with a similar matrix δr,

$$\begin{aligned} M&=A\Lambda RA^{-1}\\ &\Downarrow\\ &=A\,\delta r\Lambda Rr^{-1}\delta^{-1}A^{-1}, \end{aligned} \tag{27}$$

where r is a rotation like R and δ is a dilation like Λ. The next step is to construct invariants of the diagonalization process using A and/or A^{-1}. Let us look at the matrix $r^{-1}\delta^{-1}A^{-1}$ first:

$$r^{-1}\delta^{-1}A^{-1}=\begin{pmatrix} \delta_1^{-1}r_1^{-1} & 0 & 0\\ 0 & \delta_2^{-1}r_2^{-1} & 0\\ 0 & 0 & \delta_3^{-1}r_3^{-1}\end{pmatrix} \times\begin{pmatrix} \begin{pmatrix}A^{-1}_{11}\\ A^{-1}_{21}\end{pmatrix} & \cdots & \begin{pmatrix}A^{-1}_{16}\\ A^{-1}_{26}\end{pmatrix}\\ \vdots & \vdots & \vdots\\ \begin{pmatrix}A^{-1}_{51}\\ A^{-1}_{61}\end{pmatrix} & \cdots & \begin{pmatrix}A^{-1}_{56}\\ A^{-1}_{66}\end{pmatrix}\end{pmatrix}.$$

If we define some minivectors using the matrix A^{-1},

$$\vec{v}^{\,ij}=\begin{pmatrix} A^{-1}_{(2i-1)j}\\ A^{-1}_{(2i)j}\end{pmatrix}, \quad i=1,\ldots,3, \tag{28}$$

then the matrix $r^{-1}\delta^{-1}A^{-1}$ is composed of the minivectors

$$\delta_i^{-1}r_i^{-1}\vec{v}^{\,ij}=\delta_i^{-1}r_i^{-1}\begin{pmatrix} A^{-1}_{(2i-1)j}\\ A^{-1}_{(2i)j}\end{pmatrix}, \quad i=1,\ldots,3. \tag{29}$$

In the presence of damping it is clear that no invariants of the normalization process can be constructed out of the minivectors of Eq. (29) alone. However, consider the transpose of the matrix $A\delta r$, which is just $\delta r^{-1}A$. As before we define a set of minivectors $\vec{w}^{\,ij}$ based on this matrix:

$$\delta_i r_i^{-1}\vec{w}^{\,ij}=\delta_i r_i^{-1}\begin{pmatrix} A_{j(2i-1)}\\ A_{j(2i)}\end{pmatrix}, \quad i=1,2,3. \tag{30}$$

Now we are ready to define two sets of invariants of the diagonalization process. First we take the dot product of these minivectors

$$\begin{aligned}\eta^i_{jk}&=\delta_i^{-1}r_i^{-1}\vec{v}^{\,ij}\cdot\delta_i r_i^{-1}\vec{w}^{\,ik}=\vec{v}^{\,ij}\cdot\vec{w}^{\,ik}=A^{-1}_{(2i-1)j}A_{k(2i-1)}\\ &\quad+A^{-1}_{(2i)j}A_{k(2i)}. \end{aligned} \tag{31}$$

The damping conveniently cancels out. As for the rotation, we know that it leaves the scalar product invariant and thus η^i_{jk} is the same for all possible choices of the transformation A. We also know that the wedge or cross product is left invariant by planar rotations; therefore, we define the set of functions

$$\begin{aligned}\beta^i_{jk}&=\delta_i^{-1}r_i^{-1}\vec{v}^{\,ij}\wedge\delta_i r_i^{-1}\vec{w}^{\,ik}=\vec{v}^{\,ij}\wedge\vec{w}^{\,ik}=A^{-1}_{(2i-1)j}A_{k(2i)}\\ &\quad-A^{-1}_{(2i)j}A_{k(2i-1)}, \end{aligned} \tag{32}$$

where $(x,y)\wedge(a,b)=xb-ya$. As in the symplectic case we expect quantities that do not depend on the normalization being a function of these generalized η's and β's only.

V. RELATIONS BETWEEN BETAOIDS IN THE SYMPLECTIC CASE

As we have said the betaoids appear in two different ways. First, we know that the radii in normalized space are invariants and this leads us to contractions of A^{-1} with itself. Second, we also know that ergodic averages of the quadratic moments must also be invariants; from this emerges contractions of A with itself.

Finally, mathematical manipulations in the arbitrary nonsymplectic case forces us to consider contractions of A with its inverse only. It remains to be proved that these are all the same invariants in the symplectic case. To do this one uses the definition of a symplectic matrix given by Eq. (1). Let us introduce the following notation for an index j running from 1 to 6:

$$\text{if}\quad j=1,3,5\quad\text{then}\quad \overline{j}=2,4,6;$$

$$\text{if}\quad j=2,4,6\quad\text{then}\quad \overline{j}=1,3,5.$$

Then it follows from the symplectic condition that the betaoids can be rewritten as

$$\beta^i_{jk} = -J_{k\bar{k}}\{A^{-1}_{(2i-1)j}A^{-1}_{(2i-1)\bar{k}} + A^{-1}_{(2i)j}A^{-1}_{(2i)\bar{k}}\} = -\frac{1}{2}J_{k\bar{k}}\frac{\partial^2 \varepsilon_i}{\partial z_{\bar{k}}\partial z_j} \tag{33}$$

or as

$$\beta^i_{jk} = J_{j\bar{j}}\{A_{\bar{j}(2i-1)}A_{k(2i-1)} + A_{\bar{j}(2i)}A_{k(2i)}\} = 2J_{j\bar{j}}\frac{\partial\langle z_k z_{\bar{j}}\rangle}{\partial \varepsilon_i}. \tag{34}$$

Thus, in the symplectic case Eqs. (6) and (14) are equivalent.

Finally, before discussing the nonsymplectic case, I want to point out that a measurement of the beam envelope will lead to a measurement of the emittances and through the equivalence established in Eqs. (33) and (34). The argument will be presented in two degrees of freedom as it clearly extends to a higher dimensionality. We start by constructing the following Hamiltonian made of the ergodic envelope:

$$\begin{aligned} h(\vec{z}) = {} & \langle p_x^2\rangle x^2 + \langle x^2\rangle p_x^2 + \langle p_y^2\rangle y^2 + \langle y^2\rangle p_y^2 - 2\langle x p_x\rangle x p_x \\ & + 2\langle p_x p_y\rangle xy - 2\langle p_x y\rangle x p_y - 2\langle x p_y\rangle p_x y + 2\langle xy\rangle p_x p_y \\ & - 2\langle y p_y\rangle y p_y . \end{aligned} \tag{35}$$

From Eqs. (33) and (34) we see that this Hamiltonian is just

$$h(\vec{z}) = \frac{\varepsilon_1}{2}\varepsilon_1(\vec{z}) + \frac{\varepsilon_2}{2}\varepsilon_2(\vec{z}). \tag{36}$$

The quantities ε_1 and ε_2 are the numerical values of the emittances of the trajectory being ergodically averaged. The functions $\varepsilon_1(\vec{z})$ and $\varepsilon_2(\vec{z})$ are the Courant-Snyder invariant functions for this linear system

$$h(\vec{z}) = \frac{\varepsilon_1}{2}\varepsilon_1(\vec{z}) + \frac{\varepsilon_2}{2}\varepsilon_2(\vec{z}). \tag{37}$$

If we now perform a normal form on this Hamiltonian, the result will be

$$h_{normal}(\vec{z}) = \frac{\varepsilon_1}{2}(x^2 + p_x^2) + \frac{\varepsilon_2}{2}(y^2 + p_y^2). \tag{38}$$

The effect of the normal form will be to turn the invariant functions $\varepsilon_1(\vec{z})$ and $\varepsilon_2(\vec{z})$ into radii in phase space. Thus it follows that the numerical values of the emittances ε_1 and ε_2, can be read off easily. Once these are known the betaoids can be obtained using Eq. (14). We will now discuss the final topic of this paper, which relates to the significance of these invariants in the nonsymplectic case and to the parametrization of the one-turn map.

VI. PHYSICAL INTERPRETATION IN THE GENERAL CASE

The general case corresponds to a particle undergoing classical radiation and whose energy is restored at the rf cavities. Accelerator physicists design such systems by requiring that the eigenvalues of the matrix be inside the unit circle by a small amount. The beam will then contract until the quantum fluctuations due to the granularity of the photon become significant. The beam reaches an equilibrium. This quantum effect is totally ignored in this paper, but the descriptions of the lattice functions presented here are relevant to the computation of the equilibrium envelope $\langle z_a z_b\rangle$ defined by distribution averaging (not ergodic averaging).

In the deterministic case of a damped pseudo-harmonic-oscillator it is physically inadequate to derive the invariant betaoids or etaoids using ergodic averages. Indeed, at infinity the beam collapses to the origin and thus all averages are trivially null. Thus it is not surprising that expressions (33) and (34) are not valid invariants of the damped pseudo-harmonic-oscillator. We may be tempted to give them the following meaning: It can be shown that the Courant-Snyder invariants defined in terms of A^{-1} will shrink towards the origin and keep their shape. Indeed, if a distribution of particles depends only on the functions $\varepsilon_i(\vec{z})$, i.e., $\rho(\varepsilon_1,\varepsilon_2,\varepsilon_3)$, then the new distribution after one turn will be given by

$$\exp(2\{\alpha_1+\alpha_2+\alpha_3\})\rho(e^{(2\alpha_1)}\varepsilon_1, e^{(2\alpha_2)}\varepsilon_2, e^{(2\alpha_3)}\varepsilon_3). \tag{39}$$

For small damping, away from linear resonances, it is true that the equilibrium distribution has the form of Eq. (39) and thus one can compute the so-called equilibrium emittances and feed them into a Gaussian distribution that is a function of the Courant-Snyder functions. In the general case, we cannot talk of equilibrium emittances based of the functions ε_i and thus the formulas for the Courant-Snyder functions do not enter in any physically well-posed problem. Only the invariants computed in Eqs. (31) and (32) are potentially present in the general linear case.

Thus we may ask the following questions: What quantities, if measured by two observers, will always be the same? What quantities do not depend on the actual method or transformation A used in computing them? The answer is somewhat trivial: the one-turn matrix itself, the tune, and damping shifts due to some perturbations. Let us start with the shifts: The Sands, Chao, and envelope formalisms all give formulas for the damping as a function of the radiation field. It is remarkable that formulas for the shift of the tune (complex part of the eigenvalues) depend only on the betaoids, while formulas for the damping depend only on the etaoids.

A. Tune and damping shifts

Since we are interested in first-order perturbation theory, it suffices to see the effects of a perturbation (radiation for example) at one point around the ring. Thus, suppose we are to perturb the ring by a linear vector field $d\vec{F}$ whose action is localized. That is to say, at a given point in the ring, the phase space coordinate $\vec{z}$ is modified by a small linear impulse force $d\vec{F}$:

$$\vec{z}^{fin} = \vec{z}^{ini} + d\vec{F}, \tag{40}$$

where

$$dF_i = \sum_j dF_{ij} z_j .$$

In the language of Lie operators, which does not assume linearity, the original one-turn Lie map $\mathcal{M}$ is modified by the new impulse $d\vec{F}$ and by the normalization transformation $\mathcal{A}$ as

$$\mathcal{A}\mathcal{M}^{new}\mathcal{A}^{-1}=\mathcal{A}\mathcal{M}\exp(d\vec{F}\cdot\vec{\nabla})\mathcal{A}^{-1}$$
$$=\mathcal{A}\mathcal{M}\mathcal{A}^{-1}\mathcal{A}\exp(d\vec{F}\cdot\vec{\nabla})\mathcal{A}^{-1}$$
$$=\mathcal{R}\exp(\mathcal{A}d\vec{F}\cdot\vec{\nabla}\mathcal{A}^{-1}). \tag{41}$$

Here the map $\mathcal{R}$ is the Lie map associated with the original matrix ΛR of Eq. (22). The effect of the transformation $\mathcal{A}$ on the Lie operator $d\vec{F}\cdot\vec{\nabla}$, denoted $\mathcal{A}d\vec{F}\cdot\vec{\nabla}\mathcal{A}^{-1}$ in Eq. (41), can be computed and the answer is

$$\text{if}\quad \mathcal{A}d\vec{F}\cdot\vec{\nabla}\mathcal{A}^{-1}=d\vec{G}\cdot\vec{\nabla}\Rightarrow dG_k=\sum_{a,b,c}A^{-1}_{ka}dF_{ab}A_{bc}z_c. \tag{42}$$

The next steps, which I will omit, consist in extracting the generators of rotations in the three phase space planes as well as the generators of damping. The coefficients in front of these generators are (with some constants) the tunes and the dampings. The formulas for the shift of the complex eigenvalues $\{\pm i\mu_j-\alpha_j\}$ are

$$\mu_j^{new}=\mu_j+\frac{1}{2}\sum_{a,b}\beta^j_{ab}dF_{ab},$$
$$\alpha_j^{new}=\alpha_j+\frac{1}{2}\sum_{a,b}\eta^j_{ab}dF_{ab}. \tag{43}$$

Since the coefficients dF_{ab} are arbitrary in the general case and since the eigenvalues cannot depend on the diagonalization process, we conclude that the functions β^j_{ab} and η^j_{ab} are invariant of the diagonalization process. Of course these are the same functions we defined in Sec. IV. The formula for the damping in Eq. (43) is very famous in the context of the computation of synchrotron integrals. In particular it is customary to write the damping in the longitudinal plane only in terms of the dispersion [6]. In the transverse plane, because the longitudinal tune μ_3 is small, it is useful to derive mixed formulas involving the transverse betaoids and the usual dispersions. In Ref. [5] Ohmi, Hirata, and Oide pointed out that this can be done rigorously using a special parametrization of A. However, noticing that the etaoids and betaoids are not independent, we can actually perform such transformations in the general case without using a special parametrization. For example, in two degrees of freedom, the formula for the ergodic (or distribution) average $\langle x^2\rangle$, where $x=z_1$, can be rewritten as

$$\langle x^2\rangle=\beta_{xx}\frac{\varepsilon_x}{2}+\frac{1}{(\eta^2_{33})^2}\{\beta_{zz}\zeta_z^2+\gamma_{zz}\eta_z^2-2\alpha_{zz}\zeta_z\eta_z\}\frac{\varepsilon_z}{2}, \tag{44}$$

where $\vec{z}=(x,p_x,z,\delta)$, $\beta_{xx}=-\beta^1_{21}$, $\beta_{zz}=-\beta^2_{43}$, $\gamma_{zz}=\beta^2_{34}$, and $\alpha_{zz}=\beta^2_{44}$. This formula should be contrasted with

$$\langle x^2\rangle=\beta_{xx}\frac{\varepsilon_x}{2}+\beta_{xz}\frac{\varepsilon_z}{2}, \tag{45}$$

where $\beta_{xx}=-\beta^1_{21}$ and $\beta_{xz}=-\beta^2_{21}$, which is obtained from a ‘‘normal’’ pseudoharmonic analysis using Eq. (14), for example [2]. Biased formalisms, mixing etaoids with betaoids, are necessary for pseudo-harmonic-oscillators when one wants to exploit certain properties such the smallness of a tune. In Ref. [5] it was shown that such formalisms can rigorously diagonalize a pseudo-harmonic-oscillator. The authors constructed a special parametrization for that purpose; here I point out that there is a more fundamental link between the usual symplectic formalism (all the planes are on an equal footing) and the biased formalism. This link is realized through the interdependence of the betaoids and etaoids.

Our discussion was centered on the computation of the tunes. Of course the vector field itself and thus the one-turn matrix should be expressible in terms of our invariant functions alone. This is the topic of the next subsection.

B. The one-turn symplectic matrix

Although the comments of this section can be extended to the damped nonsymplectic system, here I will restrict the discussion to the Hamiltonian case for simplicity. In one degree of freedom, it is well known that the one-turn matrix can be expressed in terms of the tunes and the Twiss functions (one degree of freedom betaoids):

$$M=\begin{pmatrix}\cos\mu+\alpha\sin\mu & \beta\sin\mu\\ -\gamma\sin\mu & \cos\mu-\alpha\sin\mu\end{pmatrix}. \tag{46}$$

The functions β, γ, and α are, respectively, $-\beta^1_{21}$, $-\beta^1_{12}$, and β^1_{22}. It is remarkable that no etaoids enter into this formula.

The question is whether or not it is possible to extend formulas for the one-turn matrix that depend only on the tunes and the betaoids. We will discover three facts in this section.

(i) When we express the one-turn matrix in terms of the invariants, it most naturally comes in terms of a mixed betaoid-etaoid representation.

(ii) In the symplectic case, it should be possible to have a pure betaoid representation, but it must be very messy to obtain. This is why it is not seen in the ‘‘coupled’’ formalism literature.

(iii) Finally, we will give a formula that relates the etaoids to the betaoids even in the general case.

We start with the expression for the symplectic one-turn matrix in terms of A, A^{-1}, and R and then use the simple nature of the rotation R:

$$M_{ab}=\sum_{j,k=1,6}A_{aj}R_{jk}A^{-1}_{kb}$$
$$=\sum_{j=1,3}\{A_{a(2j-1)}A^{-1}_{(2j-1)b}+A_{a(2j)}A^{-1}_{(2j)b}\}\cos\mu_j$$
$$+\{A_{a(2j-1)}A^{-1}_{(2j)b}-A_{a(2j)}A^{-1}_{(2j-1)b}\}\sin\mu_j$$
$$=\sum_{j=1,3}\{\eta^j_{ba}\cos\mu_j-\beta^j_{ba}\sin\mu_j\}. \tag{47}$$

In the case of one degree of freedom, the etaoids are equal to either one or zero. It is a simple exercise to regain the famous formula (46).

In more dimensions it appears that the presence of etaoids is unavoidable in the one-turn matrix and therefore it is no big surprise that no Courant-Snyder–like formula exists in the literature for the one-turn matrix that involves only the coefficients of the invariants ε_i (betaoids) and the tunes. However, the reader familiar with Lie methods knows that the one-turn map is actually the exponential of the Poisson bracket operator associated with the function

$$-\frac{1}{2}\{\mu_1\varepsilon_1+\mu_2\varepsilon_2+\mu_3\varepsilon_3\}$$

and thus the one-turn map can in theory be a function of the betaoids only, albeit an infinite series. However, in the case of a symplectic map, it turns out that it is possible to express the etaoids in terms of *only* the betaoids using formulas (33) and (34). First, we recall that the general derivation of these invariants involves the dot and wedge product of two vectors. We know that these are related so that if

$$(x,y)\wedge(a,b)=xb-ya$$

and

$$(x,y)\cdot(a,b)=xa+yb,$$

then

$$\{x^2+y^2\}\{a^2+b^2\}=\{(x,y)\wedge(a,b)\}^2+\{(x,y)\cdot(a,b)\}^2. \tag{48}$$

This equation is applied to η^i_{jk} and β^i_{jk} with the result that

$$\{\eta^i_{jk}\}^2+\{\beta^i_{jk}\}^2=\{(A_{k(2i-1)})^2+(A_{k(2i)})^2\}\times\{(A^{-1}_{(2i-1)j})^2+(A^{-1}_{(2i)j})^2\}. \tag{49}$$

Finally, we use Eqs. (33) and (34) to rewrite the right-hand side of Eq. (49) in terms of betaoids:

$$\{\eta^i_{jk}\}^2=J_{\bar{k}k}J_{j\bar{j}}\beta^i_{j\bar{j}}\beta^i_{\bar{k}k}-\{\beta^i_{jk}\}^2. \tag{50}$$

We now substitute this result in Eq. (47),

$$\begin{aligned}M_{ab}&=\sum_{j=1,3}\eta^j_{ba}\cos\mu_j-\beta^j_{ba}\sin\mu_j\\&=\sum_{j=1,3}\mathrm{sgn}(\eta^j_{ba})\sqrt{J_{\bar{a}a}J_{b\bar{b}}\beta^j_{b\bar{b}}\beta^j_{\bar{a}a}-\{\beta^j_{ba}\}^2}\\&\quad\times\cos\mu_j-\beta^j_{ba}\sin\mu_j.\end{aligned} \tag{51}$$

This formula is somewhat impractical unless the sign of η^j_{ba} is known in advanced. Nevertheless, it is interesting to rewrite η^j_{ba} is terms of either the moments or the Courant-Snyder coefficients

$$\begin{aligned}|\eta^j_{ba}|&=2\sqrt{\frac{\partial\langle z_{\bar{b}}^2\rangle}{\partial\varepsilon_j}\frac{\partial\langle z_{\bar{a}}^2\rangle}{\partial\varepsilon_j}-\left\{\frac{\partial\langle z_{\bar{b}}z_{\bar{a}}\rangle}{\partial\varepsilon_j}\right\}^2}\\&=\frac{1}{2}\sqrt{\frac{\partial^2\varepsilon_j}{\partial z_{\bar{a}}^2}\frac{\partial^2\varepsilon_j}{\partial z_b^2}-\left\{\frac{\partial^2\varepsilon_j}{\partial z_{\bar{a}}\partial z_b}\right\}^2}.\end{aligned} \tag{52}$$

[These formulas look very much like the so-called invariant emittance defined as $\langle x^2\rangle\langle p^2\rangle-\langle xp\rangle^2$. This emittance, which is an average over an *arbitrary distribution*, is preserved by one-degree-of-freedom linear symplectic maps. In fact, it does not change even if we transport it with any linear map. It is thus a much stronger invariant and should not be confused with our betaoids and etaoids. In fact, the reader will notice that this emittance looks very much like η^1_{11}, which happens to be a trivial constant (namely, one) in the one-degree-of-freedom case.] Notice that in the symplectic case it is easy to check using Eq. (52) that $\eta^j_{a\bar{a}}=0$ using Eq. (52). Finally, in the general case, we can derive a formula for η^j_{ba} only in terms of the betaoids:

$$\eta^j_{ba}=-\sum_{c,n}\beta^j_{bc}\beta^n_{ca}. \tag{53}$$

This formula was derived by comparing Eq. (47) with the Lie representation when all the tunes are near 90°. However, it can be proved to be true by direct substitution, which implies that the formula is true for all damped pseudo-harmonic-oscillators.

ACKNOWLEDGMENTS

I would like to thank Kohji Hirata and Stefania Petracca for discussing these issues with me and Hohji Hirata for introducing me to the work of Ref. [5] and the whole beam envelope approach.

[1] E. D. Courant and H. S. Snyder, Ann. Phys. (N.Y.) **3**, 1 (1958).
[2] F. Willeke and G. Ripken, in *Proceedings of the Linear Accelerator Conference* (American Institute of Physics, New York, 1989), p. 758.
[3] M. Sands, in *Physics with Intersecting Storage Rings*, edited by B. Touschek (Academic, New York, 1971), pp. 257–409.
[4] A. W. Chao, J. Appl. Phys. **50**, 595 (1979).
[5] K. Ohmi, K. Hirata, and K. Oide, Phys. Rev. E **49**, 751 (1994).
[6] S. Petracca and K. Hirata, KEK Technical Report No. 97-28, 1997 (unpublished).