General moment invariants for linear Hamiltonian systems

Alex J. Dragt
Center for Theoretical Physics, University of Maryland, College Park, Maryland 20742

Filippo Neri
AT Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

Govindan Rangarajan
Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720
(Received 12 July 1991)

This paper studies the behavior of the moments of a particle distribution as it is transported through a Hamiltonian system. Functions of moments that remain invariant for an arbitrary linear Hamiltonian system are constructed. These functions remain approximately invariant for Hamiltonian systems that are not strongly nonlinear. Consequently, they can be used to characterize the degree of nonlinearity of the system.

PACS number(s): 41.85.−p, 03.20.+i

I. INTRODUCTION

Consider a distribution of particles being transported through a Hamiltonian system. We assume either that the particles are noninteracting or that the effects arising from particle interactions can be treated in the Vlasov approximation. An important first step toward a complete understanding of the dynamics of this transport would be to determine quantities that remain invariant under transport. Since a particle distribution is characterized by its moments, it is natural to seek invariant functions of these moments.

In this paper, we concentrate on functions of moments that remain invariant under the action of a linear but otherwise arbitrary Hamiltonian system. (This implies *inter alia* that if the particles are interacting, then their interaction can be treated in the linear Vlasov approximation.) A brief description of quadratic moment invariants has been given earlier [1, 2]. In addition to giving a more complete discussion of these invariants, this paper also describes invariants constructed out of higher-order moments. In Sec. II, equations for linear transport of particle distributions and their moments are derived. Kinematic invariants are defined in Sec. III. Invariants made out of quadratic moments are considered in detail and three functionally independent quadratic moment invariants are shown to exist for a six-dimensional phase space (an explicit proof of this fact is given in Appendix A). A systematic method of constructing higher-order invariants is then presented. Next, a compact diagrammatic representation of these invariants is derived. Finally, a procedure to determine the number of functionally independent moment invariants is outlined. Section IV contains a summary of our results. Dynamic invariants are briefly discussed in Appendix B. An alternative method of constructing kinematic invariants utilizing properties of the Lie algebra sp(6,R) is given in Appendix C.

Let $z = (q_1, p_1, q_2, p_2, q_3, p_3)$ be the six-dimensional vector describing the location of a particle in phase space. The effect of a linear Hamiltonian system on this particle can be expressed as the action of a matrix M that takes the particle from its *initial* state z^{in} to the *final* state z^{fin}:

$$z_i^{\text{fin}} = M_{ij}\, z_j^{\text{in}}. \tag{1.1}$$

Here we have used Einstein's summation convention. This convention will be used throughout the paper unless stated otherwise. It can be shown [3] that M satisfies the symplectic condition

$$\tilde{M} J M = J, \tag{1.2}$$

where $\tilde{M}$ is the transpose of M and J is an antisymmetric matrix defined as follows:

$$J = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix}. \tag{1.3}$$

Matrices M satisfying Eq. (1.2) are called symplectic matrices. The set of all such matrices can be shown to form the symplectic group Sp(6,R) [3].

II. MOMENTS AND MOMENT TRANSPORT

In this section, we define moments of a particle distribution and their transport under the action of a linear Hamiltonian system. Let $h(z)$ be a distribution function describing particle density in phase space at the point having coordinates z. Then one can define the kth-order moment $\langle z_{i_1} z_{i_2} \cdots z_{i_k} \rangle$ of the distribution as follows:

$$\langle z_{i_1} z_{i_2} \cdots z_{i_k} \rangle = \int d^6 z\, h(z)\, z_{i_1} z_{i_2} \cdots z_{i_k}. \tag{2.1}$$

This leads to the following natural definitions for the initial and final kth-order moments of the distribution before and after transport through a Hamiltonian system:

$$\langle z_{i_1} z_{i_2} \cdots z_{i_k} \rangle^{\text{in}} = \int d^6 z\, h^{\text{in}}(z)\, z_{i_1} z_{i_2} \cdots z_{i_k}, \quad (2.2)$$

$$\langle z_{i_1} z_{i_2} \cdots z_{i_k} \rangle^{\text{fin}} = \int d^6 z\, h^{\text{fin}}(z)\, z_{i_1} z_{i_2} \cdots z_{i_k}. \quad (2.3)$$

Here $h^{\text{in}}(z)$ is the initial distribution function and $h^{\text{fin}}(z)$ is the final distribution function after transport through the Hamiltonian system described by the symplectic matrix M.

We can relate the initial and final distribution functions as follows [1]. Consider an initial collection of particles around a point z^{in}. This is sent to a final collection of particles around the point z^{fin} by the action of the symplectic matrix M. By construction, the number of particles in the initial and final collections is the same. Furthermore, by Liouville's theorem, the volumes occupied by the two collections are the same. Consequently, we have the relation

$$h^{\text{fin}}(z^{\text{fin}}) = h^{\text{in}}(z^{\text{in}}). \quad (2.4)$$

Using Eq. (1.1) for z^{in}, we get

$$h^{\text{fin}}(z^{\text{fin}}) = h^{\text{in}}(M^{-1} z^{\text{fin}}). \quad (2.5)$$

Finally, since z^{fin} is an arbitrary point, the following equation is seen to be correct,

$$h^{\text{fin}}(z) = h^{\text{in}}(M^{-1} z). \quad (2.6)$$

We next relate the initial and final moments using the above relation between the initial and final distribution functions. Substituting Eq. (2.6) in Eq. (2.3) we get

$$\langle z_{i_1} z_{i_2} \cdots z_{i_k} \rangle^{\text{fin}} = \int d^6 z\, h^{\text{in}}(M^{-1} z)\, z_{i_1} z_{i_2} \cdots z_{i_k}. \quad (2.7)$$

Next make the change of variables $z = Mz'$. Since M is a symplectic matrix [3], this coordinate change leaves the volume element $d^6 z$ invariant, i.e., $d^6 z = d^6 z'$. This gives the result

$$\langle z_{i_1} z_{i_2} \cdots z_{i_k} \rangle^{\text{fin}} = \int d^6 z'\, h^{\text{in}}(z')\, (Mz')_{i_1} (Mz')_{i_2} \cdots (Mz')_{i_k}. \quad (2.8)$$

This can be rewritten as follows:

$$\langle z_{i_1} z_{i_2} \cdots z_{i_k} \rangle^{\text{fin}} = M_{i_1 j_1} M_{i_2 j_2} \cdots M_{i_k j_k} \times \int d^6 z'\, h^{\text{in}}(z')\, z'_{j_1} z'_{j_2} \cdots z'_{j_k}. \quad (2.9)$$

The integral on the right-hand side is nothing but the initial moment [cf. Eq. (2.2)]. Therefore we finally get

$$\langle z_{i_1} z_{i_2} \cdots z_{i_k} \rangle^{\text{fin}} = M_{i_1 j_1} M_{i_2 j_2} \cdots M_{i_k j_k} \langle z_{j_1} z_{j_2} \cdots z_{j_k} \rangle^{\text{in}}. \quad (2.10)$$

Using tensor notation, the above equation can be given the following compact form:

$$\langle z z \cdots z \rangle^{\text{fin}} = \left(\bigotimes^{k} M \right) \langle z z \cdots z \rangle^{\text{in}}, \quad (2.11)$$

where

$$\bigotimes^{k} M \equiv M \otimes M \otimes \cdots \otimes M \quad (2.12)$$

and the ellipses in Eqs. (2.11) and (2.12) represent terms taken k times. This equation suggests a further simplification in notation. We identify the kth-order moments with elements of a kth-rank tensor $Z^{(k)}$ as follows:

$$Z^{(k)}_{i_1 i_2 \cdots i_k} = \langle z_{i_1} z_{i_2} \cdots z_{i_k} \rangle. \quad (2.13)$$

Then Eq. (2.11) can be rewritten to give

$$Z^{(k)}_{\text{fin}} = \left(\bigotimes^{k} M \right) Z^{(k)}_{\text{in}}. \quad (2.14)$$

This is the basic equation for moment transport.

III. KINEMATIC MOMENT INVARIANTS

A. General concepts

To motivate our discussion of kinematic moment invariants, we start with a concrete example. In the field of accelerator physics, it is well known that the mean-square emittances of a beam remain invariant under transport through a linear Hamiltonian system if the various degrees of freedom are uncoupled [4]. Since these emittances are defined in terms of quadratic moments (see below), these are examples of moment invariants.

As an instructive exercise, we now go through an explicit calculation demonstrating that these emittances are indeed invariants. The mean-square emittances are defined as follows:

$$\epsilon_i^2 = \langle q_i^2 \rangle \langle p_i^2 \rangle - (\langle q_i p_i \rangle)^2, \quad i = 1, 2, 3 \quad (3.1)$$

where repeated indices are not to be summed over. Since there is assumed to be no coupling between the q_1, q_2, and q_3 degrees of freedom, we consider only ϵ_1^2 below. The invariance of ϵ_2^2 and ϵ_3^2 can be proved using similar arguments.

We define the initial and final mean-square emittances as follows:

$$(\epsilon_1^2)^{\text{in}} = \langle q_1^2 \rangle^{\text{in}} \langle p_1^2 \rangle^{\text{in}} - (\langle q_1 p_1 \rangle^{\text{in}})^2,$$
$$(\epsilon_1^2)^{\text{fin}} = \langle q_1^2 \rangle^{\text{fin}} \langle p_1^2 \rangle^{\text{fin}} - (\langle q_1 p_1 \rangle^{\text{fin}})^2. \quad (3.2)$$

Using Eq. (2.10) we find the results

$$\langle q_1^2 \rangle^{\text{fin}} = (M_{11})^2 \langle q_1^2 \rangle^{\text{in}} + 2 M_{11} M_{12} \langle q_1 p_1 \rangle^{\text{in}} + (M_{12})^2 \langle p_1^2 \rangle^{\text{in}},$$

$$\langle q_1 p_1 \rangle^{\text{fin}} = M_{11} M_{21} \langle q_1^2 \rangle^{\text{in}} + (M_{11} M_{22} + M_{12} M_{21}) \langle q_1 p_1 \rangle^{\text{in}} + M_{12} M_{22} \langle p_1^2 \rangle^{\text{in}}, \quad (3.3)$$

$$\langle p_1^2\rangle^{\text{fin}} = (M_{21})^2\langle q_1^2\rangle^{\text{in}} + 2M_{21}M_{22}\langle q_1 p_1\rangle^{\text{in}} + (M_{22})^2\langle p_1^2\rangle^{\text{in}}.$$

Substituting Eq. (3.3) in Eq. (3.2) we get

$$(\epsilon_1^2)^{\text{fin}} = (\det M)^2(\epsilon_1^2)^{\text{in}}. \tag{3.4}$$

Since the determinant of M is equal to 1 by the symplectic condition [3], we obtain the relation

$$(\epsilon_1^2)^{\text{fin}} = (\epsilon_1^2)^{\text{in}}. \tag{3.5}$$

This result proves the invariance of ϵ_1^2.

Having found functions of moments that remain invariant under uncoupled linear transport, we are led to look for generalized functions that remain invariant under the full six-dimensional transport with coupling. Before doing this, we have to define more precisely what a moment invariant is. Suppose a function $I(Z^{(k)})$ has the following property:

$$I\left(\left(\bigotimes^k M\right)Z^{(k)}\right) = I(Z^{(k)}) \quad \text{for all} \quad M \in \text{Sp}(6,\text{R}), \tag{3.6}$$

i.e., it is invariant under the set of all linear symplectic transformations. Then I is called a *kinematic moment invariant*. In contrast to dynamic invariants, which are invariant only for a particular Hamiltonian system, the kinematic invariants defined above are invariant for all linear Hamiltonian systems.

Another important concept that we will need is that of moment equivalence classes. Consider two sets of moments $Z^{(k)'}$ and $Z^{(k)}$. Suppose there exists a symplectic matrix M such that the following relation is satisfied:

$$Z^{(k)'} = \left(\bigotimes^k M\right) Z^{(k)}. \tag{3.7}$$

Then we write

$$Z^{(k)'} \sim Z^{(k)}. \tag{3.8}$$

This relation is an equivalence relation. Let $[Z^{(k)}]$ be the set of all $Z^{(k)'}$ such that $Z^{(k)'} \sim Z^{(k)}$. The set $[Z^{(k)}]$ is called the equivalence class of $Z^{(k)}$.

This leads us to the observation that a kinematic moment invariant is a class function, i.e.,

$$I(Z^{(k')}) = I(Z^{(k)}) \quad \text{if} \quad Z^{(k)'} \sim Z^{(k)}. \tag{3.9}$$

This property of the invariant can be expressed as

$$I = I([Z^{(k)}]). \tag{3.10}$$

From the above discussion, we conclude that the number of functionally independent kinematic moment invariants is equal to the dimension of the space of equivalence classes.

We are now in a position to apply the results given above to construct kinematic moment invariants. We deal first with the simpler case of quadratic moment invariants before generalizing to the case of higher-order moment invariants.

B. Quadratic moment invariants

This section deals exclusively with kinematic moment invariants constructed out of quadratic moments. Most of the results presented here will be generalized to higher-order moment invariants in the subsequent sections.

Consider the quantity [5]

$$I_2^{(n)}[Z^{(2)}] = \text{tr}[(Z^{(2)}J)^n], \tag{3.11}$$

where $Z^{(2)}$ is a second-rank tensor (i.e., a matrix) whose elements are defined by the relation [cf. Eq. (2.13)]

$$Z_{ij}^{(2)} = \langle z_i z_j\rangle. \tag{3.12}$$

We now show that $I_2^{(n)}[Z^{(2)}]$ is a kinematic invariant.

The tensor $Z^{(2)}$ transforms as follows under the action of M [cf. Eq. (2.14)]:

$$Z^{(2)} \to (M \otimes M)Z^{(2)}. \tag{3.13}$$

We therefore have to show that $I_2^{(n)}[Z^{(2)}]$ is invariant under this transformation [cf. Eq. (3.6)]. Inserting Eq. (3.13) in Eq. (3.11) we get

$$I_2^{(n)}[(M \otimes M)Z^{(2)}] = \text{tr}[(MZ^{(2)}\tilde{M}J)^n], \tag{3.14}$$

where we have used the relation

$$(M \otimes M)Z^{(2)} = MZ^{(2)}\tilde{M}. \tag{3.15}$$

We can now rewrite Eq. (3.14) as

$$I_2^{(n)}[(M \otimes M)Z^{(2)}] = \text{tr}[(Z^{(2)}\tilde{M}JM)^n]. \tag{3.16}$$

Here we have utilized the following relation satisfied by the trace of a product of matrices:

$$\text{tr}[ABC] = \text{tr}[BCA]. \tag{3.17}$$

Finally, substituting the symplectic condition Eq. (1.2) and Eq. (3.11) in Eq. (3.16), we get the desired result:

$$I_2^{(n)}[(M \otimes M)Z^{(2)}] = I_2^{(n)}[Z^{(2)}]. \tag{3.18}$$

This proves that $I_2^{(n)}[Z^{(2)}]$ is a kinematic moment invariant.

Thus there are an infinite number of quadratic moment invariants. However, $I_2^{(n)}[Z^{(2)}]$ is zero for all odd n. This is easily proved by noticing that for odd n, $I_2^{(n)}[Z^{(2)}]$ has an odd number of J's. Since $Z^{(2)}$ is a symmetric matrix whereas J is antisymmetric, $(Z^{(2)}J)^n$ contains an odd number of antisymmetric matrices for odd n. Consequently, its trace is zero.

The invariant $I_2^{(2)}[Z^{(2)}]$ can be easily calculated from Eqs. (3.11), (3.12), and (1.3). It is given by the relation [6]

$$I_2^{(2)}[Z^{(2)}] = -2\big(\langle q_1^2\rangle\langle p_1^2\rangle - \langle q_1p_1\rangle^2 + \langle q_2^2\rangle\langle p_2^2\rangle - \langle q_2p_2\rangle^2 + \langle q_3^2\rangle\langle p_3^2\rangle - \langle q_3p_3\rangle^2 +2\langle q_1q_2\rangle\langle p_1p_2\rangle - 2\langle q_1p_2\rangle\langle q_2p_1\rangle + 2\langle q_1q_3\rangle\langle p_1p_3\rangle - 2\langle q_1p_3\rangle\langle q_3p_1\rangle + 2\langle q_2q_3\rangle\langle p_2p_3\rangle - 2\langle q_2p_3\rangle\langle q_3p_2\rangle\big). \tag{3.19}$$

Obviously, $I_2^{(2)}[Z^{(2)}]$ is a generalization of the mean-square emittances ϵ_i^2 [cf. Eq. (3.1)] to the full six-dimensional phase space when coupling between the various degrees of freedom is allowed. Expressions for $I_2^{(4)}[Z^{(2)}]$, $I_2^{(6)}[Z^{(2)}]$, etc., can be calculated similarly. However, it is worth remarking that these expressions grow rapidly in length with increasing n. For example, $I_2^{(4)}$ fills a page and $I_2^{(6)}$ requires several pages.

The next step is to determine the number of functionally independent invariants. As seen earlier, this number is equal to the dimension of the space of equivalence classes of $Z^{(2)}$'s. Therefore our task is to classify the $Z^{(2)}$'s according to their equivalence classes. We recall that $Z^{(2)'}$ is said to belong to the same equivalence class as $Z^{(2)}$ (i.e., $Z^{(2)'} \sim Z^{(2)}$) if there exists a symplectic matrix M such that the following equation is satisfied [cf. Eqs. (3.7) and (3.15)]:

$$Z^{(2)'} = MZ^{(2)}\tilde{M}. \tag{3.20}$$

We now claim that given any $Z^{(2)}$, there is a "normal form" $Z^{(2)'}$, such that $Z^{(2)'} \sim Z^{(2)}$ and the entries in $Z^{(2)'}$ have a simple diagonal form.

Theorem. Given any set of quadratic moments $Z^{(2)}$, there exists a symplectic matrix M such that the transformed quadratic moments $Z^{(2)'}$ given by Eq. (3.20) have a special form. Using the notation (3.12), this form is given by the relations

$$\langle q_iq_i\rangle' = \langle p_ip_i\rangle' = \lambda_i > 0, \quad i = 1,2,3$$
$$\langle z_iz_j\rangle' = 0 \ \text{ if } i \neq j. \tag{3.21}$$

That is, all off-diagonal moments of $Z^{(2)'}$ vanish, and canonically conjugate moments are equal and positive. See Appendix A for a proof of this theorem.

As seen from Eqs. (3.21), the normal form $Z^{(2)'}$ depends only on three independent parameters. Using this result we conclude that any quadratic kinematic moment invariant is of the form

$$I_2 = I_2(\lambda_1^2, \lambda_2^2, \lambda_3^2), \tag{3.22}$$

where the *eigen mean-square emittances* $\lambda_1^2, \lambda_2^2, \lambda_3^2$ are given by the relations

$$\lambda_i^2 = \langle q_iq_i\rangle'\langle p_ip_i\rangle', \quad i = 1,2,3. \tag{3.23}$$

As we vary these three parameters, we span the entire space of equivalence classes. Therefore the dimension of this space is equal to 3. This in turn proves that there are only three functionally independent quadratic moment invariants for a six-dimensional phase space.

Suppose $Z^{(2)'}$ is used to evaluate $I_2^{(n)}$ as given by Eq. (3.11). This calculation is particularly easy since inspection of Eqs. (1.3) and (3.21) shows that $Z^{(2)'}$ and J commute. Thus the $I_2^{(n)}$ are given by the relation

$$I_2^{(n)} = \text{tr}[(Z^{(2)'}J)^n] = \text{tr}[(Z^{(2)'})^nJ^n] = 2(-1)^{n/2}[\lambda_1^n + \lambda_2^n + \lambda_3^n], \quad n \text{ even}. \tag{3.24}$$

Evidently, we can choose the eigen mean-square emittance λ_1^2, λ_2^2, and λ_3^2 themselves (up to permutations among them) to be the three functionally independent invariants. Another choice would be to take any three of the $I_2^{(n)}$. For example, one could choose $I_2^{(2)}, I_2^{(4)}$, and $I_2^{(6)}$.

We have considered only kinematic moment invariants in the above discussion. Dynamic invariants can be treated in a somewhat similar fashion. A brief discussion of these invariants is given in Appendix B.

C. Higher-order moment invariants

In this section, general expressions for kinematic invariants made out of higher-order moments will be obtained. Some will be made out of moments of a fixed order. These will be called pure invariants. Using moments of different orders, mixed invariants will also be constructed. In addition, a diagrammatic representation of these invariants will be given. Finally, the problem of finding the number of functionally independent invariants will be dealt with.

1. Pure invariants

Consider the following quantities:

$$I_{2m}^{(n)}[Z^{(2m)}] = \text{tr}\left[\left\{Z^{(2m)}\left(\bigotimes^{m} J\right)\right\}^n\right], \quad m = 1,2,\ldots,\ n = 1,2,\ldots \tag{3.25}$$

and

$$I_{2m+1}^{(2n)}[Z^{(2m+1)}] = \text{tr}\left[\left\{Z^{(2m+1)}\left(\bigotimes^{m} J\right)Z^{(2m+1)}\left(\bigotimes^{m+1} J\right)\right\}^n\right], \quad m = 1,2,\ldots,\ n = 1,2,\ldots\,. \tag{3.26}$$

The following examples clarify the notation used in the above equations:

$$\mathrm{tr}\left[\left\{Z^{(4)}\left(\bigotimes^{2} J\right)\right\}^{2}\right]$$

$$= Z^{(4)}_{i_1 i_2 i_3 i_4} J_{i_3 k_3} J_{i_4 k_4} Z^{(4)}_{k_1 k_2 k_3 k_4} J_{k_1 i_1} J_{k_2 i_2} \quad (3.27)$$

and

$$\mathrm{tr}\left[Z^{(3)}\left(\bigotimes^{1} J\right) Z^{(3)}\left(\bigotimes^{2} J\right)\right]$$

$$= Z^{(3)}_{i_1 i_2 i_3} J_{i_3 k_3} Z^{(3)}_{k_1 k_2 k_3} J_{k_1 i_1} J_{k_2 i_2}. \quad (3.28)$$

We claim that $I^{(n)}_{2m}[Z^{(2m)}]$ and $I^{(2n)}_{2m+1}[Z^{(2m+1)}]$ are kinematic invariants. Consider the quantity $I^{(n)}_{2m}[Z^{(2m)}]$. We have to show that it is invariant under the transformation [cf. Eq. (3.6)]

$$Z^{(2m)} \to \left(\bigotimes^{2m} M\right) Z^{(2m)}. \quad (3.29)$$

Substituting this into Eq. (3.25) we get

$$I^{(n)}_{2m}\left[\left(\bigotimes^{2m} M\right) Z^{(2m)}\right]$$

$$= \mathrm{tr}\left\{\left[\left(\bigotimes^{2m} M\right) Z^{(2m)}\left(\bigotimes^{m} J\right)\right]^{n}\right\}. \quad (3.30)$$

Using Eqs. (2.10), (2.11), and (3.17) this can be rewritten as

$$I^{(n)}_{2m}\left[\left(\bigotimes^{2m} M\right) Z^{(2m)}\right] = \mathrm{tr}\left\{\left[Z^{(2m)} \bigotimes^{m}(\tilde{M} J M)\right]^{n}\right\}. \quad (3.31)$$

Using the symplectic condition Eq. (1.2) and Eq. (3.25), we finally get the desired invariance relation

$$I^{(n)}_{2m}\left[\left(\bigotimes^{2m} M\right) Z^{(2m)}\right] = I^{(n)}_{2m}[Z^{(2m)}]. \quad (3.32)$$

A similar proof can be given for the invariance of $I^{(2n)}_{2m+1}[Z^{(2m+1)}]$.

Not all the invariants listed in Eqs. (3.25) and (3.26) are useful. Some are identically zero as shown below:

$$\begin{aligned} &(a)\quad I^{(n)}_{2m} = 0 \ \text{ if } m \text{ and } n \text{ are odd;}\\ &(b)\quad I^{(n)}_{4m} = 0 \ \text{ if } n = 1;\\ &(c)\quad I^{(2n)}_{2m+1} = 0 \ \text{ if } n \text{ is odd.}\end{aligned} \quad (3.33)$$

Conditions (a) and (c) can be understood by noting that $Z^{(k)}$ is a symmetric tensor and there are an odd number of J's in these two cases. Therefore the trace is zero. In case (b), even though there are an even number of J's, the antisymmetric indices of J are summed over the symmetric indices of $Z^{(k)}$ and hence $I^{(n)}_{4m}$ is zero. All these conditions can be summarized by the statement that $I^{(n)}_{m}$ is zero if n is equal to 1 or if nm is a not a multiple of 4.

Using Eqs. (3.25) and (3.26), we find up to a normalization the following expressions for the first few higher-order moment invariants in two phase-space dimensions

$$I^{(4)}_{3}[Z^{(3)}] = \langle q_1^3\rangle^2\langle p_1^3\rangle^2 - 3\langle q_1^2 p_1\rangle^2\langle q_1 p_1^2\rangle^2 + 4\langle q_1^3\rangle\langle q_1 p_1^2\rangle^3 + 4\langle q_1^2 p_1\rangle^3\langle p_1^3\rangle - 6\langle q_1^3\rangle\langle q_1^2 p_1\rangle\langle q_1 p_1^2\rangle\langle p_1^3\rangle, \quad (3.34)$$

$$I^{(2)}_{4}[Z^{(4)}] = \langle q_1^4\rangle\langle p_1^4\rangle + 3\langle q_1^2 p_1^2\rangle^2 - 4\langle q_1^3 p_1\rangle\langle q_1 p_1^3\rangle, \quad (3.35)$$

$$I^{(3)}_{4}[Z^{(4)}] = \langle q_1^4\rangle\langle p_1^4\rangle\langle q_1^2 p_1^2\rangle - \langle q_1^4\rangle\langle q_1 p_1^3\rangle^2 - \langle q_1^2 p_1^2\rangle^3 - \langle q_1^3 p_1\rangle^2\langle p_1^4\rangle + 2\langle q_1^3 p_1\rangle\langle q_1 p_1^3\rangle\langle q_1^2 p_1^2\rangle. \quad (3.36)$$

In the above expressions, the leading coefficient has been normalized to be equal to +1. An alternative derivation of such invariants using properties of the Lie algebra sp(6,R) is given in Appendix C.

2. *Mixed invariants*

Invariants can also be constructed out of moments of different orders. These are called mixed moment invariants. Using arguments similar to those given above for pure invariants, one can show that the following quantity is a kinematic invariant:

$$I^{(n_1,n_2,\ldots,n_k)}_{m_1,m_2,\ldots,m_k} = \mathrm{tr}[(W^{(m_1)})^{n_1}(W^{(m_2)})^{n_2}\cdots(W^{(m_k)})^{n_k}]. \quad (3.37)$$

For even m_j, the quantity $W^{(m_j)}$ is defined as (with $m_j = 2m$)

$$W^{(m_j)} = Z^{(2m)}\left(\bigotimes^{m} J\right). \quad (3.38)$$

If m_j is odd, the mixed invariant is zero unless the corresponding n_j is even. For odd m_j, we therefore need to define only the square of $W^{(m_j)}$ and it is given as (with $m_j = 2m+1$)

$$(W^{(m_j)})^2 = Z^{(2m+1)}\left(\bigotimes^{m} J\right) Z^{(2m+1)}\left(\bigotimes^{m+1} J\right). \quad (3.39)$$

The mixed invariants defined in Eq. (3.37) can be shown to be zero unless $\sum_{i=1}^{k}(m_i n_i)$ is a multiple of 4.

Using Eq. (3.37), the simplest example of a mixed moment invariant for a two-dimensional phase space can easily be calculated and is exhibited below with a convenient normalization:

$$I^{(2,1)}_{1,2} = \langle q_1^2\rangle\langle p_1\rangle^2 - 2\langle q_1 p_1\rangle\langle q_1\rangle\langle p_1\rangle + \langle p_1^2\rangle\langle q_1\rangle^2. \quad (3.40)$$

Such mixed moment invariants become important when the Hamiltonian contains terms linear in phase-space variables [7]. This can be seen as follows. For simplicity, let us consider only a two-dimensional phase space. The presence of linear terms in the Hamiltonian is equivalent to a shift in the origin of the phase space:

$$q_1 \rightarrow q_1 - \langle q_1 \rangle, \quad p_1 \rightarrow p_1 - \langle p_1 \rangle. \tag{3.41}$$

Under these circumstances, the $I[Z^{(k)}]$'s listed above no longer remain invariant. The new invariants can be obtained by applying the transformation given in Eq. (3.41) to these old "invariants." For example, consider $I_2^{(2)}$. For a two-dimensional phase space, we have up to a normalization the relation

$$I_2^{(2)} = \langle q_1^2 \rangle \langle p_1^2 \rangle - (\langle q_1 p_1 \rangle)^2. \tag{3.42}$$

Applying the transformation given in Eq. (3.41) we get

$$I_2^{(2)'} = \langle (q_1 - \langle q_1 \rangle)^2 \rangle \langle (p_1 - \langle p_1 \rangle)^2 \rangle - [\langle (q_1 - \langle q_1 \rangle)(p_1 - \langle p_1 \rangle) \rangle]^2. \tag{3.43}$$

Upon expanding the above expression, we find that the new invariant can be expressed as the following linear combination of the old invariants [cf. Eqs. (3.40) and (3.42)]:

$$I_2^{(2)'} = I_2^2 - I_{1,2}^{(2,1)}. \tag{3.44}$$

Therefore, when linear terms are present in the Hamiltonian, only a combination of the original moment invariants remains invariant. And, as seen in the example given above, these combinations typically involve the mixed invariants.

3. Diagrammatic representation of moment invariants

Upon a closer inspection of the formulas for kinematic moment invariants, it is seen that a compact diagrammatic representation of these invariants can be given. These are exhibited in Fig. 1. Each node represents a $Z^{(k)}_{i_1 i_2 \ldots i_k}$, where k equals the number of lines emanating from that node. And each line connecting two different nodes $Z^{(k)}_{i_1 i_2 \ldots i_k}$ and $Z^{(l)}_{j_1 j_2 \ldots j_l}$ represents a J_{rs} where $r \in \{i_1, i_2, \ldots, i_k\}$ and $s \in \{j_1, j_2, \ldots, j_l\}$. Moreover, all like indices are summed over. In this sense, a line represents not just J_{rs} but also the summation over the indices it links.

Only the nonzero moment invariants are shown. We note that all diagrams have an *even* number of lines connecting the various nodes. An odd number of lines corresponds to an odd number of J's. Therefore a diagram with an odd number of lines would correspond to a situation where one is taking a trace over a product containing an odd number of antisymmetric matrices. As seen earlier, the trace (and hence the invariant) is zero in this case. Moreover, in these diagrams, a line always connects two different nodes and not the same node (i.e., no closed loops are present). A closed loop would correspond to a situation where the antisymmetric indices of a J are being summed over the symmetric indices of a $Z^{(k)}$. This again gives zero.

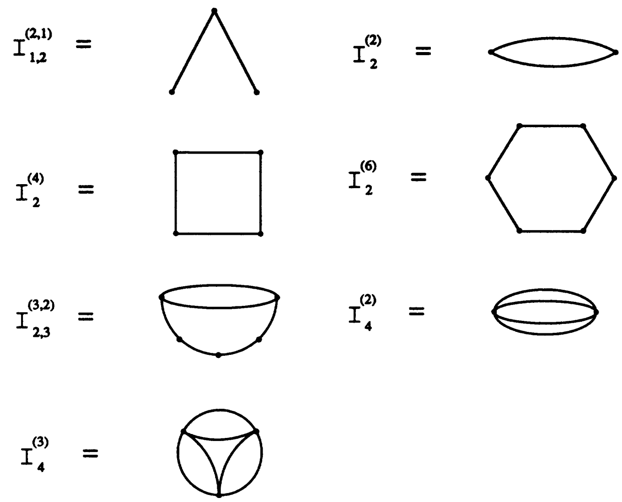

FIG. 1. Diagrammatic representation of moment invariants.

4. The number of functionally independent invariants

We finally turn to the problem of determining the number of functionally independent kinematic moment invariants. As seen earlier, this number is equal to the dimension of the space of equivalence classes. We therefore devote this section to the calculation of this dimension. For the sake of simplicity, we restrict ourselves to pure moment invariants.

Consider the set of all moments of a fixed order m. It forms an $N(m)$-dimensional vector space $V^{(m)}$ where $N(m)$ is given by the relation [8]

$$N(m) = \binom{m+5}{m}. \tag{3.45}$$

We start with a general moment vector v belonging to this vector space $V^{(m)}$. Another vector $v' \in V^{(m)}$ is said to belong to the same equivalence class as v if there exists a linear symplectic map M such that the following relation holds [cf. Eq. (2.11)]:

$$v' = \left(\bigotimes^m M \right) v. \tag{3.46}$$

We then write $v' \sim v$ and define the equivalence class $U(v)$ by the condition

$$U(v) = \{ v' \in V^{(m)} | v' \sim v \}. \tag{3.47}$$

We next pick a vector w not belonging to $U(v)$ and form its equivalence class $U(w)$. In principle, this procedure can be continued to construct the space Q of all equivalence classes. The dimension of this space would then give the number of functionally independent invariants. However, this brute-force procedure is very difficult to implement in practice except for the simplest case of first-order moments. It can be shown that any given set of first-order moments can be transformed into any other set of first-order moments using Eq. (3.46). Therefore all first-order moments belong to a single equivalence class. In other words, the space of equivalence classes is made up of a single point. Since the dimension of this space is zero, we conclude that there are no kinematic invariants constructed out of first-order moments.

For higher-order moments, we need a more practical procedure. This is obtained by realizing that we are interested only in the dimension of the space of equivalence classes. From standard algebra [9], the dimension of this space Q can be shown to have the following simple relation to the dimensions of $V^{(m)}$ and $U(v)$:

$$\dim Q = \dim V^{(m)} - \dim U(v), \tag{3.48}$$

where v is a vector belonging to $V^{(m)}$. This equation gives a practical way of calculating the number of functionally independent invariants constructed out of mth-order moments.

Since the dimension of $V^{(m)}$ is known from Eq. (3.45), all that needs to be determined is the dimension of $U(v)$. This is no easy matter if we have to first construct $U(v)$ explicitly before determining its dimension. Fortunately, this is not necessary.

To determine the dimension of $U(v)$, it is sufficient to work locally near the point v. Therefore we need to work only with the Lie algebra rather than the full group. We start with the $N(m)$-dimensional vector v and act on it with the 21 basis elements of the Lie algebra sp(6,R). We get 21 $N(m)$-dimensional vectors. We check these vectors for linear independence. The maximum number of linearly independent vectors gives the dimension of $U(v)$. We note that this dimension cannot exceed 21. Therefore the number of functionally independent invariants is at least $N(m) - 21$ [cf. Eqs. (3.45) and (3.48)]. Since $N(m)$ increases rapidly with m (it is greater than 21 for $m > 2$), the number of functionally independent invariants also increases rapidly with m.

Implementing the above procedure for a six-dimensional phase space, the dimension of $U(v)$ is found to be 6 if m is equal to 1 and 18 if m is equal to 2 [10]. It saturates at 21 if m is greater than 2. Using this information in Eq. (3.48), we see that there are no first-order moment invariants, three independent second-order moment invariants, and $N(m) - 21$ independent mth-order moment invariants for $m > 2$.

IV. SUMMARY

We have given a systematic treatment of moments and moment invariants for linear Hamiltonian systems. Three functions made of quadratic moments were constructed for a six-dimensional phase space and shown to be invariant under transport through such systems. These three invariants were shown to form a complete set. Formulas to construct higher-order moment invariants were given. Finally, a general procedure to determine the number of functionally independent invariants was outlined.

Even though the moment invariants that we have constructed are exactly invariant only for linear systems, they should remain approximately invariant for nonlinear systems if the nonlinearity is sufficiently small. These moment invariants could be used to characterize the degree of nonlinearity of a given Hamiltonian system. For example, this can be done by calculating the magnitude of variation in the values of these invariants as a particle distribution is being transported through the system.

ACKNOWLEDGMENTS

This work was supported in part by the United States Department of Energy under Contract No. DESA05-80ER10666. One of us (G.R.) was also supported by SURA/CEBAF and by the United States Department of Energy under Contract No. DE-AC03-76SFF00098 (at LBL). Finally, we are grateful to Richard K. Cooper for a careful reading of the manuscript and helpful suggestions.

APPENDIX A

In this appendix we prove the theorem of Sec. III B. Other proofs of this result can be found in Refs. [11] and [12]. However, these proofs either require advanced mathematical knowledge or are embedded in more general theorems. The proof given below is self-contained, simple, and constructive. It will be presented as a series of lemmas.

We start with the definition of the matrix Z [cf. Eqs. (2.1) and (3.12)]:

$$Z_{ij} = \langle z_i z_j \rangle = \int d^6 z\, h(z)\, z_i z_j . \tag{A1}$$

[Here, for notational simplicity, we have dropped the superscript "2" employed in Eq. (3.12).] The matrix Z is obviously symmetric. We now prove that under very general conditions the matrix Z is also positive definite.

Lemma 1. Since $h(z)$ is a density in phase space, it must be greater than or equal to zero for all z. Suppose that $h(z)$ is continuous at some point z^0 where it is nonzero. Then Z is positive definite.

Proof. Since $h(z)$ is continuous and nonzero at z^0, there exists a ball $B_\epsilon = \{z \mid \|z - z^0\| \le \epsilon\}$ such that $h(z) \ge \delta > 0$ for $z \in B_\epsilon$. Let u be any nonzero vector. We have the relation

$$\begin{aligned}(u, Zu) = u_i Z_{ij} u_j &= \int d^6 z\, h(z)\, u_i z_i u_j z_j \\ &= \int d^6 z\, h(z)\, (u_i z_i)^2 . \end{aligned} \tag{A2}$$

Using the assumption that $h(z) \ge \delta > 0$ for $z \in B_\epsilon$, we obtain the result

$$(u, Zu) \ge \delta \int_{B_\epsilon} d^6 z\, (u_i z_i)^2 \; > \; 0. \tag{A3}$$

Thus we have shown that the quadratic form (u, Zu) is greater than zero for all nonzero vectors u. Hence, by definition, the matrix Z is positive definite.

Lemma 2. Consider the "Hamiltonian" H defined by the relation

$$H(z) \equiv \frac{1}{2}(z, Zz) = \frac{1}{2} Z_{ij} z_i z_j . \tag{A4}$$

This quadratic form in z is positive definite since Z is positive definite. It follows that there exists a constant $c > 0$ such that the following result holds,

$$H(z) \geq c||z||^2 . \tag{A5}$$

Proof. Define the constant c by the relation

$$c \equiv \min_{||n||=1} H(n). \tag{A6}$$

Since the unit sphere ($\{n| \, ||n|| = 1\}$) is compact, and $H(z)$ is continuous and positive on the unit sphere, it must have a minimum c, with c greater than zero. Any nonzero vector z can be written in the form

$$z = ||z|| \frac{z}{||z||} = ||z|| n. \tag{A7}$$

Thus we have the result

$$H(z) = ||z||^2 H(n) \geq c||z||^2 . \tag{A8}$$

This proves the lemma.

Consider the matrices T and T' defined by the relations

$$T = JZ \; , \; T' = JZ'. \tag{A9}$$

Then, using Eq. (3.20) and the symplectic condition Eq. (1.2), we find that T and T' are connected by the relation

$$T' = \tilde{M}^{-1} T \tilde{M}. \tag{A10}$$

That is, the congruency relation Eq. (3.20) is converted into the similarity relation Eq. (A10) [13]. It follows that T' and T have the same eigenvalues. Consequently, we should devote our attention to the eigenvalues of T.

Since T is a real matrix, its eigenvalues must also be real or occur in complex conjugate pairs with the same multiplicity. Moreover, T cannot have zero as an eigenvalue. If it did, we would have the result

$$0 = \det(T) = \det(JZ) = \det(J)\det(Z) = \det(Z).$$

But then Z must have zero as an eigenvalue, which would violate the positive definite property. Here we have used the facts that $\det(J) = 1$ and the determinant of a matrix is the product of its eigenvalues.

Lemma 3. The spectrum of T is purely imaginary. Furthermore, the eigenvectors of T form a basis, and hence T can be diagonalized even if all its eigenvalues are not distinct.

Proof. Consider the Hamiltonian differential equation

$$\frac{dz}{dt} = [z, H] \; , \tag{A11}$$

where H is given by Eq. (A4) and $[\, ,]$ denotes the usual Poisson bracket. Equation (A11) is simply the flow generated by H. Using Eqs. (A4) and (A9), we obtain the result

$$[z, H] = Tz. \tag{A12}$$

Here, we have also made use of the relation [3]

$$[z_i, z_j] = J_{ij} . \tag{A13}$$

Substituting Eq. (A12) into Eq. (A11) and integrating over t, we obtain the result

$$z(t) = e^{tT} z(0). \tag{A14}$$

Let A be the matrix that brings T to its Jordan normal form N,

$$A^{-1} T A = N. \tag{A15}$$

Using standard properties of matrix multiplication [14], Eq. (A14) can be rewritten in the form

$$z(t) = A \exp(tA^{-1}TA) A^{-1} z(0). \tag{A16}$$

Substituting Eq. (A15) into the above equation, we obtain the result

$$z(t) = Ae^{tN} A^{-1} z(0). \tag{A17}$$

The matrix e^{tN} will also be in normal form. On its diagonal it will have the terms $e^{\sigma_j t}$ where the σ_j's are the eigenvalues of T. If T cannot be diagonalized, e^{tN} will contain terms of the form $t^m e^{\sigma_j t}$ above the diagonal [14]. Assume that σ_j is not pure imaginary for some j. Then $z(t)$ will grow exponentially as $t \to \pm\infty$ for some set of initial conditions. In this case, $||z||$ goes to infinity. Substituting this result into Eq. (A5), we find that $H(z(t)) \to \infty$ as $t \to \pm\infty$. However, since the Hamiltonian is conserved, we must also have the result

$$H(z(t)) = H(z(0)) = \text{const}. \tag{A18}$$

Thus we have a contradiction. It follows that all σ_j must be pure imaginary.

Further, if T cannot be diagonalized, $z(t)$ will grow as t^m even if all the eigenvalues are pure imaginary, for some choice of $z(0)$. Therefore we again get $||z|| \to \infty$ leading to a contradiction. Thus it must be possible to diagonalize T. This completes the proof of the lemma.

Since T can be diagonalized, it must have six linearly independent eigenvectors. Half of them will have Im $\sigma > 0$, and half will have Im $\sigma < 0$. Let ψ_j with $j = 1, 2, 3$ be the eigenvectors satisfying $\sigma_j = i\lambda_j$ with $\lambda_j > 0$. That is,

$$T\psi_j = i\lambda_j \psi_j , \quad \lambda_j > 0. \tag{A19}$$

As has been implicitly assumed in the discussion so far, let $(\, ,)$ denote the usual complex scalar product. Introduce an angular inner product $\langle \, , \rangle$ by the rule

$$\langle \chi, \phi \rangle = (\chi, K\phi) \tag{A20}$$

with K defined by the relation

$$K = -iJ. \tag{A21}$$

Here χ and ϕ are any two vectors. We note that K is Hermitian ($K = K^\dagger$) with respect to the standard scalar product $(\, ,)$. Consequently we have the relation

$$\langle \phi, \chi \rangle = \overline{\langle \chi, \phi \rangle}. \tag{A22}$$

Lemma 4. The angular inner product and the vectors

ψ_j satisfy the relations

$$\langle \psi_j , \psi_j \rangle > 0, \tag{A23}$$

$$\langle \psi_j , \psi_k \rangle = 0 \ \text{ if } \lambda_j \neq \lambda_k , \tag{A24}$$

$$\langle \psi_j , \overline{\psi_k} \rangle = 0. \tag{A25}$$

Proof. To verify Eq. (A23), rewrite Eq. (A19) in the form

$$K\psi_j = (1/\lambda_j) Z\psi_j. \tag{A26}$$

It follows that

$$\langle \psi_j , \psi_j \rangle = (1/\lambda_j)(\psi_j , Z\psi_j). \tag{A27}$$

Since Z is positive definite, it is easily verified that $(\chi, Z\chi) > 0$ for any nonzero vector χ even if χ is complex. Consequently, Eq. (A23) is correct.

To verify Eq. (A24), rewrite the first of Eqs. (A9) in the form

$$Z = -JT. \tag{A28}$$

Since Z is symmetric, we have the relation

$$\tilde{T}J + JT = 0. \tag{A29}$$

Take matrix elements of both sides of the above equation to get the result

$$(\psi_j , (\tilde{T}J + JT)\psi_k) = 0. \tag{A30}$$

From Eq. (A19) one has the results

$$(\psi_j , JT\psi_k) = i\lambda_k(\psi_j , J\psi_k) = -\lambda_k\langle \psi_j , \psi_k \rangle, \tag{A31}$$

$$\begin{aligned}(\psi_j , \tilde{T}J\psi_k) &= (T\psi_j , J\psi_k) = (i\lambda_j\psi_j , J\psi_k) \\ &= -i\lambda_j(\psi_j , J\psi_k) = \lambda_j\langle \psi_j , \psi_k \rangle.\end{aligned} \tag{A32}$$

It follows from Eq. (A30) that one has the result

$$(\lambda_j - \lambda_k)\langle \psi_j , \psi_k \rangle = 0. \tag{A33}$$

Consequently, Eq. (A24) is correct.

It remains to show that Eq. (A25) is correct. Since T is a real matrix, it follows from Eq. (A19) that

$$T\overline{\psi}_k = -i\lambda_k\overline{\psi}_k. \tag{A34}$$

Now carry out a calculation analogous to that of the preceding paragraph using the vectors ψ_j and $\overline{\psi}_k$. Doing so gives the relation

$$(\lambda_j + \lambda_k)\langle \psi_j , \overline{\psi}_k \rangle = 0. \tag{A35}$$

Thus Eq. (A25) is also correct.

Lemma 5. Starting with the vectors ψ_j, one can construct vectors ϕ_j such that

$$T\phi_j = i\lambda_j\phi_j , \tag{A36}$$

$$\langle \phi_j , \phi_k \rangle = \delta_{jk} , \tag{A37}$$

$$\langle \phi_j , \overline{\phi}_k \rangle = 0. \tag{A38}$$

Proof. Consider first the simplest case where all the λ_j are distinct. Then, inspection of the relations (A23)–(A25) shows that all that is required is a simple rescaling of the vectors ψ_j.

Next consider the case of double degeneracy, say $\lambda_1 = \lambda_2 \neq \lambda_3$. Consider the subspace spanned by ψ_1 and ψ_2. Let ψ be any nonzero vector in this subspace. Then we have the relation

$$T\psi = i\lambda\psi \quad \text{where } \lambda = \lambda_1 = \lambda_2. \tag{A39}$$

It follows, in analogy to the derivation of Eq. (A27), that we also have the relation

$$\langle \psi , \psi \rangle = (1/\lambda)(\psi , Z\psi) > 0. \tag{A40}$$

We see from Eqs. (A22) and (A40) that on this subspace the angular inner product $\langle\, , \rangle$ may also serve as a *bona fide* scalar product. Using this scalar product, apply the Gram-Schmidt procedure to the subspace spanned by ψ_1 and ψ_2 to produce two orthonormal vectors ϕ_1 and ϕ_2. Add to this orthonormal set the vector ϕ_3 defined by the relation

$$\phi_3 = \psi_3 / \sqrt{\langle \psi_3 , \psi_3 \rangle}. \tag{A41}$$

It is easily verified that these vectors satisfy the relations (A36)–(A38).

Finally, consider the case of triple degeneracy, $\lambda_1 = \lambda_2 = \lambda_3$. Now work with the subspace spanned by ψ_1, ψ_2, and ψ_3 to produce an orthonormal set. From an argument analogous to that of the preceding paragraph, the angular inner product $\langle\, , \rangle$ can also serve as a scalar product on this subspace. Consequently, using the angular inner product, the Gram-Schmidt procedure can be applied to the vectors ψ_1, ψ_2, ψ_3 to produce an orthonormal set ϕ_1, ϕ_2, ϕ_3 satisfying Eq. (A37). It is readily verified that the result of this process also satisfies Eqs. (A36) and (A38). This construction completes the proof of the lemma.

Suppose the vectors ϕ_j are decomposed into real and imaginary parts by writing the relations

$$\phi_j = \xi_j + i\eta_j , \tag{A42}$$

where the vectors ξ_j and η_j are real. By equating real and imaginary parts, it follows from Eqs. (A37) and (A38) that ξ_j and η_j obey the relations

$$(\xi_j , J\xi_k) = 0 \ , \ (\eta_j , J\eta_k) = 0, \tag{A43}$$

$$2(\xi_j , J\eta_k) = \delta_{jk} \ , \ 2(\eta_j , J\xi_k) = -\delta_{jk}. \tag{A44}$$

Similarly, the relations (A36) imply the relations

$$T\xi_j = -\lambda_j\eta_j \ , \ T\eta_j = \lambda_j\xi_j. \tag{A45}$$

Lemma 6. From the vectors ξ_j and η_j one can construct a symplectic matrix S such that the matrix $\tilde{S}ZS$ is diagonal. Specifically, $\tilde{S}ZS$ has the form

$$\tilde{S}ZS = \text{diag}(\lambda_1, \lambda_1, \lambda_2, \lambda_2, \lambda_3, \lambda_3). \tag{A46}$$

Proof. Define the matrix S by the relation

$$S = \sqrt{2}(\xi_1, \eta_1, \xi_2, \eta_2, \xi_3, \eta_3). \tag{A47}$$

Here each of the vectors ξ_j and η_j are to be viewed as column vectors so that the collection (A47) forms a 6×6 matrix. Then it is easily verified that the relations (A43) and (A44) are equivalent to the matrix relation

$$\tilde{S} J S = J. \tag{A48}$$

Thus S is symplectic as claimed.

From Eq. (A28) one finds the result

$$Z S = -J T S. \tag{A49}$$

Consequently, from Eqs. (A45) and (A47) one has the result

$$\begin{aligned} ZS = \sqrt{2}(& \lambda_1 J \eta_1, -\lambda_1 J \xi_1, \lambda_2 J \eta_2, \\ & -\lambda_2 J \xi_2, \lambda_3 J \eta_3, -\lambda_3 J \xi_3). \end{aligned} \tag{A50}$$

Finally, using Eqs. (A43), (A44), and (A50), one verifies that Eq. (A46) is also satisfied. This completes the proof of the lemma.

Evidently, the proof of the theorem we have been concerned with is now immediate. We simply take $M = \tilde{S}$. We note that if a matrix is symplectic, so is its tranpose [3]. It follows that M is symplectic and satisfies the relations (3.20) and (3.21).

We close this appendix by remarking that had we not required M to be symplectic but simply real, we could have found an M such that $Z^{(2)'}$ given by Eq. (3.20) would be the *identity* matrix. This observation follows from the fact that any positive definite matrix is real congruent to the identity matrix [15]. What is remarkable is that it is possible to achieve the simplification (3.21) using a symplectic matrix. Finally, we note that the simplification depended in a critical way on the positive definite property of Z. Indeed, the simplification may not be possible if Z is not positive definite [16].

APPENDIX B

In this appendix we discuss dynamic invariants constructed both out of moments and directly out of the phase-space coordinates. In particular we give a compact description and generalization of the Courant-Snyder invariants [17] that are widely used in the field of accelerator physics.

We call a function $\mathcal{I}(Z^{(k)})$ a *dynamic invariant* if the following equation is true *only* for a particular symplectic matrix M:

$$\mathcal{I}\left(\left(\bigotimes^{k} M\right) Z^{(k)}\right) = \mathcal{I}(Z^{(k)}). \tag{B1}$$

For simplicity, we will restrict ourselves to the case where k is equal to 2.

Consider the quantity $\mathcal{I}_2^{(n)}$ defined by the relation

$$\mathcal{I}_2^{(n)}(Z^{(2)}) = \mathrm{tr}[(Z^{(2)} J M)^n]. \tag{B2}$$

For this quantity to be kinematic invariant, it would have to remain unchanged under the following transformation [cf. Eqs. (B1) and (3.15)] for *all* $M' \in \mathrm{Sp}(6,R)$:

$$Z^{(2)} \to M' Z^{(2)} \tilde{M}'. \tag{B3}$$

However, as shown below, $\mathcal{I}_2^{(n)}$ remains unchanged in general only if M' is equal to M (where M is the same symplectic matrix that appears in the definition of $\mathcal{I}_2^{(n)}$). Substituting Eq. (B3) in Eq. (B2) we obtain the relation

$$\mathcal{I}_2^{(n)}(M' Z^{(2)} \tilde{M}') = \mathrm{tr}[(M' Z^{(2)} \tilde{M}' J M)^n]. \tag{B4}$$

If M' is equal to M, we can use the symplectic condition [Eq. (1.2)] and Eq. (3.17) to get the result

$$\mathcal{I}_2^{(n)}(M' Z^{(2)} \tilde{M}') = \mathcal{I}_2^{(n)}(Z^{(2)}). \tag{B5}$$

This proves that $\mathcal{I}_2^{(n)}$ is a dynamic invariant.

In fact, each kinematic moment invariant gives rise to a dynamic invariant by the following transformation

$$J \to J M. \tag{B6}$$

Comparing Eqs. (B2) and (3.11), we see that it was precisely this transformation that was used to obtain $\mathcal{I}_2^{(n)}$ from $I_2^{(n)}$.

The dynamic moment invariants constructed above may not be as interesting as kinematic moment invariants since they are invariant only for a given Hamiltonian system. However, dynamic invariants constructed directly using phase-space coordinates z rather than moments *do* have considerable interest. Let $\mathcal{Z}^{(k)}$ denote the kth-rank tensor constructed directly out of the phase-space coordinates z,

$$\mathcal{Z}^{(k)}_{i_1 i_2 \cdots i_k} = z_{i_1} z_{i_2} \cdots z_{i_k}. \tag{B7}$$

Again, for simplicity, we will consider only the case $k = 2$. Then the quantities $\mathrm{tr}[(\mathcal{Z}^{(2)} J)^n]$, which are the analogs of $I_2^{(2)}[Z^{(2)}]$, are all zero [18]. However, the analogs of the dynamic invariants defined by Eq. (B2) are not trivial. Call them $\mathcal{J}_2^{(n)}$. They are defined by the relations

$$\mathcal{J}_2^{(n)}(z) = \mathrm{tr}[(\mathcal{Z}^{(2)} J M)^n]. \tag{B8}$$

Direct calculation shows that all the $\mathcal{J}_2^{(n)}$ are related to $\mathcal{J}_2^{(1)}$ by the equation [19]

$$\mathcal{J}_2^{(n)}(z) = [\mathcal{J}_2^{(1)}(z)]^n. \tag{B9}$$

The quantity $\mathcal{J}_2^{(1)}$ is given in turn by the relation

$$\begin{aligned} \mathcal{J}_2^{(1)}(z) = \mathrm{tr}(\mathcal{Z}^{(2)} J M) &= \mathcal{Z}^{(2)}_{ij} J_{jk} M_{ki} \\ &= z_i z_j J_{jk} M_{ki} = (z, J M z). \end{aligned} \tag{B10}$$

Evidently $\mathcal{J}_2^{(1)}$, unlike $I_2^{(1)}$, is not zero. Moreover, direct calculation shows that it is invariant under the action of M as expected,

$$\begin{aligned} \mathcal{J}_2^{(1)}(Mz) = (Mz, JMMz) &= (z, \tilde{M} J M M z) \\ &= (z, JMz) = \mathcal{J}_2^{(1)}(z). \end{aligned} \tag{B11}$$

Here use has again been made of the symplectic condition Eq. (1.2). It follows from Eq. (B9) that all the $\mathcal{J}_2^{(n)}$ are dynamic invariants. However, they are all functionally dependent on $\mathcal{J}_2^{(1)}$.

We can obtain additional functionally independent dynamic invariants by slightly modifying (B10). Consider the quantities $\Xi^{(n)}(z)$ defined by the relations

$$\Xi^{(n)}(z) = (z, JM^n z). \tag{B12}$$

We note that $\Xi^{(1)}$ is identical to $\mathcal{J}_2^{(1)}$. By a calculation analogous to (B11), one sees that all the $\Xi^{(n)}$ are dynamic invariants. Evidently they also are all nonzero. We will next see that the $\Xi^{(n)}$ are related to generalized Courant-Snyder invariants in a case of particular physical interest.

Let $P^{(1)}, P^{(2)}$, and $P^{(3)}$ denote 6×6 matrices that project out the $(q_1, p_1), (q_2, p_2)$, and (q_3, p_3) subspaces, respectively. They have the form

$$P^{(1)} = \begin{pmatrix} I & & \\ & 0 & \\ & & 0 \end{pmatrix}, \quad P^{(2)} = \begin{pmatrix} 0 & & \\ & I & \\ & & 0 \end{pmatrix}, \quad P^{(3)} = \begin{pmatrix} 0 & & \\ & 0 & \\ & & I \end{pmatrix}. \tag{B13}$$

Here I denotes the 2×2 identity matrix, 0 denotes the 2×2 null matrix, and all other entries in the $P^{(i)}$ are zero. Define analogous matrices $J^{(i)}$ by the rule

$$J^{(i)} = JP^{(i)}. \tag{B14}$$

Suppose the matrix M can be brought to normal form N. That is, we assume there is a *symplectic* matrix A such that M is given by the relation

$$M = ANA^{-1} \tag{B15}$$

with N written in the form

$$N = \exp\left(\sum_{i=1}^{3} \mu_i J^{(i)}\right). \tag{B16}$$

Here the quantities μ_i are the (eigen) phase advances of M. Consider the quadratic forms $\Gamma_i(z)$ defined by the relations

$$\Gamma_i(z) = (A^{-1}z, P^{(i)}A^{-1}z). \tag{B17}$$

The quantities $\Gamma_i(z)$ are dynamic invariants that generalize the Courant-Snyder invariants [20].

To see that this is so, we compute $\Gamma_i(Mz)$ as shown below:

$$\begin{aligned} \Gamma_i(Mz) &= (A^{-1}Mz, P^{(i)}A^{-1}Mz) \\ &= (NA^{-1}z, P^{(i)}NA^{-1}z) \\ &= (A^{-1}z, \tilde{N}P^{(i)}NA^{-1}z) = (A^{-1}z, P^{(i)}A^{-1}z) \\ &= \Gamma_i(z). \end{aligned} \tag{B18}$$

Here we have used the relations

$$A^{-1}M = NA^{-1}, \tag{B19}$$

$$\tilde{N}P^{(i)}N = P^{(i)}\tilde{N}N = P^{(i)}, \tag{B20}$$

which follow from Eqs. (B15) and (B16), respectively.

Now insert Eq. (B15) into Eq. (B12). Doing so gives the result

$$\begin{aligned} \Xi^{(n)}(z) &= (z, JAN^nA^{-1}z) = (z, \tilde{A}^{-1}JN^nA^{-1}z) \\ &= (A^{-1}z, JN^nA^{-1}z). \end{aligned} \tag{B21}$$

Here we have used the symplectic condition for A written in the form

$$JA = \tilde{A}^{-1}J. \tag{B22}$$

Next we observe that JN^n can be written in the form

$$\begin{aligned} JN^n &= \sum_{i=1}^{3} JP^{(i)}N^n \\ &= \sum_{i=1}^{3} JP^{(i)}\exp(n\mu_i J^{(i)}) \\ &= \sum_{i=1}^{3} [J^{(i)}\cosh(n\mu_i J^{(i)}) + J^{(i)}\sinh(n\mu_i J^{(i)})] \\ &= \sum_{i=1}^{3} [J^{(i)}\cosh(n\mu_i J^{(i)}) - P^{(i)}\sin(n\mu_i)]. \end{aligned} \tag{B23}$$

Note that the matrices $J^{(i)}\cosh(n\mu_i J^{(i)})$ are antisymmetric, and therefore cannot contribute to the diagonal matrix element of the form $(A^{-1}z, *A^{-1}z)$ involved in Eq. (B21). Consequently, upon combining Eqs. (B21) and (B23), we see that the $\Xi^{(n)}$ can be written in the final form

$$\Xi^{(n)}(z) = -\sum_{i=1}^{3} \sin(n\mu_i)\Gamma_i(z). \tag{B24}$$

Thus, if M can be brought to the assumed normal form, the $\Xi^{(n)}$ are linear combinations of the generalized Courant-Synder invariants.

APPENDIX C

This appendix gives an alternative method for calculating the kinematic moment invariants. It presupposes a more detailed knowledge of the representation theory of Lie algebras on the part of the reader than was assumed in the main text.

It can be shown that Lie operators [3] $:w_i:$ corresponding to the 21 basis monomials w_i quadratic in the variables $q_1, p_1, q_2, p_2, q_3, p_3$, form a basis for the Lie algebra sp(6,R). It can also be shown that the Cartan subalgebra of sp(6,R) in this basis is given as [8]

$$\mathcal{H} = \{:q_1p_1:, :q_2p_2:, :q_3p_3:\}. \tag{C1}$$

Further, the basis elements $:p_1^2:$, $:p_2^2:$, and $:p_3^2:$ correspond to raising operators while $:q_1^2:$, $:q_2^2:$, and $:q_3^2:$ correspond to lowering operators.

Let $\{P_\alpha(z)\}$ $[\alpha = 1, 2, \ldots, N(m)]$ denote the set of basis monomials of a fixed degree m in the six phase-space variables. Here $N(m)$ is given by Eq. (3.45). These basis monomials give an N-dimensional irreducible representation $d^m(w_i)$ of sp(6,R) by [8]

$$:w_i: P_\alpha(z) = d^m(w_i)_\alpha^\beta P_\beta(z), \quad i = 1, 2, \ldots, 21. \tag{C2}$$

The set $\{P_\alpha(z)\}$ is said to form a basis for the carrier space of the representation $d^m(w_i)$.

In contrast to the more conventional choice for the basis elements of sp(6,R), the basis elements chosen above (i.e., $:w_i:$) have the special property that each $P_\alpha(z)$ corresponds to one unique weight vector λ_α. If the ordering of elements in $\mathcal{H}$ is fixed as given in Eq. (C1), this weight vector is given by the relation

$$\lambda_\alpha = (a_\alpha, b_\alpha, c_\alpha). \tag{C3}$$

Here a_α, b_α, and c_α are specified by the equations

$$\begin{aligned} :q_1p_1: P_\alpha(z) &= a_\alpha P_\alpha(z), \\ :q_2p_2: P_\alpha(z) &= b_\alpha P_\alpha(z), \\ :q_3p_3: P_\alpha(z) &= c_\alpha P_\alpha(z). \end{aligned} \tag{C4}$$

Consider a general element of $\{P_\alpha(z)\}$ given by the relation

$$P_\alpha(z) = q_1^{r_1} p_1^{r_2} q_2^{r_3} p_2^{r_4} q_3^{r_5} p_3^{r_6}, \tag{C5}$$

where $r_1 + r_2 + r_3 + r_4 + r_5 + r_6 = m$. Then, using Eq. (C4), we find the weight vector corresponding to $P_\alpha(z)$ to be as follows:

$$\lambda_\alpha = (r_2 - r_1, r_4 - r_3, r_6 - r_5). \tag{C6}$$

For example, it is seen that p_1^m, p_2^m, and $p_3^m \in \{P_\alpha(z)\}$ have weight vectors given by $(m,0,0)$, $(0,m,0)$, and $(0,0,m)$, respectively. Therefore the monomial p_1^m corresponds to the "highest" weight vector.

The above concepts can be applied to moments as follows. Each $P_\alpha(z)$ corresponds to a unique mth-order basis moment $\langle P_\alpha(z)\rangle$ through the relation [cf. Eq. (2.1)]

$$\langle P_\alpha(z)\rangle = \int d^6z'\, h(z')\, P_\alpha(z'). \tag{C7}$$

Further, the action of the basis elements of sp(6,R) on these basis moments is given as

$$:w_i: \langle P_\alpha(z)\rangle = \langle :w_i: P_\alpha(z)\rangle. \tag{C8}$$

Therefore we can assign to $\langle P_\alpha(z)\rangle$ the same weight vector λ_α associated with $P_\alpha(z)$. Moreover, if one has a product of moments $\langle P_\alpha\rangle\langle P_\beta\rangle \ldots \langle P_\omega\rangle$, its weight vector is given by $\lambda_\alpha + \lambda_\beta + \cdots + \lambda_\omega$.

With this brief background, we are now in a position to construct kinematic moment invariants. The construction goes as follows. The most general $I_m^{(n)}$ is given by a relation of the form

$$I_m^{(n)} = C_{\alpha_1\alpha_2\cdots\alpha_n}\langle P_{\alpha_1}\rangle\langle P_{\alpha_2}\rangle \cdots \langle P_{\alpha_n}\rangle \tag{C9}$$

with

$$\alpha_1 \le \alpha_2 \le \cdots \le \alpha_n. \tag{C10}$$

Being an invariant implies that $I_m^{(n)}$ satisfies the following set of equations:

$$:w_i: I_m^{(n)} = 0, \quad i = 1,2,\ldots,21. \tag{C11}$$

It is these annihilation conditions that uniquely determine $I_m^{(n)}$ (up to an overall scale factor).

We first consider the action of the Cartan basis elements $:q_1p_1:$, $:q_2p_2:$, and $:q_3p_3:$ on $I_m^{(n)}$. This gives the following results [cf. Eqs. (C8) and (C4)]:

$$:q_ip_i: I_m^{(n)} = C_{\alpha_1\alpha_2\ldots\alpha_n}\lambda^i_{\alpha_1\alpha_2\ldots\alpha_n}\langle P_{\alpha_1}\rangle\langle P_{\alpha_2}\rangle \ldots \langle P_{\alpha_n}\rangle, \quad i = 1,2,3 \tag{C12}$$

where the weight vector $\lambda_{\alpha_1\alpha_2\cdots\alpha_n}$ satisfies the relation [cf. the discussion following Eq. (C8)]

$$\lambda_{\alpha_1\alpha_2\cdots\alpha_n} = \lambda_{\alpha_1} + \lambda_{\alpha_2} + \cdots + \lambda_{\alpha_n}, \tag{C13}$$

and $\lambda^i_{\alpha_1\alpha_2\cdots\alpha_n}$ is the ith component of this weight vector. Setting Eq. (C12) to zero, we see that $C_{\alpha_1\alpha_2\ldots\alpha_n}$ is zero unless all three components of the corresponding weight vector $\lambda^i_{\alpha_1\alpha_2\ldots\alpha_n}$ happen to be zero. Therefore we are led to the following result:

$$I_m^{(n)} = \sum_i A_i W_i^{(m)}, \tag{C14}$$

where each $W_i^{(m)}$ is a product of n mth-order basis moments such that its weight vector is zero. This implies that each $W_i^{(m)}$ has equal powers of q's and p's [cf. Eq. (C6)]. Now the remaining annihilation conditions [cf. Eq. (C11)] can be used to determine the coefficients A_i (up to an overall scale factor).

As an example, let us calculate $I_4^{(2)}$ for a two-dimensional phase space. Using Eq. (C14) we get

$$I_4^{(2)} = A_1\langle q_1^4\rangle\langle p_1^4\rangle + A_2\langle q_1^3p_1\rangle\langle q_1p_1^3\rangle + A_3\langle q_1^2p_1^2\rangle^2. \tag{C15}$$

Next, we need to evaluate Eq. (C11) for w_i equal to q_1^2 and p_1^2. Since each term in $I_2^{(4)}$ has equal powers of q_1 and p_1, both these annihilation conditions give exactly the same result. Therefore we need to evaluate only one of them. Consider the following condition:

$$:q_1^2: I_4^{(2)} = 0. \tag{C16}$$

Substituting Eq. (C15) in the above equation, we obtain the result [cf. Eq. (C8)]

$$\begin{aligned} :q_1^2: I_4^{(2)} = {} & (8A_1 + 2A_2)\langle q_1^4\rangle\langle q_1p_1^3\rangle \\ & +(6A_2 + 8A_3)\langle q_1^3p_1\rangle\langle q_1^2p_1^2\rangle = 0. \end{aligned} \tag{C17}$$

Setting the coefficient of each independent term to zero, we get the following relations:

$$A_2 = -4A_1, \quad A_3 = 3A_1. \tag{C18}$$

Choosing the normalization $A_1 = 1$, the final expression for the invariant is found to be

$$I_4^{(2)} = \langle q_1^4\rangle\langle p_1^4\rangle + 3\langle q_1^2p_1^2\rangle^2 - 4\langle q_1^3p_1\rangle\langle q_1p_1^3\rangle. \tag{C19}$$

This is seen to agree with the expression given in Eq. (3.35).

For working out invariants by hand, the construction method outlined above is sometimes quite convenient. Moreover, it brings out certain properties of the invariants that were not obvious in our earlier construction. For example, we see now that each term in the expression for a kinematic moment invariant should have equal powers of q's and p's.

Finally, we remark that the sum of the coefficients A_i must be equal to zero for these invariants. This can be seen from the following argument. Suppose the distribution function $h^{in}(z)$ is a δ function whose support is such that one has the relation

$$\langle z_1 \rangle = \langle z_2 \rangle = \cdots \langle z_6 \rangle = 1. \tag{C20}$$

For such a distribution one also has the relation

$$\langle z_{i_1} z_{i_2} \cdots z_{i_k} \rangle = \langle z_{i_1} \rangle \langle z_{i_2} \rangle \cdots \langle z_{i_k} \rangle = 1. \tag{C21}$$

Consequently, for such a distribution, the invariant $I_m^{(n)}$ must have the value

$$I_m^{(n)} = \sum_i A_i. \tag{C22}$$

Observe that a δ-function distribution corresponds to the case of a single-particle distribution. Thus, as already seen, all higher-order moments can be calculated in terms of first-order moments. Moreover, any given set of first-order moments can be transformed into any other set of first order moments using Eq. (3.46). This result follows from the fact that Sp(6,R) acts transitively on phase space. That is, under the action of the symplectic group Sp(6,R), any nonzero point in phase space can be sent to any other nonzero point. Consequently, any set of first-order moments is equivalent to the set $\langle z_i \rangle'$ where

$$\begin{gathered} \langle z_i \rangle' = 0, \quad i = 1, 2, \ldots, 5 \\ \langle z_6 \rangle' \text{ is equal to any desired nonzero number.} \end{gathered} \tag{C23}$$

But $I_m^{(n)}$ is a class function. Therefore, by the arguments above, its value depends only on $\langle z_6 \rangle'$. Moreover, since $\langle z_6 \rangle'$ can be made arbitrarily small, $I_m^{(n)}$ must vanish for the δ-function distribution. It follows from Eq. (C22) that the sum of the coefficients A_i must be zero.

---

[1] A. J. Dragt, R. L. Gluckstern, F. Neri, and G. Rangarajan, in *Frontiers of Particle Beams; Observation, Diagnosis and Correction*, edited by M. Month and S. Turner, Lecture Notes in Physics Vol. 343 (Springer-Verlag, Berlin, 1989), p. 94.

[2] F. Neri and G. Rangarajan, Phys. Rev. Lett. **64**, 1073 (1990); W.P. Kells, *ibid.* **65**, 2205 (1990).

[3] For a review of the properties of symplectic matrices, see A. J. Dragt, in *Physics of High Energy Particle Accelerators*, edited by R. A. Carrigan, F. R. Huson, and M. Month, AIP Conf. Proc. No. 87 (American Institute of Physics, New York, 1982), p. 147; see also Ref. [20] below.

[4] P. Lapostolle, IEEE Trans. Nucl. Sci. **18**, 1101 (1971); F. Sacherer, *ibid.* **18**, 1105 (1971).

[5] A similar expression for this invariant was discovered independently by D. Holm, W. Lysenko, and C. Scovel, J. Math. Phys. **31**, 1610 (1990).

[6] The kinematic invariants $I_2^{(2)}$ and $I_4^{(2)}$ (to be defined later) were first discovered by W. Lysenko using a different method. See W. Lysenko and M. Overly, in *Proceedings of the Workshop on Linear Accelerator and Beam Optics Codes, San Diego, CA, 1988*, edited by C. R. Eminhizer, AIP Conf. Proc. No. 177 (American Institute of Physics, New York, 1988), p. 323; see also W. Lysenko, Los Alamos National Laboratory Technical Report No. LA-UR-88-165, 1988 (unpublished).

[7] In the field of accelerator physics, such linear terms in the Hamiltonian correspond to the presence of misaligned, displaced, or mispowered optical elements in the beam transport system. See, for example, L. M. Healy, A. J. Dragt, and I. Gjaja, J. Math. Phys. (to be published).

[8] G. Rangarajan, Ph. D. thesis, University of Maryland, 1990.

[9] G. Birkhoff and S. Mac Lane, *A Survey of Modern Algebra*, 3rd ed. (Macmillan, New York, 1965), pp. 213 and 214.

[10] To make these calculations, we chose v at random using the results of a random number generator for its entries. We then determined the dimension of $U(v)$ numerically by counting the number of linearly independent vectors resulting from the action of sp(6,R) on v. This was done repeatedly for several different vectors v obtained from different random number seeds, and identical results were obtained in each case. Note that by this procedure we are able to determine only the dimension of what might be called a typical or generic equivalence class. That is, by this procedure we are unable to determine whether or not there may also exist additional atypical equivalence classes arising from possibly atypical vectors $\overline{v} \in V^{(m)}$ having smaller measure in $V^{(m)}$. One atypical case is that for which all entries in $\overline{v}$ are equal. In this case we found that the dimension of $U(\overline{v})$ is 6 for all m. This result could have been anticipated. Moreover, it can be shown that all invariants vanish for this case. See Appendix C. It remains to be determined whether there are or are not other atypical vectors of greater interest.

[11] J. Moser, Commun. Pure Appl. Math. **11**, 81 (1958).

[12] A. Weinstein, Bull. Am. Math. Soc. **75**, 814 (1971).

[13] This correspondence between symplectic congruency and symplectic similarity has been used previously by other authors. See Ref. [11] above and the paper of E.D. Courant that appeared in *Perspectives in Modern Physics, Essays in Honor of H.A. Bethe*, edited by R.E. Marshak (Interscience, New York, 1966), p. 257.

[14] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed. (Johns Hopkins University Press, Baltimore, 1989).

[15] This follows from Sylvester's law of inertia. See pp. 251, 254, and 259 of Ref. [9] above. See also Eqs. (2.24) and (2.25) of Ref. [1] above.

[16] For a discussion of what is possible in the general case, see Refs. [11] and [12] above and the following additional references: J. Williamson, Am. J. Math. **58**, 141 (1936); **59**, 599 (1937); **61**, 897 (1939); A. Wintner, Ann. Mat. **13**, 105 (1934); N. Burgoyne and R. Cushman, Celes. Mech. **8**, 435 (1974); N. Burgoyne and R. Cushman, in

Proceedings of the 1976 Ames Research Center (NASA) Conference on Geometric Control Theory, edited by C. Martin and R. Hermann (Math-Sci, Brookline, 1977); A. J. Laub and K. Meyer, Celes. Mech. **9**, 213 (1974); R. Abraham and J. E. Marsden, *Foundations of Mechanics*, 2nd ed. (Benjamin/Cummings, Reading, MA, 1978), pp. 491 and 492.

[17] E. Courant and H. Snyder, Ann. Phys. (N.Y.) **3**, 1 (1958).

[18] This can be seen in at least two ways. First, by direct computation, one finds that each J_{ij} is contracted with a term of the form $z_i z_j$. Since J is antisymmetric and $z_i z_j$ is symmetric, the result is zero. A more profound explanation consists of observing that the symplectic group acts transitively on phase space, and hence there cannot be any nontrivial kinematic invariants constructed only out of phase-space variables. See Appendix C.

[19] This calculation is simplified by the use of dyadic notation. In this notation, $\mathcal{Z}^{(2)}$ is given by the expression $\mathcal{Z}^{(2)} = |z)(z|$. Correspondingly, one finds the result $\mathrm{tr}[(\mathcal{Z}^{(2)}JM)^n] = \mathrm{tr}[|z)(z|JM|z)(z|JM \cdots |z)(z|JM] = [(z|JM|z)]^{n-1}\,\mathrm{tr}[|z)(z|JM] = [(z|JM|z)]^n = [\mathcal{J}_2^{(1)}(z)]^n$.

[20] A. J. Dragt *et al.*, Annu. Rev. Nucl. Part. Sci. **38**, 455 (1988).