Separation of Motions of Many-Body Systems into Dynamically Independent Parts by Projection onto Equilibrium Varieties in Phase Space. II

G. CARMI*
Belfer Graduate School of Science, Yeshiva University, New York, New York

AND

D. BOHM
Birkbeck College, London, England

(Received 13 August 1962; revised manuscript received 10 July 1963)

We continue the development started in the preceding paper, in which we treated the many-body problem by separating the motion into an oscillatory part $\delta\mathbf{x}^i, \delta\mathbf{p}_i$, and a nonoscillatory part $\mathbf{X}^i$, $\mathbf{P}_i$, the latter being obtained by a noncanonical transformation from $\mathbf{x}^i$, $\mathbf{p}_i$ which is just so tailored as to project out the oscillatory features from $\mathbf{x}^i$, $\mathbf{p}_i$, and thereby "projecting" $\mathbf{x}^i$, $\mathbf{p}_i$ onto the equilibrium variety $J_{\mathbf{k}}(\mathbf{x}^i, \mathbf{p}_i)=0$ (where $J_{\mathbf{k}}$ is the oscillatory action variable) in phase space. In this paper, we first develop a condensed notation in phase space which facilitates calculations. With the aid of this notation, we then give our results a simple geometrical interpretation in phase space by introducing a certain canonically invariant metrical tensor O_{ij}. This tensor (which is antisymmetric) does not yield the usual orthogonal or pseudo-orthogonal metric, but rather, what is called a "symplectic metric" (i.e., invariant to the symplectic group of transformations). One then sees that the projections that we make are "orthogonal," in the symplectic sense, to the equilibrium varieties. Likewise, one can see quite generally, that the entire canonical formalism, including the Poisson brackets and the Hamiltonian equations of motion, reduces to simple geometrical relations in phase space, the form of which is suggestive for possible further developments, especially with regard to the treatment in higher approximations. We apply our ideas to the electron gas, and illustrate the dynamics of the plasma with the aid of a comparison with a simple two-dimensional model, possessing all the essential features described above. In this way, we are able to understand many of the basic features of the plasma motions, in terms of concepts such as the generalization of the notion of centrifugal force and Coriolis force to phase space. By going over to a local geodetic frame in the equilibrium variety, we are led in a natural way to the concept of a set of "quasiparticles" for the plasma. If the number of collective oscillatory coordinates is s, then there will be $3N-s$ of these "quasiparticle" coordinates. The latter do not represent any of the actual original particles out of which the system is constituted, but rather, they represent effective pulse-like distributions of charge, which move together in a correlated way so as to resemble an actual particle in many respects.

1. INTRODUCTION

IN a previous paper[1] (to be denoted by I), we have developed a method for drawing conclusions about the over-all dynamical properties of a many-body system on the basis of a knowledge of oscillatory or collective variables. In particular, we have shown that if there are canonical pairs $Q_{\mathbf{k}}$ and $P_{\mathbf{k}}$, oscillating with angular frequencies $\omega_{\mathbf{k}}$, then there exists a separation of the motion $\mathbf{x}^i(t)$, $\mathbf{p}_i(t)$ into two parts. One part, $\delta\mathbf{x}^i$, $\delta\mathbf{p}_i$ given by Eq. (3.6) of I, is purely oscillatory, and the other part $\mathbf{X}^i$, $\mathbf{P}_i$ given by Eq. (3.5) of the same paper, is the purely nonoscillatory part, which was seen to correspond to a special equilibrium solution, so chosen that $\delta\mathbf{x}^i$, $\delta\mathbf{p}_i$ will never contain secular perturbations.

We shall begin by reformulating the results of I in a condensed notation which, firstly, will make it easier to treat the dynamical problem and secondly, will be needed for the discussion of the geometrical ideas to come later. We will show in terms of this notation that the equations of motion of the purely nonoscillatory part $\mathbf{X}^i$, $\mathbf{P}_i$ are the same as those of an actual motion of the systems, thus justifying our use of $\mathbf{X}^i$, $\mathbf{P}_i$ as a "comparison motion" which remains near to the actual motion.

We then go on to give an interpretation of our projective method in terms of the geometry of phase space. By a consideration of the expressions for the Poisson brackets and the equations of motion, we are led to introduce into phase space a certain metrical tensor, which is invariant to a general canonical transformation. This tensor does not yield the usual kind of orthogonal or pseudo-orthogonal metric, but rather, what is called a "symplectic"[2] metric (i.e., a metric invariant to the group of symplectic transformations). With the aid of this tensor, our method of separating variables into dynamically independent parts can be interpreted as the dropping of a "perpendicular," in the symplectic sense, from an arbitrary phase point $\mathbf{x}^i$, $\mathbf{p}_i$ to the "projected point" $\mathbf{X}^i$, $\mathbf{P}_i$ in the equilibrium variety. Thus, the metric specializes the projection from a general one to a "perpendicular" projection in the symplectic sense. This procedure, which is canonically invariant (and, therefore, consistent with the time

* This research was supported by the Department of Scientific and Industrial Research of the British Government while the author was at Bristol University, England, and by the U. S. Air Force Office of Scientific Research and the National Science Foundation.

[1] D. Bohm and G. Carmi, Phys. Rev. 133, A319 (1964), hereafter referred to as I.

[2] F. D. Murnaghan, Comm. Dublin Institute Advanced Studies, Ser. A, No. 13, 1958 (unpublished).

evolution of the system) can thus constitute a simple geometrical interpretation of the split in the motion.

In Secs. 4, 5, 6, we then apply these ideas to a discussion of the electron gas. In particular, we show that our geometrical notions allow a simple and intuitively highly suggestive interpretation of the dynamics of the electron plasma in terms of concepts such as Coriolis and centrifugal forces in phase space, stable spiraling motion on a "hypercylinder," about an equilibrium variety as "axis," etc. A comparison with our simple two-dimensional model of Sec. 2 of I shows that the latter reflects to quite an appreciable degree the essential physical features of a many-body problem, as looked upon from the point of view of collective coordinates (or conservation rules).

2. DISCUSSION OF SEPARATION OF THE MOTION IN PHASE-SPACE NOTATION

Let us begin by recalling the definition of oscillatory and nonoscillatory parts of the motion, which correspond to the existence of s canonical pairs Q_k, P_k of oscillatory variables. [See Eqs. (I-3.5) and (I-3.6)]:

$$\delta\mathbf{x}^i=\sum_\mathrm{k}\left(\frac{\partial P_\mathrm{k}}{\partial\mathbf{p}_i}Q_\mathrm{k}-\frac{\partial Q_\mathrm{k}}{\partial\mathbf{p}_i}P_\mathrm{k}\right),\qquad \delta\mathbf{p}_i=\sum_\mathrm{k}\left(\frac{\partial Q_\mathrm{k}}{\partial\mathbf{x}^i}P_\mathrm{k}-\frac{\partial P_\mathrm{k}}{\partial\mathbf{x}^i}Q_\mathrm{k}\right),\tag{2.1a}$$

$$\delta F=\sum_\mathrm{k}\left([F,P_\mathrm{k}]Q_\mathrm{k}-[F,Q_\mathrm{k}]P_\mathrm{k}\right),\tag{2.1b}$$

$$\mathbf{X}^i=\mathbf{x}^i+\sum_\mathrm{k}\left(\frac{\partial Q_\mathrm{k}}{\partial\mathbf{p}_i}P_\mathrm{k}-\frac{\partial P_\mathrm{k}}{\partial\mathbf{p}_i}Q_\mathrm{k}\right),\qquad \mathbf{P}_i=\mathbf{p}_i+\sum_\mathrm{k}\left(\frac{\partial P_\mathrm{k}}{\partial\mathbf{x}^i}Q_\mathrm{k}-\frac{\partial Q_\mathrm{k}}{\partial\mathbf{x}^i}P_\mathrm{k}\right).\tag{2.1c}$$

In order to avoid the need to write out $\mathbf{x}^i$, $\mathbf{p}_i$, P_k, and Q_k separately, we now introduce another notation, in which we represent the phase point by a single symbol z^i, having twice as many dimensions as there are degrees of freedom. The first set of indices (up to $3N$, the total number of degrees of freedom) will be taken to represent the coordinates $\mathbf{x}^i$ while the second set (equal in number) will represent the momenta $\mathbf{p}_i$. Similarly, we represent the oscillatory variables Q_k, P_k, by the symbol Q^μ, where the index μ has $2s$ values, the first s values representing the Q_k and the second s values representing the P_k. In this notation, the Poisson brackets of any two variables F and G have the simplified form

$$[F,G]=\sum_{i=1}^{3N}\left(\frac{\partial F}{\partial\mathbf{x}^i}\frac{\partial G}{\partial\mathbf{p}_i}-\frac{\partial F}{\partial\mathbf{p}_i}\frac{\partial G}{\partial\mathbf{x}^i}\right)=\sum_{i,j=1}^{6N}O^{ij}\frac{\partial F}{\partial z^i}\frac{\partial G}{\partial z^j},\tag{2.2}$$

where O^{ij} is a matrix, whose elements are all zero, except those connecting $\mathbf{x}^i$ and $\mathbf{p}_i$. These latter are ±1, being $+1$ if the index i is associated with a coordinate $\mathbf{x}^i$ and j with momentum $\mathbf{p}_j$, and -1 if the association is the other way. Or, in matrix form,

$$O=\begin{bmatrix} & & & 1 & 0 & 0\cdots \\ & & & 0 & 1 & 0 \\ & 0 & & 0 & 0 & 1 \\ & & & \cdot & & \cdot \\ & & & \cdot & & \cdot \\ & & & \cdot & & \cdot \\ -1 & 0 & 0\cdots & & & \\ 0 & -1 & 0 & & & \\ 0 & 0 & -1 & & 0 & \\ \cdot & & \cdot & & & \\ \cdot & & \cdot & & & \\ \cdot & & \cdot & & & \end{bmatrix}.\tag{2.2a}$$

Henceforth, we shall make use of the summation convention, viz., "dummy" indices appearing both in an upper indexed quantity and a lower indexed quantity are to be summed over. Thus,

$$[F,G]=O^{ij}\frac{\partial F}{\partial z^i}\frac{\partial G}{\partial z^j}.\tag{2.2b}$$

Here, it should be noted that previously only configuration space could be endowed with an invariant geometrical meaning, and hence $\mathbf{p}_i$ had to be visualized as a vector imbedded in configuration space, and in that case, it was necessary to regard $\mathbf{p}_i$ as a covariant vector, to be represented by a lower index. But now, $\mathbf{p}_i$ is being taken on the same footing as $\mathbf{x}^i$, the two together forming a $6N$-component vector z^i in phase space, the invariant geometrical meaning of which will be discussed in Sec. 3. z^i is then the contravariant vector in phase space, while a typical covariant vector is given by $(\partial F)/(\partial z^i)$, where F is a "scalar" function.

We shall also have at times to express sums of upper indexed quantities, and in order to facilitate this, we shall introduce the symbol O_{ij}, defined by

$$O_{ij}O^{jk}=\delta_i{}^k.\tag{2.3}$$

By an elementary calculation, one obtains

$$O_{ij}=-O^{ij}.\tag{2.3a}$$

In this notation, the equations of motion for a general dynamical system take the form

$$\dot{z}^i=O^{ij}(\partial H/\partial z^j),\tag{2.4a}$$

$$\dot{F}=O^{ij}(\partial F/\partial z^i)(\partial H/\partial z^j),\tag{2.4b}$$

where F is an arbitrary function of z^i.

We now apply this notation to our formulas of separation of the motion into parts. First, we apply (2.2b) to Q^μ, obtaining

$$[Q^\mu,Q^\nu]=O^{ij}(\partial Q^\mu/\partial z^i)(\partial Q^\nu/\partial z^j)=O^{\mu\nu},\tag{2.5}$$

where $O^{\mu\nu}$ is the same kind of antisymmetric matrix as O^{ij}, except that its indices apply to the set of canonical variables Q^μ instead of the set z^i and, therefore, the matrix $O^{\mu\nu}$ is a $2s\times 2s$ matrix only. In a similar way, we can introduce the lower index symbol

$$O_{\mu\nu}=-O^{\mu\nu}, \tag{2.5a}$$

which fulfills

$$O_{\mu\nu}O^{\nu\alpha}=\delta_\mu{}^\alpha. \tag{2.5b}$$

With this notation, the equations (2.1) take the more abbreviated form

$$\delta z^i=O^{ij}O_{\mu\nu}Q^\mu(\partial Q^\nu/\partial z^j), \tag{2.6a}$$

$$\delta F=O^{ij}O_{\mu\nu}Q^\mu(\partial F/\partial z^i)(\partial Q^\nu/\partial z^j), \tag{2.6b}$$

$$Z^i=z^i-O^{ij}O_{\mu\nu}Q^\mu(\partial Q^\nu/\partial z^j). \tag{2.6c}$$

The above is, of course, correct only in the linear approximation. To go on to higher orders, we note that an infinitesimal canonical transformation with generating function V can be expressed, in the present notation, as

$$(z^i)'-z^i=\lambda O^{ij}(\partial V/\partial z^j). \tag{2.7}$$

By repeated application of this transformation and by finally allowing the parameter λ to vary with z^i, we can show that the nonlinear transformation (I-4.6) is

$$Z^i=z^i+O^{ij}O_{\mu\nu}Q^\mu\frac{\partial Q^\nu}{\partial z^j}+\tfrac{1}{2}O^{ij}O_{\mu\nu}O^{kl}O_{\alpha\beta}Q^\mu Q^\alpha\frac{\partial Q^\beta}{\partial z^l}\frac{\partial^2 Q^\alpha}{\partial z^k\partial z^j}+\cdots, \tag{2.8a}$$

$$[\![F]\!]=F(Z^i)=F(z^i)+O^{ij}O_{\mu\nu}Q^\mu\frac{\partial Q^\nu}{\partial z^j}\frac{\partial F}{\partial \mathbf{z}^i}+\tfrac{1}{2}O^{ij}O_{\mu\nu}O^{kl}O_{\alpha\beta}Q^\mu Q^\alpha\frac{\partial Q^\beta}{\partial \mathbf{z}^l}\frac{\partial}{\partial \mathbf{z}^k}\left(\frac{\partial Q^\nu}{\partial z^j}\frac{\partial F}{\partial z^i}\right)+\cdots. \tag{2.8b}$$

We also could have obtained (2.8), as in I, by means of a Taylor expansion in Q_k, P_k around the point z^i. Thus, the coefficients of the various powers of Q^μ in these expressions should be functions evaluated at the expansion-point z^i. However, it is often convenient to have an expansion in which the coefficients are evaluated at Z^i. If (as is usually the case) we need only go to second order in Q^μ, then the coefficients of $Q^\mu Q^\alpha$ can already be evaluated at Z^i [and thus be replaced, as in the notation introduced in I (Sec. 3), by "$[\![\ ]\!]$" quantities]. However, the coefficients of Q^μ may contain first-order terms. Thus, in (2.8b) we can take $G_{ij}{}^\nu=(\partial Q^\nu/\partial z^j)(\partial F/\partial z^i)$ and expand it by applying (2.8b) again, setting $F=G$. Here, it will be sufficient to go to first order in Q, because G is already multiplied by Q^μ. We have

$$[\![G]\!]\equiv G(Z^i)=G(z^i)+O^{ij}O_{\mu\nu}Q^\mu\frac{\partial Q^\nu}{\partial z^j}\frac{\partial G}{\partial z^i}\equiv G(z^i)+O^{ij}O_{\mu\nu}Q^\mu\left[\!\!\left[\frac{\partial Q^\nu}{\partial z^j}\right]\!\!\right]\left[\!\!\left[\frac{\partial G}{\partial z^i}\right]\!\!\right]. \tag{2.9a}$$

We see that the effect of inserting the value of $G(z^i)$ obtained from this equation into the first-order term of (2.8b) is just to reverse the sign of the second-order term and to take the "$[\![\ ]\!]$" of all coefficients. Thus,

$$F(z^i)=[\![F]\!]-O^{ij}O_{\mu\nu}Q^\nu\left[\!\!\left[\frac{\partial Q^\mu}{\partial z^j}\right]\!\!\right]\left[\!\!\left[\frac{\partial F}{\partial z^i}\right]\!\!\right]+\tfrac{1}{2}O^{ij}O_{\mu\nu}O^{kl}O_{\alpha\beta}Q^\mu Q^\alpha\left[\!\!\left[\frac{\partial Q^\beta}{\partial z^l}\right]\!\!\right]\left[\!\!\left[\frac{\partial G_{ij}{}^\nu}{\partial z^k}\right]\!\!\right]. \tag{2.9b}$$

The application of the above to the case $F(z^i)=z^i$ yields

$$z^i=Z^i-O^{ij}O^{\mu\nu}Q_\nu\left[\!\!\left[\frac{\partial Q_\mu}{\partial z^j}\right]\!\!\right]+\tfrac{1}{2}O^{ij}O_{\mu\nu}O^{kl}O_{\alpha\beta}Q^\mu Q^\alpha\left[\!\!\left[\frac{\partial Q^\beta}{\partial z^l}\right]\!\!\right]\frac{\partial^2 Q^\nu}{\partial z^k\partial z^j}. \tag{2.9c}$$

As a special case, we can apply the above equations to the Hamiltonian $H(z^i)$. We begin by applying (2.9b), with $F=H$. The first-order term of the resulting expansion will be proportional to $O^{ij}[\![(\partial Q^\mu/\partial z^i)(\partial H/\partial z^j)]\!]$. By Eq. (2.4a), this is equal to $[\![(dz^i/dt)(\partial Q^\mu/\partial z^i)]\!]\equiv[\![dQ^\mu/dt]\!]$, and by definition $[\![dQ^\mu/dt]\!]=0$, since the equilibrium hypersurface $Q^\mu=0$ is, by hypothesis, a constant of motion. Thus, the first-order term in the expansion (2.9b) of H must vanish. As a result [as we have already indicated in I (Sec. 4)], the expansion of the Hamiltonian begins with second-order terms. We have (to second order):

$$H(z^i)\cong H(Z^i)+\tfrac{1}{2}O^{ij}O_{\mu\nu}O^{kl}O_{\alpha\beta}\times\left[\!\!\left[\frac{\partial Q^\beta}{\partial z^l}\frac{\partial}{\partial z^k}\left(\frac{\partial Q^\nu}{\partial z^j}\frac{\partial H}{\partial z^i}\right)\right]\!\!\right]Q^\mu Q^\alpha. \tag{2.10}$$

The coefficients are, in general, slowly varying functions of Z^i (in many cases of physical interest they do, in fact, turn out to be constants).

From the above (approximate) Hamiltonian one can, in the manner described in I (Sec. 4), obtain the equations of motion, provided that one first evaluates the Poisson brackets $[Z^a,Z^b]$. Thus, the equations of motion of the Z^a can be obtained from $\dot{Z}^a=[Z^a,H]$ by using (2.2b) and (2.10). Since Z^a has a zero Poisson bracket with Q^μ, the second part in (2.10) does not contribute in the linear approximation, and we have

$$\dot{Z}^a = O^{ij}\frac{\partial H(Z^b)}{\partial z^j}\frac{\partial Z^a}{\partial z^i}$$

$$= O^{ij}\left[\!\left[\frac{\partial H}{\partial Z^b}\right]\!\right]\frac{\partial Z^b}{\partial z^j}\frac{\partial Z^a}{\partial z^i} = \frac{\partial [\![H]\!]}{\partial Z^b}[Z^b,Z^a]. \quad (2.11)$$

In Appendix A it is shown that

$$[Z^b,Z^a] = O^{ij}\frac{\partial Z^b}{\partial z^i}\frac{\partial Z^a}{\partial z^j} = O^{ij}P_i{}^b P_j{}^a, \quad (2.12)$$

where $P_j{}^a = (\partial Z^a)/\partial z^j$ is a projection matrix which projects to zero any vector u_a which is normal to any of the surfaces $Q^\alpha = 0$, and leaves unchanged any vector which is normal to $\xi = \text{const.}$, ξ being any of the residual (noncollective) variables. Since $H(Z)$ is a function of the ξ only, $\partial H(Z)/\partial Z^b$ will be unchanged by $P_i{}^b$, i.e.,

$$\dot{Z}^a = O^{ij}P_j{}^a P_j{}^b P_i{}^b \frac{\partial [\![H]\!]}{\partial Z^b} = O^{ij}P_j{}^a \frac{\partial [\![H]\!]}{\partial Z^i}.$$

It is further shown in Appendix A that the matrix $O^{ij}P_j{}^a$ (contravariant in its two indices i and a) projects to zero any covectors which are normal to the surfaces $Q^\alpha = 0$, and turns any vector which is normal to $\xi = \text{const.}$, into its corresponding contravector. Thus,

$$\dot{Z}^a = O^{ai}(\partial [\![H]\!]/\partial Z^i). \quad (2.13)$$

Remembering that $[\![H]\!] = H(Z)$, i.e., the original Hamiltonian evaluated at the projected point Z, and that (2.4a) are Hamilton's equations of motion in our condensed notation, we see from (2.13) that the projected motion $Z^i(t)$ *is itself a possible motion of the system*. This was shown here by expanding the Hamiltonian (2.10) about the equilibrium surface and using the correct Poisson brackets. As was already emphasized in I (Sec. 4), this result could have been foreseen by noting that $Q^\alpha = 0$ will be a possible solution for the oscillatory variables, in which case we will have $Z^i = z^i$, so that Z^i is a possible solution.

The advantage of writing the equations of motion in the form (2.11) will become clear once the full power of the geometrical interpretation (to be developed in Sec. 3) is utilized for further developments of the theory.

3. GEOMETRICAL INTERPRETATION OF THE SEPARATION METHOD IN PHASE SPACE

In this section, we shall show that our method of separating dynamical variables into oscillatory and nonoscillatory parts can be interpreted as a factorization of phase space into a direct product of mutually orthogonal subspaces, provided that the orthogonality is defined on the basis of what may be called a "symplectic" metric, whose (canonically invariant) metrical tensor is given by the (antisymmetric) O^{ij} matrix defined in (2.2a).

A very simple special case of an orthogonality relationship between the two components of the motion was encountered in the discussion [at the end of I (Sec. 3)], in which the separation of variables (I-3.5) and (I-3.6) was applied to the example given in I (Sec. 2), for case (A) (the case in which the angular momentum p_θ is small enough so that centrifugal and Coriolis forces can be neglected, while the equilibrium variety, $r = r_0$, is independent of the state of motion). It was shown there that the transformation (I-3.6) is equivalent to a projection from a point x^i (in the configuration space) along a normal to the equilibrium variety down to the point, X^i, of intersection of the normal in question with this variety (with a similar interpretation for the relation between p^i and P_i, which need not, however, be discussed in detail for the purposes of the present treatment). The projection was orthogonal in the ordinary Euclidean sense because in this simple case, the oscillatory motion δx^i is normal to the equilibrium variety, which latter can be described in configuration space only. On the other hand, in the more general case (B) (for which p_θ is large enough so that centrifugal and Coriolis forces can no longer be neglected), the motion ceases to be orthogonal to the equilibrium variety (because of coupling induced by the Coriolis force between radial and angular parts of the motion), so that the notion of ordinary orthogonal projection can no longer be applied. Indeed, because the equilibrium radius r_e now depends on the angular momentum p_θ (as a result of the effects of centrifugal force), it is no longer even possible to represent the equilibrium variety in configuration space alone, but rather, it must be represented in *phase space*. Nevertheless, the idea of projecting the phase point $(\mathbf{x}^i,\mathbf{p}_i)$, along a line in the direction of a *purely oscillatory* part $\delta\mathbf{x}^i$, $\delta\mathbf{p}_i$ of the motion, down to the point $\mathbf{X}^i$, $\mathbf{P}_i$ of intersection of this line with the equilibrium hypersurface can be still applied, using the non-Euclidean metric which is suggested by the equations of motion themselves. As a result, all of the basic properties of ordinary perpendicular projections (in particular, as we shall see, the notion of analyzing a motion as the sum of all its projected parts) can be retained so as to make possible the use of a simple geometrical description, affording considerable insight into the separation of the motion that we are studying. We saw in Sec. 2 that the Poisson brackets of two functions F and G can be expressed as

$$[F,G] = (\partial F/\partial z^i)O^{ij}(\partial G/\partial z^j), \quad (3.1)$$

and Hamilton's equations for any function f as

$$\dot{f} = [f,H] = (\partial f/\partial z^i)O^{ij}(\partial H/\partial z^j). \quad (3.2)$$

It is well known that the Poisson brackets are invariant to an arbitrary canonical transformation, or in other words, that $[F,G]$ is a scalar in such a transformation. Since $\partial F/\partial z^i$ and $\partial G/\partial z^i$ are covariant vectors in z^i space, this shows that O^{ij} can be regarded as playing the role of a metrical tensor, with the aid of which invariant scalar products can be defined in the way that is usual in geometry.

The metric O^{ij} applies for covariant vectors, such as $\partial F/\partial z^i$. To define scalar products of contravariant vectors Δz^i we use the inverse of the matrix O^{ij}, given by (2.3), viz.,

$$O_{ij}=-O^{ij}, \tag{3.3}$$

which is the matrix fulfilling

$$O_{ij}O^{jk}=\delta_j{}^k. \tag{3.4}$$

The associated bilinear form for the scalar product of two contravariant vectors δz^i and Δz^i is

$$E(\delta z^i,\Delta z^i)=O_{ij}\delta z^i\Delta z^j. \tag{3.5}$$

This form is, like (3.1), invariant under arbitrary canonical transformations. Evidently, if we choose $\Delta z^i=\delta z^i$ we obtain $E(\delta z^i,\delta z^i)\equiv 0$, because of the antisymmetry of O_{ij}. Thus, the "length" defined with the aid of this form is zero for every vector. Metrics that give zero length for certain vectors are known (e.g., the pseudo-Euclidean-Minkowski metric of relativity, which gives zero length to light vectors). The fact that in our case the length given by the form (3.5) is zero for *all* vectors will not, as we shall see, interfere with our discussion, since the dynamical problem actually involves only the concept of projection, and, therefore, requires only the scalar product of one vector with another, and not the "length" in the usual metrical sense.

Although the invariance of the metrical tensors O_{ij} and O^{ij} is evident from the above discussion, we present in Appendix B a direct proof of this invariance, through which also their relation to the symplectic group is clarified.

Thus, the Poisson brackets (3.1) represent the "symplectic" scalar product of two covariant vectors $\partial F/\partial z^i$ and $\partial G/\partial z^j$, while the equations of motion (3.2) assert that the rate of change of any dynamical variable f is equal to the symplectic scalar product of $\partial f/\partial z^i$ with $\partial H/\partial z^i$. In particular, the velocity $\dot{z}^i$ of the "phase fluid" at the point z^i is given by $O^{ij}(\partial H/\partial z^j)$, i.e., the contravector obtained from the covector $\partial H/\partial z^i$ by "index raising." At first sight it might therefore seem that the phase fluid moves in a (symplectically) normal direction to the surfaces $H=$const. However, in Appendix C it is shown that (contrary to what one would expect on basis of our acquaintance with *symmetric* metrics) $O^{ij}\partial A(z)/\partial z^j$ is a vector *in* the surface $A=$const., and that (essentially) the only way of invariantly associating a normal vector n_A with a given surface $A=$const. is by

$$n_A{}^i=O^{ij}\partial B/\partial z^j, \quad \text{where} \quad [A,B]_{\text{P.B.}}=1, \tag{3.6}$$

i.e., the normal to $A=$const. is a vector which lies in the surface canonically *conjugate* to A.

It is further shown in Appendix C that while the symplectic scalar product of a vector with itself is zero, one can nevertheless associate with a displacement vector an invariant measure; in particular, the displacement vector normal to the surface $A=0$, leading from a point Z^i on that surface to a nearby point z^i at which A has a certain (small) value $A(z^i)$, will be

$$\delta z_A{}^i=O^{ij}A(z)(\partial B/\partial z^j). \tag{3.7}$$

Thus, the direction defined by (3.6) can be "normalized" so as to give a displacement vector of definite "extent."[3] If A and B belong to a set of functions F^μ in involution (i.e., a set of canonically conjugate pairs, having vanishing Poisson brackets for members belonging to different pairs) such that $A=F^\mu$ and $B=O_{\mu\nu}F^\nu$, (3.7) can also be written as[4]

$$\delta z_\mu{}^i=O^{ij}O_{[\mu]\nu}(\partial F^{[\mu]}/\partial z^j)F^\nu, \tag{3.8}$$

where $[\mu]$ indicates that the repeated index μ is *not* to be summed over.

As shown in Appendix C, the involutionary character of the F^μ, i.e.,

$$[F^\mu,F^\nu]=O^{\mu\nu},$$

ensures that $\sum_\mu \delta z_\mu{}^i$, the vector sum of the symplectic perpendiculars dropped from the point z^i to the surfaces $F^\mu=0$, is equal to the perpendicular dropped from z^i to the surface of *intersection* of these surfaces. Calling this perpendicular δz^i, we thus have

$$\delta z^i=\textstyle\sum_\mu \delta z_\mu{}^i=O^{ij}O_{\mu\nu}(\partial F^\mu/\partial z^j)F^\nu. \tag{3.9}$$

It is clear that the projections thus defined are invariant to an arbitrary canonical transformation. Such a transformation will change z^i into $z'^i(z^j)$, and at the same time $\partial F^\mu/\partial z^i$ will undergo the contragredient transformation.

So far, the index μ has been regarded as fixed. We can however make another kind of canonical transformation in which the z^i are fixed, while the F^μ (being some set of functions) are changed into another set of functions.

This, in fact, is what is often done in perturbation theory: one obtains a set of first-order functions by $F_1{}^\mu(z^i)$, and then one starts from these as a basis to go to a second-order set $F_2{}^\alpha(F_1{}^\mu)$ by a canonical transformation on the $F_1{}^\mu$.

---

[3] The term "extent" is here used instead of "length," because it is measured in terms of the surfaces $A(z^i)=$const. on which the end points of the displacement vector lie, rather than by a scalar product of vector with itself. The surfaces may belong to a family of surfaces which have transformationl laws of their own (independent of the canonical transformations considered), which associate with "extent" a broader sense of invariance.

[4] Again z^i is supposed to be sufficiently near to the surfaces $F^\mu=0$ for higher powers of $F^\mu(z^i)$ (and hence also of $\delta z_\mu{}^i$) to be negligible.

The important point to notice here is that our projection method is invariant to *both* kinds of canonical transformation. This property will be significant if one wishes to carry out a series of successive approximations. Thus, one will make a certain split by projecting z^i onto the surface $F_1{}^\mu(z^i)=0$ in the first approximation, then $F_1{}^\mu$ onto $F_2{}^\alpha=0$ in $F_1{}^\mu$ space in the second approximation, etc. (such a problem may arise, for example, in the effort to improve the random-phase approximation in plasma theory).

Equation (3.9) is identical with (2.6a), the projection formula that we obtained for the separation of the dynamics into oscillatory and nonoscillatory parts, provided that the F^μ are identified with the s canonical pairs of oscillatory variables (labeled Q^μ in Sec. 2).

The formulas given above can be expressed in an alternative way that is often useful. We introduce the complex variables,

$$R_k=\left(\frac{m\omega_k}{2}\right)^{1/2}Q_k-i\left(\frac{1}{2m\omega_k}\right)^{1/2}P_k\,, \tag{3.10}$$

which oscillate harmonically according to the relation

$$dR_k/dt=-i\omega_k R_k\,. \tag{3.11}$$

As can readily be verified, the Poisson brackets for these variables are

$$[R_k{}^*,R_{k'}]=i\delta_{k,k'}\,, \tag{3.12}$$

while the remaining Poisson brackets $[R_k{}^*,R_{k'}{}^*]=[R_k,R_{k'}]=0$. The action and angle variables are

$$J_k=R_k{}^*R_k,\quad \Phi_k=\frac{1}{2i}\ln\frac{R_k{}^*}{R_k}, \tag{3.13}$$

so that

$$R_k=(J_k)^{1/2}\exp(i\Phi_k)\,. \tag{3.14}$$

By means of a simple calculation, one then obtains from (3.9) (with only one pair of F^ν, corresponding to Q_k, P_k, being considered)

$$\delta z_k{}^i=iO^{ij}\left[R_k{}^*\frac{\partial R_k}{\partial z^j}-\frac{\partial R_k{}^*}{\partial z^j}R_k\right], \tag{3.15}$$

and summing up over all the oscillatory variables:

$$\delta z^i=iO^{ij}\textstyle\sum_k[R_k{}^*(\partial R_k/\partial z^j)-(\partial R_k{}^*/\partial z^j)R_k]. \tag{3.16}$$

For a general function $f(z^i)$, we have

$$[\![f]\!]\equiv f(Z^i)=f(z^i)-i\textstyle\sum_k R_k(f,R_k{}^*)+i\sum R_k{}^*(f,R_k)+\cdots, \tag{3.17}$$

where $[\![f]\!]$ again means f evaluated at the projected point Z^i.

If we note that in terms of action and angle variables

$$R_k{}^*R_k=J_k\,,\quad \frac{1}{2i}\ln\frac{R_k{}^*}{R_k}=\Phi_k\,,$$

we obtain

$$\delta z^i=\textstyle\sum_k O^{ij}J_k(\partial\Phi_k/\partial z^j)+\cdots, \tag{3.18}$$

and

$$[\![f]\!]=f(z^i)-\sum_k O^{ij}J_k\frac{\partial\Phi}{\partial z^j}\frac{\partial f}{\partial z^i}+\cdots. \tag{3.19}$$

By expressing z^i in terms of x^i, p_i we see that these are essentially the same expression as (I-3.5) and (I-3.6).

In the example of I (Sec. 2) there was a special case (A) in which the velocities were small enough so that the equilibrium variety and the oscillatory coordinate Q were functions of the x^i alone and did not involve the p_i. In such a case the momenta $P_k=m\dot{Q}_k$ are given by $P_k=\sum_i m\partial Q_k/(\partial x^i)\dot{x}^i=\sum_i(\partial Q_k/\partial x^i)p_i$. As a result,

$$(\partial P_k/\partial p_i)=(\partial Q_k/\partial x^i)\,,\quad (\partial Q_k/\partial p_i)=0\,. \tag{3.20}$$

Therefore, the Poisson bracket in this case reduced to

$$[Q_k,P_{k'}]=\textstyle\sum_i(\partial Q_k/\partial x^i)(\partial Q_{k'}/\partial x^i)=0\,, \tag{3.21}$$

and the requirement that $[Q_k,P_{k'}]=\delta_{kk'}$ reduces to

$$\textstyle\sum_i{}^2(\partial Q_k/\partial x^i)(\partial Q_{k'}/\partial x^i)=\delta_{kk'}\,. \tag{3.22}$$

This means that the surfaces whose normals are $\partial Q_k/\partial x^i$ form an orthonormal set *in configuration space*. The projection formulas then reduce to

$$\begin{aligned}\delta x^i&=\textstyle\sum_i(\partial Q_k/\partial x^i)Q_k\,,\\ \delta p^i&=\textstyle\sum_k[(\partial Q_k/\partial x^i)P_k-(\partial P_k/\partial x^i)Q_k].\end{aligned}\quad (i=1,2) \tag{3.23}$$

The second term $Q_k(\partial P_k/\partial x_i)=Q_k(\partial^2Q_k/\partial x^i\partial x^j)p_j$ will be small in the approximations used, viz., that $v^i=p_i/m$ is being taken to be small [as explained in I (Sec. 2) and as we shall see in Sec. 4]. Equations (3.23) then represent a projection along the (Euclidean) normal in configuration space, i.e., $\{\partial Q_k/\partial x^i\}$, of the equilibrium variety, $Q_k=0$, down to the intersection, $X^i=x^i-\delta x^i$, of this normal with the equilibrium variety in question. The vector δp_i regarded as a covariant vector in configuration space, is then the component of the total momentum p_i in the direction normal to the equilibrium variety, so that $P_i=p_i-\delta p_i$, the remainder must be its projection onto this variety (i.e., tangential to it) [and as seen from $P_k=\sum(\partial Q_k/\partial x^i)p_i$, P_k is just the projection of the total momentum into the above described normal].

The above discussion shows the relation between symplectic projection in *phase space* and orthogonal projection *in configuration space* that can arise in certain limiting cases of the former.

The geometric description of the collective motion can be developed further. To begin with, let us consider the case in which only a single pair of oscillatory variables (say, $Q^\alpha=Q_k$, $Q^\beta=P_k$) is excited. Then, the action variable $J_k=P_k{}^2/(2m\omega_k)+(Q_k{}^2\omega_k m)/2$ is a constant of the motion, while the angle variable $\Phi=\tan^{-1}[(m\omega_kQ_k)/P_k]$ increases linearly with time ($\Phi=\Phi_0+\omega_kt$), so that $Q_k=[(2J_k)/m\omega_k]^{1/2}\cos\Phi_k$, $P_k=(2m\omega_kJ_k)^{1/2}\sin\Phi_k$ oscillate harmonically with frequency $\omega_{\mathbf{k}}$. If we map the motion in $Q_{\mathbf{k}}$, $P_{\mathbf{k}}$ space, it will be an ellipse, with semimajor and semiminor axes $(2J_{\mathbf{k}}/m\omega_{\mathbf{k}})^{1/2}$, $(2m\omega_{\mathbf{k}}J_{\mathbf{k}})^{1/2}$. On the other hand, the nonoscillatory variables ξ will tend to increase more or less linearly with the time [e.g., as the angle θ did in the example given in I (Sec. 2), and as we shall see later, the particle variables $\mathbf{X}^i$, $\mathbf{P}_i$ do in the plasma case].

To simplify the problem, suppose that only a single one of the ξ variables differs from a constant. Then the motion will be a spiral, going round an elliptical cylinder with center $J_{\mathbf{k}}=0$ and carried into a third dimension by the ξ motion. When this motion is mapped into z^i space, it will of course, have to take place in the variety $J_{\mathbf{k}}=$const., which will (for $J_{\mathbf{k}}$) still be a kind of elliptical cylinder, its "axis" being the variety $J_{\mathbf{k}}=0$ (which is the intersection of $Q_{\mathbf{k}}=0$, $P_{\mathbf{k}}=0$, and therefore of dimension $6N-2$). (See Fig. 1.) The formula (3.14) for $\delta z_{\mathbf{k}}{}^i$ then represents elliptical motion on this cylinder with all the ξ and all the other Q^μ set equal to zero. The actual motion will be $z^i(t)=Z^i(t)+\delta z^i(t)$, and since $Z^i(t)$ represents the nonoscillatory variables, the elliptical motion will have a more or less linear motion in another dimension superposed on it, so that the total is again seen to be a spiral, but this time we have described it as mapped into a z^i space, with the aid of our projection formulas, rather than into $Q_{\mathbf{k}}$, $P_{\mathbf{k}}$, ξ space.

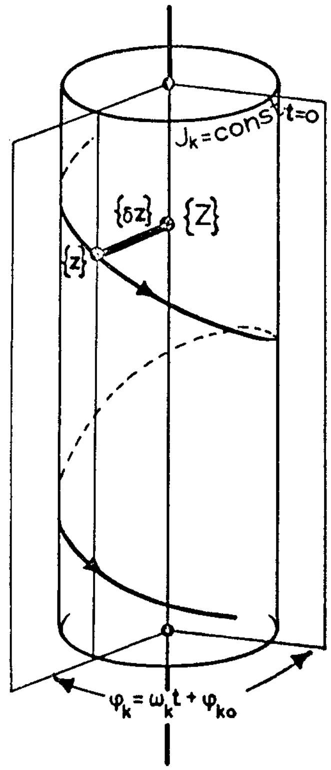

FIG. 1. The collective oscillation as a spiral motion on the "cylinder" $J_k=$const. about the "axis surface" of $6N-2s$ dimensions in phase space.

4. CLASSICAL THEORY OF THE ELECTRON GAS

Thus far, we have been considering an idealized problem in which the separation of the whole set of variables into a set of $Q_{\mathbf{k}}$ ($2s$ in number) which oscillate harmonically, and a set of (ξ) ($6N-2s$ in number) which do not, was simply assumed at the outset as one of the conditions of the problem. In the actual many-body problem it is necessary to study how this separation comes about, to see what determines it, and to find its limitations. We shall now proceed to see how this is done for the electron gas, which furnishes a typical case of this kind of problem. After these aspects of the problem of separation have been discussed, we shall then, in Sec. 5, apply our projection method, and treat the oscillations and the residual nonoscillatory part of the motion in some detail. This treatment will both illustrate our method and throw further light on the problem of the electron gas.

We shall start in this section by summarizing the essential features of the dynamics of the electron gas, as it has been treated thus far, and in the next section, we discuss how the above described separation comes about. We consider such a gas (a plasma) as consisting of N electrons in a box of unit dimensions, containing a uniform distribution of positive charge, leading to over-all neutrality. For this system, it has been found[5] that for small enough wave number $\mathbf{k}$, the Fourier coefficients

$$\rho_{\mathbf{k}}=\textstyle\sum_i \exp(-i\mathbf{k}\cdot\mathbf{x}^i) \tag{4.1}$$

of the electron density

$$\rho(\mathbf{x})=\textstyle\sum_i \delta(\mathbf{x}-\mathbf{x}^i) \tag{4.2}$$

constitute approximate collective coordinates, which oscillate harmonically around $\rho_{\mathbf{k}}=0$ according to the equation

$$\ddot{\rho}_{\mathbf{k}}+\omega_p{}^2\rho_{\mathbf{k}}=0\,, \tag{4.3}$$

where $\omega_p=(4\pi Ne^2/m)^{1/2}$ is the well known "plasma frequency."

This is proved by differentiating $\rho_{\mathbf{k}}$ twice with respect to time; utilizing the equations of motion of the electrons[6]

$$m\ddot{\mathbf{x}}^i=-4\pi e^2 i\sum_j\sum_{\mathbf{k}'\neq0}\frac{\mathbf{k}'}{\mathbf{k}'^2}\exp[i\mathbf{k}'\cdot(\mathbf{x}^i-\mathbf{x}^j)]\,, \tag{4.4}$$

one obtains

$$\begin{aligned}\ddot{\rho}_{\mathbf{k}}+\sum_i(\mathbf{k}\cdot\mathbf{v}_i)^2\exp(-i\mathbf{k}\cdot\mathbf{x}^i)&\\ +\omega_p{}^2\rho_{\mathbf{k}}+(\omega_p{}^2/N)\sum_{\mathbf{k}'\neq\mathbf{k}}(\mathbf{k}\cdot\mathbf{k}'/k'^2)&\\ \times\sum_{ij}\exp[-i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}^i-i\mathbf{k}'\cdot\mathbf{x}^j]&=0.\end{aligned} \tag{4.5}$$

[5] D. Pines and D. Bohm, Phys. Rev. **85**, 338 (1952).

[6] Strictly speaking, the term with $j=i$ should be left out, but since it adds up to zero in a spherical distribution over $\mathbf{k}$, with a cutoff at very high $\mathbf{k}$, k_m, we can include it without changing the value of the force on an electron. This cutoff is equivalent to neglecting the very short range of the force, which is significant only in close collisions.

If the temperature T is small enough, there will be an appreciable number of k vectors for which

$$k^2\langle v^2\rangle=k^2(KT/m)\ll\omega_p^2;\quad\text{or}\quad k<k_D,\quad\text{where}\quad k_D^2\langle v^2\rangle=\omega_p^2 \tag{4.6}$$

($\lambda_D=2\pi/k_D$ is the well-known Debye length). If this is the case, and if the velocity distribution is close to the usual Maxwellian distribution, then the second term in Eq. (4.5) can be neglected in comparison to the third. This we shall call the "Low-Velocity Approximation" (L.V.A.), meaning thereby that the lower the mean speeds, the greater the number of $\mathbf{k}$ vectors that satisfy (4.6) and, therefore, the more useful the method. In addition, in order to obtain collective coordinates, it has been necessary to make what is called the "Random Phase Approximation" (R.P.A.), which consists in neglecting the last terms in Eq. (4.5). (It can be seen that to the extent that the particles are distributed at random, the last term will tend to be negligible in comparison to the third.) With these approximations, one then obtains, as is evident, simple harmonic motion of $\rho_{\mathbf{k}}$ with the plasma frequency ω_p.

As pointed out above, the conclusion that the $\rho_{\mathbf{k}}$ will oscillate is valid only for $\mathbf{k}$'s satisfying (4.6). Let s be the number of such $\mathbf{k}$'s. The essential point for our purposes here is that these s degrees of freedom have, so to speak, been "liberated" from dependence on the detailed behavior of the individual particles, and instead behave in a self-determined way; i.e., they oscillate harmonically, indifferent to the detailed states of the various particles. Associated to each $\rho_{\mathbf{k}}$ (which corresponds to the $Q_{\mathbf{k}}$ discussed in the previous sections and henceforth be thus denoted) we can define a canonical momentum

$$P_{\mathbf{k}}=m\beta_k(d/dt)\rho_{\mathbf{k}}^*=-i\beta_k\textstyle\sum_i(\mathbf{k}\cdot\mathbf{p}_i)\exp[i(\mathbf{k}\cdot\mathbf{x}^i)], \tag{4.7}$$

where α_k, β_k are normalizing factors;

$$\alpha_k=\left(\frac{4\pi e^2}{k^2}\right)^{1/2}\frac{1}{(\omega_{\mathbf{k}})^{1/2}},\quad Q_{\mathbf{k}}=\alpha_k\rho_{\mathbf{k}}=\alpha_k\sum\exp(-i\mathbf{k}\cdot\mathbf{x}^i),$$
$$\beta_k=\left(\frac{\omega_{\mathbf{k}}}{4\pi e^2k^2N^2}\right)^{1/2}. \tag{4.7a}$$

α_k and β_k fulfil $Nk^2\alpha_k\beta_k=1$, and are so chosen as to obtain the Poisson bracket relations

$$[Q_{\mathbf{k}},P_{\mathbf{k}'}]=(1/N)\textstyle\sum_i\exp[i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{x}^i].$$

In the R.P.A., the sum on the right-hand side of this expression vanishes when $\mathbf{k}\neq\mathbf{k}'$, and is unity for $\mathbf{k}'=\mathbf{k}$. Thus,

$$[Q_{\mathbf{k}},P_{\mathbf{k}'}]=\delta_{\mathbf{k}\mathbf{k}'}\text{ (in the R.P.A.)}. \tag{4.8}$$

Now, evidently $[Q_{\mathbf{k}},Q_{\mathbf{k}'}]=0$, and

$$[P_{\mathbf{k}},P_{\mathbf{k}'}]=i\beta_k\beta_{k'}\textstyle\sum_i(\mathbf{k}\cdot\mathbf{k}')[\exp(i(\mathbf{k}'+\mathbf{k})\cdot\mathbf{x}^i)](\mathbf{k}'-\mathbf{k})\cdot\mathbf{p}_i$$

and this vanishes in the R.P.A., provided that $\sum_i\mathbf{p}_i=0$ (which will, in general, be taken to be the case).

It follows then that in the R.P.A., the $Q_{\mathbf{k}}P_{\mathbf{k}}$ constitute a canonical set of $2s$ functions, in involution.

From $Q_{\mathbf{k}}$ and $P_{\mathbf{k}}$ we can obtain the constants of the motion

$$J_{\mathbf{k}}=\tfrac{1}{2}P_{\mathbf{k}}^*P_{\mathbf{k}}+\tfrac{1}{2}Q_{\mathbf{k}}^*Q_{\mathbf{k}}, \tag{4.9a}$$

and the angle variables

$$\Phi_{\mathbf{k}}=\tan^{-1}(P_{\mathbf{k}}/Q_{\mathbf{k}}). \tag{4.9b}$$

From the fact that $Q_{\mathbf{k}}$ and $P_{\mathbf{k}}$ oscillate harmonically, it follows that

$$\Phi_{\mathbf{k}}=\Phi_{\mathbf{k}0}+\omega_pt. \tag{4.9c}$$

The Hamiltonian of the plasma oscillations themselves evidently must be

$$H_p=\sum\omega_{\mathbf{k}}J_{\mathbf{k}}=\sum_{\mathbf{k}}\tfrac{1}{2}\omega_{\mathbf{k}}(Q_{\mathbf{k}}^*Q_{\mathbf{k}}+P_{\mathbf{k}}^*P_{\mathbf{k}}). \tag{4.9d}$$

An improved collective coordinate can be obtained, which takes into account [to a higher order in $((\mathbf{k}\cdot\mathbf{v}^i)/\omega_p)$] the term $\sum_i(\mathbf{k}\cdot\mathbf{v}^i)^2\exp(-i\mathbf{k}\cdot\mathbf{x}^i)$ in Eq. (4.5), but which still uses the R.P.A. This coordinate $q_{\mathbf{k}}$ and its corresponding canonical momentum $p_{\mathbf{k}}$ have been shown[5] to be

$$q_{\mathbf{k}}=\alpha_k\sum_i\frac{\omega_p^2\exp(-i\mathbf{k}\cdot\mathbf{x}^i)}{\omega_{\mathbf{k}}^2-(\mathbf{k}\cdot\mathbf{v}^i)^2},$$

and

$$p_{\mathbf{k}}=-i\beta_k\sum_i\frac{\omega_p^2}{\omega_{\mathbf{k}}^2}\frac{(\mathbf{k}\cdot\mathbf{v}^i)\exp(-i\mathbf{k}\cdot\mathbf{x}^i)}{\omega^2-(\mathbf{k}\cdot\mathbf{v}^i)^2}, \tag{4.10}$$

which oscillate with the frequency

$$\omega_k^2\cong\omega_p^2+k^2\langle v^2\rangle. \tag{4.10a}$$

(The factors in front of $q_{\mathbf{k}}$ and $p_{\mathbf{k}}$ have been chosen for convenience in further calculations.)

It has further been shown that $q_{\mathbf{k}}$ can be regarded as constituting an "oscillatory" (or collective) part of the total Fourier coefficient $\rho_{\mathbf{k}}$ of the particle density, such that

$$\rho_{\mathbf{k}}=q_{\mathbf{k}}+\eta_{\mathbf{k}}, \tag{4.11}$$

where

$$\eta_{\mathbf{k}}=\rho_{\mathbf{k}}-q_{\mathbf{k}}=\textstyle\sum_i\eta_{\mathbf{k}i},$$

with

$$\eta_{\mathbf{k}i}=\frac{\omega_p^2-\omega_{\mathbf{k}}^2-(\mathbf{k}\cdot\mathbf{v}^i)^2}{\omega^2-(\mathbf{k}\cdot\mathbf{v}^i)^2}\exp(-i\mathbf{k}\cdot\mathbf{x}^i).$$

The function $\eta_{\mathbf{k}}$ satisfies the equation

$$\ddot{\eta}_{\mathbf{k}}=-\textstyle\sum_i(\mathbf{k}\cdot\mathbf{v}^i)^2\eta_{\mathbf{k}i}. \tag{4.12}$$

This means that each particle moving at a velocity $\dot{\mathbf{x}}^i=\mathbf{v}^i$ contributes its own effective frequency $\omega_i=\mathbf{k}\cdot\mathbf{v}^i$ to the fluctuation of $\eta_{\mathbf{k}}$, and because of the noncoherence of these frequencies, $\eta_{\mathbf{k}}$ will represent a randomly fluctuating part of the charge density which is evidently associated only with the irregular thermal part of the motions of the individual particles, and not with the regular oscillatory part.

The collective coordinates $q_{\mathbf{k}}$, while being accurate to a higher order in the expansion parameter $(\mathbf{k}\cdot\mathbf{v}^i)/\omega_{\mathbf{k}}$ than the $\rho_{\mathbf{k}}$, have, in general, the same range of $\mathbf{k}$ for validity as does $\rho_{\mathbf{k}}$ itself. In other words, there will be essentially the same number s of oscillating $q_{\mathbf{k}}$ as there are of $\rho_{\mathbf{k}}$. We shall, therefore, call this the "Improved Low-Velocity Approximation" or the I.L.V.A. As we shall see in the next section, this range of validity is determined by the condition that regular oscillations are possible for those $\mathbf{k}$ for which the collective effects of the forces are great enough to overcome the effects of random thermal motions.

5. THE EQUILIBRIUM VARIETIES FOR THE ELECTRON GAS

In the previous section, we gave a brief summary of some of the main properties of the electron gas that are relevant to the separation of the motion into a collective oscillatory part, and an individual nonoscillatory part. As we have indicated, the approximations leading to oscillatory behavior of the collective coordinates $\rho_{\mathbf{k}}$ break down for $\mathbf{k}$ greater than a certain limiting k_D, which is $(\omega_p^2/\langle v^2\rangle)^{1/2}$. The question then arises as to whether this limitation on collective oscillation is a result of the special choice of $\rho_{\mathbf{k}}$ (or $q_{\mathbf{k}}$) for the oscillatory variables (so that it is due only to a limitation in our mathematical methods of describing the motion), or whether it is an inherent dynamical feature of the system. We shall now see that the second of these alternatives is true, and that the $\mathbf{k}$ reflects a real physical limitation on oscillatory collective behavior. In the course of the discussion, it will also become clear that the use of $\rho_{\mathbf{k}}$ (and $q_{\mathbf{k}}$) as collective coordinates is not an arbitrary choice, but that it is necessitated by the form of the interactions.

Finally, we shall see that the use of the improved collective coordinate $q_{\mathbf{k}}$ instead of $\rho_{\mathbf{k}}$, corresponds to taking $r_e(p_\theta)$ instead of r_0 for the equilibrium radius in the example discussed in I (Sec. 2) (i.e., it is equivalent to taking into account "centrifugal" forces due to the curvature of the equilibrium variety).

In order to facilitate the analysis of the problem, we shall make a minor simplification, which, however, does not change the essential character of the separation of the motion into parts that we are studying.

In the expression of the Coulomb potential, $\varphi(x) = 4\pi e^2 \sum_{\mathbf{k}}^{k_m} \rho_{\mathbf{k}} \exp(i\mathbf{k}\cdot\mathbf{x})/k^2$ [see footnote to Eq. (4.4) for a definition of the cutoff k_m], we choose k_m such that the total number of $\rho_{\mathbf{k}}$'s is equal to $3N$, the number of degrees of freedom. This will cut off the Coulomb force at the mean interparticle spacing; and since for typical cases $k_s \ll k_m$, the effect on collective oscillation will be negligible. (As we shall see, this simplification will make possible what is, in principle, an exact discussion of the statistical mechanics of the system in terms of the $\rho_{\mathbf{k}}$ instead of the $\mathbf{x}^i$.) The potential energy of the system is then

$$W = 2\pi e^2 \textstyle\sum_{k<k_m} (\rho_{\mathbf{k}}^* \rho_{\mathbf{k}}/k^2)\,. \tag{5.1}$$

Let us start with the case of very low temperature ($T \cong 0$), and increase the temperature gradually. Since the electrical force on the ith particle is

$$\mathbf{F}_i = -4\pi e^2 i \sum_{\mathbf{k}} (\mathbf{k}/k^2)\rho_{\mathbf{k}} \exp(i\mathbf{k}\cdot\mathbf{x}^i)\,, \tag{5.2}$$

the system will be in equilibrium if all the $\rho_{\mathbf{k}}$ are zero. Because there are now $3N$ of the $\rho_{\mathbf{k}}$, the conditions of equilibrium will, in principle, determine the location of all the particles. Evidently (remembering the uniform positive background which neutralizes the whole system), one possibility for such an equilibrium is a crystal, i.e., a periodic array of electrons.

If this system is excited to a low degree, $(KT/(m\omega_p^2))^{1/2} < 1/k_m$, where K is Boltzmann's constant, all the $\rho_{\mathbf{k}}$ will oscillate harmonically in the manner described in Sec. 4. The system is, therefore, oscillating around an equilibrium variety in configuration space, of dimension zero; i.e., around a stable point (similarly, the momenta $\mathbf{p}_i$ oscillate around $\mathbf{p}_i = 0$, so that there is a stable point in phase space, also).

As the temperature is raised, the oscillations increase in amplitude, and eventually there will be an appreciable probability for the particles to escape their fixed equilibrium positions and to move about in a relatively free way (i.e., the system becomes a fluid).

When this takes place, then, as we have already indicated, the arguments leading to the conclusion that all the $\rho_{\mathbf{k}}$ oscillate harmonically and in a self-determined way, break down. This breakdown occurs first, as is well known,[7] for the highest values of $\mathbf{k}$; i.e., those for which the mean drift speed of the particles sends an electron, within a period of an oscillation, over an appreciable part of a wavelength $\lambda = 2\pi/k$, so that the potential is no longer appropriate for automatically correcting excesses or deficiencies of charge. Then, as the temperature is raised still further, more and more of the $\rho_{\mathbf{k}}$ lose their oscillatory character, the dividing line being roughly at $k = k_D = ((m\omega_p^2)/KT)^{1/2}$.

In order to discuss the above described breakdown of oscillatory character more precisely, we shall now consider the problem from the point of view of statistical mechanics. According to the Boltzmann distribution, the probability in configuration space after integration over the momentum part of phase space is proportional to

$$\exp(-W/KT)d\mathbf{x}^1\cdots d\mathbf{x}^N = \exp\left(-\sum_{\mathbf{k}}^{k_m} \frac{2\pi e^2}{KT}\rho_{\mathbf{k}}^*\rho_{\mathbf{k}}\right) J\left(\frac{\partial \mathbf{x}^i}{\partial \rho_{\mathbf{k}}}\right)(d\rho_{\mathbf{k}}\cdots)\,, \tag{5.3}$$

[7] D. Bohm and E. P. Gross, Phys. Rev. **75**, 1851, 1864 (1949).

where J is the Jacobian of the transformation from $\mathbf{x}^i$ space to $\rho_\mathbf{k}$ space.[8] Now, from the special form of $\rho_\mathbf{k} = \sum_i \exp(-i\mathbf{k}\cdot\mathbf{x}^i)$ as a sum of N terms, each of absolute value unity, and with a variable phase, it follows from the central limit theorem of statistics that in the mapping from $\mathbf{x}^i$ space into $\rho_\mathbf{k}$ space, almost all of the $\mathbf{x}^i$ space will be carried into a region in which $|\rho_\mathbf{k}|$ is not very much greater than its root-mean-square $[\langle\rho_\mathbf{k}^2\rangle^{1/2} = (N)^{1/2}]$, although a very small fraction of the total $\mathbf{x}^i$ space will go into the regions between $|\rho_\mathbf{k}| \cong (N)^{1/2}$ and $|\rho_\mathbf{k}| = N$. Roughly speaking, this means that for almost every physical state of the system, $\rho_\mathbf{k}$ is really free to oscillate only between certain bounds, which are essentially $|\rho_\mathbf{k}|_{\max} \cong (N)^{1/2}$. As Yevick and Percus have shown,[9] the determinant is regular and nonvanishing in this region, becoming infinite only at the effective bounds[10] of $\rho_\mathbf{k}$, as described above, while it remains nearly constant away from the bounds. Hence, at points that are not near these bounds, the variations of the probability function (5.3) with $\rho_\mathbf{k}$ is determined mainly by the Boltzmann factor

$$\exp[-\sum(2\pi e^2\rho_\mathbf{k}^*\rho_\mathbf{k})/KT].$$

This factor, however, becomes negligible when $|\rho_\mathbf{k}| > |\rho_{\mathbf{k}_0}| = [KT/(2\pi e^2)]^{1/2}$. From the ratio

$$\frac{|\rho_{\mathbf{k}_0}|}{|\rho_\mathbf{k}|_{\max}} = [KT/(2\pi N e^2)]^{1/2} = \frac{k_m}{k_D}, \tag{5.4}$$

we see that as long as the temperature is so low that $k < k_m \ll k_D$, the random thermal excitation of the $\rho_\mathbf{k}$ degrees of freedom will not bring any of the $\rho_\mathbf{k}$ near the above described bounds in $\rho_\mathbf{k}$ space. We may picture this result as due to the "restoring force," viz., $-\omega_p^2\rho_\mathbf{k}$ [see Eq. (4.3)] which acts so as to bring $\rho_\mathbf{k}$ back to zero, thus giving rise to harmonic oscillations according to the equation $\ddot{\rho}_\mathbf{k} = -\omega_p^2\rho_\mathbf{k}$. [From Eq. (4.3) it is evident that the restoring force represents the effect of the Coulomb potential.] At low temperatures, this "force" is able to keep the system so near to $\rho_\mathbf{k} = 0$ that in its random thermal motion, it practically never comes near to the upper bound of $|\rho_\mathbf{k}|$. The fact that there is such an upper bound can, therefore, be ignored, and the resulting oscillations are, to all intents and purposes, free and indifferent to each other, as well as to the detailed behavior of the individual particle motions.

As the temperature is raised, however, then eventually the $|\rho_\mathbf{k}|$ in their random thermal motions will begin to reach their upper bounds $|\rho_\mathbf{k}| \cong (N)^{1/2}$, as follows from (5.4). This will happen first for the largest values of $\mathbf{k}$, and as the temperature is raised still further, for smaller values of $\mathbf{k}$, the dividing line being, of course, at a value of k of the order of k_D (which drops like $T^{-1/2}$). For the values of $\mathbf{k}$ for which this happens, the exponential factor in the probability function (5.3) will become practically a constant. The constancy of this term (which latter represents the sole effect of the potential energy on the probability function) means that these $\rho_\mathbf{k}$ are behaving as if there were no interaction whatsoever between the particles. Since the collective oscillations are a consequence of the Coulomb interactions, this also means that the "restoring forces" (which we have seen to arise from the force term in the equations of motion of the particles) are overwhelmed by the dispersing effects of the random thermal motions. Hence, for $k > k_D$, the $\rho_\mathbf{k}$ behave as they would if the particles were perfectly free. For $k < k_D$, however, the probability distribution of the $\rho_\mathbf{k}$ coordinates is still being limited in $\rho_\mathbf{k}$ space by the Boltzmann factor $\exp[-\sum(2\pi e^2/k^2)\rho_\mathbf{k}^*\rho_\mathbf{k}]$ to a region which is significantly smaller than the upper bound in $\rho_\mathbf{k}$ space. This means that for these $\rho_\mathbf{k}$, the potential energy plays a major part in determining the behavior of the $\rho_\mathbf{k}$, and in terms of the equations of motion (4.5) one sees that this must come about through the fact that these $\rho_\mathbf{k}$ are oscillating harmonically in a self-determined way near $\rho_\mathbf{k} = 0$.

If we go back to the phase space of the $\mathbf{x}^i$, $\mathbf{p}_i$ then (for $k < k_D$) the oscillations of the phase point is near an equilibrium variety given by

$$Q_\mathbf{k}(\mathbf{x}^i) = 0, \quad P_\mathbf{k}(\mathbf{x}^i, \mathbf{p}_i) = 0. \tag{5.5}$$

This variety has $6N - 2s$ dimensions.

The remainder of the degrees of freedom $\rho_\mathbf{k}$ (corresponding to $k > k_D$) are strongly coupled because the $\rho_\mathbf{k}$ become as large as $(N)^{1/2}$. Therefore, as can be seen from the equations of motion (4.5), the nonlinear terms neglected in the R.P.A. become significant. As a result, for $k > k_D$, the $\rho_\mathbf{k}$ are not a relevant set of coordinates, and it is more useful to express these degrees of freedom in terms of the $\mathbf{x}^i$ and $\mathbf{p}_i$. However, since there are $6N$ of these, and only $6N - 2s$ particle-like degrees of freedom to be thus described, we will use the quantities $\mathbf{X}^i$, $\mathbf{P}_i$, satisfying the identities (for $k < k_D$)

$$Q_\mathbf{k}(\mathbf{X}^i) \equiv 0, \quad P_\mathbf{k}(\mathbf{X}^i, \mathbf{P}_i) \equiv 0, \tag{5.6}$$

in order to describe these degrees of freedom. This will be done by projecting the particle coordinates and momenta $\mathbf{x}^i$ and $\mathbf{p}_i$ into the points $\mathbf{X}^i$, $\mathbf{P}_i$ in the equilibrium variety, with the aid of the method developed previously in this paper. As will be seen, the $\mathbf{X}^i$, $\mathbf{P}_i$ (restricted by the above identities) describe $3N - s$ "quasiparticles" which move very nearly in straight lines at constant velocity. It must be emphasized, however, that these quasiparticles are abstractions and do not correspond to any of the original individual particles at all. Rather (as will be seen by projecting the

[8] This expression is exact, and an equation of this kind holds, indeed, for very general types of systems.

[9] G. J. Yevick and J. K. Percus, Phys. Rev. 101, 1186 (1956).

[10] Although the determinant becomes infinite at the bounds, it does so in such a way that the integral of (5.3) over a small region of $\rho_\mathbf{k}$ near these bounds not only remains finite, but also goes to zero as the size of this region goes to zero. Thus, the region near the bounds is not very important in statistical averages.

particle coordinates and momenta into the equilibrium variety), they are associated to complicated and correlated motions of the whole system in which the charge density fluctuations $\rho_{\mathbf{k}}$ are restricted to a short range (i.e., $k>k_D$). (The fluctuations of $\rho_{\mathbf{k}}$ for $k<k_D$ are of course the regular plasma oscillations.)

We have thus seen that as the temperature of the system is raised from zero, a real physical change in the nature of the motion comes about, and this conclusion is not just a consequence of the specific conjecture $\rho_{\mathbf{k}}$ for the oscillatory collective coordinates. This change is a consequence of the variation of the balance between the tendency of the potential energy to stabilize the system near a fixed point in configuration space, and that of the kinetic energy to disperse the system through the whole configuration space. At any given temperature, the balance of these two tendencies will be such that in $2s$ dimensions (corresponding to the $Q_{\mathbf{k}}$, $P_{\mathbf{k}}$, for $k<k_D$) the forces of interaction bring about an oscillation about a $6N-2s$ dimensional variety, in which latter the random fluctuations have overwhelmed the effects of the forces. As T is raised from 0 to ∞, this equilibrium variety is increased in dimensionality from 0 to $6N$.

As in the transiton from case (A) to case (B) in the example of I (Sec. 2), this equilibrium variety will, upon increase of temperature, be represented by momentum-dependent functions, so that it has to be described in phase space rather than in configuration space. In the example, this was seen in the circumstance that the equilibrium circle $r(x^i)-r_0=0$ had to be replaced by $r(x^i)-r_e(p_\theta)=0$; i.e., the centrifugal forces effected a velocity-dependent shift in this circle. Similarly, in the plasma problem, the equilibrium surface $\rho_{\mathbf{k}}(\mathbf{x}^i)=0$ will be shifted to

$$\rho_{\mathbf{k}}(\mathbf{x}^i)-\eta_{\mathbf{k}}(\mathbf{x}^i,\mathbf{p}_i)=q_{\mathbf{k}}(\mathbf{x}^i,\mathbf{p}_i)=0 \tag{5.7}$$

as a result of corresponding "centrifugal" forces due to the curvature of the surface $\rho_{\mathbf{k}}(\mathbf{x}^i)=0$. To see how the η term comes about, we write the equations of motion (4.5) for $\rho_{\mathbf{k}}$ once again. We still use the R.P.A., but we no longer neglect $\sum(\mathbf{k}\cdot\mathbf{v}^i)^2\exp(-i\mathbf{k}\cdot\mathbf{x}^i)$. The result is

$$\ddot{\rho}_{\mathbf{k}}+\omega_p^2\rho_{\mathbf{k}}=-\sum(\mathbf{k}\cdot\mathbf{v}^i)^2\exp(-i\mathbf{k}\cdot\mathbf{x}^i). \tag{5.8}$$

We now recall that the equations of motion of the $\rho_{\mathbf{k}}$, (4.5), were obtained by differentiating $\rho_{\mathbf{k}}$ twice with respect to time. In other words, we took

$$\frac{d^2\rho_{\mathbf{k}}}{dt^2}=\frac{d}{dt}\left(\frac{\partial\rho_{\mathbf{k}}}{\partial\mathbf{x}^i}\dot{\mathbf{x}}^i\right)=\frac{\partial^2\rho_{\mathbf{k}}}{\partial\mathbf{x}^i\partial\mathbf{x}^j}\dot{\mathbf{x}}^i\dot{\mathbf{x}}^j+\frac{\partial\rho_{\mathbf{k}}}{\partial\mathbf{x}^i}\ddot{\mathbf{x}}^i.$$

The term on the right-hand side of (5.8) is just $(\partial^2\rho_{\mathbf{k}})/(\partial\mathbf{x}^i\partial\mathbf{x}^j)\dot{\mathbf{x}}^i\dot{\mathbf{x}}^j$; but $(\partial^2\rho_{\mathbf{k}})/(\partial\mathbf{x}^i\partial\mathbf{x}^j)$ evidently represents the curvature of the hypersurface in configuration space, $\rho_{\mathbf{k}}=$constant, and $(\partial^2\rho_{\mathbf{k}})/(\partial\mathbf{x}^i\partial\mathbf{x}^j)\dot{\mathbf{x}}^i\dot{\mathbf{x}}^j$ represents the "centrifugal force" arising from the movement of the configuration point in this curved surface.

The above analogy of the plasma with the example of I (Sec. 2) suggests that there will have to be a shift of frequency as well as a shift of the equilibrium variety, dependent on the momenta. Therefore, we rewrite Eq. (5.8) as

$$\begin{aligned}\ddot{\rho}_{\mathbf{k}}+\omega_{\mathbf{k}}^2\rho_{\mathbf{k}}&=(\omega_{\mathbf{k}}^2-\omega_p^2)\rho_{\mathbf{k}}-\textstyle\sum_i(\mathbf{k}\cdot\mathbf{v}^i)^2\exp(-i\mathbf{k}\cdot\mathbf{x}^i)\\&=\textstyle\sum_i[\omega_{\mathbf{k}}^2-\omega_p^2-\sum_i(\mathbf{k}\cdot\mathbf{v}^i)^2]\exp(-i\mathbf{k}\cdot\mathbf{x}^i).\end{aligned} \tag{5.9}$$

It can now be seen that our definition for $\eta_{\mathbf{k}}$ is nothing but a special solution of the inhomogeneous Eq. (5.9) for $\rho_{\mathbf{k}}$, in which there is no part oscillating with the plasma frequency. $q_{\mathbf{k}}$ is then a solution of the homogeneous equation $\ddot{q}_{\mathbf{k}}+\omega_{\mathbf{k}}^2q_{\mathbf{k}}=0$ with the corrected $\omega_{\mathbf{k}}$ being given by Eq. (4.10a). It should be clear that the procedure (5.9) is essentially equivalent to that given originally by Pines and Bohm,[5] and therefore, it has the same results and limitations. The present formulation, however, emphasizes the relative role of the η and q quantities from the mathematical point of view.

6. APPLICATION OF THE PROJECTION METHOD TO THE ELECTRON GAS

In this section, we shall apply the projection method to the electron gas in order to carry out explicitly the separation of the motion into an oscillatory collective part and a nonoscillatory, individual, particle-like part. Our first object is to show that the results of the Bohm-Pines[11] treatment can be obtained in a simpler way and with less calculations. By doing it this way, we also obtain further insight into the problem, especially with the aid of the geometrical concepts that we have introduced throughout this article. Then, by comparing the electron gas with the simple example of I (Sec. 2), we further develop the suggestive interpretations started towards the end of Sec. 5 of the present paper, in terms of a generalized notion of "centrifugal" and "Coriolis" forces in configuration space. With the aid of concepts of the kind that we have been describing, one obtains a way of thinking about the many-body problem in general, which may be of help in further developments and applications. As an example of this type of concept, we introduce the notion of a set of $3N$ "quasiparticles," describing the nonoscillatory features of the motion.

The oscillatory collective coordinates for this problem (in the I.L.V.A. as explained in Sec. 4) are given in Eqs. (4.10) and (4.10a). Various combinations of these variables can be taken as the canonical set. Let us adopt the set

$$\begin{aligned}R_{\mathbf{k}}&=\frac{1}{2\omega_{\mathbf{k}}}\left(\frac{\omega_{\mathbf{k}}^2}{\omega_p^2}q_{\mathbf{k}}-i\omega_{\mathbf{k}}p_{\mathbf{k}}^*\right)=\frac{\alpha_k}{2}\sum_i\frac{\exp(-i\mathbf{k}\cdot\mathbf{x}^i)}{1-(\mathbf{k}\cdot\mathbf{p}_i/m\omega_{\mathbf{k}})},\\R_{\mathbf{k}}^*&=\frac{1}{2\omega_{\mathbf{k}}}\left(\frac{\omega_{\mathbf{k}}^2}{\omega_p^2}q_{\mathbf{k}}^*+i\omega_{\mathbf{k}}p_{\mathbf{k}}\right)=\frac{\alpha_k}{2}\sum_i\frac{\exp(i\mathbf{k}\cdot\mathbf{x}^i)}{1-(\mathbf{k}\cdot\mathbf{p}_i/m\omega_{\mathbf{k}})},\end{aligned} \tag{6.1}$$

[11] D. Bohm and D. Pines, Phys. Rev. **92**, 609 (1953).

where α_k is defined in (4.7a) as

$$\alpha_k=(4\pi e^2/k^2\omega_p)^{1/2}. \tag{6.2}$$

We are, thus, using the complex oscillatory variables introduced in (3.10) which satisfy a *first*-order equation,

$$\dot{R}_{\mathbf{k}}=-i\omega_{\mathbf{k}}R_{\mathbf{k}},\quad \dot{R}_{\mathbf{k}}^*=i\omega_{\mathbf{k}}R_{\mathbf{k}}^*. \tag{6.3}$$

The Poisson brackets are [as in (3.12)]

$$\begin{aligned}&[R_{\mathbf{k}}^*,R_{\mathbf{k}'}]=i\delta_{\mathbf{k}\mathbf{k}'},\\ &[R_{\mathbf{k}}^*,R_{\mathbf{k}'}^*]=[R_{\mathbf{k}},R_{\mathbf{k}'}]=0.\end{aligned} \tag{6.4}$$

In terms of these variables, the projection formula for the part of the motion corresponding to oscillatory collective behavior is given by Eq. (3.16). In order to show explicitly what these formulas mean, we shall return from the z^i notation to the use of $\mathbf{x}^i$, $\mathbf{p}_i$. We have

$$\begin{aligned}\delta\mathbf{x}^i&=i\sum_{\mathbf{k}}\left(R_{\mathbf{k}}\frac{\partial R_{\mathbf{k}}^*}{\partial\mathbf{p}_i}-\frac{\partial R_{\mathbf{k}}}{\partial\mathbf{p}_i}R_{\mathbf{k}}^*\right)\\ &=i\sum_{\mathbf{k}}\frac{\alpha_k\mathbf{k}}{m\omega_{\mathbf{k}}\sqrt{2}}\frac{R_{\mathbf{k}}\exp(i\mathbf{k}\cdot\mathbf{x}^i)-R_{\mathbf{k}}^*\exp(-i\mathbf{k}\cdot\mathbf{x}^i)}{(1-\gamma^i)^2},\\ \delta\mathbf{p}_i&=-i\sum_{\mathbf{k}}\left(R_{\mathbf{k}}\frac{\partial R_{\mathbf{k}}^*}{\partial\mathbf{x}^i}-\frac{\partial R_{\mathbf{k}}}{\partial\mathbf{x}^i}R_{\mathbf{k}}^*\right)\\ &=\sum_{\mathbf{k}}\frac{\alpha_k\mathbf{k}}{\sqrt{2}}\frac{R_{\mathbf{k}}\exp(i\mathbf{k}\cdot\mathbf{x}^i)+R_{\mathbf{k}}^*\exp(-i\mathbf{k}\cdot\mathbf{x}^i)}{1-\gamma^i},\end{aligned} \tag{6.5}$$

where we have used the abbreviation

$$\gamma^i=(\mathbf{k}\cdot\mathbf{p}_i/m\omega_{\mathbf{k}}). \tag{6.6}$$

The "projected" variables $\mathbf{X}^i$, $\mathbf{P}_i$, corresponding to the nonoscillatory (noncollective) part of the motion, are then

$$\mathbf{X}^i=\mathbf{x}^i-\delta\mathbf{x}^i,\quad \mathbf{P}_i=\mathbf{p}_i-\delta\mathbf{p}_i, \tag{6.7}$$

and as shown generally in I (Secs. 3 and 4) these satisfy the identities

$$R_{\mathbf{k}}(\mathbf{X}^i,\mathbf{P}_i)=0,\quad R_{\mathbf{k}}^*(\mathbf{X}^i,\mathbf{P}_i)=0. \tag{6.8}$$

We recall also that $\delta z^i=(\delta\mathbf{x}^i,\delta\mathbf{p}_i)$ is "symplectically" normal to the intersection of the hypersurfaces, $R_k(\mathbf{x}^i,\mathbf{p}_i)=0$, $R_k^*(\mathbf{x}^i,\mathbf{p}_i)=0$.

In order to study the equations of motion, we should first express the Hamiltonian in terms of the $R_{\mathbf{k}}$ and $\mathbf{X}^i$, $\mathbf{P}_i$ variables. We have seen in (2.10) how this is done for the general case. The Hamiltonian, expressed to second order in $R_{\mathbf{k}}$ can be evaluated, and in the approximations that we are using, it turns out to be (as a simple calculation shows)

$$H=H(\mathbf{X}^i,\mathbf{P}_i)+\sum_{\mathbf{k}}\tfrac{1}{2}\omega_{\mathbf{k}}R_{\mathbf{k}}^*R_{\mathbf{k}}, \tag{6.9}$$

where $H(\mathbf{X}^i,\mathbf{P}_i)$ is, as pointed out in Sec. 2, just the original Hamiltonian expressed in terms of $\mathbf{X}^i$, $\mathbf{P}_i$ variables. This is

$$H(\mathbf{X}^i,\mathbf{P}_i)=\sum_i\frac{P_i^2}{2m}+\sum_{\mathbf{k}}\frac{2\pi e^2}{k^2}[\![\rho_{\mathbf{k}}^*\rho_{\mathbf{k}}]\!],$$

where $[\![\rho_{\mathbf{k}}]\!]$ represents the value of $\rho_{\mathbf{k}}$ on the equilibrium hypersurface; viz., according to Eq. (4.11): $[\![\rho_{\mathbf{k}}]\!]=[\![q_{\mathbf{k}}+\eta_{\mathbf{k}}]\!]=[\![\eta_{\mathbf{k}}]\!]$ (since, by hypothesis, $[\![q_{\mathbf{k}}]\!]\sim[\![\frac{1}{2}(R_{\mathbf{k}}+R_{\mathbf{k}}^*)]\!]=0$). Therefore, this part of the Hamiltonian reduces to

$$H(\mathbf{X}^i,\mathbf{P}_i)=\sum\frac{P_i^2}{2m}+\sum_{\mathbf{k}}^{k_s}\frac{2\pi e^2}{k^2}[\![\eta_{\mathbf{k}}^*\eta_{\mathbf{k}}]\!]+H_{\mathrm{s.r.}}. \tag{6.9a}$$

Here $H_{\mathrm{s.r.}}$ houses all the terms with $k>k_s$, cut off from the second sum. These evidently correspond to short-range interactions, and will be neglected here (see Bohm-Pines[11] for discussion).

To obtain the equations of motion, we need the Poisson brackets. We recall, however,[1] that while the variables have to be defined to second order in $R_{\mathbf{k}}$ (in order to show that $R_{\mathbf{k}}$ has zero Poisson brackets with $\mathbf{X}^i$, $\mathbf{P}_i$ while the Poisson brackets of the latter are functions only of $\mathbf{X}^i$ and $\mathbf{P}_i$) these second-order expressions are never actually needed explicitly, provided that one simply uses the above properties of the Poisson brackets. We conclude then from (6.9) that the $R_{\mathbf{k}}$ oscillate harmonically, with

$$\dot{R}_{\mathbf{k}}=[R_{\mathbf{k}},H]=-i\omega_{\mathbf{k}}R_{\mathbf{k}},\quad \dot{R}_{\mathbf{k}}^*=[R_{\mathbf{k}}^*,H]=i\omega_{\mathbf{k}}R_{\mathbf{k}}^*,$$

in agreement with (6.3). As for the $\mathbf{X}^i$, $\mathbf{P}_i$, we could obtain the equations of motion in a similar way with the aid of the Poisson brackets. For example,

$$[\mathbf{X}^i,\mathbf{P}_j]=\delta_{ij}-i\sum_{\mathbf{k}}\left(\frac{\partial R_{\mathbf{k}}}{\partial\mathbf{x}^i}\frac{\partial R_{\mathbf{k}}^*}{\partial\mathbf{p}_j}-\frac{\partial R_{\mathbf{k}}}{\partial\mathbf{p}_j}\frac{\partial R_{\mathbf{k}}^*}{\partial\mathbf{x}^i}\right)+\text{c.c.} \tag{6.10}$$

However, we may instead utilize the theorem proved in Sec. 2; viz., that the projected variables $\mathbf{X}^i$, $\mathbf{P}_i$ are special solutions of the original equations of motion; so that

$$\begin{aligned}\dot{\mathbf{X}}^i&=\frac{\partial H(\mathbf{X}^i,\mathbf{P}_i)}{\partial\mathbf{P}_i}=\frac{\mathbf{P}_i}{m},\\ \dot{\mathbf{P}}_i&=-\frac{\partial H(\mathbf{X}^i,\mathbf{P}_i)}{\partial\mathbf{X}^i}=-i\sum_k\frac{2\pi e^2}{k^2}\mathbf{k}[\![\eta_{\mathbf{k}}]\!]\exp(i\mathbf{k}\cdot\mathbf{X}^i).\end{aligned} \tag{6.11}$$

$e\eta_{\mathbf{k}}$ is just the kth Fourier coefficient of the part of the charge density that is not associated to plasma oscillations, but rather to random fluctuations. Equation (6.11) can then be written in the form

$$md^2\mathbf{X}^i/dt^2=-e[\nabla\varphi_\eta(\mathbf{x})]_{\mathbf{x}=\mathbf{X}^i}, \tag{6.12}$$

where $\varphi_\eta(\mathbf{x})$ is the potential arising from the charge distribution which is the Fourier transform of $e\eta_{\mathbf{k}}$. This is just the "screened" Coulomb potential as discussed in the Bohm-Pines treatment. Thus, the essential results of this treatment are reproduced.

In the absence of oscillation, the long-range Coulomb potential is then, of course, screened out. The residual part of this potential, appearing in (6.12) implies, however, that the particles will still deflect each other somewhat, although much less than would have occurred without the screening (since $\eta_{\mathbf{k}}$ is in general only a small fraction of $\rho_{\mathbf{k}}$ when $\mathbf{k}\cdot\mathbf{v}^i/\omega_p$ is small). The residual deflection is, in fact, necessary for the consistency of the relations $R_{\mathbf{k}}(\mathbf{x}^i,\mathbf{p}_i)=0$, $R_{\mathbf{k}}^*(\mathbf{x}^i,\mathbf{p}_i)=0$, which must hold in the absence of oscillation. For these relations imply a curved surface in the configuration space, and if the configuration point is to remain on such a surface, then [in analogy with the example given in I (Sec. 2)] there must be a force which opposes the "centrifugal" force, and prevents the configuration point from shooting off tangentially from the surface. Therefore, not all of the Coulomb potential can be screened; a part must survive to provide this deflecting force, and this part is just $\varphi_\eta(\mathbf{x})$.

There will also be [as in case (B) of the example] a "Coriolis" force coupling the oscillatory motion perpendicular to the equilibrium variety in *configuration* space to the tangential motion. (It is only when one goes to phase space and uses the "symplectic" normal in phase space that the oscillatory motion is purely "perpendicular" to the equilibrium variety.) However, if we go to the L.V.A. [analogous to case (A) of our simple example], in which the oscillatory coordinates are taken as $Q_{\mathbf{k}}=\alpha_k\sum\exp[-i(\mathbf{k}\cdot\mathbf{x}^i)]$ and $P_{\mathbf{k}}=(m\dot{Q}_{\mathbf{k}})$ [see Eq. (4.7)] then, as shown in Eq. (3.22), the oscillatory motion will be orthogonal to the equilibrium variety in *configuration* space (in the usual Euclidean sense), being decoupled from the tangential motion. This approximation means that the "Coriolis" and "centrifugal" forces are being neglected. In this case, we obtain from (2.1a) and (3.22)

$$\begin{aligned}\delta\mathbf{x}^i&=\sum_{\mathbf{k}}^{k_s}\frac{\partial Q_{\mathbf{k}}}{\partial\mathbf{x}^i}Q_{\mathbf{k}}=\sum_{\mathbf{k}}^{k_s}\frac{i\beta_k\mathbf{k}}{(\omega_{\mathbf{k}})^{1/2}}\rho_{\mathbf{k}}\exp(i\mathbf{k}\cdot\mathbf{x}^i)\,,\\ \delta\mathbf{p}_i&=\sum_{\mathbf{k}}\left(\frac{\partial Q_{\mathbf{k}}}{\partial\mathbf{x}^i}P_{\mathbf{k}}-\frac{\partial P_{\mathbf{k}}}{\partial\mathbf{x}^i}Q_{\mathbf{k}}\right).\end{aligned}\tag{6.13}$$

Now, it can be seen that the term $Q_{\mathbf{k}}(\partial P_{\mathbf{k}}/\partial\mathbf{x}^i)=(\omega_{\mathbf{k}})^{-1/2}\beta_k\mathbf{k}(\mathbf{k}\cdot\mathbf{p}_i)\exp(-i\mathbf{k}\cdot\mathbf{x}^i)Q_{\mathbf{k}}$ is of order $(\mathbf{k}\cdot\mathbf{v}^i)/\omega_p$ times $\sum_{\mathbf{k}}(\partial Q_{\mathbf{k}}/\partial\mathbf{x}^i)P_{\mathbf{k}}$, so that in this approximation [where $(\mathbf{k}\cdot\mathbf{v}^i/\omega_p)$ is assumed to be very small] we can neglect it.

Therefore,

$$\delta\mathbf{p}_i=\sum_{\mathbf{k}}\frac{\partial Q_{\mathbf{k}}}{\partial\mathbf{x}^i}P_{\mathbf{k}}=\sum_{\mathbf{k}}^{k_s}\frac{i\alpha_k\mathbf{k}}{(\omega_{\mathbf{k}})^{1/2}}P_{\mathbf{k}}\exp(-i\mathbf{k}\cdot\mathbf{x}^i)\,.\tag{6.14}$$

The Hamiltonian is

$$H=\sum(P_i^2/2m)+\tfrac{1}{2}\sum_{\mathbf{k}}(Q_{\mathbf{k}}^*Q_{\mathbf{k}}+\omega_p^2P_{\mathbf{k}}^*P_{\mathbf{k}})\,,\tag{6.15}$$

where

$$\mathbf{P}_i=\mathbf{p}_i-\delta\mathbf{p}_i\,.\tag{6.16}$$

From Eq. (3.22), we see that the $(\partial Q_{\mathbf{k}})/(\partial\mathbf{x}^i)$ should be orthogonal to each other (in the usual, Euclidean sense); viz., that

$$\sum_i\frac{\partial Q_{\mathbf{k}}^*}{\partial\mathbf{x}^i}\frac{\partial Q_{\mathbf{k}'}}{\partial\mathbf{x}^i}=\delta_{\mathbf{k}\mathbf{k}'}\,,\tag{6.17}$$

and if we write $Q_{\mathbf{k}}=\alpha_k\sum\exp[-i(\mathbf{k}\cdot\mathbf{x}^i)]$, $(\partial Q_{\mathbf{k}})/(\partial\mathbf{x}^i)=-i\mathbf{k}\alpha_k\exp[-i(\mathbf{k}\cdot\mathbf{x}^i)]$, we obtain

$$\sum_i\frac{\partial Q_{\mathbf{k}}^*}{\partial\mathbf{x}^i}\frac{\partial Q_{\mathbf{k}'}}{\partial\mathbf{x}^i}=\alpha_k\alpha_{k'}(\mathbf{k}\cdot\mathbf{k}')\sum_i\exp(i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}^i)\,,$$

and in the R.P.A., the above is equal to $\delta_{\mathbf{k},\mathbf{k}'}$.

Now, consider a certain point $\mathbf{x}_0^i$, $\mathbf{p}_{i0}$ on the equilibrium hypersurfaces $Q_{\mathbf{k}}(\mathbf{x}^i)=0$, $P_{\mathbf{k}}(\mathbf{x}^i,\mathbf{p}_i)=0$. In the case that we are considering [$(\mathbf{k}\cdot\mathbf{v}^i)/\omega_p$ very small], the configuration space point $\mathbf{x}_0^i$ will move only a very short distance compared to the wavelength of oscillation during several periods of collective oscillation. As a result, during this time, the normal to the above hypersurfaces changes negligibly in direction, so that, to a first approximation, we can replace the surface by its tangent hyperplane.

The vectors $\delta\mathbf{x}_{\mathbf{k}}^i=(\partial Q_{\mathbf{k}}^*)/(\partial\mathbf{x}^i)Q_{\mathbf{k}}$ and $\delta\mathbf{p}_{i\mathbf{k}}=(\partial Q_{\mathbf{k}}^*)/(\partial\mathbf{x}^i)P_{\mathbf{k}}$ evidently form an orthogonal set which is normal to the intersection of the hypersurfaces in configuration space $Q_{\mathbf{k}}=0$. But by their very definition, $\delta\mathbf{x}^i=\sum_{\mathbf{k}}\delta\mathbf{x}_{\mathbf{k}}^i$ and $\delta\mathbf{p}_i=\sum_{\mathbf{k}}\delta\mathbf{p}_{i\mathbf{k}}$ are the parts of $\mathbf{x}^i$ and $\mathbf{p}_i$ that are normal to this intersection of the hypersurfaces. Therefore, the remainder $\mathbf{X}^i=\mathbf{x}^i-\delta\mathbf{x}^i$, and $\mathbf{P}^i=\mathbf{p}_i-\delta\mathbf{p}_i$, is the part of $\mathbf{x}^i$, $\mathbf{p}^i$ which is *in* the intersection of these hypersurfaces.

The above relationships suggest that it will be worthwhile to consider an orthogonal transformation[12] (i.e., a generalized rotation in configuration space), with the point $\mathbf{x}_0^i$ taken as a fixed center. This rotation can, in principle, be chosen so that each of the s new axes (labeled $\zeta_{\mathbf{k}}$) will be parallel to one of the $\delta\mathbf{x}_{\mathbf{k}}^i$, and therefore, normal to the intersection of the surfaces $Q_{\mathbf{k}}(\mathbf{x}^i)=0$. The remainder of the new axes (labeled ξ_l) will then lie *in* the intersection of these surfaces. Because of the orthonormality of the $(\partial Q_{\mathbf{k}})/(\partial\mathbf{x}^i)$, it follows that

$$\zeta_{\mathbf{k}}=Q_{\mathbf{k}}\,.\tag{6.17a}$$

The momenta, being regarded as covariant vectors embedded in $\mathbf{x}^i$ space will undergo the same transformation as the $\mathbf{x}^i$. Thus, $p_{\mathbf{k}}$ the momenta canonically conjugate to $\zeta_{\mathbf{k}}$, will be

$$p_{\mathbf{k}}=P_{\mathbf{k}}\,.\tag{6.17b}$$

[12] Such a transformation was considered in D. Bohm [D. Bohm, in *The Many Body Problem*, edited by C. DeWitt (John Wiley & Sons, Inc., New York, 1959)] for the L. V. A. case, but on the basis of a more specialized line of reasoning.

The Hamiltonian (6.15) becomes

$$H=\sum_{t=1}^{3N-s}\left[(p_t)^2/2m\right]+\tfrac{1}{2}(\zeta_{\mathbf{k}}{}^2+p_{\mathbf{k}}{}^2)\,, \qquad (6.18)$$

where p_t is the momentum conjugate to ξ_t.

From the above Hamiltonian, we see that (as was to be expected, of course) the $\zeta_{\mathbf{k}}$ quantities oscillate harmonically (i.e., the motion normal to the equilibrium hypersurface is stable). However, in this approximation the p_t will be constants of the motion, and their corresponding coordinate will satisfy the equation

$$\xi_t=(p_t/m)t+\xi_{t0}. \qquad (6.19)$$

The motion in the equilibrium variety is, therefore, not stable, because the ξ_t increase without limit.

Of course, after some time, the approximation of replacing the equilibrium variety by its tangent variety will break down. In other words, there is no simple integrable separation of the kind given in Eq. (6.19). Nevertheless, this equation will hold for a number of periods of oscillation, so that one obtains a good qualitative idea of the nature of the motion.

The ξ_t, p_t then behave effectively as if they were free particles. We, therefore, called them "quasiparticles." These particles are purely abstractions, and are evidently not among the original constituent particles of the system. This is clear, because there are only $N-s/3$ of them. Moreover, if one evaluates carefully the orthogonal transformation leading to (6.17a) and (6.17b), one will see that when one of the ξ_t moves in a linear way, independently of the others, all of the original particles move with very complicated correlated motions [indeed, these correlations are just what are needed to maintain the relations $Q_{\mathbf{k}}(\mathbf{X}^i)=0$ and $P_{\mathbf{k}}(\mathbf{X}^i,\mathbf{P}_i)=0$]. Associated with such ξ_t motions there will be fluctuations of the $P_{\mathbf{k}}$, but only for $k>k_D$ (for $k<k_D$, $\rho_{\mathbf{k}}\sim Q_{\mathbf{k}}$ is by hypothesis independent of the $\mathbf{X}^i$, $\mathbf{P}_i$, and therefore of ξ_t). The ξ_t, therefore, correspond to a random background motion of the charge, in which only short-range fluctuations can appear.

7. CONCLUSION

Very little can be said about the behavior of a many-particle system unless it possesses some symmetry properties, other than merely those associated with the elementary conservation laws. Conceptually, a symmetry property shows up in several equivalent ways: in the existence of a *uniform constant* of the motion, of a *conservation law*, of a *stable equilibrium variety* in phase space (which should *truly* be of lower dimensionality than that of the energy hypersurface of the system, i.e., it should not pervade the latter in a *filamentary, quasiergodic* way[13]) or finally, in the existence of *collective coordinates* (or, as one could equivalently say, in the possibility *in principle* of at least partially transforming the system to *normal modes*).

The question of *finding* these symmetry properties (or these constants of the motion, collective coordinates, etc.) in a systematic way, is probably one of the deepest and most difficult problems of physics. All that so far has been done in this direction, from Poincare's and Birkhoff's formal series to partial summations of Feynman diagrams, merely seems to lift some corners of the blanket but provides us with not much more of a *general, systematic* solution than some lucky guesses of the form of collective (or of the remaining, individual) coordinates, e.g., the plasmons in an electron gas or the quasiparticles in superconductivity.

On the other hand, even if the symmetry properties of a system are given, one is still confronted with the problem of how to exploit this information for a description of the motions of the particles and, generally, of the dynamics of various dynamical variables of interest. Thus, we know that plasma waves observed in an electron gas must have their origin in some sort of coherent vibration of the individual electrons "on top" of their random thermal motion, but it is only when we systematically separate the dynamics into the two corresponding parts that we can answer some more refined questions, e.g., the explanation, from first dynamical principles, of the Cooper[14] pairing correlations,[15] or of the vortices in liquid He.[15] In dynamical considerations of this kind, the availability of a geometric-intuitive visualization, in phase space, seems to be of great heuristic help. Thus,[15] the geometric interpretation of the Poisson brackets seems to lead to a very elegant clue as to how, under certain circumstances, the R.P.A. (Sec. 4) can be circumvented and improved collective modes be defined.

The treatment in these papers was entirely classical. From the point of view of the questions of principle which we sought to clarify—the interrelationship between constants of motion, equilibrium varieties, collective coordinates, and nonsecular, stable oscillations—this is an advantage. These questions—except for features which depend on quantum statistics—are all inherent in the dynamical formalism, whether the latter be of q numbers or of c numbers, and are certainly easier to handle if the quantities involved are all commutative. For practical applications, however, the quantum counterpart is necessary. At first sight, the use of Poisson brackets $[\mathbf{X}^i,\mathbf{P}_j]\neq\delta_j{}^i$, $[\mathbf{P}_i,\mathbf{P}_j]\neq 0$ etc., seems to present significant difficulties, and one might feel inclined to return to the Bohm-Pines formalism of redundant variables[11] (in which $[\mathbf{X}^i,\mathbf{P}_j]=\delta_j{}^i$ etc.), even at the cost of allotting to the $\mathbf{X}^i$, $\mathbf{P}_j$ variables (which describe only *part* of the particle motion and therefore, in reality, occupy only a $6N$-$2s$-dimensional space) the

[13] E. Fermi, Z. Physik **24**, 261 (1923).

[14] L. N. Cooper, Phys. Rev. **104**, 1189 (1956); J. Bardeen, L. N. Cooper, and J. R. Schrieffer, *ibid.* **108**, 1175 (1957).

[15] Work now in progress by one of us (G. C.).

unnatural extension of a $6N$-dimensional space. At closer inspection, however, the apparent difficulty turns out to be beneficial,[15] and by not evading it one remains closer to the physical reality of the problem, while none of the facilities of the other methods, for cases which are solvable in some approximation, are lost. This is hoped to be dealt with in subsequent papers.

ACKNOWLEDGMENTS

The authors wish to acknowledge the hospitality of Professor M. H. L. Pryce and of the Department of Physics at Bristol University, England, where the basic part of this research was performed. One of us (G.C.) also wishes to thank Professor Y. Aharonov for his constant interest and for clarifying discussions.

APPENDIX A: THE GEOMETRICAL MEANING OF THE POISSON BRACKETS $[Z^i,Z^j]$

In Sec. 2, Eq. (2.11), we derived the following expression for the rate of change of the projected point (i.e., its velocity):

$$\dot{Z}^a=(\partial H/\partial Z^b)[Z^b,Z^a]. \tag{A.1}$$

We now proceed to interpret geometrically the Poisson brackets appearing in this equation. This will be of interest in itself, and will also help to simplify the expression for $\dot{Z}^a$ further.

We recall (Sec. 2) that since the Poisson brackets $[Z^b,Z^a]$ can be evaluated on the equilibrium variety, $Q^\alpha=0$, only the first-order expansion (2.6c) for Z^a is needed to evaluate them. We obtain

$$\frac{\partial Z^a}{\partial z^i}=\delta_i{}^a-O^{ab}O_{\mu\nu}\left[\!\!\left[\frac{\partial Q^\mu}{\partial z^i}\right]\!\!\right]\cdot\left[\!\!\left[\frac{\partial Q^\nu}{\partial z^b}\right]\!\!\right]. \tag{A.2}$$

To work out the meaning of the equations of motion (A.1) for Z^a, it is convenient first to define

$$O^{ab}O_{\mu\nu}\left[\!\!\left[\frac{\partial Q^\mu}{\partial z^i}\right]\!\!\right]\cdot\left[\!\!\left[\frac{\partial Q^\nu}{\partial z^b}\right]\!\!\right]=L_i{}^a. \tag{A.3}$$

Let us now consider the effect of the matrix, $L_i{}^a$ on a (covariant) vector u in phase space. First of all, let us choose

$$u_a=[\![(\partial Q^\alpha/\partial x^a)]\!], \tag{A.4}$$

which represents the direction of the normal to the hypersurface, $Q^\alpha=0$. We have

$$L_i{}^a u_a=O^{ab}O_{\mu\nu}\left[\!\!\left[\frac{\partial Q^\mu}{\partial z^i}\right]\!\!\right]\cdot\left[\!\!\left[\frac{\partial Q^\nu}{\partial z^b}\right]\!\!\right]\cdot\left[\!\!\left[\frac{\partial Q^\alpha}{\partial z^a}\right]\!\!\right].$$

But according to (2.2b) and (2.5) $O^{ab}[\![(\partial Q^\nu/\partial z^b)]\!]\times[\![(\partial Q^\alpha/\partial z^a)]\!]=O^{\alpha\nu}$, and by (2.5a), we obtain

$$L_i{}^a u_a=-\delta_\mu{}^\alpha[\![(\partial Q^\mu/\partial z^i)]\!]=-u_i. \tag{A.5}$$

It follows then that u_a is an *eigenvector* of the matrix, $L_i{}^a$, belonging to the eigenvalue (-1). Since this result holds for all the Q^α's, it follows that any vector normal to the hypersurface $Q^\alpha=0$ must likewise be such an eigenvector.

On the other hand, if we choose

$$u_a'=\partial\xi/\partial z^a, \tag{A.6}$$

where ξ is a member of the set of variables that complement the Q^α and are canonically independent of Q^α [as introduced in I (Sec. 3)], then by a similar calculation we obtain

$$L_i{}^a u_a'=O^{ab}O_{\mu\nu}\left[\!\!\left[\frac{\partial Q^\mu}{\partial z^i}\right]\!\!\right]\cdot\left[\!\!\left[\frac{\partial Q^\nu}{\partial z^b}\right]\!\!\right]\cdot\left[\!\!\left[\frac{\partial \xi}{\partial z^a}\right]\!\!\right].$$

Because of the canonical independence of ξ and Q^ν, $[Q^\nu,\xi]=O^{ab}[\![(\partial Q^\nu/\partial z^b)]\!](\partial\xi/\partial z^a)=0$. Therefore,

$$La_i u_a'=0. \tag{A.7}$$

The above results show that $L_i{}^a$ is a matrix which projects out to zero any (covariant) vector normal to the surfaces $\xi=0$, and multiplies by (-1) any vector normal to $Q^\alpha=0$. From this, it follows that $P_i{}^a=(\partial Z^a)/(\partial z^i)=\delta_i{}^a+L_i{}^a$ is a projection matrix, which projects to zero any vector normal to $Q^\alpha=0$, and leaves unchanged any vector normal to $\xi=0$.

Now, it is clear that the Poisson brackets $[Z^b,Z^a]$ are closely related to the above projection matrices. Thus,

$$[Z^b,Z^a]=O^{ij}\frac{\partial Z^b}{\partial z^i}\frac{\partial Z^a}{\partial z^j}=O^{ij}P_i{}^bP_j{}^a. \tag{A.8}$$

Thus, the equations of motion (2.11) can be written as

$$\dot{Z}^a=O^{ij}P_j{}^aP_j{}^b[\![(\partial H/\partial z^b)]\!].$$

In order to complete the evaluation of $\dot{Z}^a$, let us consider $P_i{}^b[\![\partial H/(\partial z^b)]\!]=v_i$, which is, of course, another covariant vector in the phase space, and let us study the quantity:

$$O^{ij}P_j{}^a v_i=O^{ij}(\delta_j{}^a+L_j{}^a)v_i.$$

[Note that this matrix product is different from the one given after (A.4) because of the introduction of O^{ij} and the summation over the i, j indices instead of the a, b, indices.] We have [from (A.3)]

$$O^{ij}L_j{}^a=O^{ij}O^{ab}O_{\mu\nu}\left[\!\!\left[\frac{\partial Q^\mu}{\partial z^j}\right]\!\!\right]\cdot\left[\!\!\left[\frac{\partial Q^\nu}{\partial z^b}\right]\!\!\right].$$

As in the previous case, we apply this matrix to two vectors. First, if $v_i=[\![(\partial Q^\alpha)/(\partial z^i)]\!]$, we obtain

$$\begin{aligned}O^{ij}L_j{}^a v_i&=O^{ij}O^{ab}O_{\mu\nu}\left[\!\!\left[\frac{\partial Q^\mu}{\partial z^i}\right]\!\!\right]\cdot\left[\!\!\left[\frac{\partial Q^\nu}{\partial z^b}\right]\!\!\right]\cdot\left[\!\!\left[\frac{\partial Q^\alpha}{\partial z^i}\right]\!\!\right]\\&=O^{ab}O_{\mu\nu}O^{\mu\alpha}\left[\!\!\left[\frac{\partial Q^\nu}{\partial z^b}\right]\!\!\right]\\&=\delta_\nu{}^\alpha O^{ab}\left[\!\!\left[\frac{\partial Q^\nu}{\partial z^b}\right]\!\!\right]=O^{ab}\left[\!\!\left[\frac{\partial Q^\alpha}{\partial z^b}\right]\!\!\right]=O^{ai}v_i.\end{aligned}$$

In a similar way, it can be shown that if we choose $v_i'=\partial\xi/(\partial z^i)$, then $O^{ij}L_j{}^a v_i'=0$. Remembering that $O^{ia}=-O^{ai}$, it follows then that

$$O^{ij}P_j{}^a=O^{ij}(\delta_j{}^a+L_j{}^a) \tag{A.9}$$

is a projection matrix contravariant in its two indices, that yields zero for covectors normal to the surfaces $Q^\alpha=0$, and leaves unchanged covectors that are normal to $\xi=0$, except for turning them into contravectors.

Now from (2.10), $\partial H/(\partial z^i)$, will in general, have two contributions, one not containing Q, and the second proportional to Q, so that the contribution of the latter will vanish on the equilibrium surface. We conclude then that $[\![(\partial H)/(\partial z^i)]\!]$ has contributions from $(\partial H(Z^a))/(\partial z^i)$ only. And since, as shown in I (Sec. 4), $[Z^a,Q]$ vanishes, it follows that Z^a and $H(Z^a)$ will be functions of the complementary variables ξ only. Therefore, the vector $u_i=[\![\partial H/(\partial z^i)]\!]$ will be an eigenvector of the projection operators that we have discussed, belonging to the eigenvalue unity. As a result, Eq. (2.16) reduces to

$$\dot{Z}^a=O^{ai}[\![(\partial H/\partial Z^i)]\!]. \tag{A.10}$$

In this way, we have demonstrated directly that when the equations of motion of the Z are obtained from the expansion for the Hamiltonian (2.10), with the correct Poisson bracket relations, they reduce to the original equations of motion evaluated on the equilibrium surface. This means that the $Z^i(t)$ *represent a possible motion of the system* (which is in fact an equilibrium motion).

APPENDIX B: THE CANONICAL INVARIANCE OF THE METRICAL TENSOR O^{ij} AND THE RELATION TO THE SYMPLECTIC GROUP

Although the invariance of the metrical tensors O_{ij} and O^{ij} is evident from the discussion in Sec. 3, (3.1)–(3.3), it is instructive to demonstrate this directly. To do this, we note that any canonical transformation can be built from a series of infinitesimal canonical transformations, so that we may restrict ourselves to a discussion of the latter only. In the condensed phase-space notation, such an infinitesimal canonical transformation reads

$$z'^i-z^i\equiv Dz^i=O^{ij}\lambda(\partial V/\partial z^j), \tag{B.1}$$

where $\lambda V(z^j)$ is the generating function, λ being a small constant. By differentiating (B.1) we obtain

$$\Delta Dz^i\equiv D\Delta z^i=O^{ij}\lambda\frac{\partial^2 V}{\partial z^j\partial z^k}\Delta z^k. \tag{B.2}$$

This shows that when the z^i space undergoes the *over-all* infinitesimal canonical transformation (B.1), then this induces at any point z^i a *local* affine transformation on the vectors Δz^i (which is, of course, also infinitesimal), and whose coefficients are given in (B.2). To bring this out more sharply, let us write $(\partial^2 V)/(\partial z^j\partial z^k)=a_{jk}$. We then have

$$D\Delta z^i=O^{ij}\lambda a_{jk}\Delta z^k. \tag{B.3}$$

The coefficients of the affine transformations are then

$$b_k{}^i=O^{ij}\lambda a_{jk}, \tag{B.4}$$

so that

$$\Delta z'^i-\Delta z^i\equiv D\Delta z^i=b_k{}^i\Delta z^k. \tag{B.5}$$

We shall now apply this transformation to the scalar product (3.5) expressed in terms of the *new* variables $(z')^i$, viz.:

$$E(\delta z'^i,\Delta z'^i)=O_{ij}\Delta z'^i\delta z'^j. \tag{B.6}$$

We have for the change of E under the transformation (to first order in λ)

$$\begin{aligned}D(E)&=O_{ij}D(\Delta z^i)\delta z^j+O_{ij}(\Delta z^i)D(\delta z^j)\\&=\lambda(O_{ij}O^{jk}(\partial V/\partial z^k)\Delta z^k\delta z^j+O_{ij}\Delta z^j O^{jk}(\partial V/\partial z^k)\delta z^k).\end{aligned}$$

Using (3.4) and the antisymmetry of O_{ij} and O^{ij}, we obtain

$$D(E)=\lambda\left(-\frac{\partial V}{\partial z^j}\Delta z^j\delta z^j+\frac{\partial V}{\partial z^k}\Delta z^k\delta z^k\right)\equiv 0.$$

It follows then that the scalar product (3.5) is invariant to an arbitrary infinitesimal canonical transformation. But by going through the above detailed proof of this invariance, we have gained further insight into the problem. For the local affine transformation (B.5) induced by the over-all transformation (B.1) in z^i space is not an arbitrary affine transformation. Rather, it is restricted by the relation

$$b_k{}^i=\lambda O^{ij}a_{jk},$$

where $a_{jk}=(\partial^2 V)/(\partial z^j\partial z^k)$ must be a *symmetrical* matrix. On the other hand, general affine transformation would allow a_{jk} to be an *arbitrary* matrix. Thus, the set (B.5) of local affine transformations has only $\frac{1}{2}(6N)(6N+1)$ parameters and not $(6N)^2$ as an arbitrary affine transformation would have.

It is readily verified that the set of affine transformations restricted in the above way (i.e., by a_{jk} being a symmetric matrix) form a group (as they must, since the local affine transformations are induced by the over-all canonical transformations, which form a group). This local group is isomorphic to what is known as the *symplectic group*.[2,16] This symplectic group is in fact defined as essentially the one which keeps the bilinear form $O_{ij}u^iv^j$ invariant.

It is basically because O_{ij} is an *antisymmetrical* tensor that the scalar product of $O_{ij}u^iv^j$ is invariant to a

[16] It should be noted that M. Schönberg [M. Schönberg, Ac. Brasil de Ciênc. **29**, 473 (1958); **30**, 1, 117, 259, 429 (1958)], mentions the symplectic character of classical phase space, while noting that the Jordan-Klein quasialgebra of the boson creation and annihilation operators is related to the symplectic geometry in the same way as the Jordan-Wigner algebra of fermion creation and annihilation operators is related to affine geometry of twice the dimensionality (namely, by being its Clifford algebra).

symplectic transformation. The usual symmetrical metric tensor g_{ij} leads in fact (when diagonalized), to an invariance of the corresponding form $g_{ij}u^iv^j=\sum u^iv^i$ to an orthogonal transformation (or more generally to a pseudo-orthogonal transformation, if the eigenvalues of g_{ij} are not all positive definite).

APPENDIX C: REVIEW OF SOME BASIC PROJECTIVE, AFFINE, AND METRIC CONCEPTS AS APPLIED TO A SYMPLECTIC SPACE

In simple affine geometry (without a metric), one can obtain scalar products only of a vector v^i with another one w_i that is contragredient to it; viz., $E=w_iu^i$ [e.g., if $u^i=\Delta z^i, w_i=\partial F/(\partial z^i)$, then $E=\partial F/(\partial z^i)\Delta z^i=\Delta F$ is a scalar, if F is a scalar]. Such scalar products do in fact have a simple geometrical interpretation. For the contravariant vector w_i is in a one-to-one correspondence with the hyperplanes passing through the point z^i, these hyperplanes being defined by the equation

$$w_i\Delta z^i=0. \tag{C.1}$$

Thus, it may be said that w_i determines a unique hyperplane. For the general point $z^i+\Delta z^i$ (which is in the neighborhood of z^i), we have

$$w_i\Delta z^i=E.$$

As a result, the scalar product $E=w_i\Delta z^i$ is a measure of the separation of the point $z^i+\Delta z^i$ from the hyperplane (C.1) (i.e., $E=0$ if the point is incident in the hyperplane, and $|E|$ increases if one goes out from it on any given line).

Now, what a general metrical tensor m does is to make possible the definition of scalar products of a vector v^i with another vector w^i that is *cogredient* with it. How does this possibility come about? The answer is basically that to each covariant vector w_i, the general metrical tensor associates a *unique* contravariant vector $w^i=m^{ij}w_i$, and to each contravariant vector w^i, a unique covariant vector $w_i=m_{ij}w^j$. In view of (3.10), this means that to each contragredient vector is associated a unique hyperplane and to each hyperplane a unique contragredient vector. And if m_{ij} is either symmetric ($m_{ij}=g_{ij}$) or antisymmetric ($m_{ij}=O_{ij}$), then the correspondence is not only one-one, but also symmetrical, in the sense that if the correspondence is applied twice (i.e., vector to hyperplane and hyperplane to vector), then the original vector is obtained again. Such a correspondence is called a *polarity* in projective geometry.[17]

It follows then that one can regard a *metrical* scalar product $m_{ij}u^iv^j$ as a simple *affine* scalar product of either the vector u^i with the hyperplane $m_{ij}v^j$, polar to the vector v_j, or of the vector v^j with the hyperplane $m_{ij}u^i$ polar to u_i. In either case, the metrical scalar product is a measure of the separation of one vector from the polar plane associated to the other. Since this separation is to represent an invariant geometrical relationship between the two vectors u and v, it is necessary that the metrical tensor m_{ij} be so defined that the bilinear form $m_{ij}u^iv^j$ is invariant (as is evidently the case with our tensor $m_{ij}=O_{ij}$).

With an ordinary Euclidean metric ($m_{ij}=\delta_{ij}$) which is invariant only to orthogonal transformations, the polarity defined by m_{ij} is very simple. For to each vector w^i is associated a plane coordinate $w_i=\delta_{ij}w^j=w^i$ (the upper and lower index vectors being the same in Euclidean metric). The plane $w_i\Delta x^i=\sum_i w^i\Delta x^i=0$ is simply the plane through the origin (taken here to be z^i) which is perpendicular (in the simple Euclidean sense) to the vector w^i, and when $\sum_i w^i\Delta x^i=E\neq 0$, then, as is well known, E is the simple Euclidean distance of the point $x^i+\Delta x^i$ to the plane $\sum_i w^i\Delta x^i=0$. The scalar product of w^i with v^i is then just the distance from the point $\Delta x^i=v^i$ to the plane $\sum_i w^i\Delta x^i=0$.

In the phase-space problem for which $m_{ij}=O_{ij}$, we find that the plane $w_i=O_{ij}w^i$ *contains* the vector w^i, as is evident from the fact that $w_iw^i=w^iO_{ij}w^j=0$. Thus, the phase-space metric, O_{ij}, associates to each vector a polar hyperplane passing through it and this provides a further interpretation of the fact that the length of every vector is zero.

Vice versa, given a covariant vector w_i one can associate to it a polar contravariant vector $w^i=O^{ij}w_j$ which is evidently *in* the hyperplane $w_i\Delta x^i=0$ (because $w_iO^{ij}w_j=0$). For example, consider any function $F(z^i)$. The surface $F=0$ defines (in the neighborhood of each point z^i) the hyperplane $\partial F/(\partial z^i)\Delta z^i=0$. Associated to the covariant vector $w_i=\partial F/(\partial z^i)$ (the local plane coordinates of the hypersurface) is the contravariant vector w^i, which is polar to it; viz.,

$$w^i=O^{ij}(\partial F/\partial z^j), \tag{C.2}$$

and from $\partial F/(\partial z^i)O^{ij}(\partial F)/(\partial z^j)\equiv 0$ it follows, of course, that w^i is *in* the hypersurface $Q=0$. As a special case of this relation, consider the equation of motion

$$(dz^i/dt)=O^{ij}(\partial H/\partial z^j).$$

This equation then means that the velocity $(dz^i)/dt$ is the vector in phase space which is polar to the hypersurface $H=c$ at the point in question (and which, of course, lies in this hypersurface, as it should be).

Thus far, we have not really introduced any metrical concepts, such as perpendicularity and perpendicular projection. All our discussions until this point have been done in terms of the projective relationships, such as incidence of lines in hyperplanes, and polarities. In order to work out the genuinely metrical implications of the tensor O_{ij} we shall first consider the related problem of perpendicular projections. Now the general notion of projection arises in the problem of decomposing a vector Δz^i into a sum of parts,

$$\Delta z^i=\textstyle\sum_\mu \delta z_\mu{}^i. \tag{C.3}$$

[17] See e.g., J. A. Todd, *Projective and Analytic Geometry* (Pitman Publishing Corporation, London, 1948).

Each part $\delta z_\mu{}^i$ should represent a displacement from the point $z^i+\Delta z^i$ [where $F^\mu \cong (\partial F^\mu)/(\partial z^i)\Delta z^i$] to the surface $F^\mu=0$ (as in fact was done in our separation of the motion into oscillatory and nonoscillatory parts in the previous sections). The question then arises as to the proper definition of the *direction* in which this displacement is to be carried out. Evidently, this direction must be chosen in a way that is invariant to all transformations under consideration in the problem of interest (e.g., rotations in ordinary space, canonical transformations in phase space, etc.). From the projective point of view, such an invariance requires that this direction be defined as that of a contravariant vector polar to some plane, which latter is invariantly associated to the problem (e.g., $F^\mu=0$).

If the metrical tensor is symmetric ($m_{ij}=g_{ij}$), then the usual procedure of defining invariant directions of projections onto the plane $F=0$, is to choose the vector $w^{i\mu}=g^{ij}(\partial F^\mu)/(\partial z^j)$ which is polar to that plane itself. Since a symmetrical metric can always be reduced to a Euclidean metric ($g^{ij}=\delta^{ij}$) by a suitable linear transformation, the meaning of this procedure can be seen by going to a local Euclidean frame. We then have $w^{i\mu}=g^{ij}(\partial F^\mu)/(\partial z^j)=(\partial F^\mu)/(\partial z^i)$, which is just a vector in the direction of the Euclidean normal to the surface $F^\mu=0$, and the projection thus defined is indeed a perpendicular projection in the usual sense.

To obtain the magnitude of $\delta z^{i\mu}$, the displacement from Δz^i to the surface $F^\mu=0$, we note that this is evidently proportional to $\Delta F^\mu=(\partial F^\mu)/(\partial z^i)\Delta z^i$ itself. In this way we arrive at the definition of $\delta z^{i\mu}$, viz.,

$$\delta z^{i\mu}=g^{ij}\frac{\partial F^\mu}{\partial z^j}\Delta F^\mu=g^{ij}\frac{\partial F^\mu}{\partial z^j}\frac{\partial F^\mu}{\partial z^k}\Delta z^k. \qquad \text{(C.4)}$$

Returning to (C.3), which expresses the requirement that the projected parts add up to give the original vector for an arbitrary Δz^i, we obtain a restriction in the functions F^μ, viz., that the $(\partial Q_\mu)/(\partial z^i)$ form an orthogonal set. For

$$\Delta z^i=\sum_\mu \delta z^{i\mu}=\sum_\mu g^{ij}\frac{\partial F^\mu}{\partial z^j}\frac{\partial F^\mu}{\partial z^j}\Delta z^k,$$

and if this is true for arbitrary Δz^k, we have

$$\sum_\mu g^{ij}\frac{\partial F^\mu}{\partial z^j}\frac{\partial F^\mu}{\partial z^k}=\delta_k{}^i. \qquad \text{(C.5)}$$

If we had gone in the dual way of defining z^i as a function of F^μ and projecting ΔF^μ as a sum of projections, we would have the dual relation

$$\sum\frac{\partial F^\mu}{\partial z^i}\frac{\partial F^\nu}{\partial z^j}g^{ij}=\delta^{\mu\nu}. \qquad \text{(C.6)}$$

For an antisymmetric metric, however, such as O_{ij}, the above described procedure of defining the direction of projection from $z^i+\Delta z^i$ onto the hyperplane $F^\mu=0$ by the vector polar to $F^\mu=0$ itself, is not adequate. For, as we have seen, the vector $O^{ij}(\partial F^\mu)/(\partial z^j)$ polar to $(\partial F^\mu)/(\partial z^j)$ is now in the plane $F^\mu=0$ itself, rather than pointing outside of it.

In order to deal with this problem, it is essential to recall that in the general projective point of view that we are adopting, the only role of the metrical tensor in defining the projection onto $Q^\mu=0$ is to determine a unique direction of projection, as a vector polar to *some* plane. This plane need not, in general, be the plane $F^\mu=0$ itself, *but can be any plane that is invariantly related to $F^\mu=0$.* In short, in order to project onto a plane $F^\mu=0$, we may use as the direction a vector that is polar to some other plane $F^\nu=0$. But if we start with a complete set of functions F^μ which are in involution [i.e., $O^{ij}(\partial F^\mu/\partial z^i)(\partial F^\nu/\partial z^j)=O^{\mu\nu}$ as defined in Eq. (2.5)], then to each surface $F^\mu=0$, there is only one other F^ν which defines a polar vector $O^{ij}(\partial F^\nu)/(\partial z^j)$ that is *not* in the surface $F^\mu=0$; viz., the $F^\nu(=F^\alpha)$ which is *canonically conjugate* to F^μ. Let us therefore define the direction of projection into $F^\mu=0$ as along that vector $O^{ij}(\partial F^\alpha)/(\partial z^j)$ which is polar to the canonical conjugate of F^μ. As in the case of the symmetrical metric, the measure of the displacement is $\Delta F^\mu=(\partial F^\mu)/(\partial z^k)\Delta z^k$. Thus, the projected part, $\delta z_\mu{}^i$, will be equal to $O^{ij}(\partial F^\alpha)/(\partial z^j)F^\mu=O^{ij}(\partial F^\alpha/\partial z^j)(\partial F^\mu/\partial z^l)\Delta z^l$ (recalling that F^α is canonically conjugate to F^μ). Using the symbol $O_{\mu\nu}$ defined in (2.5), we obtain for the above

$$\delta z_\mu{}^i=O^{ij}O_{[\mu]\nu}\frac{\partial F^{[\mu]}}{\partial z^j}F^\nu=O^{ij}F_{[\mu]\nu}\frac{\partial F^{[\mu]}}{\partial z^j}\frac{\partial F^\nu}{\partial z^l}\Delta z^l \qquad \text{(C.7)}$$

(where the sum over $[\mu]$ is not carried out, even though $[\mu]$ is a dummy index).

Returning to the requirement (C.3) that the sum of the projected parts shall equal the original vector, we must have (with μ being summed over)

$$\sum_\mu \delta z_\mu{}^i=\Delta z^i=O^{ij}O_{\mu\nu}\frac{\partial F^\mu}{\partial z^j}\frac{\partial F^\nu}{\partial z^l}\Delta z^l. \qquad \text{(C.8)}$$

Since this is true for arbitrary z^i, we have

$$O^{ij}O_{\mu\nu}\frac{\partial F^\mu}{\partial z^j}\frac{\partial F^\nu}{\partial z^l}=\delta_l{}^i.$$

To see what this means, we write it out in terms of the $(Q_{\text{k}},P_{\text{k}})$ notation (i.e., $F^\mu=Q_{\text{k}}$ for $\mu=1\cdots 3N$, and $F^\mu=P_{\text{k}}$ for $\mu=3N+1\cdots 6N$) viz.,

$$\sum_{\text{k}}\left(\frac{\partial Q_{\text{k}}}{\partial x^l}\frac{\partial P_{\text{k}}}{\partial p_i}-\frac{\partial Q_{\text{k}}}{\partial p_i}\frac{\partial P_{\text{k}}}{\partial x^l}\right)=\delta^i{}_l. \qquad \text{(C.9)}$$

This is just the Lagrange bracket of x^l with p_i, and the above requirement is satisfied if the Q_k, P_k form a canonical set (in involution).

If we had expressed Z^i as a function of Q_k, and made a similar projection of $\Delta Q^\mu = \sum_\mu \delta Q^\mu$, we would have obtained instead of the Lagrange bracket relation, the Poisson bracket relations

$$\sum_i \left(\frac{\partial Q_k}{\partial \mathbf{x}^i} \frac{\partial P_{k'}}{\partial \mathbf{p}_i} - \frac{\partial Q_k}{\partial \mathbf{p}_i} \frac{\partial P_{k'}}{\partial \mathbf{x}^i} \right) = \delta_{kk'}, \qquad \text{(C.10)}$$

which also are satisfied if Q_k, P_k, form a canonical set of functions.

PHYSICAL REVIEW VOLUME 133, NUMBER 2A 20 JANUARY 1964

Meissner Effect and Flux Quantization in the Quasiparticle Picture*

DIETRICH A. UHLENBROCK AND BRUNO ZUMINO
Department of Physics, New York University, Washington Square, New York

(Received 29 July 1963)

A superconductor in a magnetic field is studied by means of a generalized quasiparticle transformation. Using the minimum principle for the thermodynamic grand potential, nonlinear equations are derived for the superconductor, which can be considered valid at finite temperature. For small magnetic potential these equations are linearized and shown to imply the London formula and the Meissner effect. For a multiply-connected superconductor the nonlinear equations are shown to be consistent with Maxwell's equation only if the magnetic flux is quantized in the units predicted by Onsager.

1. INTRODUCTION

In the basic work of Bardeen, Cooper, and Schrieffer,[1] the energy spectrum of a superconductor is described in terms of independent quasiparticle excitations. The natural mathematical tool for this description is the quasiparticle canonical transformation of Bogolyubov[2] and Valatin.[3] The presence of external electromagnetic fields, however, invalidates the simple pairing of the original Bogolyubov-Valatin transformation. In order to study this situation one can introduce a generalized quasiparticle transformation which makes no such pairing assumptions. This was done again by Bogolyubov[2,4] and also, in slightly different forms, by a number of other authors.[5] Using these methods a fully gauge invariant treatment of the Meissner effect has been achieved.

These discussions of the Meissner effect are usually limited to the case of absolute zero temperature. The generalization to finite temperatures is obviously desirable in order to study electromagnetic properties up to the transition temperature. However, it is far from obvious how one should proceed in order to generalize to finite temperatures some of the above mentioned discussions. It is convenient to follow rather closely Bogolyubov's work[4] which can be generalized to finite temperature without too much difficulty.[6] We make an ansatz for the density matrix in terms of independent quasiparticles and use the principle of minimization for the grand potential. In this way we obtain a set of nonlinear equations for a superconductor in a magnetic field which are valid at finite temperature. The Meissner effect is discussed in a fully gauge invariant way by linearizing the equations with respect to the vector potential. The current is seen to vanish at the critical temperature except for residual Landau-type diamagnetism. In the simple special case of a factorizable interaction one recovers the temperature dependence of the penetration depth given originally by BCS.[1]

Once the Meissner effect is established, one can use the nonlinear equations to study the phenomenon of magnetic flux quantization in a multiply-connected superconductor. This phenomenon, which was predicted theoretically by London[7] and Onsager,[8] has been verified experimentally by Deaver and Fairbank and by Doll and Näbauer.[9] For mathematical simplicity

* The research reported in this paper was supported in part by the National Science Foundation under Grand No. NSF-G20964 and in part by the Army Research Office (Durham) under Contract No. DA-ARO-(D)-31-124-G356.

[1] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).

[2] N. N. Bogolyubov, J. Phys. (U.S.S.R.) **11**, 23 (1947); Zh. Eksperim. i Teor. Fiz. **34**, 58 (1958) [translation: Soviet Phys.—JETP **7**, 41 (1958)].

[3] J. Valatin, Nuovo Cimento **7**, 843 (1958).

[4] N. N. Bogolyubov, Usp. Fiz. Nauk **67**, 549 (1959) [translation: Soviet Phys.—Usp. **2**, 236 (1959)].

[5] See e.g., J. Valatin, Phys. Rev. **122**, 1012 (1961); and J. M. Blatt, Progr. Theoret. Phys. (Kyoto) **24**, 851 (1960).

[6] Actually we can make some simplifications, since we do not seek the energy spectrum of collective excitations. This latter problem can be studied separately and was actually solved by E. R. Velibekov, Dokl. Akademii Nauk SSSR **142**, 1265 (1962) [translation: Soviet Phys.—Doklady **7**, 134 (1962)].

[7] F. London, *Superfluids* (Dover Publications, Inc., New York, 1961), Vol. I.

[8] L. Onsager, Phys. Rev. Letters **7**, 50 (1961).

[9] B. S. Deaver and W. M. Fairbank, Phys. Rev. Letters **7**, 43 (1961); R. Doll and M. Näbauer, Phys. Rev. Letters **7**, 51 (1961); Z. Physik **169**, 526 (1962).