THE
PHYSICAL REVIEW

A journal of experimental and theoretical physics established by E. L. Nichols in 1893

SECOND SERIES, VOL. 133, NO. 2A 20 JANUARY 1964

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

D. BOHM
Birkbeck College, London, England

AND

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

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

In two papers (of which this is the first) our central concern is to draw conclusions about the over-all dynamical properties of a many-body system. This is done without trying to solve the equations of motion, but rather, on the basis of our knowledge of oscillatory or collective variables (or more generally, from the existence of conservation rules and of the uniform constants of the motion). Our main result is that, corresponding to the collective coordinates (or the uniform constants of the motion) there exists a separation of the motions into two parts, one of which is collective or oscillatory, and regular, and the other of which is noncollective, nonoscillatory, and irregular. This separation is here obtained by a (canonically invariant) method of projections in phase space, from the actual phase point $\mathbf{x}^i$, $\mathbf{p}_i$ along a certain line, which is the direction of a "pure" de-excitation of an oscillation, down to a certain "projected point" $\mathbf{X}^i$, $\mathbf{P}_i$, which is the intersection of the line with an equilibrium subspace (or variety), the latter consisting of all the points in the phase space for which the collective excitation is zero. We apply this separation to an illustrative example consisting of a simple two-dimensional model, possessing all the essential features of the general problem under discussion. We obtain the results corresponding to the Bohm-Pines theory, as applied to this case, in a very simple way, without having to introduce supernumerary variables or subsidiary conditions (our results being generalized to the plasma case in the following paper). Instead of subsidiary conditions, we have a corresponding number of identities among the "projected motions" $\mathbf{X}^i$, $\mathbf{P}_i$, so that in effect, $\mathbf{X}^i$, $\mathbf{P}_i$, together with the collective oscillatory variables, span a space of $6N$ dimensions (where N is the number of particles). This definition of the $\mathbf{X}^i$, $\mathbf{P}_i$ replaces the two canonical transformations of Bohm-Pines, and is equivalent to a certain noncanonical transformation, which removes the collective part of the motion. Our method may also be regarded as a systematic generalization of that of Tomonaga; firstly, being an extension of the latter's method from configuration space to phase space, and to collective variables that are momentum dependent, and secondly, being the development of a general separation method for arbitrary variables, which contains Tomonaga's Taylor expansion of the Hamiltonian as a special case. The projection method associates to each actual motion $\mathbf{x}^i(t)$, $\mathbf{p}_i(t)$ a unique equilibrium motion $\mathbf{X}^i(t)$, $\mathbf{P}_i(t)$, about which it oscillates. This association is such that, from the very way in which it is defined, the possibility of an indefinitely large increase of $\delta\mathbf{x}^i=\mathbf{x}^i-\mathbf{X}^i$, $\delta\mathbf{p}_i=\mathbf{p}_i-\mathbf{P}_i$ with time is avoided, so that the $\delta\mathbf{x}^i$, $\delta\mathbf{p}_i$, will oscillate stably in every order of approximation, without the need for special precautions to avoid secular terms, as is necessary in the usual perturbation treatments (e.g., in celestial mechanics).

1. INTRODUCTION

THE most general way of stating the central problem of many-body dynamics is that we have a large number (10^{23}, say) of interacting particles, for which, in principle, we know the equations of motion, but whose solution is, evidently, impossible in practice. Moreover, even if we did know the solutions, they would be of no use directly, because we would be lost in the huge mass of data required to express them. The central question is, therefore, to discover over-all properties, which enable us to draw conclusions about the general behavior of the system without our having to solve the problem in all detail.

* 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 Air Force Office of Scientific Research and the National Science Foundation, U.S.A.

A familiar example of such procedure is afforded by the thermodynamic properties of a large scale system, which are treated by statistical methods. These properties are not sufficient, however, for a discussion of the dynamics of the system, because they refer only to quasiequilibrium conditions. On the other hand, it is well known that there are conservation rules and associated constants of the motion, which enable some conclusions about the dynamics of the system to be drawn, without a detailed solution for the motion of all the particles. In fact, if one had a sufficient number of constants of the motion one could in principle use these to solve completely for the behavior of the particles, provided that these constants of the motions were all *uniform*. By a uniform constant of the motion, one means a function $f(\mathbf{x}^i,\mathbf{p}_i)$ of the positions and the momenta of all the particles, which determines a "regular" hypersurface in phase space if f is given a certain value. Generally, most constants of the motion of the system are not uniform, being represented by surfaces which fill a region of phase space quasiergodically, so that a small change of $\mathbf{p}$ and $\mathbf{x}$ can correspond to a large change of f. If a function represents a nonuniform constant of the motion, it will not be of much use for drawing conclusions about the system which are independent of the details of the movements, because the determination of f provides no real restrictions on the location of the system in phase space.

The only known exact uniform constants of the motion valid for all isolated systems are the total energy, momentum, and angular momentum. However, there is a wide range of systems having a large number of uniform functions that are approximately constants of the motion. In particular, systems with a collective behavior (such as the electron-ion plasma) can quite easily be seen to possess uniform constants of motion to the same approximation in which the behavior is collective. For example, if a system has a collective coordinate Q_k and a corresponding canonical momentum P_k which oscillate harmonically with frequency ω_k, then by means of a canonical transformation, we obtain an action variable

$$J_k=\tfrac{1}{2}(m\omega_k Q_k^2)+(P_k^2/m\omega_k),$$

and an angle variable

$$\phi_k=\tan^{-1}(P_k/m\omega_k Q_k),$$

with

$$\phi_k=\phi_{k0}+\omega_k t.$$

The action variable J_k is everywhere a regular function of Q_k and P_k, and, being a constant of the motion, is therefore a uniform constant of this kind.[1] The constancy of J_k expresses the conservation of the energy of a single plasma mode (of course, only approximately, because such modes are really damped, usually after a fairly large number of oscillations). On the other hand, the constant of the motion ϕ_{k0} is multivalued, and has an irregularity at $Q_k=P_k=0$, so that it is not a uniform constant of the motion. Therefore, the only constant that is relevant for the separation of the motion into two dynamically independent parts is J_k.

It is evidently desirable to obtain as many uniform constants of the motion as possible, whether exact or approximate. As we have already pointed out, however, we can in general obtain only some fraction of the total number of degrees of freedom in this form.[2] If there were a complete set of uniform constants of the motion, then, as we have already indicated, the intersection of the associated surfaces in phase space would determine the phase orbit completely, so that the description in terms of particle coordinates could be discarded altogether. If (as is true in general) we do not have a complete set of uniform constants of the motion, we cannot discard the particle description altogether. It will be the main object of these papers to develop a systematic method of dividing the motion of the system into two parts, one of which is associated in a natural way to the constants of motion and the other which is the remainder. In this way, as we shall see, one is able to draw many conclusions concerning the dynamics of the system, without actually solving the equations of motion fully, and in a way that is independent of the details of individual particle movements.

We shall exemplify this separation by the elementary case of conservation of total momentum $\boldsymbol{\pi}=\Sigma\mathbf{p}_i$ of a system of N particles (which, in fact, has already been treated in several ways by many authors). The canonically conjugate coordinate is the center of mass, $\boldsymbol{\xi}=(1/N)\sum_i \mathbf{x}^i$. In this case, the separation suggested above can be expressed by writing the coordinates and momenta of each particle in the form[3]

$$\begin{aligned}\mathbf{x}_i&=\mathbf{X}^i(t)+\boldsymbol{\xi}(t)\\ \mathbf{p}_i&=\mathbf{P}_i(t)+\frac{1}{N}\boldsymbol{\pi}(t).\end{aligned}\tag{1.1}$$

If $V=\sum V(\mathbf{x}^i-\mathbf{x}^j)=\sum V(\mathbf{X}^i-\mathbf{X}^j)$, the equations of

[1] In all cases in literature when a system is said to possess collective modes, these modes are implicitely supposed to be uniform functions of $\mathbf{x}^i$, $\mathbf{p}_i$.

[2] e.g., in the electron plasma, the collective coordinates ceases to be collective if the associated wavelength is smaller than the Debye length, which is generally much larger than the interparticle spacing. See, for example, D. Pines and D. Bohm, Phys. Rev. **85**, 338 (1952).

[3] The collective parts of $\mathbf{x}^i$, $\mathbf{p}_i$—here $\boldsymbol{\xi}$ and $(1/N)\boldsymbol{\pi}$, and in general $\delta\mathbf{x}^i$ and $\delta\mathbf{p}_i$—do *not* have to be canonically conjugate to each other; rather, they are quantities *derived* from the collective canonical pair (here $\boldsymbol{\xi}$ and $\boldsymbol{\pi}$) in such a way as to give $(\mathbf{X}^i,\boldsymbol{\xi})=(\mathbf{P}_i,\boldsymbol{\xi})=(\mathbf{X}^i,\boldsymbol{\pi})=(\mathbf{P}_i,\boldsymbol{\pi})=0$. This insures the separation of the equations of motion into independent sets.

motion split correspondingly into two independent sets,

$$\frac{d\xi}{dt}=\frac{\pi}{m}, \qquad \frac{d\mathbf{X}^i}{dt}=\frac{\mathbf{P}_i}{m}$$

$$\frac{d\pi}{dt}=0, \qquad \frac{d\mathbf{P}_i}{dt}=-\frac{\partial}{\partial \mathbf{X}^i}\sum_{ij} V(\mathbf{X}^i-\mathbf{X}^j).$$

What will be relevant to the subsequent discussion is the (for this case trivial) fact that the two parts of the motion have a physically different character, viz., that the ξ part increases linearly with time, while the remaining part has complicated changes of momentum resulting from the interactions. In the general case (e.g., collective coordinates), the validity of such a separation will similarly depend on the fact that the two parts differ physically in various ways, such as that their characteristic frequencies will be very different; that one will be ordered and the other disordered; that one may be mainly collective and the other mainly individual; that one may be stable, the other unstable, etc. Altogether the limitation on the number of available uniform constants of motion will be seen to be an inherent physical limitation which splits the degree of freedom into two groups of different physical character.

The example afforded by the center-of-mass motion is so simple that not much is gained by the separation described above. But in other cases (such as the electron plasma) where there is a fairly large number of uniform approximate constants of motion, the part associated with these constants (the collective part) will separate out as a self-determining group of components of the motion which contain a significant amount of information about the behavior of this system. A systematic separation of the dynamics may, therefore, provide considerable additional insight into the behavior of the system.

The generalization of (1.1) will be seen to be (see 3.6)

$$\begin{aligned}\mathbf{x}^i&=\mathbf{X}^i-\textstyle\sum_{\mathrm{k}} J_{\mathrm{k}}\dfrac{\partial\phi_{\mathrm{k}}}{\partial\mathbf{p}_i}+\text{higher order terms in } J_{\mathrm{k}},\\ \mathbf{p}_i&=\mathbf{P}_i+\textstyle\sum_{\mathrm{k}} J_{\mathrm{k}}\dfrac{\partial\phi_{\mathrm{k}}}{\partial\mathbf{x}^i}+\text{higher order terms in } J_{\mathrm{k}},\end{aligned} \tag{1.2}$$

where the J_{k} are the uniform constants of the motion and the ϕ_{k} are their canonical conjugates. Generally, for any dynamical variable $F(\mathbf{x}^i,\mathbf{p}_i)$, a similar split can be made, the result being

$$F(\mathbf{x}^i,\mathbf{p}_i)=F(\mathbf{X}^i,\mathbf{P}_i)-\textstyle\sum_{\mathrm{k}} J_{\mathrm{k}}[F,\phi_{\mathrm{k}}]+\text{higher order terms.} \tag{1.2a}$$

Here $[F,\phi_{\mathrm{k}}]$ represents the Poisson bracket of F and ϕ_{k}. $\delta\mathbf{x}^i\equiv\mathbf{x}^i-\mathbf{X}^i$ and $\delta\mathbf{p}_i=\mathbf{p}_i-\mathbf{P}_i$ are the parts of the motion of each particle associated with the uniform constants of motion, while $\mathbf{X}^i,\mathbf{P}_i$ represent the remainder for which the $J_{\mathrm{k}}(\mathbf{X}^i,\mathbf{P}_i)$ are zero. Alternatively, the same separation can be expressed in terms of the oscillatory variables

$$Q_{\mathrm{k}}=(2J_{\mathrm{k}}/\omega_{\mathrm{k}})^{1/2}\cos\phi_{\mathrm{k}}, \quad P_{\mathrm{k}}=(2J_{\mathrm{k}}\omega_{\mathrm{k}})^{1/2}\sin\phi_{\mathrm{k}},$$

by means of similar equations which will be given in Secs. 2 and 3 [the identities $J_{\mathrm{k}}(\mathbf{X}^i,\mathbf{P}_i)=0$ being replaced by $P_{\mathrm{k}}(\mathbf{X}^i,\mathbf{P}_i)=0$, $Q_{\mathrm{k}}(\mathbf{X}^i,\mathbf{P}_i)=0$].

As in the case of center-of-mass coordinates [Eq. (1)], the equations of motion for $\mathbf{X}^i,\mathbf{P}_i$ will be seen to separate completely from those for $\delta\mathbf{x}^i$, $\delta\mathbf{p}_i$; i.e., the two parts of the motion will prove to be dynamically independent. It is important, however, to express this separation in terms of the canonical formalism. The problems that arise in doing this can be illustrated in terms of the example of center-of-mass coordinates. Let the original Hamiltonian of the system be

$$H=\sum_{i=1}^{N}\frac{p_i^2}{2m}+\sum_{ij} V(\mathbf{x}^i-\mathbf{x}^j).$$

This can evidently be written as

$$H=\frac{1}{2m}\sum_i\Big(\mathbf{p}_i-\frac{1}{N}\sum_j\mathbf{p}_j\Big)^2+\frac{N}{2m}\Big(\frac{1}{N}\sum_j\mathbf{p}_j\Big)^2+\sum_{ij} V(\mathbf{x}^i-\mathbf{x}^j). \tag{1.3}$$

By going over to the variables defined by the separation (1), we obtain

$$H=\sum_i\frac{P_i^2}{2m}+V(\mathbf{X}^i-\mathbf{X}^j)+\frac{\pi^2}{2mN}=H_{\mathrm{I}}(\mathbf{X}^i,\mathbf{P}_i)+H_\pi(\pi). \tag{1.4}$$

We see that the Hamiltonian has in fact split into two parts, corresponding to our split in the coordinates. It is important to stress, however, that the variables $\mathbf{P}_i$ and $\mathbf{X}^i$ satisfy the identities

$$\sum\mathbf{P}_i\equiv 0, \quad \sum\mathbf{X}^i\equiv 0. \tag{1.5}$$

This means that there are in reality only $3N-3$ independent pairs of canonical variables among the $\mathbf{X}^i$, $\mathbf{P}_i$. Indeed, the relevant Poisson brackets,

$$[\mathbf{X}^i,\mathbf{P}_j]=\delta_j{}^i-1/N, \tag{1.6}$$

express the fact that the $\mathbf{X}^i$, $\mathbf{P}_i$ do not form an independent set of canonical variables.

There are two ways in which one can now proceed. The first is to ignore the noncanonical character of the $\mathbf{X}^i$, $\mathbf{P}_i$, i.e., one changes the formulation so as to have $[\mathbf{X}^i,\mathbf{P}_j]=\delta_j{}^i$ (which is the form most easily cast into quantum-theory). To achieve that, Eqs. (1.5) can no

longer hold as identities but rather as subsidiary conditions for singling certain solutions $\mathbf{X}^i(t)$ out of the totality of solutions $\mathbf{x}^i(t)$ (the consistency of such a subsidiary condition with the time-development of the system is gauranteed ultimately by J_k being a constant of the motion and, therefore, $J_k=0$ holding for all times if it holds for $t=0$). Thus, the $\mathbf{X}^i$, $\mathbf{P}_i$ are nothing but $\mathbf{x}^i$, $\mathbf{p}_i$'s which obey certain initial conditions $J_k=0$ and, therefore, they are obviously canonical variables for which $[\mathbf{X}^i,\mathbf{P}_j]=\delta_j{}^i$. However, this means that the total momentum variable π can no longer be identically equal to $\sum_i \mathbf{p}_i$ (because if it were, the $\mathbf{P}_i$, for which we want to keep the definition $\mathbf{P}_i=\mathbf{p}_i-(1/N)\pi$, would fulfill $\sum_i \mathbf{P}_i=0$ *identically*). Thus, π (and, similarly, its canonical conjugate ξ) must be additional, "redundant" variables which span together with the $\mathbf{X}^i$, $\mathbf{P}_i$, a $(6N+2)$-dimensional phase space. The relations $\pi=\sum \mathbf{p}_i$ must then be interpreted as subsidiary conditions in this extended space, i.e., conditions which single out certain motions in the extended domain of variables. Again, the consistency of these conditions is guaranteed by the constancy of π and the problem reduces to an initial-values problem. ξ and π remain canonically conjugate to each other in this scheme, although they are now variables independent of the original particle variables. To obtain the correct equations of motion for these "redundant" variables, the original Hamiltonian must be extended by a π-dependent part and, when the Poisson bracket relationships are applied, one obtains "extended" equations of motions for all variables involved. By applying the subsidiary conditions, however, one returns to the original "restricted problem," which is physically the correct one. The principal difficulty in this procedure (which is essentially the one adopted by Bohm and Pines in their treatment of the plasma oscillation variables[4]) is that upon transition to quantum theory, the subsidiary condition eliminates the redundant degrees of freedom but not their zero-quantum fluctuations. The latter give rise to divergencies (see Ref. 5) which, although they can be shown to be harmless (Ref. 6, Chap. IV), nevertheless make it desirable also to consider another and more direct approach.

The second approach is in regarding (1.5) as constraints which are *identically* satisfied and, using the corresponding Poisson bracket relation (1.6), $\mathbf{X}^i$, $\mathbf{P}_i$, π, ξ are now all explicit functions of x^i and p_i whose variation with t will follow, of course, from the formula $\dot{f}=[f,H]$ which holds for any function $f(x,p)$. We then note, as can easily be verified, that

$$[\mathbf{X}^i,H_{\mathrm{II}}]=[\mathbf{P}_i,H_{\mathrm{II}}]=[\xi,H_{\mathrm{I}}]=[\pi,H_{\mathrm{I}}]=0. \quad (1.7)$$

Therefore,

$$d\mathbf{X}^i/dt=[\mathbf{X}^i,H]=[\mathbf{X}^i,H_{\mathrm{I}}]=\frac{1}{m}\sum_j [\mathbf{X}^i,\mathbf{P}_j]\mathbf{P}_j$$

$$=\frac{1}{m}\left[\mathbf{P}_i-\frac{1}{N}\sum_j \mathbf{P}_j\right]=\frac{1}{m}\mathbf{P}_i,$$

$$d\mathbf{P}^i/dt=[\mathbf{P}_i,H]=[\mathbf{P}_i,H_{\mathrm{I}}]=\sum \frac{\partial V}{\partial \mathbf{X}_j}[\mathbf{P}_i,\mathbf{X}_j]$$

$$=-\sum \frac{\partial V}{\partial \mathbf{X}^j}\left(\delta_i{}^j-\frac{1}{N}\right)=-\frac{\partial V}{\partial \mathbf{X}^i},$$

$$d\xi/dt=[\xi,H]=[\xi,H_{\mathrm{II}}]=\frac{\pi}{\mathbf{m}\mathbf{N}},$$

$$d\pi/dt=(\pi,H)=(\pi,H_{\mathrm{II}})=0.$$

The essential point is that the separated parts, $H_{\mathrm{I}}(\mathbf{X}^i,\mathbf{P}_i)$ and $H_{\mathrm{II}}(\pi)$, of the Hamiltonian, serve as effective Hamiltonians in the derivation of the equations of motions of $\mathbf{X}^i$, $\mathbf{P}_i$ and ξ, π, respectively, provided that the correct (noncanonical) Poisson bracket relations are used. Thus, the separation in the behavior of the dynamical variables is matched by the corresponding separation in the Hamiltonian.

In the above case, we could evidently have foreseen the separation in the Hamiltonian (as well as in the variables themselves) without any special method. In these papers we shall show that our general method for effecting a separation for any dynamical variables, indicated in the discussion leading up to Eq. (1.2), will also lead to a separation of the Hamiltonian into parts which will assume similar roles as in the example discussed. This result will be valid for any system for which one knows a certain number of uniform constants of motion J_k and their canonical conjugates ϕ_k (alternatively, the oscillating variables Q_k and P_k). Indeed, the essential results are implicit in Eq. (1.2a). For if (1.2a) is applied to the Hamiltonian function, one obtains just the desired separation. Moreover, as a result of the way which $\mathbf{X}^i$ and $\mathbf{P}_i$ are defined, their Poisson brackets with the J_k vanish, while the Poisson brackets of the $\mathbf{X}^i$, $\mathbf{P}_i$ within themselves are not functions of the J_k, ϕ_k. This, combined with the separation of the Hamiltonian, will guarantee the separation of the motion into two dynamically independent parts.

This method of separation constitutes a general way of formulating the many-body problem canonically, requiring no redundant variables or subsidiary conditions (and may therefore be considered as an alternative to the Bohm-Pines method).

On the other hand, our treatment can be regarded as a consistent generalization of the method of Tomonaga,[7]

[4] D. Bohm and D. Pines, Phys. Rev. **92**, 609 (1953).
[5] C. Kuper, Proc. Phys. Soc. (London) **A69**, 492 (1956); E. Adams, Phys. Rev. **98**, 1130 (1955).
[6] D. Bohm, *The Many-Body Problem*, edited by C. DeWitt (John Wiley & Sons, Inc., New York, 1959).
[7] S. Tomonaga, Progr. Theoret. Phys. (Kyoto) **13**, 464, 481 (1955).

who discussed the part $\delta\mathbf{x}^i$ of the motion associated with a given collective oscillation and gave the collective part of the Hamiltonian, corresponding to our separation described above. However, he did not introduce the $\mathbf{X}^i$, $\mathbf{P}_i$, nor did he discuss any properties of the noncollective part of the dynamics. The procedure used by Tomonaga for the separation of the Hamiltonian was formulated for this variable alone, as a rather special application of Taylor's theorem; no attempt was made to affect a separation in the $\mathbf{x}^i$, $\mathbf{p}_i$ themselves or in any other dynamic variable. Furthermore, Tomonaga's method is applicable only to the rather special case of collective coordinates which are functions of the $\mathbf{x}^i$ only (and not of the $\mathbf{p}_i$), which means that, as we shall see, his method amounts to a projection procedure in configuration space rather than in phase space. It is only in the latter that the full power of the method makes itself available and that conceptual and intuitive clarity is achieved.

A fundamental difference between the method described above and the subsidiary condition method is that (as we have already indicated), in the latter, the Poisson brackets satisfy the relations $[\mathbf{X}^i,\mathbf{P}_j]=\delta_j{}^i$, while in the former, $[\mathbf{X}^i,\mathbf{P}_j]\neq\delta_j{}^i$. In the transition to quantum mechanics, the problem is therefore quite straightforward for the subsidiary condition method, because it is necessary merely to replace the dynamical variable $\mathbf{P}_i$ by the operator $(\hbar/i)(\partial/\partial\mathbf{X}^i)$. In our method, however, this evidently cannot be done. Methods like those of Dirac[8] for attacking dynamical problems with constraints would only lead back to the subsidiary condition formulation. A way of treating this problem is being investigated now and is expected to be a subject of a later paper. At present we shall restrict ourselves to the classical case, in which most of the essential characteristics of the many-body problem emerge more clearly and simply than in the quantum formulation anyway.

Conceptually, one obtains the separation by considering the relation between equilibrium and nonequilibrium states. The hypersurfaces associated with the zero values of the uniform constants of motion J_k can in general be seen to be surfaces of "equilibrium" in a sense which is the direct generalization of the usual definition of this concept. The noncollective part $\mathbf{X}^i(t)$ of any motion $\mathbf{x}^j(t)$, may thus be considered as a kind of zero motion "about" which the actual motion oscillates in a stable way. This zero motion fulfills, as we have seen, $J_k(\mathbf{X}^i,\mathbf{P}_i)=0$, i.e., it is represented by a phase point moving on the equilibrium surface $J_k=0$ in phase space (or rather, on the intersection of all J_k surfaces considered). One may, therefore, say that the $\mathbf{X}^i$ motion is obtained from the $\mathbf{x}^i$ motion by extracting the collective part of the motion from the latter; in other words, $\mathbf{X}^i$ is obtained by "de-exciting" the original $\mathbf{x}^i$ motion into a state in which the collective constants of motion J_k are zero. Of the infinitely many ways of relating an $\mathbf{X}^i$ on the hypersurface to a given $\mathbf{x}^i$ outside the surface, only the one given by Eq. (1.2) will result in a complete separation of the dynamics into independent parts; it is conceptually important to note that only then will the $\mathbf{x}^i$ motion be stably related to the $\mathbf{X}^i$ motion. We shall see that this particular choice of $\mathbf{X}^i$ for a given $\mathbf{x}^i$ admits a geometrical interpretation: $\mathbf{X}^i$, $\mathbf{P}_i$ is the perpendicular projection of $\mathbf{x}^i$, $\mathbf{p}_i$ onto the hypersurface. For this to have a well-defined meaning which is, furthermore, invariant under the canonical evolution of the motion, a canonically invariant metric must first be defined in phase space (the "symplectic" metric). The introduction of these geometrical ideas provides a rather powerful tool for intuitive thinking on the many-body problem and their applications will be discussed elsewhere.

The idea of relating a motion to a nearby equilibrium motion about which it oscillates is of course already well known, and has indeed been applied very widely in the study of many-body systems (i.e., celestial mechanics, plasma theory, collective motion in nuclei). The simplest case of this kind is the one in which the forces tend to restore the system to a certain fixed equilibrium point. Small oscillations about such points will be stable, in the sense that the system will never move far away from it. As the kinetic energy is raised, the system may of course eventually become unstable. However, it frequently turns out (e.g., in the case of collective motion, as we shall see in Sec. 5) that the restoring forces are much weaker in certain directions in phase space than in others. As a result, the system will be able to escape and to move relatively freely in the direction of weak forces, while the components of the motion in the other directions will still execute stable oscillations. In this way, the system proceeds through a set of neighborhoods in phase space which generally all have the same character. Thus, instead of oscillating around a fixed point, it will oscillate around an equilibrium hypersurface (or variety), which is a continuous set of points in phase space possessing strong restoring forces in the directions "normal" to the corresponding hypersurface, while in the direction of the hypersurface itself the restoring force is weak enough so that it is overcome by the kinetic energy. As the kinetic energy is raised still further, and if (as is usually the case) the hypersurface is curved, "centrifugal" and "Coriolis" forces come into play, with the result that the equilibrium hypersurface is in itself altered, so that the surface depends on the general state of motion of the system (in the case of collective coordinates this situation will be seen to arise when the effects of random thermal motions on the collective oscillations are taken into account).

The fact that the motion is stable only in the direction normal to the equilibrium variety and not in the direction tangential, means that a small perturbation can in

[8] P. A. M. Dirac, Can. J. Math. **2**, 129 (1950).

general cause unstable transitions between the various possible motions *within* this variety (i.e., transitions between the various motions in the hypersurface), which, however, will never lead the system far away from the variety. In this way, one obtains a generalization of the concept of stable motions around an equilibrium orbit (e.g., as in planetary motions). In the latter case, the equilibrium variety is just simply an orbit (an ellipse) which is a one dimensional set of points. Such an orbit can be stable in the sense that if there is a small deviation (e.g., due to a perturbation), the system will still oscillate around the orbit in question. Nevertheless, even if the *orbit* is stable, the *motion* need not, in general, be so. For a small perturbation of the linear momentum *in* the direction of the orbit may cause a change of position along the orbit which accumulates with time, so that the perturbed *motion* would no longer remain close to the unperturbed one, although the orbits of the two motions coincide. In the more general case of a higher dimensional equilibrium variety (e.g., the plasma, where this variety consists of a hypersurface of $3N-s$ dimensions), the possible instabilities of this kind are of course far more complex, because there can be unstable transtions which change the "directions" of the equilibrium orbit as well as its linear momentum, while the motions in directions "normal" to this variety remain stable.

A common method of treating such problems has been to begin with the actual orbit $\mathbf{x}^i(t)$, $\mathbf{p}_i(t)$ and compare this with some equilibrium orbit $\mathbf{x}_0{}^i(t)$, $\mathbf{p}_{0_i}(t)$ such that the difference $\mathbf{x}^i(t)-\mathbf{x}_0{}^i(t)$ is small enough so that perturbation theory may be applied (the "difference method"). However, because the motion is unstable to transitions within the equilibrium variety, there arises the well-known problem of "secular" perturbations. Basically this problem has its origin in the fact that, unless the initial conditions of the comparison motion $\mathbf{x}_0{}^i(t)$, $\mathbf{p}_{0_i}(t)$ are very carefully chosen in relation to those for the actual motion $\mathbf{x}^i(t)$, $\mathbf{p}_i(t)$, the difference $\delta\mathbf{x}^i(t)$, $\delta\mathbf{p}_i(t)$ will eventually increase "secularly" (i.e., unstably and without limit) thus invalidating the assumptions on which the perturbation theory is based. The proper choice of initial conditions for $\mathbf{x}_0{}^i(t)$, $\mathbf{p}_{0_i}(t)$ can be a fairly complicated problem, even in those cases where one can actually solve for the unstable features of the motion. For even if the initial conditions of $\mathbf{x}_0{}^i(t)$, $\mathbf{p}_{0_i}(t)$ are adjusted properly for an unperturbed system, any perturbation may cause transitions in the "unstable" directions, in which the change of the actual motion $\mathbf{x}^i(t)$, $\mathbf{p}_i(t)$ is somewhat different from that in the comparison motion $\mathbf{x}_0{}^i(t)$, $\mathbf{p}_{i_0}(t)$, this leading to differences $\delta\mathbf{x}^i(t)$, $\delta\mathbf{p}_i(t)$, which grow secularly. This means that at each stage of the calculation the initial conditions of $\mathbf{x}_0{}^i(t)$, $\mathbf{p}_{0_i}(t)$ must be given a special adjustment which is different for each case, so that no generally valid expressions can be obtained.

A treatment of this kind is, therefore, practicable only in simple problems in mechanics (such as in planetary motion) where a detailed solution of the unstable part of the motion is possible and indeed of considerable interest in itself. However, in cases where there are very many degrees of freedom (e.g., where collective coordinates assume an important role), the motion within the equilibrium variety is not only too complicated to be calculated, but its detailed behavior is of little interest in itself. It follows then, that for these problems, the difficulties connected with the proper definitions of the comparison motion $\mathbf{x}_0{}^i(t)$, $\mathbf{p}_{0_i}(t)$ are largely formal, and one could attempt to exploit this additional freedom to relate to each actual motion a comparison motion, in a way which is automatically free of secular instabilities for all stages of the calculation.

It is evident that what is needed here is a systematic and general way of associating to each actual motion $\mathbf{x}^i(t)$, $\mathbf{p}_i(t)$ a *particular* equilibrium motion $\mathbf{X}^i(t)$, $\mathbf{P}_i(t)$, so defined that the differences $\delta\mathbf{x}^i=\mathbf{x}^i-\mathbf{X}^i$, $\delta\mathbf{p}_i=\mathbf{p}_i-\mathbf{P}_i$ *never* become large. But as can be seen by an inspection of Eq. (2), this is just what our projection method does. Thus, for each phase point $\mathbf{x}^i$, $\mathbf{p}_i$ another phase point $\mathbf{X}^i$, $\mathbf{P}_i$ is associated, and (as will be shown in this paper) if $\mathbf{x}^i(t)$, $\mathbf{p}_i(t)$ represent a solution of the equations of motion, $\mathbf{X}^i(t)$, $\mathbf{P}_i(t)$ will be another solution of these equations. Moreover, it is evident that by that very mode of definition, $\mathbf{x}^i$ and $\mathbf{p}_i$ cannot become large; for, as can be seen from Eq. (1.2), they are proportional to J_k which latter will remain constantly small.

It follows then, that our projection method does, in fact, solve our problem of giving a proper choice of the comparison motion $\mathbf{X}^i(t)$, $\mathbf{P}_i(t)$ in such a way that secular increases of $\mathbf{x}^i$, $\mathbf{p}_i$ are automatically obviated. This is done through expressions which have general validity and which are indifferent both to the physical situation considered and to the type of perturbation used. Moreover, this choice requires no detailed solution for the unstable parts of the motion. Thus, the purely formal problems associated with the proper definition of the equilibrium motion are avoided.

2. SIMPLE MODEL OF A SYSTEM WITH AN EQUILIBRIUM VARIETY

In order to make our discussion of oscillations about equilibrium varieties more concrete, we shall in this section present a simple example of a system possessing such a variety. This example has the advantage that both the oscillatory and the nonoscillatory variables can be solved for exactly within the approximation of small oscillations. It can be, therefore, used for comparision with the results of our method which latter is, however, designed to apply to cases for which no exact solutions are possible for the nonoscillatory variables.

The model that we shall consider is that of a single particle moving in a two-dimensional potential field which is such as to possess a one-dimensional equi-

librium variety. A simple case of such a variety is a circle. Consider, for example, the potential given by

$$V=\tfrac{1}{2}\alpha(r-r_0)^2, \tag{2.1}$$

for which the equilibrium curve is the circle of radius $r=r_0$, and which implies a radial restoring force $F=-\alpha(r-r_0)$ towards the circle, $r=r_0$. If the particle is displaced a small distance away from $r=r_0$, it will tend to oscillate about this circle. To study these oscillations, let us express the equations of motion in polar coordinates. With the Hamiltonian

$$H=\frac{p_r^2}{2m}+\frac{p_\theta^2}{2mr^2}+\frac{\alpha}{2}(r-r_0)^2, \tag{2.2}$$

the equations of motions are

$$\dot{r}=\frac{p_r}{m},\quad \dot{p}_r-\frac{p_\theta^2}{mr^3}=-\alpha(r-r_0),$$
$$\dot{p}_\theta=0,\qquad \dot{\theta}=\frac{p_\theta}{mr^2}, \tag{2.3}$$

which yield

$$m\ddot{r}=-\alpha(r-r_0)+p_\theta^2/mr^3, \tag{2.4a}$$
$$p_\theta=\text{const}, \tag{2.4b}$$
$$m\ddot{\theta}=-(2p_\theta/r^3)\dot{r}. \tag{2.4c}$$

The term p_θ^2/mr^3 is of course the centrifugal force, while $-(2p_\theta/r^3)\dot{r}$ is the Coriolis force.

To treat small oscillations, we first note that there is an equilibrium radius $r=r_e$, which is determined by the equation

$$\frac{p_\theta^2}{mr_e^3}=\alpha(r_e-r_0), \tag{2.5}$$

representing a balance of restoring and centrifugal forces. We then define a small displacement $\delta r=r-r_e$. The equation for δr is obtained by expanding (2.4a) up to first-order terms in δr:

$$m\delta\ddot{r}=-\left(\alpha+\frac{3p_\theta^2}{mr_e^4}\right)\delta r. \tag{2.6}$$

As was to be expected, δr oscillates harmonically, with a frequency given by

$$\omega^2=\frac{\alpha}{m}+\frac{3p_\theta^2}{m^2r_e^4}=\omega^2(p_\theta), \tag{2.7}$$

which is a function of the angular momentum only, and, therefore, a constant in time.

The equation for θ follows from $\dot{\theta}=p_\theta/mr^2$. Expanding in δr, we obtain

$$\dot{\theta}=\frac{p_\theta}{mr_e^2}\left[1-2\frac{\delta r}{r_e}+3\left(\frac{\delta r}{r_e}\right)^2-\cdots\right]. \tag{2.8}$$

In the linear approximation the first term in δr will be adequate. Since $\delta r=A\cos(\omega t+\varphi_0)$, we have

$$\theta=\frac{p_\theta}{mr_e^2}\left(t-\frac{2}{\omega}\frac{A\sin(\omega t+\varphi_0)}{r_e}\right). \tag{2.9a}$$

Since $p_r=m\delta\dot{r}=-m\omega A\sin(\omega t+\varphi_0)$, the above can be written

$$\theta=\frac{p_\theta}{mr_e^2}t+\frac{2p_\theta}{m^2r_e^3\omega^2}p_r. \tag{2.9b}$$

The above equation shows that on top of a uniform motion in θ, proportional to the angular momentum p_θ, there is an oscillation of θ which is 90° out of phase with the oscillation in r.

Finally, let us expand the Hamiltonian (2.2) to the second order in r, so that it will give the equation of motions accurate to the first order:

$$H=\frac{p_r^2}{2m}+\frac{m}{2}\omega^2(\delta r)^2+\frac{p_\theta^2}{2mr_e^2}+\frac{\alpha}{2}(r_e-r_0)^2. \tag{2.10}$$

The physical aspects of the motion that will be relevant for us can easily be seen from the above equation. Let us begin with the case of very small p_θ [case (A)]. The equilibrium orbit can then be approximated by the circle $r=r_0$, and the frequency of oscillation of δr by $(\alpha/m)^{1/2}$. Although p_θ is small, $p_\theta t$ will eventually become appreciable. However, the oscillatory part of θ can be neglected so that the only significant oscillatory part of the motion is in the radial directions alone.

This approximation is evidently equivalent to the neglect of centrifugal and Coriolis forces, as can be seen from the equations of motion (2.4). Over a period of an oscillation, the angle changes only slightly, so that the motion along the direction of the circle may be considered as effectively rectilinear during that time. As a result of the slow movement in the circle, however, the particle experiences a slowly varying direction of the restoring force, and this will turn the direction of oscillation, so as to remain always normal to the surface. Thus, the decoupling of normal and tangential motions is not purely local, but continues in the large.

If p_θ is raised, the centrifugal and Coriolis forces will begin to play a role [case (B)]. First of all, there will be a coupling between tangential and normal motions, with the result that the oscillations cease to be normal to the equilibrium variety [as implied by Eq. (2.9)]. Secondly, the centrifugal force will cause a shift of the equilibrium orbit from $r=r_0$ to $r=r_e$, the latter being determined by the condition given earlier, viz., that the restoring force balances the centrifugal force. As a result, the equilibrium variety will depend on $p_\theta=x^1p_2-x^2p_1$, and is, therefore, a function of the momenta as well as the coordinates (i.e., it will be represented by a surface in phase space rather than in configuration space). Finally,

there is the shift of frequency of oscillation which depends, as we have seen, on p_θ.

Since we wish to use this example to discuss the separation of oscillatory and nonoscillatory parts of motion within the framework of the canonical formalism, we shall now proceed to give a canonical transformation to a new set of variables, in terms of which this separation is accomplished. Denoting the new variables by primed quantities, we write for the generating function of this transformation

$$S=p_r'[r-r_e(p_\theta')]+p_\theta'\theta. \tag{2.11}$$

The transformation itself is given by

$$\begin{aligned} p_\theta&=\frac{\partial S}{\partial\theta}=p_\theta', \\ p_r&=\frac{\partial S}{\partial r}=p_r', \\ \delta r'&=\frac{\partial S}{\partial p_r'}=r-r_e(p_\theta')=r-r_e(p_\theta), \\ \theta'&=\frac{\partial S}{\partial p_\theta'}=\theta-p_r'\frac{\partial r_e}{\partial p_\theta'}=\theta-p_r\frac{\partial r_e}{\partial p_\theta}. \end{aligned} \tag{2.12}$$

In the above transformation, the angular momentum and radial momentum are left unchanged. The new radial coordinate $\delta r'$ is just $r-r_e(p_\theta)$, the amplitude of radial oscillation which has now been transformed into one of the canonical variables. The new angle θ' differs from θ by $-p_r(\partial r_e/\partial p_\theta)$, which is equal to the second term on the right-hand side of (2.9b). We conclude, therefore, that θ' simply increases linearly with the time, because it is obtained by removing the oscillatory part from θ.

Thus, the oscillations are described by the variables $\delta r'$, p' or equivalently, by the uniform constant of motion $J_r=p_r'^2/2m\omega+\frac{1}{2}m\omega(\delta r')^2$ and its canonical conjugate $\phi=\tan^{-1}(\sqrt{m\omega}p_r/\delta r')$. These oscillatory variables are completely decoupled from the variables θ', p_θ' which have no oscillatory behavior whatsoever. This decoupling can be expressed in another way by rewriting the approximate Hamiltonian (2.10) in terms of our new set of variables;

$$H=\frac{p_r'^2}{2m}+\tfrac{1}{2}m\omega^2(p_\theta')(\delta r')^2 + \frac{p_\theta'^2}{2mr_e^2(p_\theta')}+\frac{\alpha}{2}[r_e(p_\theta')-r_0]^2. \tag{2.10a}$$

The Hamiltonian separates into a sum of two terms, one representing the energy of the oscillatory part of the motion, and the other the nonoscillatory part.[9] The role of these terms as effective Hamiltonians for determining the equations of motion for the corresponding groups of variables, will be discussed towards the end of Sec. 4.

3. DISCUSSION OF THE BASIC DYNAMICAL CONCEPTS INVOLVED IN THE SEPARATION OF THE MOTION INTO DYNAMICAL INDEPENDENT PARTS

In Sec. 2, we treated a simple example by means of straightforward methods, showing the separation of the motion into oscillatory and nonoscillatory parts by solving explicitly for the variables associated with these parts. In the many-body problem, however, it is usually possible to obtain explicit expressions only for the oscillatory variables (e.g., the collective coordinates). Indeed, as already indicated in Sec. 1, an explicit solution for the nonoscillatory part, even if it were possible (which it generally is not) would be of little use, because the full details of these complicated motions are hardly relevant for any problem of interest.

The oscillatory variables can be thought as given in the form of a certain number s of canonical pairs $Q_k(\mathbf{x}^i,\mathbf{p}_i)$, $P_k(\mathbf{x}^i,\mathbf{p}_i)$, which fulfill

$$[Q_k,P_{k'}]=\delta_{kk'},\quad [Q_k,Q_{k'}]=[P_k,P_{k'}]=0.$$

This leaves $3N-s$ degrees of freedom—the nonoscillatory or the noncollective degrees of freedom, of which only general dynamical features but no detailed solution will interest us.

In the example of Sec. 2 there is only a single pair of oscillatory variables [see Eqs. (2.12)] which we shall denote by

$$Q=\delta r'=r-r_e,\quad P=p_r'=p_r. \tag{3.1}$$

In accordance with the program outlined in Sec. 1 [see Eqs. (1.2)], we first define $\delta\mathbf{x}^i$, $\delta\mathbf{p}_i$, the purely oscillatory part of the motion. We note that there is an equilibrium hypersurface in phase space, given by the intersection of the hypersurfaces

$$Q_k(\mathbf{x}^i,\mathbf{p}_i)=0,\quad P_k(\mathbf{x}^i,\mathbf{p}_i)=0 \tag{3.2}$$

[or alternatively, by $J_k(\mathbf{x}^i,\mathbf{p}_i)=0$].

If the system is in a state of no oscillation, then the phase point $\mathbf{x}^i$, $\mathbf{p}_i$ will be moving in this intersection-surface. On the other hand, if the system is in a state of oscillation of some small amplitude, $\mathbf{x}^i$, $\mathbf{p}_i$ will be on a nearby surface, given by the (stationary) intersection of the $2s$ (moving) hypersurfaces

$$Q_k(\mathbf{x}^i,\mathbf{p}_i)=Q_{0k}\exp(-i\omega_k t),\quad P_k(\mathbf{x}^i,\mathbf{p}_i)=P_{0k}\exp(-i\omega_k t),$$

[9] At first sight, the term $\omega^2(p_\theta')(\delta r')^2/2$ seems to give a coupling between the two parts. However, because p_θ' is a constant of the motion, the only effect is to make the frequency of oscillation p_θ'-dependent, since the effect of $\partial\omega/\partial p_\theta'\neq 0$ on the equations of motion, $\dot\theta=\partial H/\partial p_\theta$, is of second order in $\delta r'$ and therefore can be neglected in our present linear treatment.

(where Q_{0k}, P_{0k} are constants) or, equivalently, of the

$$J_k(\mathbf{x}^i,\mathbf{p}_i)=\tfrac{1}{2}m\omega_k Q_{k0}{}^2+\frac{P_{k0}{}^2}{2m\omega_k}=\text{const.}$$

In attempting to separate the motion $\mathbf{x}^i(t)$, $\mathbf{p}_i(t)$ into a part $\mathbf{X}^i(t)$, $\mathbf{P}_i(t)$ which contains no contribution to the oscillation [and hence fulfills $Q_k(\mathbf{X}^i,\mathbf{P}_i)=P_k(\mathbf{X}^i,\mathbf{P}_i)=0$] and a part $\delta\mathbf{x}^i(t)$, $\delta\mathbf{p}_i(t)$ which is responsible for the oscillatory character of the $Q_k(\mathbf{x}^i,\mathbf{p}_i)=Q_k(\mathbf{X}^i+\delta\mathbf{x}^i,\ \mathbf{P}_i+\delta\mathbf{p}_i)$ and $P_k(\mathbf{X}^i+\delta\mathbf{x}^i,\ \mathbf{P}_i+\delta\mathbf{p}_i)$, care should be taken that no nonoscillatory part enter $\delta\mathbf{x}^i$, $\delta\mathbf{p}_i$ (such a part would correspond to the "secular parts" mentioned in Sec. 1); i.e., in the separation

$$\mathbf{x}^i=\mathbf{X}^i+\delta\mathbf{x}^i,$$
$$\mathbf{p}_i=\mathbf{P}_i+\delta\mathbf{p}_i,$$

$\mathbf{X}^i$, $\mathbf{P}_i$ should be "purely nonoscillatory," and $\delta\mathbf{x}^i$, $\delta\mathbf{p}_i$ "purely oscillatory." To do this without the explicit knowledge of the residual degrees of freedom (which, together with the Q_k, P_k, span the full dimensionality of phase space), we proceed in a manner which is a generalization of that used by Tomonaga[7] in the separation of the Hamiltonian. Let ξ_l, $\mathbf{p}_{\xi l}$ represent the $3N-s$ residual nonoscillatory degrees of freedom, then $\mathbf{x}^i$, $\mathbf{p}_i$ are *in principle* expressible as

$$\mathbf{x}^i(Q_k,P_k;\ \xi_l,\mathbf{p}_{\xi l}),$$
$$\mathbf{p}_i(Q_k,P_k;\ \xi_l,\mathbf{p}_{\xi l}). \tag{3.3}$$

The nonoscillatory part of $\mathbf{x}^i$, $\mathbf{p}_i$ will now obviously be

$$\mathbf{X}^i=\mathbf{x}^i(0,0;\ \xi_l,\mathbf{p}_{\xi l}),$$
$$\mathbf{P}_i=\mathbf{p}_i(0,0;\ \xi_l,\mathbf{p}_{\xi l}), \tag{3.4}$$

which can be further expressed as

$$\mathbf{X}^i=\mathbf{x}^i-\sum_k Q_k\frac{\partial\mathbf{x}^i}{\partial Q_k}-\sum P_k\frac{\partial\mathbf{x}^i}{\partial P_k}$$
$$-\text{higher order terms in } Q_k,\ P_k$$
$$=\mathbf{x}^i-\sum_k Q_k[\mathbf{x}^i,P_k]_{\text{P.B.}}+\sum_k P_k[\mathbf{x}^i,Q_k]_{\text{P.B.}}$$
$$+\text{higher order terms},$$
$$\mathbf{P}_i=\mathbf{p}_i+\sum_k P_k[\mathbf{p}_i,Q_k]_{\text{P.B.}}-\sum_k Q_k[\mathbf{p}_i,P_k]_{\text{P.B.}}$$
$$+\text{higher order terms}, \tag{3.5a}$$

and hence,

$$\delta\mathbf{x}^i=\mathbf{x}^i-\mathbf{X}^i=-\sum_k Q_k[\mathbf{x}^i,P_k]_{\text{P.B.}}+\sum_k P_k[\mathbf{x}^i,Q_k]_{\text{P.B.}}$$
$$+\text{higher order terms},$$
$$\delta\mathbf{p}_i=\mathbf{p}_i-\mathbf{P}_i=\sum_k P_k[\mathbf{p}_i,Q_k]_{\text{P.B.}}-\sum_k Q_k[\mathbf{p}_i,P_k]_{\text{P.B.}} \tag{3.6a}$$
$$+\text{higher order terms}.$$

In these definitions, only the explicit dependence of Q_k, P_k on $\mathbf{x}^i$, $\mathbf{p}_i$ need be known; it is not required to know how the $\mathbf{x}^i$, $\mathbf{p}_i$ depend on Q_k, P_k, ξ_l, $\mathbf{p}_{\xi l}$. Evidently, $\mathbf{X}^i$, $\mathbf{P}_i$ fulfill $Q_k(\mathbf{X}^i,\mathbf{P}_i)=P_k(\mathbf{X}^i,\mathbf{P}_i)=0$, to the order in which they are defined, as a glance at the definition of these entities in terms of the underlying variables Q_k, P_k, ξ_l, $\mathbf{p}_{\xi l}$ shows. The $\mathbf{X}^i$, $\mathbf{P}_i$ can therefore be represented as a point in phase space on the intersection surface $Q_k=0$, $P_k=0$, i.e., on the subspace of no oscillation. To each point $\mathbf{x}^i$, $\mathbf{p}_i$—which, in general, represents some state containing a certain degree of oscillatory (collective) excitation—thus corresponds, via (3.5a) a "projected" point $\mathbf{X}^i$, $\mathbf{P}_i$. Physically, $\mathbf{X}^i$, $\mathbf{P}_i$ is the momentary state obtained from the momentary state $\mathbf{x}^i$, $\mathbf{p}_i$, if the latter is "collectively de-excited," without any change in the residual variables.

If $f(\mathbf{x}^i,\mathbf{p}_i)$ is any dynamical variable of the system (the energy or the momentum, say) we can, in a similar fashion, effect a separation into "purely nonoscillatory" and "purely oscillatory" parts

$$f(\mathbf{x}^i,\mathbf{p}_i)=f(\mathbf{X}^i,\mathbf{P}_i)+\delta f,$$

by again imagining f as a function $f(Q_k,P_k;\ \xi_l,\mathbf{p}_{\xi l})$ and defining the "purely nonoscillatory" part as $[\![f]\!]\equiv f(0,0;\ \xi_l,\mathbf{p}_{\xi l})$; evidently

$$[\![f]\!]=f(\mathbf{X}^i,\mathbf{P}_i),$$

i.e., the same function f evaluated at the projected point $\mathbf{X}^i$, $\mathbf{P}_i$. In a fashion similar to that which led to (3.6a) we obtain

$$\delta f=\sum_k Q_k[f,P_k]_{\text{P.B.}}-\sum_k P_k[f,Q_k]_{\text{P.B.}}$$
$$+\text{higher order terms} \tag{3.6b}$$

and

$$[\![f]\!]=f(\mathbf{X}^i,\mathbf{P}_i)=f(\mathbf{x}^i,\mathbf{P}_i)$$
$$-\sum_k Q_k[f,P_k]+\sum_k P_k[f,Q_k]$$
$$+\text{higher order terms}. \tag{3.5b}$$

To the first order, it does not make any difference whether the coefficients such as $\partial Q_k/\partial\mathbf{p}_i$ and $[F,Q]$ in the above equations are evaluated at the actual point $(\mathbf{x}^i,\mathbf{p}_i)$, or at the projected point $(\mathbf{X}^i,\mathbf{P}_i)$. We shall at times find it convenient to adopt one procedure or the other. It must be emphasized that Eqs. (3.5–3.6) are not canonical transformations, because the coefficients of the derivatives in (3.5), (3.6) are functions of $\mathbf{x}^i$, $\mathbf{p}_i$ and not constants.[10] In fact, it may appropriately be called a "Clebschian" transformation because the coefficients $P(\mathbf{x}^i,\mathbf{p}_i)$ and $Q(\mathbf{x}^i,\mathbf{p}_i)$ are here playing a part analogous to that of Clebschian multipliers[11] ξ in the

[10] It is clear for example that a transformation which changes certain dynamical variables P and Q into zero must have a vanishing Jacobian determinant; thus it cannot be a canonical transformation, which later has a Jacobian equal to unity.

[11] H. Lamb, *Hydrodynamics* (Dover Publications Inc., New York), sixth ed., p. 248.

expression $\mathbf{v}=\nabla\varphi+\xi\nabla\eta$, for the velocity in hydrodynamics. The meaning of the Clebschian transformation (3.5–3.6) can be brought out more clearly by going over to the constants of the motion (the action variables) and their canonical conjugates (the angle variables). These are given by

$$Q_k=(2J_k/\omega_k)^{1/2}\cos\phi_k\,,\quad P_k=(2J_k\omega_k)^{1/2}\sin\phi_k\,.$$

Since $(\phi_k,J_{k'})=\delta_{kk'}$, $\phi_k J_k$ constitute an alternative set of canonical variables, in terms of which the above transformations can be carried out. Inserting these expressions into (3.6a), one obtains (to first order)

$$\delta\mathbf{x}^i=-\sum_k J_k\frac{\partial\phi_k}{\partial\mathbf{p}_i},\quad \delta\mathbf{p}_i=\sum_k J_k\frac{\partial\phi_k}{\partial\mathbf{x}^i}. \tag{3.6c}$$

We also obtain

$$\delta f=-\sum_k J_k[f,\phi_k], \tag{3.6d}$$

and

$$\mathbf{X}^i=\mathbf{x}^i+\sum_k J_k\frac{\partial\phi_k}{\partial\mathbf{p}_i},\quad \mathbf{P}_i=\mathbf{p}_i-\sum_k J_k\frac{\partial\phi_k}{\partial\mathbf{x}^i}, \tag{3.5c}$$

$$[\![f]\!]=f+\sum_k J_k[f_k,\phi_k]. \tag{3.5d}$$

Applying (3.6d) to $f=\phi_k$, we obtain $[\![\phi_k]\!]=\phi_k$ so that ϕ_k is unchanged. With $f=J_k$, we obtain $[\![J_k]\!]=J_k-J_k=0$. Noting that J_k, ϕ_k, are just the polar coordinates of Q_k, P_k space, we see that our displacement takes an arbitrary point $\mathbf{x}^i$, $\mathbf{p}_i$ (having excitation variables Q_k, P_k or, alternatively, J_k, ϕ_k) and moves it on a "radius" in Q_k, P_k space down to the "origin" $Q_k=0$, $P_k=0$ (or alternatively $J_k=0$). This displacement is, therefore, nothing but a de-excitation of the oscillation along a line of constant phase ϕ_k, and in such a way that all other variables (ξ_l,$\mathbf{P}_{\xi l}$, or θ', P_θ' in the example of Sec. 2) are left unchanged. That is to say, it is a projection along the line $\phi=$constant, which does nothing but to project out *all* the oscillatory part of the motion in the most direct possible way.

Returning to the simple example introduced in Sec. 2, we had [Eq. (2.12)]

$$Q=r-r_e(p_\theta)=(x^{1^2}+x^{2^2})^{1/2}-r_e(x^1p_2-x^2p_1)\,,$$

$$P=p_1=\frac{x^1p_1+x^2p_2}{r}.$$

In the calculation of the variables δx^i and X^i, we shall need the following quantities (evaluated on the equilibrium surface, $r=r_e$):

$$\left[\!\!\left[\frac{\partial P}{\partial p_i}\right]\!\!\right]=\left[\!\!\left[\frac{x^i}{r}\right]\!\!\right];\quad \left[\!\!\left[\frac{\partial P}{\partial x^i}\right]\!\!\right]=\left[\!\!\left[\frac{p_i}{r}-\frac{p_r}{r}x^i\right]\!\!\right]$$

$$=\left[\!\!\left[\frac{p_i}{r}\right]\!\!\right]-\left[\!\!\left[\frac{Px^i}{r}\right]\!\!\right]=\left[\!\!\left[\frac{p_i}{r}\right]\!\!\right], \tag{3.7}$$

where we have used $P=0$ on the equilibrium surface, and

$$\left[\!\!\left[\frac{\partial Q}{\partial x^i}\right]\!\!\right]=\left[\!\!\left[\frac{x^i}{r}\right]\!\!\right]-\left[\!\!\left[\frac{\partial r_e}{\partial p_\theta}\frac{\partial p_\theta}{\partial x^i}\right]\!\!\right]$$

$$=\left[\!\!\left[\frac{x^i}{r}\right]\!\!\right]-\left[\!\!\left[\frac{\partial r_e}{\partial p_\theta}\right]\!\!\right][\![\sum_j\epsilon_j{}^ip_j]\!], \tag{3.8}$$

$$\left[\!\!\left[\frac{\partial Q}{\partial p_i}\right]\!\!\right]=-\left[\!\!\left[\frac{\partial r_e}{\partial p_\theta}\right]\!\!\right][\![\sum_j\epsilon_j{}^ix^j]\!],$$

where $\epsilon_j{}^i$ is a two by two matrix which is zero for $i=j$, $+1$ for $j=1$, $i=2$ and -1 for $j=2$, $i=1$. We then calculate δx^i, δp_i to first order in Q and P. According to (3.5a) we have

$$\delta x^i=\left[\!\!\left[\frac{\partial P}{\partial p_i}\right]\!\!\right]Q-\left[\!\!\left[\frac{\partial Q}{\partial p_i}\right]\!\!\right]P$$

$$=\left[\!\!\left[\frac{x^i}{r}\right]\!\!\right]Q+\left[\!\!\left[\frac{\partial r_e}{\partial p_\theta}\right]\!\!\right][\![\sum_j\epsilon_j{}^ix^j]\!], \tag{3.9a}$$

$$\delta p_i=\left[\!\!\left[\frac{\partial Q}{\partial x^i}\right]\!\!\right]P-\left[\!\!\left[\frac{\partial P}{\partial x^i}\right]\!\!\right]Q$$

$$=\left[\!\!\left[\frac{x^i}{r}\right]\!\!\right]P-\left[\!\!\left[\frac{\partial r_e}{\partial p_\theta}\right]\!\!\right][\![\sum\epsilon_j{}^ip_j]\!]P-\left[\!\!\left[\frac{p_i}{r}\right]\!\!\right]Q. \tag{3.9b}$$

To see what these equations mean, let us first consider case (A), for which, as we recall, p_θ is so small that centrifugal and Coriolis forces could be neglected, so that, in this approximation, $r_e\cong r_0=$constant. In this case, $Q=r-r_0$ is a function $Q(x)$,only of x^i and not of p_i. Thus, $\partial r_e/\partial p_\theta=0$ and the above equations simplify to

$$\delta x^i=\left[\!\!\left[\frac{x^i}{r}\right]\!\!\right]Q,\quad \delta p_i=\left[\!\!\left[\frac{x_i}{r}\right]\!\!\right]P. \tag{3.10}$$

These equations imply that δx^i and δp_i are both directed along a radius, or in other words, that the oscillatory part of the motion is normal to the equilibrium variety $r=r_0$ (a conclusion which is trivial in this approximation).

We now solve for the variables X^i, P_i. These are

$$X^i=x^i-\delta x^i=x^i-\left[\!\!\left[\frac{x_i}{r}\right]\!\!\right]Q,$$

$$P_i=p_i-\delta p_i=p_i-\left[\!\!\left[\frac{x^i}{r}\right]\!\!\right]P. \tag{3.11}$$

It is easiest to see what this means by going to polar coordinates. Thus, writing $Q=r-r_0$, $x^1=r\cos\theta$, $x^2=r\sin\theta$, and replacing $[\![x^i/r]\!]$ by x^i/r (both being

equivalent in the linear approximation) we obtain

$$X^1=\cos\theta[r-(r-r_0)]=r_0\cos\theta,$$
$$X^2=\sin[r-(r-r_0)]=r_0\sin\theta. \tag{3.12}$$

Thus, X^i is a point on the equilibrium variety, $r=r_0$; and it has the same value of θ as x^i.

Similarly we see that the vector P_i is nothing but the total momentum, less the radial part of the momentum, or in other words, just the part of the momentum $mr\dot{\theta}=P_\theta/mr$ in the angular direction [and since P_θ is so small in case (A), we can replace r by r_0 to approximate for this part of the momentum by P_θ/mr_0]. Thus, P_i is on the surface $P_r=(x^1p_1+x^2p_2)/r=0$ in phase space.

The fact that X^i, P_i is on the equilibrium variety [i.e., $Q(X^i)=0$ and $P(X^i,P_i)=0$] is evidently a consequence of the way in which it was defined. This relation can, however, be verified for the general case (to the first order) by substituting $f=Q$ and $f=P$, respectively into Eq. (3.5b). One obtains

$$Q(X^i,P_i)=Q(x^i,p_i)-Q(x^i,p_i)\equiv 0,$$
$$P(X^i,P_i)=P(x^i,p_i)-P(x^i,p_i)\equiv 0. \tag{3.13}$$

If we go on to case (B), where p_θ is large enough so that centrifugal and Coriolis forces cannot be neglected, then, as Eqs. (3.9a) and (3.9b) show, the motion is no longer normal to the equilibrium circle $r=r_e$ (which now depends on p_θ). Indeed, as can be shown by a simple calculation, the additional terms in these equations correspond to the terms in (2.9) and (2.12) which imply that radial oscillations and angular motion are coupled. Nevertheless, as we shall see in the next paper, one can obtain a systematic geometrical interpretation along these lines, and show that if a certain "symplectic" metric is introduced in phase space, then the projection is "symplectically" normal to the equilibrium variety. In this way we obtain a simple and instructive geometrical interpretation of our method, which we shall develop systematically in the following paper.

4. THE POISSON BRACKETS AND THE EQUATIONS OF MOTION

As in the example of the center-of-mass variables in Sec. 1, we now proceed to demonstrate the dynamical independence of the oscillatory and nonoscillatory parts defined by Eqs. (3.5) and (3.6). We first note that $\delta\mathbf{x}^i$, $\delta\mathbf{p}_i$, $\mathbf{X}^i$, $\mathbf{P}_i$ were defined as functions of the $\mathbf{x}^i$, $\mathbf{p}_i$, (and hence their equations of motion can be obtained,[12] as for any function $f(\mathbf{x}^i,\mathbf{p}_i)$, from the equation $\dot{f}=[f,H]$).

In order to obtain the separation of the equations of motion into two dynamically independent parts, we shall show that (1) the Poisson brackets of the set Q_k, P_k with the set $\mathbf{X}^i$, $\mathbf{P}_i$ is zero to the necessary order, (2) the Poisson brackets $[\mathbf{X}^i,\mathbf{X}^j]$, $[\mathbf{X}^i,\mathbf{P}_j]$, $[\mathbf{P}^i,\mathbf{P}_j]$ are functions only of the $\mathbf{X}^i$, $\mathbf{P}_i$ and not of Q_k, P_k to the necessary order, (3) the Hamiltonian splits into a sum of two terms

$$H=H_{\mathrm{I}}+H_{\mathrm{II}}, \tag{4.1}$$

such that, in a linear treatment, H_{II} is a quadratic function of the Q_k, P_k (with coefficients which may, however, depend on the $\mathbf{X}^i$ and $\mathbf{P}_i$), while H_{I} is a function of the $\mathbf{X}^i$, $\mathbf{P}_i$ only.

As a result, the equations of motion will then separate into

$$\frac{dQ_k}{dt}=[Q_k,H_{\mathrm{II}}] \tag{4.2}$$
$$\frac{dP_k}{dt}=[P_k,H_{\mathrm{II}}], \tag{4.3}$$
$$\frac{d\mathbf{X}^i}{dt}=[\mathbf{X}^i,H_{\mathrm{I}}] \tag{4.4}$$
$$\frac{d\mathbf{P}_i}{dt}=[\mathbf{P}_i,H_{\mathrm{I}}]. \tag{4.5}$$

The equations for Q_k, P_k will contain the $\mathbf{X}^i$, $\mathbf{P}_i$ at most through the coefficients of the $|Q_k|^2$ and $|P_k|^2$ in H_{II}. In all relevant cases, these coefficients are either constants (as in the case of the example treated in Sec. 2, and, as we shall see in Paper II, in the plasma case), or slowly varying functions of the time (as such to effect an adiabatic shift in the character of the oscillation). Thus, H_{II} will be the effective Hamiltonian for the oscillatory variables, and H_{I} for the nonoscillatory variables.

[12] At this point, one might be tempted to adopt the following simple point of view: $\mathbf{X}^i$, $\mathbf{P}_i$—and, in general, $[\![f]\!]=f(\mathbf{X}^i,\mathbf{P}_i)$, for any dynamical variable $f(\mathbf{x}^i,\mathbf{p}_i)$—were obtained from $\mathbf{x}^i,\mathbf{p}_i$ and f by setting the constants of motion J_k equal to zero (without changing the residual variables), and hence, all one did was to go over from certain solutions of the Hamiltonian system under consideration to other solutions (namely to the "nearest" collectively unexcited ones). The question of writing down the equations of motion for $\mathbf{X}^i$, $\mathbf{P}_i$, $[\![f]\!]$ would, therefore, seem to have reduced to a triviality, as they just fulfill the original Hamiltonian equations. Similarly, one could, by subtracting the two systems of equations, quite easily write down the equations for $\delta\mathbf{x}^i$, $\delta\mathbf{p}_i$, δf [the resulting equations would be the well known "equations of variation"—see, e.g., E. T. Whittaker, *Analytical Dynamics*, (Dover Publications, Inc., New York, 1944), third ed., p. 269].

However, one can immediately see that this point of view, if carried consistently through, would just lead back to the subsidiary-condition method of Bohm and Pines. The $\mathbf{X}^i(t)$, $\mathbf{P}_i(t)$, considered as possible motions of the system should, as any other motion, have Poisson brackets $(\mathbf{X}^i,\mathbf{P}_j)=\delta_j{}^i$ associated with them, and the fact that they fulfill $J_k=0$ would then be a matter of proper choice of the *initial* conditions; i.e., $J_k=0$ would have to be considered as *subsidiary* conditions rather than *identities*; *dynamically* they must be considered as spanning a phase space of $6N$ dimensions [as is implied by the rank of the $6N\times 6N$ matrix $(\mathbf{X}^i,\mathbf{P}_j)$], and since they constitute only the noncollective part of the actual motion $\mathbf{x}^i(t)$, $\mathbf{p}_i(t)$, we would have to extend this phase space by introducing $2s$ "redundant" variables, as explained in Sec. 1. In our procedure, we consider the equations $J_k=0$ as *identities*; as a result, $(\mathbf{X}^i,\mathbf{P}_j)\neq\delta_j{}^i$, the latter being a matrix of rank $6N$-$2s$ only. Hence, the restriction of $\mathbf{X}^i$, $\mathbf{P}_i$, to represent the noncollective part of $\mathbf{x}^i$, $\mathbf{p}_i$ is incorporated into the dynamics of $\mathbf{X}^i$, $\mathbf{P}_i$ and will continue to hold even in more general cases when the $\mathbf{X}^i$, $\mathbf{P}_i$ can no longer be interpreted as possible motions of the system (e.g., if the latter is subjected to random external perturbations).

As the equations for the $\mathbf{X}^i$, $\mathbf{P}_i$ will have to be formulated in terms of the Poisson brackets $(\mathbf{X}^i,\mathbf{P}_j)\neq\delta_j{}^i$, a statement that $\mathbf{X}^i(t)$, $\mathbf{P}_i(t)$ is a possible motion of the original system would not be sufficient, and their equations of motion must be considered in detail.

To verify (1), (2), and (3), it must be noted that although we are treating the dynamical variables and the equations of motion only to the *first* order in Q_k and P_k, we must start with the expression of H and all other variables to the *second* order in Q_k and P_k, in order to obtain the Poisson Brackets relations to the first order, which latter are necessary to obtain the equations of motion to the same order. This is only a formal requirement, however, resulting from our insistence on a canonical formalism. After all differentiations have been carried out so as to give the equations of motion, everything need be expressed to the first order only.

It is readily verified that to second order, (3.6d) would read

$$F(\mathbf{X}^i,\mathbf{P}_i)=F(\mathbf{x}^i,\mathbf{p}_i)+\sum_k\left(J_k[F,\phi_k]+\frac{J_k^2}{2}[[F,\phi_k]\phi,_k]\right). \quad (4.6a)$$

With $F=J_k(\mathbf{x}^i,\mathbf{p}_i)$ and $F=\phi_k(\mathbf{x}^i,\mathbf{p}_i)$, we again obtain

$$J_k(\mathbf{X}^i,\mathbf{P}_i)=J_k(\mathbf{x}^i,\mathbf{p}_i)-J_k(\mathbf{x}^i,\mathbf{p}_i)\equiv 0,$$
$$\phi_k(\mathbf{X}^i,\mathbf{P}_i)=\phi_k(\mathbf{x}^i,\mathbf{p}_i),$$

and with $F=\mathbf{x}^i$ or $F=\mathbf{p}_i$,

$$\begin{aligned}\mathbf{X}^i&=\mathbf{x}^i+\sum_k J_k\frac{\partial\phi_k}{\partial\mathbf{p}_i}+\sum_k\frac{J_k^2}{2}\left[\frac{\partial\phi_k}{\partial\mathbf{p}^i},\phi_k\right],\\ \mathbf{P}_i&=\mathbf{p}_i-\sum_k J_k\frac{\partial\phi_k}{\partial\mathbf{x}^i}-\sum_k\frac{J_k^2}{2}\left[\frac{\partial\phi_k}{\partial\mathbf{x}^i},\phi_k\right].\end{aligned} \quad (4.6b)$$

This represents a displacement in phase space which carries a point $\mathbf{x}^i$, $\mathbf{p}_i$, with certain values of

$$J_k=\frac{\omega_k}{2}|Q_k|^2+\frac{1}{2\omega_k}|P_k|^2$$

into a point $\mathbf{X}^i$, $\mathbf{P}_i$, with $Q_k=0$, $P_k=0$. This "collective de-excitation" is along the "radius" $\phi_k=$const, which is now no longer a straight line element but, in general, curved (as indicated by the second-order terms).

In verifying (1), (2), and (3), we again consider $\mathbf{x}^i$ and $\mathbf{p}_i$ everywhere as functions of the Q_k, P_k and of the $6N-2s$ complementary canonical variables ξ, $\mathbf{p}_\xi$ (which have vanishing Poisson brackets with the Q_k, P_k). Thus, for an arbitrary F, $F(\mathbf{x}^i,\mathbf{p}_i)=G(Q_k,P_k,\xi,\mathbf{p}_\xi)$, with a suitable G. As long as G is analytic in Q_k, P_k, the transformation (4.6) operating on this function, will reduce the Q_k and P_k to zero and leave us with $G(0,0,\xi,\mathbf{p}_\xi)$, up to terms of third or higher order in Q_k and P_k. Moreover, $G(0,0,\xi,\mathbf{p}_\xi)$ is evidently equal to $F(\mathbf{X}^i,\mathbf{P}_i)$ to that order. Thus, it is clear that (taking $F=\mathbf{x}^i$ or $\mathbf{p}_i$)

$$\begin{aligned}&[\mathbf{X}^i,Q_k]=0,\quad [\mathbf{X}^i,P_k]=0,\\ &[\mathbf{P}_i,Q_k]=0,\quad [\mathbf{P}_i,P_k]=0,\end{aligned} \quad (4.7)$$

to second order in Q_k, P_k.

Thus, in this approximation, $\mathbf{X}^i$ and $\mathbf{P}_i$ do not depend on Q_k, P_k, and therefore their Poisson bracket $(\mathbf{X}^i,\mathbf{P}_i)$ does not depend on Q_k, P_k either, up to terms of third and higher order. This means that it makes no difference whether the $(\mathbf{X}^i,\mathbf{P}_i)$ are calculated on or off the equilibrium surface $Q_k=0$, $P_k=0$ (a fact of which we shall make use presently).

After the equations of motion have been obtained there is no longer any need to express $\mathbf{X}^i$, $\mathbf{P}_i$ or their Poisson brackets to second order. *They can be evaluated to first order only.* We shall never, in fact, encounter the need to use the second-order expansion for $\mathbf{X}^i$, $\mathbf{P}_i$ (not even in order to obtain their Poisson brackets). The introduction of the higher order expansion for $\mathbf{X}^i$, $\mathbf{P}_i$ was only a formal step, needed to show the canonical independence of the $\mathbf{X}^i$, $\mathbf{P}_i$ and Q_k, P_k.

With regard to the Hamiltonian, however, it will, of course, be important to express it correctly to the second order in Q_k and P_k, in order to obtain the correct first-order equations for Q_k and P_k. To expand the Hamiltonian, we can use Eq. (4.6) writing

$$H=H(\mathbf{X}^i,\mathbf{P}_i)+\delta^1H+\delta^2H, \quad (4.7a)$$

where δ^1H and δ^2H are of first and second order, respectively. Such an expansion will be carried out explicitly in Paper II. For the present we shall only draw some conclusions concerning the general properties of this expansion, by supposing that the Hamiltonian $H(\mathbf{x}^i,\mathbf{p}_i)$ has been expressed in terms of Q, P, ξ, $\mathbf{p}_\xi$, viz. $H(\mathbf{x},\mathbf{p})=K(Q,P,\xi,\mathbf{p}_\xi)$. For the simple example given in Sec. 2, it was, in fact, possible to obtain such an expression from Eq. (2.10a)

$$H=\frac{p^2}{2m}+m\omega^2(p_\theta)\frac{Q^2}{2}+\frac{p_\theta^2}{2mr_e(p_\theta)^2}+\frac{\alpha}{2}(r_e-r_0)^2, \quad (2.10b)$$

where we have set $p_r=P$, $r-r_e(p_\theta)=Q$. As we have indicated earlier, it is not possible in practice to obtain such an expression for the general problem. Nevertheless, we may imagine the function K to be expanded to second order in Q_k and P_k

$$K(Q_k,P_k,\xi,p_\xi)=K(0,0,\xi,p_\xi)+\sum_k\left(a_k(\xi,p_\xi)\frac{Q_k^2}{2}+b_k(\xi,p_\xi)\frac{P_k^2}{2}\right). \quad (4.7b)$$

(We know that the terms linear in Q_k and P_k must vanish because by hypothesis, we are considering functions Q_k and P_k that oscillate harmonically about $Q_k=0$, $P_k=0$.)

When (4.7b) is expressed in terms of $\mathbf{x}^i$, $\mathbf{p}_i$ and $\mathbf{X}^i$, $\mathbf{P}_i$, we are able to translate it into terms which do not

require us to know ξ, p_ξ. We obtain

$$H(\mathbf{x}^i,\mathbf{p}_i)=H(\mathbf{X}^i,\mathbf{P}_i)+A(\mathbf{X}^i,\mathbf{P}_i)\frac{Q_k^2}{2}+B(\mathbf{X}^i,\mathbf{P}_i)\frac{P_k^2}{2}$$
$$(\text{where } A(\mathbf{X}^i,\mathbf{P}_i)\equiv[A(\mathbf{x}^i,\mathbf{p}_i)]_{Q_k,P_k=0}$$
$$=[a(Q_k,P_k,\xi,p_\xi)]_{Q_k,P_k=0}). \quad (4.7c)$$

The coefficients $A(\mathbf{X}^i,\mathbf{P}_i)$ and $B(\mathbf{X}^i,\mathbf{P}_i)$ will, in all relevant cases, either be constants or slowly varying functions of the time, in which latter case the oscillations can be treated by the adiabatic approximation.

To obtain the equations of motion from an expression such as (4.7) for the Hamiltonian, we would need an explicit expression for the Poisson brackets of the $\mathbf{X}^i$ and $\mathbf{P}_i$. Let us recall that these can be evaluated on the equilibrium surface, $P_k=0$, $Q_k=0$ (which is indicated by brackets "$[\![\ ]\!]$" around the expressions). As a typical case, consider

$$[[\![\mathbf{X}^a,\mathbf{P}_b]\!]]=\sum_i\left[\!\!\left[\left(\frac{\partial\mathbf{X}^a}{\partial\mathbf{x}^i}\frac{\partial\mathbf{P}_b}{\partial\mathbf{p}_i}-\frac{\partial\mathbf{X}^a}{\partial\mathbf{p}_i}\frac{\partial\mathbf{P}_b}{\partial\mathbf{x}^i}\right)\right]\!\!\right]. \quad (4.8)$$

To evaluate this Poisson bracket we consider the relevant quantities

$$\left[\!\!\left[\frac{\partial\mathbf{X}^c}{\partial\mathbf{x}^i}\right]\!\!\right],\quad \left[\!\!\left[\frac{\partial\mathbf{X}^c}{\partial\mathbf{p}_i}\right]\!\!\right],\quad \left[\!\!\left[\frac{\partial\mathbf{P}_c}{\partial\mathbf{x}^i}\right]\!\!\right],\quad \left[\!\!\left[\frac{\partial\mathbf{P}_c}{\partial\mathbf{p}_i}\right]\!\!\right]. \quad (4.8a)$$

In the differentiations involved here, second-order terms in Q_k and P_k will, evidently, make no contribution (even after differentiation, such contributions will vanish, as everything is evaluated on the equilibrium surface). Therefore, the first-order expansion (3.5a) will (as we have already indicated before) be sufficient to calculate the Poisson brackets (4.8). When (3.5a) is substituted in (4.8a), one obtains

$$\left[\!\!\left[\frac{\partial\mathbf{X}^c}{\partial\mathbf{x}^i}\right]\!\!\right]=\delta_i{}^c+\sum_k\left[\!\!\left[\frac{\partial Q_k}{\partial\mathbf{p}_c}\right]\!\!\right]\left[\!\!\left[\frac{\partial P_k}{\partial\mathbf{x}^i}\right]\!\!\right]-\sum_k\left[\!\!\left[\frac{\partial P_k}{\partial\mathbf{p}_c}\right]\!\!\right]\left[\!\!\left[\frac{\partial Q_k}{\partial\mathbf{x}^i}\right]\!\!\right]$$
$$\left[\!\!\left[\frac{\partial\mathbf{X}^c}{\partial\mathbf{p}^i}\right]\!\!\right]=\sum_k\left[\!\!\left[\frac{\partial Q_k}{\partial\mathbf{p}_c}\right]\!\!\right]\left[\!\!\left[\frac{\partial P_k}{\partial\mathbf{p}_i}\right]\!\!\right]-\sum_k\left[\!\!\left[\frac{\partial P_k}{\partial\mathbf{p}^i}\right]\!\!\right]\left[\!\!\left[\frac{\partial Q_k}{\partial\mathbf{p}_i}\right]\!\!\right],\ \text{etc.} \quad (4.9)$$

Although the rest of the calculations are straightforward, they would prove to be very unwieldy if the present notation is used (not only because the expressions for the Poisson brackets are lengthy, but also because the second-order expansions for H are quite complicated). In Paper II, however, we shall develop a more condensed notation, with the aid of which the results can be obtained more easily. The advantage of being able to write down these equations of motion with the aid of the noncanonical Poisson brackets $[\mathbf{X}^i,\mathbf{P}_j]$ will, it is hoped, become evident in subsequent papers. In the present paper, however, we want to stress only one property of the $\mathbf{X}^i(t)$, $\mathbf{P}_i(t)$ motion, a property which is indifferent to whether its equations of motion are derived in the way indicated above or through the argument of footnote 13; namely, that $d\mathbf{X}^i/dt=[\![\partial H/\mathbf{p}_i]\!]$, $\partial\mathbf{P}_i/dt=-[\![\partial H/\partial\mathbf{x}^i]\!]$. This will become evident through the condensed notation of Paper II, and can also be understood directly as indicated in footnote 13, since it just means that $\mathbf{X}^i(t)$, $\mathbf{P}_i(t)$ is a possible solution of the original equations of motion. In other words, to each actual solution, $\mathbf{x}^i(t)$, $\mathbf{p}_i(t)$ of the equations of motion, the transformation (3.6) and (4.6) associates a *special* comparison solution $\mathbf{X}^i(t)$, $\mathbf{P}_i(t)$ in the equilibrium variety. This comparison solution is special in the sense that as a result of the way in which it is defined, the differences $\delta\mathbf{x}^i=\mathbf{x}^i-\mathbf{X}^i$, $\delta\mathbf{p}_i=\mathbf{p}_i-\mathbf{P}_i$ will evidently not increase without limit in a secular way, because they are proportional to Q_k and P_k, which oscillate harmonically and remain small.

To illustrate the meaning of this property of the $\mathbf{X}^i$, $\mathbf{P}_i$, let us return to the example given in Sec. 2. We shall consider case (A), in which p_θ is small enough so that the equilibrium variety can be approximated by the circle $r=r_0$. In the description of small oscillations about the equilibrium circle, we introduced a set of canonical variables (2.12). With their aid, we associated to each point $r(t)=r_0+\delta r(t)$, $p_r(t)$, $\theta'=\theta-p_r(\partial r_e/\partial p_\theta)$, $p_\theta'=p_\theta$, a corresponding equilibrium point of the variety $Q=r-r_0=0$, $P=p_r=0$ with coordinates $\theta_0'=\theta'$ and $p_\theta'=p_\theta$. In other words, the values of the nonoscillatory variables θ' and p_θ', in the comparison orbit have, in this way, been chosen to be the same as in the actual orbit. If such a choice had not been made, then there would have been an increase of $\theta'-\theta_0'=(p_\theta'-p_{\theta 0}')t/mr_0^2$ without limit, with the passage of time.

Of course, in the above simple example it was possible to choose conditions on the comparison orbit in the equilibrium variety so as to avoid a secular increase of $\delta\theta$ with time because there was available an explicit expression for the nonoscillatory variables θ' and p_θ', by means of the canonical transformation (2.12). More generally, however, such an explicit expression cannot, as we have already pointed out, be found. Thus, if an arbitrary motion in the equilibrium variety is taken as a comparison motion, then the difference $\delta\mathbf{x}^i$, $\delta\mathbf{p}_i$ will, in general, undergo a secular increase with time. The special way of defining the comparison motion by means of the transformation (3.6) and (4.6) leading to the "projected" point $\mathbf{X}^i$, $\mathbf{P}_j$ is therefore, in effect, a means of choosing initial conditions of the comparison orbit, such that secular increases of $\delta\mathbf{x}^i$, $\delta\mathbf{p}_i$ are certain not to occur. As a result, the conditions needed for application of perturbation theory will always be satisfied.