IFS #194

Plasma Kinetic Theory in Action-Angle Variables

S.M. Mahajan

and

C.Y. Chen

Institute for Fusion Studies

The University of Texas at Austin

Austin, Texas 78712

Abstract

An appropriate canonical perturbation theory to correctly deal with general electromagnetic field perturbation is developed, and is used to set up plasma kinetic theory in action-angle variables. A variety of test problems are solved to show the unifying power of the method. Basic linear, quasilinear, and nonlinear equations are derived which can serve as the starting point for a whole range of plasma problems.

I. Introduction

Action and angle variables have often provided a very convenient coordinate system in which the problems of classical physics can be formulated and solved. In celestial mechanics, for example, this approach has been extensively used to determine the effects of small perturbations on the motion of the planets. It is a bit strange that this obviously powerful tool has not found much currency amongst researchers in plasma physics. Apart from a few initial studies,[1-2] the use of action invariants to solve the Vlasov equation in complicated geometry was attempted only recently by Kaufman,[3] and then by other workers[4] who primarily used and extended Kaufman's formalism. Unfortunately, the basic formalism is fundamentally flawed, and when applied to deal with time-dependent magnetic field perturbations, it leads to incorrect results even for the linear Vlasov theory! (see Sec. IV for details.) The formalism, however, yields correct results for strictly electrostatic perturbations which defines the domain of validity of the analysis of Refs. 3-4.

It is surprising that a method based on the well-known time-dependent canonical perturbation theory (Kaufman's) would lead to such inconsistencies. In this paper, we show why a mechanical application of the standard perturbation theory is not possible in the presence of the electromagnetic perturbed fields (EMPF). We also develop a modified perturbation theory which can correctly handle the EMPF. Arguments leading to, and the formalism, are presented in Sec. II.

The remaining sections of this paper are devoted to the formulation of plasma kinetic theory in terms of action-angle variables using our modified perturbation theory. The Vlasov equation in the new phase space is derived in Sec. III, while some examples in the linear theory are worked out in Sec. IV, where we show that our results are identical with standard results, and are different from the results of the earlier theory.[3] After having shown that our theory has corrected the inadequacies of the earlier theory, we go ahead and develop quasilinear and nonlinear aspects of plasma kinetic theory in action-angle variables in Secs. V and VI. Expressions for the quasilinear diffusion coefficient, and the convection coefficient are derived in Sec. V, while Sec. VI is essentially devoted to a formal description of the renormalized nonlinear theory. A brief discussion, which includes some comments on the importance of this work, is given in Sec. VII.

II. Modified Canonical Theory for EMPF

An essential step in setting up the time-dependent or the canonical perturbation theory[5] is to obtain motion invariants of the corresponding unperturbed system described by the Hamiltonian H_0. Provided that the found invariants α's and their conjugates β's constitute a complete set to form a new phase space, a generating function G must exist which can perform a canonical transformation (p and q are the original phase space coordinates)

$$(\mathbf{p}, \mathbf{q}) \overset{G(\mathbf{q},\vec{\alpha})}{\longrightarrow} (\vec{\alpha}, \vec{\beta}) \tag{1}$$

with the result that the motion of the total system is simply along a constant line in the new phase space $(\vec{\alpha}, \vec{\beta})$, i.e., the Hamiltonian H_0 is independent of the β's.

For the perturbed system, the transformation Eq. (1) is still employed, and then the Hamiltonian H can be expressed as a sum of two parts; the formally unchanged, β independent H_0, and h which is a function of $(\mathbf{p}, \mathbf{q})$ as well as $(\vec{\alpha}, \vec{\beta})$. Clearly, the α's change only due to h,

$$\dot{\alpha}_i = -\frac{\partial H}{\partial \beta_i} = -\partial h(\alpha_i, \beta_i, t)/\partial \beta_i. \tag{2}$$

Equation (2) provides a basis for a perturbation theory. The first step in the approximation scheme is to substitute the unperturbed constant values of the $\alpha_i = \alpha_{i0}$ in the right-hand side of Eq. (2) (after taking the β_i derivatives). The equation of motion then could be integrated to obtain the perturbed time dependent α_i's. Notice that the invariant actions $\mathbf{J}$ or J_i are to be identified with α_i, and the conjugate angles $\vec{\theta}$ or θ_i with β_i in the action-angle formalism.

Although this procedure is standard and widely used, it runs into considerable difficulty when applied to the treatment of magnetic field perturbations (EMPF). We must point out that there is nothing theoretically incorrect with the method. In fact, the phase space $(\vec{\alpha}, \vec{\beta})$ obtained from transformation Eq. (1) is canonical, and does reduce to its counterpart in the corresponding unperturbed system, which seemingly fulfills the requirements for a perturbation theory. However, an important technical problem arises when magnetic field perturbation is present; the generalized momentum (e, m and $\mathbf{v}$ are the charge, mass and velocity of the particle respectively, and $\mathbf{A}$ is the vector potential)

$$\mathbf{p} = m\mathbf{v} + \frac{e}{c}\mathbf{A} \tag{3}$$

contains the perturbed vector potential explicitly. Consequently, the coordinate system $(\vec{\alpha}, \vec{\beta})$ obtained from $(\mathbf{p}, \mathbf{q})$ by a canonical transformation also contains the EMPF (in the rest of this paper, the abbreviation EMPF will be used for the perturbed magnetic field, perturbed vector potential etc.). The use of such a coordinate system to describe the motion of any dynamical system must necessarily suffer from severe intrinsic disadvantages.

1) Generally, the object of a calculation will be to determine the EMPF. Thus the meaning of the coordinates will remain quite obscure till the problem is solved.
2) The coordinate frame is, by definition, not the same as its counterpart in the unperturbed system. This confuses the situation because a cornerstone of the perturbation theory is that we do have the knowledge of the unperturbed system and its phase-space variables α_{i0} and β_{i0}. The confusion between (α_i, β_i) and $(\alpha_{i0}$ and $\beta_{i0})$ can lead to serious errors as in the linearized solution of Ref. 3. We discuss this point in detail in the next section.
3) Because the coordinate system has fast variation caused by the varying EMPF, it is no longer practical to solve systems like Vlasov equation by decomposing the perturbation into Fourier harmonics; the superposition is no longer valid, and the resultant Fourier transform of the equation will have convolutions even in the linear analysis.

It is thus strongly indicated that we must look for a different coordinate system in order to exploit the powerful and elegant machinery of canonical perturbation theory. A logical choice will be the coordinate system which contains only the equilibrium electromagnetic fields, given by the vector potential $\mathbf{A}_0$, and the scalar potential Φ_0, i.e., the generalized canonical momentum

$$\mathbf{p}_0 = m\mathbf{v} + \frac{e}{c}\mathbf{A}_0 \tag{4}$$

should replace $\mathbf{p}$. It is from this new set $\mathbf{p}_0$ and $\mathbf{q}_0 \equiv \mathbf{q}$ that we obtain the appropriate

$\vec{\alpha}_0$ and $\vec{\beta}_0$ by a generating function G,

$$\mathbf{p}_0, \mathbf{q}_0 \overset{G(\mathbf{q}_0,\, \vec{\alpha}_0)}{\longrightarrow} (\vec{\alpha}_0,\ \vec{\beta}_0). \tag{5}$$

Notice that the transformation given in Eq. (5) is exactly equivalent to solving the equations of motion of a particle in the equilibrium fields. We, of course, assume that this problem is solved for all cases under consideration. For example, Kaufman[3] has implicitly solved the equilibrium motion of a guiding center plasma in a torus, while Hazeltine, Mahajan and Hitchcock[4] have obtained explicit expressions for the special case of a high aspect-ratio torus.

We wish to point out here that in the absence of the EMPF ($\mathbf{A} = 0$), $\mathbf{p}_0$ is identically equal to $\mathbf{p}$, and thus the two coordinate systems will be equivalent, and the standard perturbation theory will yield correct results. Clearly, the results of all the previous papers following this approach are correct for purely electrostatic perturbations ($\mathbf{A} = 0,\ \Phi \neq 0$).

These new variables, however, do not form a canonical conjugate pair with respect to the total Hamiltonian H, ($\mathbf{A} \neq 0$) i.e., $\dot{\alpha}_{0i} \neq \partial H/\partial\beta_{0i}$. This fact is of crucial importance, because it is the identification of $\dot{\alpha}_{0i}$ with $\partial H/\partial\beta_{0i}$ that constitutes the principal mistake of Refs. 3-4. For our case, we shall have to find appropriate expressions for the rate of change of actions and angles in the presence of perturbations. We must remark that although $\vec{\alpha}_0$ and $\vec{\beta}_0$ are not canonical (for H), they are obtained as a canonical transformation from $\mathbf{p}_0$, $\mathbf{q}_0$ which label the unperturbed state described by the equilibrium Hamiltonian H_0. Thus ($\vec{\alpha}_0,\ \vec{\beta}_0$) will retain some of the crucial properties of canonical variables.

The nonrelativistic Hamiltonian for a charged particle in an electromagnetic field is given by

$$H = \frac{1}{2m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}_t\right)^2 + e\Phi_t \tag{6}$$

where the scalar and the vector potentials Φ_t and $\mathbf{A}_t$ include both the unperturbed ($\Phi_0,\ \mathbf{A}_0$) and the perturbed ($\Phi,\ \mathbf{A}$) components, i.e.,

$$\Phi_t = \Phi_0 + \Phi,$$

$$\mathbf{A}_t = \mathbf{A}_0 + \mathbf{A}.$$

By making use of the relation

$$\mathbf{p}_0 = \mathbf{p} - \frac{e}{c}\mathbf{A} \tag{7}$$

which follows from Eqs. (3) and (4), the Hamiltonian H can be written in the form

$$H = H_0 + h$$

where

$$H_0 = \frac{1}{2m}\left(\mathbf{p}_0 - \frac{e}{c}\mathbf{A}_0\right)^2 + e\Phi_0, \tag{8}$$
$$h = e\Phi.$$

It may seem peculiar that the troublesome perturbed vector potential $\mathbf{A}$ has completely disappeared from the scene; it appears explicitly in neither H_0 nor h. The effects of $\mathbf{A}$, however, will reappear when we make the transformation from $(\mathbf{p}, \mathbf{q})$ to $\mathbf{p}_0$, $\mathbf{q}_0$.

As we remarked earlier, the first step in the development of the canonical perturbation theory is the solution of the unperturbed problem. We assume that the unperturbed problem has been completely solved: the invariants $\vec{\alpha}_0$, the conjugate angles $\vec{\beta}_0$, and the generating function $G(\mathbf{q}_0, \vec{\alpha}_0)$ which mediates the canonical transformation of Eq. (5), have all been obtained. Since $\vec{\alpha}_0$ are the invariants of the unperturbed system, the Hamiltonian H_0 is a function of $\vec{\alpha}_0$ alone, i.e.,

$$H_0 = H_0(\vec{\alpha}_0) \tag{9}$$

a property which will be extensively used later, and which is a major source of simplification.

The equation of evolution of any dynamical quantity Q (for example $\vec{\alpha}_0$ and $\vec{\beta}_0$ labelling the trajectory of the charged particle) is governed by the total Hamiltonian H, and is

$$\dot{Q} = \left.\frac{\partial Q}{\partial t}\right|_{(\mathbf{p},\mathbf{q})} + [Q, H]_{(\mathbf{p},\mathbf{q})} \tag{10a}$$

where

$$[f, g] = \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} \tag{10b}$$

is the Poisson bracket. In Eq. (10b), we have used Einstein's summation convention over repeated indices; this convention will be used throughout the paper.

Our aim is now to transform Eq. (10a) to the coordinate system defined by the equilibrium invariants, and their conjugate angles $(\vec{\alpha}_0, \vec{\beta}_0)$. To do this, we must first transform the equation to the coordinate system $\mathbf{p}_0$, $\mathbf{q}_0$ because it is only these (and not $(\mathbf{p}, \mathbf{q})$) which are related to $(\vec{\alpha}_0, \vec{\beta}_0)$ by a known canonical transformation [Eq. (5)]. This step is the distinguishing feature of our analysis. The primary source of error in the earlier analysis (Refs. 3-4) is that they failed to distinguish between $(\mathbf{p}, \mathbf{q})$ and $\mathbf{p}_0$, $\mathbf{q}_0$.

The details of the transformations $(\mathbf{p}, \mathbf{q}) \rightarrow \mathbf{p}_0, \mathbf{q}_0 \rightarrow (\vec{\alpha}_0, \vec{\beta}_0)$ are worked out in Appendix A. Here we simply give the trajectories of the charged particle in terms of the action-angle variables, $\vec{\alpha}_0 \equiv \mathbf{J}$, $\vec{\beta}_0 \equiv \vec{\theta}$. The evolution equations are

$$\begin{aligned}\dot{\mathbf{J}} &= -e\frac{\partial \Phi}{\partial \vec{\theta}} - \frac{e}{c}\frac{\partial \mathbf{J}}{\partial p_{0j}}\frac{\partial A_j}{\partial t} - \frac{e}{c}\left[\frac{\partial \mathbf{J}}{\partial p_{0j}}\frac{\partial H_0}{\partial \mathbf{J}} \cdot \frac{\partial}{\partial \vec{\theta}} - \left(\frac{\partial \mathbf{J}}{\partial p_{0j}} \cdot \frac{\partial H_0}{\partial \mathbf{J}}\right)\frac{\partial}{\partial \vec{\theta}}\right]A_j \\ &= -e\frac{\partial \Phi}{\partial \vec{\theta}} - \frac{e}{c}\frac{\partial \mathbf{J}}{\partial p_{0j}}\frac{\partial A_j}{\partial t} - \frac{e}{c}\vec{\Omega} \times \left(\frac{\partial \mathbf{J}}{\partial p_{0j}} \times \frac{\partial A_j}{\partial \vec{\theta}}\right),\end{aligned} \tag{11}$$

$$\begin{aligned}(\delta\dot{\vec{\theta}}) \equiv \dot{\vec{\theta}} - \vec{\Omega} = +e\frac{\partial \Phi}{\partial \mathbf{J}} - \frac{e}{c}\frac{\partial \vec{\theta}}{\partial p_{0j}}\frac{\partial A_j}{\partial t} - \frac{e}{c}\Bigg[\frac{\partial \vec{\theta}}{\partial p_{0j}}\vec{\Omega} \cdot \frac{\partial}{\partial \vec{\theta}} \\ + \left(\frac{\partial \mathbf{J}}{\partial p_{0j}} \cdot \vec{\Omega}\right)\frac{\partial}{\partial \mathbf{J}}\Bigg]A_j.\end{aligned} \tag{12}$$

where $\Phi = \Phi(\mathbf{J}, \vec{\theta}, t)$, and $\mathbf{A} = \mathbf{A}(\mathbf{J}, \vec{\theta}, t)$. The quantities $(\partial \mathbf{J}/\partial p_{0j})$, $(\partial \vec{\theta}/\partial p_{0j})$ and $\partial H_0/\partial \mathbf{J} \equiv \vec{\Omega}(\mathbf{J})$ are known functions of $\mathbf{J}$ and $\vec{\theta}$ from the solution of the unperturbed problem. Since $\mathbf{J}$'s are the invariants in the equilibrium fields, their evolution is due to the perturbed fields only; $\dot{\theta}$, however, has the additional equilibrium frequencies $\vec{\Omega}(\mathbf{J})$. We remind the reader that H_0 is independent of $\vec{\theta}$.

It can be easily recognized that the first two terms on the right-hand side of Eqs. (11) and (12) reflect the effect of the perturbed electric field (both static and inductive), while the last term is due to the Lorentz force. The effects of the unperturbed fields is contained in the equilibrium quantities $(\partial \mathbf{J}/\partial p_{0i})$, etc..

A very important feature of Eqs. (11) and (12) is the linear dependence of $\dot{\mathbf{J}}$ and $\dot{\vec{\theta}}$ on the perturbed fields Φ and $\mathbf{A}$. Notice that this would not be the case if the coefficients of transformation $(\partial \mathbf{J}/\partial p_{0j})$ etc., depended upon the perturbed fields. This feature assures the validity of the superposition rule.

The rest of this paper is devoted to the application of this method to plasma kinetic theory.

III. Vlasov Equation in Action-Angle Variables

In this section, we set up the Vlasov equation in action-angle variables. The treatment of this section is quite general with the added feature that actions $\mathbf{J}$ are the invariants of the equilibrium system. In fact, the invariance (or adiabatic invariance) is the only reason to prefer $(\mathbf{J}, \vec{\theta})$ to any other systems of coordinates.

To put things in perspective, we deal with a guiding-center plasma in an axisymmetric toroidal configuration. Kaufman has very elegantly solved the equilibrium problem for this system. We request the reader to consult Ref. 3 for details. The essential results are that the guiding-center motion can be described in terms of three adiabatic invariants: $M = (mc^2/e)\mu$ proportional to the magnetic moment μ; P_φ, the canonical angular momentum (this action is an absolute invariant for an axisymmetric system); and J_p which is proportional to the flux enclosed by a drift surface. These actions have associated conjugate angles θ_g, φ and θ respectively.

Throughout this paper, we shall use the condensed notation

$$\mathbf{J} = (M, P_\varphi, J_p)$$

and

$$\vec{\theta} = (\theta_g,\ \varphi,\ \theta) \tag{13}$$

to denote the action-angle variables. It is also important to define the triad of frequencies

$$\vec{\Omega} = \frac{\partial H_0}{\partial \mathbf{J}} \equiv (\omega_g, \omega_\varphi, \omega_\theta) \tag{14}$$

where $H_0 = mv^2/2 + e\Phi_0$ is the equilibrium Hamiltonian. Notice that for a toroidal plasma, $\theta_g(\omega_g)$ is like a gyro angle (average gyro-frequency) $\varphi(\omega_\varphi)$ is a toroidal-like angle (average toroidal frequency), and $\theta(\omega_\theta)$ is a poloidal-like angle (average poloidal frequency). The generating function G which mediates the transformation from $\mathbf{p}_0$, $\mathbf{q}_0$ to $(\mathbf{J},\ \vec{\theta})$ is given and discussed in Kaufman's paper. Having briefly delineated the solution of a typical equilibrium problem, we proceed to deal with the Vlasov Equation.

In the new variable system, the Vlasov equation becomes

$$\frac{dF}{dt}(\mathbf{J},\vec{\theta},t) = \frac{\partial F}{\partial t} + \dot{\vec{\theta}} \cdot \frac{\partial F}{\partial \vec{\theta}} + \dot{\mathbf{J}} \cdot \frac{\partial F}{\partial \mathbf{J}} = 0, \tag{15a}$$

or equivalently

$$\frac{\partial F}{\partial t} + \frac{\partial}{\partial \mathbf{J}} \cdot (\dot{\mathbf{J}} f) + \frac{\partial}{\partial \vec{\theta}} \cdot (\dot{\vec{\theta}}) = 0 \tag{15b}$$

which follows from the divergence identity

$$\frac{\partial}{\partial \mathbf{J}} \cdot (\dot{\mathbf{J}}) + \frac{\partial}{\partial \vec{\theta}} \cdot (\dot{\vec{\theta}}) = 0$$

derived in Appendix A. Note that for purely canonical variables, the relation

$$\frac{\partial}{\partial J}\dot{\mathbf{J}}_i + \frac{\partial}{\partial \theta_i}\dot{\theta}_i = 0. \qquad \text{[No summation convention]}$$

holds for each component i. For these non-canonical variables J and θ, it is only the total divergences which cancel. The Hamiltonian takes the form

$$H = H_0(\mathbf{J}) + e\Phi(\mathbf{J},\vec{\theta},t). \tag{16}$$

The distribution function is to be decomposed into its equilibrium part f_0, and the fluctuating part f

$$F = f_0 + f(\mathbf{J},\vec{\theta},t). \tag{17}$$

Clearly, the equilibrium part f_0 satisfies the equation

$$\vec{\Omega} \cdot \frac{\partial f_0}{\partial \vec{\theta}} = 0 \tag{18}$$

implying the simple solution

$$f_0 = f_0(\mathbf{J}) \tag{19}$$

which is clearly a result of our proper choice of the coordinate system. The equation governing the fluctuating or the perturbed distribution function f is obtained from Eqs. (11)-(19),

$$\frac{\partial f}{\partial t} + \vec{\Omega} \cdot \frac{\partial f}{\partial \mathbf{J}} + \dot{\mathbf{J}} \cdot \frac{\partial f_0}{\partial \mathbf{J}} = -\left[(\delta\dot{\vec{\theta}}) \cdot \frac{\partial f}{\partial \vec{\theta}} + \dot{\mathbf{J}} \cdot \frac{\partial f}{\partial \mathbf{J}}\right] \tag{20}$$

where we have not explicitly substituted expressions for $\dot{\mathbf{J}}$ and $(\delta\dot{\vec{\theta}})$. We shall continue developing the formal theory implicitly, and use detailed forms only where we deal with particular cases. In Eq. (20), all the terms on the left-hand side (right-hand side) are linear (nonlinear) in the perturbed quantities f, Φ and $\mathbf{A}$.

Now we exploit another important characteristic of our coordinate system, the cyclic nature of the $\vec{\theta}$'s, to expand all the perturbed quantities as Fourier series in $\vec{\theta}$. A typical perturbed quantity g is decomposed as

$$g(\mathbf{J},\vec{\theta},t) = \sum_{\vec{\ell},\omega} g_{\vec{\ell},\omega}(\mathbf{J}) \exp(-i\omega t + i\vec{\ell}\cdot\vec{\theta}) \tag{21a}$$

where the Fourier transform $g_{\vec{\ell},\omega}(\mathbf{J})$ is given by

$$g_{\vec{\ell},\omega}(\mathbf{J}) = \frac{1}{(2\pi)^4}\int d\vec{\theta}\, dt \;\; g(\mathbf{J},\vec{\theta},t)\exp(+i\omega t - i\vec{\ell}\cdot\vec{\theta}), \tag{21b}$$

and $\vec{\ell}$ is a triad of integers labelling the Fourier harmonics. Making use of Eqs. (20)-(22), we obtain $[g_{\vec{\ell},\omega}(\mathbf{J}) \equiv g_{\vec{\ell}}$, the index ω, and the argument $\mathbf{J}$ will be generally suppressed in the rest of the paper$]$,

$$-i(\omega - \vec{\ell}\cdot\vec{\Omega})f_{\vec{\ell}} + \dot{\mathbf{J}}_{\vec{\ell}}\cdot\frac{\partial f_0}{\partial \mathbf{J}} = -N_{\vec{\ell}} \tag{22}$$

where the nonlinear term $N_{\vec{\ell}}$ is the convolution

$$N_{\vec{\ell}} = \sum_{\vec{\ell}'}\left[i(\vec{\ell}-\vec{\ell}')\cdot(\delta\dot{\vec{\theta}})_{\vec{\ell}'} \;\; f_{\vec{\ell}-\vec{\ell}'} + \dot{\mathbf{J}}_{\vec{\ell}'}\cdot\frac{\partial}{\partial\mathbf{J}}f_{\vec{\ell}-\vec{\ell}'}\right]. \tag{23}$$

Notice that both $\dot{\mathbf{J}}_{\vec{\ell}}$ and $(\delta\dot{\vec{\theta}})_{\vec{\ell}}$ are extremely complicated terms obtained by taking the Fourier transforms of Eqs. (11) and (12), and are given by

$$\dot{\mathbf{J}}_{\vec{\ell}} = -ie\left[\vec{\ell}\Phi_{\vec{\ell}} + \sum_{\vec{\ell}'}\mathbf{K}^j(\vec{\ell}',\vec{\ell}-\vec{\ell}')A^j_{\vec{\ell}'}\right], \tag{24}$$

$$(\delta\dot{\vec{\theta}})_{\vec{\ell}} = -ie\left[-i\frac{\partial}{\partial\mathbf{J}}\Phi_{\vec{\ell}} + \sum_{\vec{\ell}'}\mathbf{L}^j(\vec{\ell}',\vec{\ell}-\vec{\ell}')A^j_{\vec{\ell}'}\right], \tag{25}$$

with the vector operators $\mathbf{K}^j$ and $\mathbf{L}^j$ defined by

$$\mathbf{K}^j = \frac{1}{c}\left[-\omega\mathbf{T}^j_{\vec{\ell}-\vec{\ell}'} + \vec{\Omega}\times\left(\mathbf{T}^j_{\vec{\ell}-\vec{\ell}'}\times\vec{\ell}'\right)\right], \tag{26a}$$

$$\mathbf{L}^j = \frac{1}{c}\left[-\omega \mathbf{S}^j_{\vec{\ell}-\vec{\ell}'} + (\vec{\ell}' \cdot \vec{\Omega}) S^j_{\vec{\ell}-\vec{\ell}'} - i(\vec{\Omega} \cdot \mathbf{T}^j_{\vec{\ell}-\vec{\ell}'}) \frac{\partial}{\partial \mathbf{J}}\right] \tag{26b}$$

where $\mathbf{T}^j_{\vec{\ell}}$ and $\mathbf{S}^j_{\vec{\ell}}$ are the Fourier transforms of the known quantities. $\partial \mathbf{J}/\partial p_{0j}$ and $\partial \vec{\theta}/\partial p_{0j}$ respectively. Equation (22) is the main result of this section, and is an expression of the exact, nonlinear, Fourier transformed Vlasov equation in the presence of fully electromagnetic perturbations in a general magnetic field geometry. The formally simple structure of this equation shows the power and elegance of the action-angle variable approach which allows the Vlasov equation in a complicated geometry to look exactly like the Vlasov equation describing a field-free plasma. This formal equivalence follows from the fact that in the invariant action-angle space, the particle trajectories are always straight lines (as in the field-free case). The result is a unified formalism to deal with a whole class of plasma problems; we do not have to begin with a different looking Vlasov equation every time we change the equilibrium geometry. All formal manipulations can be carried out on Eq. (22), and depending on what $(\mathbf{J}, \vec{\theta})$ we use, it could, for example, describe the infinite homogeneous field-free plasma, or the response of trapped particles in a tokamak. The translation $(\mathbf{J}, \vec{\theta}) \leftrightarrow (\mathbf{p}_0, \mathbf{q}_0)$ is, of course, given by the solution of the equilibrium problem.

IV. Linearized Vlasov Equation

Setting $N_{\vec{\ell}} = 0$ in Eq. (22) leads to the linearized Vlasov equation

$$-i(\omega - \vec{\ell} \cdot \vec{\Omega}) f_{\vec{\ell}} + \dot{\mathbf{J}} \cdot \frac{\partial f_0}{\partial \mathbf{J}} = 0 \tag{27}$$

which is readily solved to obtain

$$f_{\vec{\ell}} = \frac{\dot{\mathbf{J}}_{\vec{\ell}} \cdot \frac{\partial f_0}{\partial \mathbf{J}}}{i(\omega - \vec{\ell} \cdot \vec{\Omega})}. \tag{28}$$

Notice that our clever choice of the coordinate has resulted in making the linear solution $f_{\vec{\ell}}$ independent of $\delta\dot{\vec{\theta}}_{\vec{\ell}}$.

A. Free Field Limit

One of the main reasons for this paper was to correct the errors in the earlier formulation of action-angle variable theories. Let us see how our new linear results compare with known standard results. The simplest check is to obtain the field-free ($\mathbf{A}_0 = 0$, $\Phi_0 = 0$) limit of Eq. (28). This is easily accomplished by the transformations $\mathbf{J} \to \mathbf{p}$, the ordinary conserved momentum, $\vec{\theta} \to \mathbf{x}$, $\vec{\ell} \to \mathbf{k}(\partial/\partial\vec{\theta} \to \partial/\partial\mathbf{x})$, and $\vec{\Omega} \to \mathbf{v}$. The results are

$$\mathbf{T}^j_{\vec{\ell}} \to \hat{e}_i \delta_{i,j} \delta_{\mathbf{k},0} \tag{29}$$

where $\hat{e}_i$ is the unit vector along the direction i, the operator,

$$\mathbf{K}^j \to \frac{1}{c}\left[-\omega \hat{e}_i \delta_{ij} + \mathbf{v} \times (\hat{e}_i \times \mathbf{k})\delta_{ij}\right]\delta_{\mathbf{k},\mathbf{k}'}, \tag{30}$$

$$\dot{\mathbf{J}}_{\vec{\ell}} \to -ie\left\{\mathbf{k}\Phi_{\mathbf{k}} + \frac{1}{c}\left[-\omega \mathbf{A}_{\mathbf{k}} - \mathbf{v} \times (\mathbf{k} \times \mathbf{A}_{\mathbf{k}})\right]\right\}, \tag{31}$$

and finally ($\mathbf{p} = m\mathbf{v}$, $E_{\mathbf{k}}$ and $B_{\mathbf{k}}$ are respectively the perturbed electric and magnetic fields)

$$f_{\mathbf{k}} = \frac{e}{m}\frac{-i\left[\mathbf{k}\Phi_{\mathbf{k}} - \frac{\omega}{c}\mathbf{A}_{\mathbf{k}} - \frac{1}{c}\mathbf{v} \times (\mathbf{k} \times \mathbf{A}_{\mathbf{k}})\right]}{i(\omega - \mathbf{k}\cdot\mathbf{v})} \cdot \frac{\partial f_0}{\partial \mathbf{v}} \tag{32a}$$

$$\equiv \frac{e}{m}\frac{\left(\mathbf{E}_{\mathbf{k}} + \frac{1}{c}\mathbf{v} \times \mathbf{B}_{\mathbf{k}}\right)}{i(\omega - \mathbf{k}\cdot\mathbf{v})} \cdot \frac{\partial f_0}{\partial \mathbf{v}} \tag{32b}$$

which is precisely the required result.[6] By using Faraday's law $\mathbf{B}_k = (c/\omega)(\mathbf{k} \times \mathbf{E}_{\mathbf{k}})$, we can rewrite Eq. (32b) purely in terms of the electric field

$$\begin{aligned} f_{\mathbf{k}} &= \frac{e}{im\omega}\frac{\left[\omega \mathbf{E}_{\mathbf{k}} + \mathbf{v} \times (\mathbf{k} \times \mathbf{E}_{\mathbf{k}})\right]}{\omega - \mathbf{k}\cdot\mathbf{v}} \cdot \frac{\partial f_0}{\partial \mathbf{v}} \\ &\equiv \frac{e}{im\omega}\left[\mathbf{E}_{\mathbf{k}} + \frac{(\mathbf{v}\cdot\mathbf{E}_{\mathbf{k}})\mathbf{k}}{\omega - \mathbf{k}\cdot\mathbf{v}}\right] \cdot \frac{\partial f_0}{\partial \mathbf{v}}. \end{aligned} \tag{32c}$$

We have derived this form of $f_{\mathbf{k}}$ to compare it with the linear result, Eq. (26),

$$\delta f_{\vec{\ell}}(\mathbf{J}, \omega) = \frac{\delta H_{\vec{\ell}}(\mathbf{J}, \omega)\vec{\ell} \cdot \frac{\partial f_0}{\partial \mathbf{J}}}{\omega - \vec{\ell}\cdot\vec{\Omega}}$$

of Ref. 3. Using the standard prescription to go to the field-free case, we obtain

$$\delta f_{\mathbf{k}} = \frac{e}{im\omega}\frac{(\mathbf{v}\cdot\mathbf{E}_k)\left(\mathbf{k}\cdot\frac{\partial f_0}{\partial \mathbf{v}}\right)}{(\omega - \mathbf{k}\cdot\mathbf{v})} \tag{33}$$

which does not agree with Eq. (32c); Eq. (33) lacks the first term in the square brackets in Eq. (32). The mistake was made in Eq. (25) of Ref. 3 where $\delta\dot{\mathbf{J}} = -\partial H/\partial\vec{\theta} =$

$-\partial(H_0 + \delta H)/\partial\vec{\theta} = -\partial\delta H/\partial\theta$ was used. As pointed out earlier, $(\mathbf{J}, \vec{\theta})$ are not canonical variables for the total Hamiltonian H, and therefore $\delta\dot{\mathbf{J}}$ is not equal to $-\partial\delta H/\partial\boldsymbol{\theta}$. By using the correctly derived expressions for $\dot{\mathbf{J}}$ and $\dot{\theta}$ [Eqs. (11) and (12)], we do indeed reproduce the standard results. After this demonstration, we shall no longer belabor the point that the earlier treatments using action-angle variables were incorrect.

B. Low-Frequency Trapped Particle Response in Tokamaks

We now show how Eq. (28) can be readily used to obtain low-frequency gyro-averaged, bounce-averaged, response for deeply trapped particles in a tokamak. For simplicity, we derive only the electrostatic limit for which [see Eq. (24)]

$$\dot{\mathbf{J}}_{\vec{\ell}} = -ie\vec{\ell}\Phi_{\vec{\ell}} \tag{34}$$

leading to

$$f_{\vec{\ell}} = \frac{-e\Phi_{\vec{\ell}}\,\vec{\ell}\cdot\frac{\partial f_0}{\partial \mathbf{J}}}{\omega - \vec{\ell}\cdot\vec{\Omega}} \tag{35}$$

where $\vec{\ell} = (\ell_g, \ell_\varphi, \ell_\theta)$ with ℓ_g, ℓ_φ, and ℓ_θ the gyro, toroidal and bounce harmonic numbers respectively. The gyro and bounce averaged response is obtained by simply setting $\ell_g = 0 = \ell_\theta$,

$$f_\ell = -e\Phi_\ell\frac{\ell\frac{\partial f_0}{\partial p_\varphi}}{\omega - \ell\Omega_\varphi} \tag{36}$$

where $\ell_\varphi \equiv \ell$. We must remind the reader that these harmonic numbers are in action-angle space and not in real space. Since Maxwell's equations are simple in real space, we will need to convert f_ℓ's to real space harmonics before we can use them to calculate perturbed current and density. This translation mechanism is adequately discussed in Ref. 4.

Before ending this section we would like to point out that tremendous calculational simplification occurs for a very important special class of distribution functions $f_0(\mathbf{J})$ which depend on $\mathbf{J}$ only through the Hamiltonian H_0; the Maxwellian distribution belongs to this class. For $f_0(\mathbf{J}) = f_0[H_0(\mathbf{J})]$,

$$\frac{\partial f_0}{\partial \mathbf{J}} = \frac{\partial f_0}{\partial H_0}\frac{\partial H_0}{\partial \mathbf{J}} = \vec{\Omega}\frac{\partial f_0}{\partial H_0} \tag{37}$$

which results in the simplification

$$\dot{\mathbf{J}}_{\vec{\ell}}\cdot\frac{\partial f_0}{\partial \mathbf{J}} = -ie\vec{\Omega}\cdot\left[\vec{\ell}\Phi_{\vec{\ell}} - \frac{\omega}{c}\sum_{\vec{\ell}'}\mathbf{T}^j_{\vec{\ell}-\vec{\ell}'}A^j_{\vec{\ell}'}\right]\frac{\partial f_0}{\partial H_0}. \tag{38}$$

Further simplification is possible only when one knows $\mathbf{T}^{j}_{\vec{\ell}-\vec{\ell}'}$ which depends upon the particular problem being investigated.

V. Quasilinear Theory

A. General Theory

A very important application of the action-angle formalism is the development of the quasilinear transport theory for complicated geometries. In fact, Kaufman's original paper was precisely intended for this purpose.

The principal object of the quasilinear transport theory is to obtain an equation for the slow evolution of the equilibrium (or averaged) distribution function f_0 in response to the perturbing electromagnetic fields. The calculation is carried out in two distinct steps.

1) The equilibrium distribution function f_0 is given a time dependence $f_0 = f_0(\mathbf{J}, t)$ so that Eq. (15b) becomes

$$\frac{\partial f_0}{\partial t} + \frac{\partial f}{\partial t} + \vec{\Omega} \cdot \frac{\partial f}{\partial \vec{\theta}} = -\frac{\partial}{\partial \mathbf{J}} \cdot (\dot{\mathbf{J}} f_0 + \dot{\mathbf{J}} f) - \frac{\partial}{\partial \vec{\theta}} \cdot [(\delta \dot{\vec{\theta}}) f]. \tag{39}$$

Averaging over the angles $\vec{\theta}$ yields ($\langle F \rangle \equiv f_0$)

$$\frac{\partial f_0}{\partial t} = -\frac{\partial}{\partial \mathbf{J}} \cdot \langle \dot{\mathbf{J}} f_0 + \dot{\mathbf{J}} f \rangle \tag{40}$$

because the fast $\vec{\theta}$ variation gives zero average for the terms linear in the fluctuating part f. Notice that although $\dot{\mathbf{J}}$ is linear in the fluctuating fields Φ and $\mathbf{A}$, it has been retained in the equation because the coefficients $\partial \mathbf{J}/\partial p_{0j}$ can also have $\vec{\theta}$ dependence which could give nonzero average $\dot{\mathbf{J}}$. However, we must state that the factors $(\partial \mathbf{J}/\partial p_{0j})$ can have only slow equilibrium dependence in a quiescent plasma, and therefore the term proportional to $\langle \dot{\mathbf{J}} \rangle$ will contribute only for very low $\vec{\ell}$ number part of the fluctuation spectrum. Equation (40) can be written in the form

$$\frac{\partial f_0}{\partial t} = -\frac{\partial}{\partial \mathbf{J}} \cdot [\langle \dot{\mathbf{J}} f \rangle + \langle \dot{\mathbf{J}} \rangle f_0], \tag{41}$$

where f is still the exact fluctuating distribution function, i.e., the solution of Eq. (20).

2) Approximating the total fluctuating f by its linear value f^L is the essential assumption of the quasilinear theory, and leads to

$$\frac{\partial f_0}{\partial t} = -\frac{\partial}{\partial \dot{\mathbf{J}}} \cdot \left[\langle \dot{\mathbf{J}} f^L \rangle + \langle \dot{\mathbf{J}} \rangle f_0\right]. \tag{42}$$

Using the definition

$$f^L = \sum_{\vec{\ell}} f_{\vec{\ell}} \quad e^{i\vec{\ell}\cdot\vec{\theta}},$$

the definition of $\dot{\mathbf{J}}_{-\vec{\ell}}$, and Eq. (28), we can write Eq. (42) in the equivalent forms

$$\frac{\partial f_0}{\partial t} = \frac{\partial}{\partial \mathbf{J}} \cdot \left[\overleftrightarrow{D} \cdot \frac{\partial f_0}{\partial \mathbf{J}} \langle \dot{\mathbf{J}} \rangle f_0\right], \tag{43a}$$

or

$$\frac{\partial f_0}{\partial t} = \frac{\partial}{\partial \dot{\mathbf{J}}} \overleftrightarrow{D} \cdot \frac{\partial f_0}{\partial \dot{\mathbf{J}}} + \langle \dot{\mathbf{J}} \rangle \cdot \frac{\partial f_0}{\partial \mathbf{J}} \tag{43b}$$

which requires the use of

$$\left\langle \frac{\partial}{\partial \mathbf{J}} \cdot \dot{\mathbf{J}} \right\rangle + \left\langle \frac{\partial}{\partial \vec{\theta}} \cdot \dot{\vec{\theta}} \right\rangle = \frac{\partial}{\partial \mathbf{J}} \cdot \langle \dot{\mathbf{J}} \rangle = 0.$$

In Eqs. (43a) and (43b),

$$\overleftrightarrow{D} = i \sum_{\vec{\ell}} \frac{\dot{\mathbf{J}}_{\vec{\ell}} \dot{\mathbf{J}}_{-\vec{\ell}}}{\omega - \vec{\ell} \cdot \vec{\Omega}} \tag{44}$$

is the coefficient of quasilinear diffusion, and

$$\langle \dot{\mathbf{J}} \rangle = \dot{\mathbf{J}}_{\vec{\ell}=0} = -\frac{ie}{c} \sum_{\vec{\ell}'} \left[-\omega T^j_{-\vec{\ell}'} + \vec{\Omega} \times \left(T^j_{-\vec{\ell}'} \times \vec{\ell}'\right)\right] A^j_{\vec{\ell}'} \tag{45}$$

denotes convection in the action-angle space. Notice that the origin of the convective term is entirely due to the electromagnetic perturbation A^j.

At this stage, it is pertinent to point out that the quasilinear equation (43) preserves the particle conserving property of the exact Vlasov equation, i.e,

$$\frac{\partial N_0}{\partial t} = \frac{\partial}{\partial t} \int d\mathbf{J} f_0 = \int d\mathbf{J} \frac{\partial}{\partial \mathbf{J}} \cdot \left[D \cdot \frac{\partial f_0}{\mathbf{J}} + \langle \dot{\mathbf{J}} \rangle f_0\right] = 0. \tag{46}$$

Other conservation laws can be derived by the standard procedure.

If the perturbation were pure electrostatic ($\mathbf{A} = 0$), the convective term vanishes, and the evolution of f_0 in action space becomes purely diffusive,

$$\frac{\partial f_0}{\partial t} = \frac{\partial}{\partial \mathbf{J}} \cdot \overleftrightarrow{D}_{es} \cdot \frac{\partial}{\partial \mathbf{J}} f_0 \tag{47}$$

where [use of Eq. (24) for $\dot{\mathbf{J}}_{es}$]

$$\overleftrightarrow{D}_{es} = ie^2 \sum_{\vec{\ell}} \vec{\ell}\,\vec{\ell}\,\frac{\Phi_{\vec{\ell}}\,\Phi_{-\vec{\ell}}}{\omega - \vec{\ell}\cdot\vec{\Omega}}. \tag{48}$$

Since it is the presence of the magnetic perturbation $\mathbf{A}$ that destroys the canonical nature of $\mathbf{J}$ and $\vec{\theta}$, the convection term $\langle \dot{\mathbf{J}} \rangle$ will be a functional of $\mathbf{A}$ alone.

From the analysis in the preceding section, it is quite clear that the quasilinear transport theory becomes considerably more complicated when electromagnetic perturbations are present. The change is not only quantitative, but qualitative also in the sense that the evolution of f_0 is no more purely diffusive in the action-angle space. The relative importance of diffusion, and convection terms would clearly depend upon the details of the situation. Equations (43) and (44), the main results of this section, provide the starting equations to develop, for example, a theory of radial transport in tokamaks caused by electromagnetic fluctuations.

Nonlinear Theory

Equation (22) is the fully nonlinear Fourier transformed equation for the fluctuating distribution function. It can serve as a basis for developing a general renormalized turbulence theory for a Vlasov plasma in arbitrary geometry.

The standard renormalization procedure[7] consists in breaking up the nonlinearity $N_{\vec{\ell}}$ into three terms; the first is proportional to (coherent with) $f_{\vec{\ell}}$, the second is proportional to the perturbed electromagnetic fields, and the third is called the fully incoherent term. The effects of the electromagnetic fields are contained in $(\delta\dot{\vec{\theta}})$ and $\dot{\mathbf{J}}$ in our formulation. Thus we seek the decomposition of the nonlinear term as

$$N_{\vec{\ell}} = D_{\vec{\ell}} f_{\vec{\ell}} + \mathbf{C}_{1\vec{\ell}} \cdot (\delta\dot{\vec{\theta}})_{\vec{\ell}} + \mathbf{C}_{2\vec{\ell}} \cdot \dot{\mathbf{J}}_{\ell} + \tilde{N}_{\vec{\ell}} \tag{49}$$

where $\tilde{N}_{\vec{\ell}}$ is the remaining incoherent part of $N_{\vec{\ell}}$ after the coherent parts have been subtracted. Clearly, the term $D_{\vec{\ell}} f_{\vec{\ell}}$ will renormalize the linear propagator on the left-hand side of Eq. (22).

Detailed expressions for $D_{\vec{\ell}}$, $C_{1\vec{\ell}}$, and $\mathbf{C}_{2\vec{\ell}}$ are obtained by the following simple procedure. We add $D_{\vec{\ell}} f_{\vec{\ell}}$ to both sides of Eq. (22) to obtain the formal solution

$$f_{\vec{\ell}} = g_{\vec{\ell}} \left[+ \dot{\mathbf{J}}_{\vec{\ell}} \cdot \frac{\partial f_0}{\partial \mathbf{J}} + N_{\vec{\ell}} - D_{\vec{\ell}} f_{\vec{\ell}} \right] \tag{50}$$

where

$$g_{\vec{\ell}} = \frac{1}{i(\omega - \vec{\ell} \cdot \vec{\Omega} + iD_{\vec{\ell}})} \tag{51}$$

is the renormalized propagator. Substituting Eq. (50) into the nonlinear term Eq. (23) leads to

$$N_{\vec{\ell}} = \sum_{\vec{\ell}'} \left[i(\vec{\ell} - \vec{\ell}') \cdot (\delta \dot{\vec{\theta}})_{\vec{\ell}'} + \dot{\mathbf{J}}_p \cdot \frac{\partial}{\partial \mathbf{J}} \right] g_{\vec{\ell}-\vec{\ell}'} \left[N_{\vec{\ell}-\vec{\ell}'} + \dot{\mathbf{J}}_{\vec{\ell}-\vec{\ell}'} \cdot \frac{\partial f_0}{\partial \mathbf{J}} - D_{\vec{\ell}-\vec{\ell}'} f_{\vec{\ell}-\vec{\ell}'} \right] \tag{52}$$

where $\frac{\partial}{\partial \mathbf{J}}$ operates on all quantities on its right. It is straightforward to see that the terms coherent with $f_{\vec{\ell}}$ and $\mathbf{A}_{\vec{\ell}}$ can come only from $N_{\vec{\ell}-\vec{\ell}'}$, because $\vec{\ell}' = 0$ is not permitted. Denoting this part by $\bar{N}_{\vec{\ell}}$, and substituting Eq. (23) into Eq. (52), we obtain

$$\begin{aligned} \bar{N}_{\vec{\ell}} = \sum_{\vec{\ell}', \vec{\ell}''} & \left[i(\vec{\ell} - \vec{\ell}') \cdot (\delta \dot{\vec{\theta}})_{\vec{\ell}'} + \dot{\mathbf{J}}_{\vec{\ell}'} \cdot \frac{\partial}{\partial \mathbf{J}} \right] g_{\vec{\ell}-\vec{\ell}'} \\ & \left[i(\vec{\ell} - \vec{\ell}' - \vec{\ell}'') \cdot (\delta \dot{\vec{\theta}})_{\vec{\ell}''} f_{\vec{\ell}-\vec{\ell}'-\vec{\ell}''} + \dot{\mathbf{J}}_{\vec{\ell}''} \cdot \frac{\partial}{\partial \mathbf{J}} f_{\vec{\ell}-\vec{\ell}'-\vec{\ell}''} \right] \end{aligned} \tag{53}$$

from which we can readily extract

$$\begin{aligned} D_{\vec{\ell}} f_{\vec{\ell}} = \sum_{\vec{\ell}'} & \left[i(\vec{\ell} - \vec{\ell}') \cdot (\delta \dot{\vec{\theta}})_{\vec{\ell}'} + \dot{\mathbf{J}}_{\vec{\ell}'} \cdot \frac{\partial}{\partial \mathbf{J}} \right] g_{\vec{\ell}-\vec{\ell}'} \\ & \left[i\vec{\ell} \cdot (\delta \dot{\vec{\theta}})_{-\vec{\ell}'} + \dot{\mathbf{J}}_{-\vec{\ell}'} \cdot \frac{\partial}{\partial \mathbf{J}} \right] f_{\vec{\ell}} \, , \end{aligned} \tag{54}$$

$$\mathbf{C}_{1\vec{\ell}} \cdot (\delta \dot{\vec{\theta}})_{\vec{\ell}} = -i \sum_{\vec{\ell}'} \left[i(\vec{\ell} - \vec{\ell}') \cdot (\delta \dot{\vec{\theta}})_{\vec{\ell}'} + \dot{\mathbf{J}}_{\vec{\ell}'} \cdot \frac{\partial}{\partial \mathbf{J}} \right] \left(g_{\vec{\ell}-\vec{\ell}'} f_{-\vec{\ell}'} \vec{\ell}' \cdot (\delta \dot{\vec{\theta}})_{\vec{\ell}} \right) \tag{55}$$

and

$$\mathbf{C}_{2\vec{\ell}} \cdot \dot{\mathbf{J}}_{\vec{\ell}} = \sum_{\vec{\ell}'} \left[i(\vec{\ell} - \vec{\ell}') \cdot (\delta \dot{\vec{\theta}})_{\vec{\ell}'} + \dot{\mathbf{J}}_{\vec{\ell}'} \cdot \frac{\partial}{\partial \mathbf{J}} \right] \left(g_{\vec{\ell}-\vec{\ell}'} \frac{\partial f_{-\vec{\ell}'}}{\partial \mathbf{J}} \cdot \dot{\mathbf{J}}_{\vec{\ell}} \right). \tag{56}$$

By making use of expression for $(\delta \dot{\vec{\theta}})_{\vec{\ell}}$ and $\dot{\mathbf{J}}_{\vec{\ell}}$, one can easily express $\mathbf{C}_{1\vec{\ell}} \cdot (\delta \dot{\vec{\theta}})_{\vec{\ell}} + \mathbf{C}_{2\vec{\ell}} \cdot (\dot{\mathbf{J}})_{\vec{\ell}}$ as $a_{\vec{\ell}} \Phi_{\vec{\ell}} + \mathbf{b}_{\vec{\ell}} \cdot \mathbf{A}_{\vec{\ell}}$. Renormalized Eq. (22) reads

$$f_{\vec{\ell}} = g_{\vec{\ell}} \left\{ \left[\left(\frac{\partial f_0}{\partial \mathbf{J}} \right) + \mathbf{C}_{2\vec{\ell}} \right] \cdot \dot{\mathbf{J}}_{\ell} + \mathbf{C}_{1\ell} \cdot (\delta \dot{\vec{\theta}})_{\vec{\ell}} + \tilde{N}_{\ell} \right\}. \tag{57}$$

In the coherent renormalization theories, $\bar{N}_{\vec{\ell}} = 0$, and Eq. (57) serves as a formal nonlinear solution of the Vlasov equation.

VII. Discussion and Conclusions

We have shown that a mechanical application of the canonical perturbation theory can lead to erroneous results when magnetic field perturbations (**A**) are present. The principal reason is that for $\mathbf{A} \neq 0$, the invariant actions and their associated angles (derived for the unperturbed system) cease to be canonical variables of the total Hamiltonian H, i.e., $\dot{\mathbf{J}} \neq -\partial H/\partial\vec{\theta}$, and $\dot{\vec{\theta}} \neq \partial H/\partial \mathbf{J}$. We have also derived approximate expression for the orbits $\mathbf{J}$ and $\vec{\theta}$ for a charged particle moving in general electromagnetic fields (perturbed and unperturbed) to develop a modified perturbation theory which is used to formulate the kinetic theory of plasmas in the action-angle variables. Simple tests of the correctness of our theory are provided by comparing our results with standard known results. The general formalism is used to derive basic kinetic equation for the study of linear, quasilinear, and renormalized nonlinear theories.

As stated earlier, the great advantage of the action-angle variable approach is to unify the treatment of such diverse plasma problems as the determination of fluctuating distribution function for a field free plasma, and for trapped particles in a tokamak. The entire formal structure is the same, because in the invariant action-angle space, the particle trajectories are always straight lines. We believe that our Eqs. (28), (43)-(44), and (54)-(57) can be used as starting points for a wide variety of kinetic plasma problems.

The application of this formalism to specific problems of interest will be the subject of a later paper; this paper is intended to be a general delineation of the theory.

Acknowledgment

We thank Professor H.L. Berk for useful suggestions. This work was supported by the U.S. Department of Energy, contract no. DE-FG05-80ET-53088.

Appendix A – Equations of Motion and Particle Conservation Law

We first derive a general perturbation formalism for the evolution of equilibrium invariants of a charged particle. The Hamiltonian H of the particle is given by Eqs. (6) and (8), and the variables $(\mathbf{p}_0, \mathbf{q}_0)$ and $(\vec{\alpha}_0, \vec{\beta}_0)$ are described by Eqs. (4) and (5) and (7). The equation of motion of α_{0i} is [Eq. 10a]

$$\dot{\alpha}_{0i} = \left.\frac{\partial \alpha_{0i}}{\partial t}\right|_{(\mathbf{p},\mathbf{q})} + [\alpha_{0i}, H]_{(\mathbf{p},\mathbf{q})}. \qquad (A-1)$$

From Eq. (7), and $q = q_0$, we derive the transformation coefficients,

$$\frac{\partial p_{0j}}{\partial p_i} = \delta_{ij} \quad , \quad \frac{\partial q_{0j}}{\partial p_i} = 0$$

$$\frac{\partial p_{0j}}{\partial q_i} = -\frac{e}{c}\frac{\partial A_j}{\partial q_i} \quad , \quad \frac{\partial q_{0j}}{\partial q_i} = \delta_{ij} \qquad (A-2)$$

which are used to obtain

$$\left.\frac{\partial \alpha_{0i}}{\partial t}\right|_{(\mathbf{p},\mathbf{q})} = -\frac{e}{c}\frac{\partial \alpha_{0i}}{\partial p_{0j}}\frac{\partial A^j}{\partial t}, \qquad (A-3)$$

$$[\alpha_{0i,H}]_{(\mathbf{p},\mathbf{q})} = [\alpha_{0i,H}]_{\mathbf{p}_0,\mathbf{q}_0} + \langle \alpha_{0i,H} \rangle_{\mathbf{p}_0,\mathbf{q}_0} \qquad (A-4)$$

where the bracket $\langle \quad \rangle$ $\mathbf{p}_0$, $\mathbf{q}_0$ is defined by

$$\langle f, g \rangle_{\mathbf{p}_0,\mathbf{q}_0} \equiv -\frac{e}{c}\left\{ [A^j, g]_{\mathbf{p}_0,\mathbf{q}_0}\frac{\partial f}{\partial p_{0j}} - [A^j, f]_{\mathbf{p}_0,\mathbf{q}_0}\frac{\partial g}{\partial p_{0j}} \right\}. \qquad (A-5)$$

Since both the brackets [ ] and $\langle \quad \rangle$ are linear in their arguments, Eq. (A4) can be split into four terms

$$[\alpha_{0i}, H]_{(\mathbf{p},\mathbf{q})} = [\alpha_{0i}, H_0]_{\mathbf{p}_0,\mathbf{q}_0} + [\alpha_{0i,h}]_{\mathbf{p}_0,\mathbf{q}_0}$$

$$+ \langle \alpha_{0i}, H_0 \rangle_{\mathbf{p}_0,\mathbf{q}_0} + \langle \alpha_{0i}, h \rangle_{\mathbf{p}_0,\mathbf{q}_0}. \qquad (A-6)$$

Since the perturbed Hamiltonian $h \equiv e\Phi$ is independent of p_{0i}, the fourth term of Eq. (A-6) vanishes

$$\langle \alpha_{0i}, h \rangle_{\mathbf{p}_0,\mathbf{q}_0} = 0.$$

Since $(\vec{\alpha}_0, \vec{\beta}_0)$, though not canonical variables for the total Hamiltonian H, are obtained as a canonical transformation from $\mathbf{p}_0$, $\mathbf{q}_0$, the Poisson bracket remains invariant, i.e.

$$[f, g]_{\mathbf{p}_0,\mathbf{q}_0} = [f, g]_{\vec{\alpha}_0,\vec{\beta}_0}$$

which immediately leads to

$$[\alpha_{0i}, H_0]_{\mathbf{p}_0,\mathbf{q}_0} = [\alpha_{0i}, H_0(\vec{\alpha}_0)]_{\vec{\alpha}_0,\vec{\beta}_0} = 0.$$

In order to apply the invariance of Poisson brackets to other terms, we need probably do a formal treatment of both the electrostatic field Φ and electromagnetic field $\mathbf{A}$. We assume that we have obtained the explicit solutions from the transformation Eq. (5),

$$\mathbf{q}_0 = \mathbf{q}_0(\vec{\alpha}_0, \vec{\beta}_0)$$

$$\mathbf{p}_0 = \mathbf{p}_0(\vec{\alpha}_0, \vec{\beta}_0). \qquad (A-7)$$

Since the α_0's are motion invariants of the particle in the unperturbed situation, we use them as labels to mark the particle. The fields acting on the particle are expressed as

$$\Phi_{\vec{\alpha}0}(\vec{\beta}_0, t) = \Phi(\mathbf{q}_0(\vec{\alpha}_0, \vec{\beta}_0), t)$$

$$\mathbf{A}_{\vec{\alpha}0}(\vec{\beta}, t) = \mathbf{A}(\mathbf{p}_0(\vec{\alpha}_0, \vec{\beta}_0), t). \qquad (A-8)$$

With this understanding, Eq. (A-6) can be further simplified:

$$[\alpha_{0i}, H_0]_{\mathbf{p}_0,\mathbf{q}_0} = -e\frac{\partial \Phi}{\partial \beta_{0i}}$$

$$\begin{aligned}\langle \alpha_{0i}, H_0\rangle_{\mathbf{p}_0,\mathbf{q}_0} &= -\frac{e}{c}\left\{ \left[A^j, H_0\right]_{\alpha_0,\beta_0} \frac{\partial \alpha_{0i}}{\partial p_{0j}} - \left[A^j, \alpha_{0i}\right]_{\alpha_0,\beta_0} \frac{\partial H_0}{\partial p_{0j}} \right\} \\ &= -\frac{e}{c}\left\{ \frac{\partial A^j}{\partial \beta_{0k}} \cdot \frac{\partial H_0}{\partial \alpha_{0k}} \cdot \frac{\partial \alpha_{0i}}{\partial p_{0j}} - \frac{\partial A^j}{\partial \beta_{0i}} \cdot \frac{\partial H_0}{\partial \alpha_{0k}} \cdot \frac{\partial \alpha_{0k}}{\partial p_{0j}} \right\}\end{aligned}$$

where Einstein's rule is obeyed (and the same in the rest of the paper). Finally, we find

$$\begin{aligned}\dot{\alpha}_{0i} = -e\frac{\partial \Phi}{\partial \beta_{0i}} - \frac{e}{c}\frac{\partial \alpha_{0i}}{\partial p_{0j}} \cdot \frac{\partial A^j}{\partial t} - \frac{e}{c}\left\{ \frac{\partial \alpha_{0i}}{\partial p_{0j}} \cdot \frac{\partial H_0}{\partial \alpha_{0k}} \frac{\partial}{\partial \beta_{0k}} \right. \\ \left. - \frac{\partial \alpha_{0k}}{\partial p_{0j}} \cdot \frac{\partial H_0}{\partial \alpha_{0k}} \cdot \frac{\partial}{\partial \beta_{0i}} \right\} A^j \end{aligned} \qquad (A-9)$$

or equivalently (the subscript 0 if suppressed for α and β)

$$\dot{\vec{\alpha}} = -e\frac{\partial \Phi}{\partial \vec{\beta}} - \frac{e}{c}\frac{\partial \vec{\alpha}}{\partial p_{0j}}\frac{\partial A_j}{\partial t} - \frac{e}{c}\left[\frac{\partial \vec{\alpha}}{\partial p_{0j}} \cdot \left(\frac{\partial H_0}{\partial \vec{\alpha}} \cdot \frac{\partial}{\partial \vec{\beta}} \right) - \left(\frac{\partial \vec{\alpha}}{\partial p_{0j}} \cdot \frac{\partial H_0}{\partial \vec{\alpha}} \right) \frac{\partial}{\partial \vec{\beta}} \right] A^j.$$

Equation (A-9) is the evolution equation for an invariant under the influence of electrostatic and electromagnetic perturbed field.

Following a similar procedure, an equation for the conjugate β_{0i} is readily obtained

$$\dot{\vec{\beta}} = \frac{\partial H_0}{\partial \vec{\alpha}} + e\frac{\partial \Phi}{\partial \vec{\alpha}} - \frac{e}{c}\frac{\partial \vec{\beta}}{\partial p_{0j}} \cdot \frac{\partial A^j}{\partial t}$$
$$- \frac{e}{c}\left\{ \frac{\partial \vec{\beta}}{\partial p_{0j}} \left(\frac{\partial H_0}{\partial \vec{\alpha}} \cdot \frac{\partial}{\partial \vec{\beta}} \right) + \left(\frac{\partial \vec{\alpha}}{\partial p_{0j}} \cdot \frac{\partial H_0}{\partial \vec{\alpha}} \right) \frac{\partial}{\partial \vec{\alpha}} \right\} A^j. \qquad (A-10)$$

Although the variables $\vec{\alpha} \equiv \mathbf{J}$, and $\vec{\beta} \equiv \vec{\theta}$ are not canonical with respect to the total Hamiltonian H, the equations of motion A-9 and A-10 [or equivalently Eqs. (11) and (12)] satisfy the extremely important divergence relation

$$S = \frac{\partial}{\partial \mathbf{J}} \cdot (\dot{\mathbf{J}}) + \frac{\partial}{\partial \vec{\theta}} \cdot (\dot{\vec{\theta}}) \equiv \frac{\partial \dot{\mathbf{J}}_i}{\partial J_i} + \frac{\partial \dot{\theta}_i}{\partial \theta_i} = 0. \qquad (A-11)$$

The proof of (A-11) is a bit tedius, and requires making use of several relationships between $\mathbf{p}_0$, $\mathbf{q}_0$ and $\mathbf{J}$, $\vec{\theta}$. Two of these useful identities are

$$\frac{\partial \mathbf{J}}{\partial p_{0j}} = \frac{\partial q_{0j}}{\partial \vec{\theta}} \qquad (A-12)$$

$$\frac{\partial \vec{\theta}}{\partial p_{0j}} = -\frac{p q_{0j}}{\partial \mathbf{J}} \qquad (A-13)$$

which can be readily derived by remembering that $\dot{\theta} = \partial H_0 / \partial \mathbf{J}$, and $\dot{\mathbf{J}}\Big|_{H_0} = 0$. Taking the appropriate divergence of Eqs. 11 and 12, and using (A-12) and (A-13), we obtain

$$\frac{c}{e} S = S_1 + S_2 \qquad (A-14)$$

where

$$S_1 = -\frac{p}{\partial J_i}\left[\frac{\partial q_{0j}}{\partial \theta_i} \frac{\partial A_j}{\partial t}\right] + \frac{\partial}{\partial \theta_i}\left[\frac{\partial q_{0i}}{\partial J_i} \frac{\partial A_j}{\partial t}\right]$$
$$= -\left[q_{0j} \; , \; \frac{\partial A_j}{\partial t}\right]_{\vec{\theta},\mathbf{J}} = \left[q_{0j} \; , \; \frac{\partial A_j}{\partial t}\right]_{\mathbf{q}_0,\mathbf{p}_0} = 0 \qquad (A-15)$$

because A_j is independent of $\mathbf{p}_0$, and

$$S_2 = -\frac{\partial}{\partial J_i}\left[\frac{\partial q_{0j}}{\partial \theta_i}\left(\Omega \cdot \frac{\partial A_j}{\partial \vec{\theta}}\right) - \left(\vec{\Omega} \cdot \frac{\partial q_{0j}}{\partial \vec{\theta}}\right)\frac{\partial A_j}{\partial \theta_i}\right]$$
$$+ \frac{\partial}{\partial \theta_i}\left[\frac{\partial q_{0j}}{\partial \mathbf{J}}\left(\vec{\Omega} \cdot \frac{p A_j}{\partial \vec{\theta}}\right) - \left(\Omega \cdot \frac{\partial q_{0j}}{\partial \vec{\theta}}\right)\frac{\partial A_j}{\partial \mathbf{J}}\right].$$

Using the facts

$$\frac{\partial A_j}{\partial \vec{\theta}} = \frac{\partial A_j}{\partial q_{0m}} \frac{\partial q_{0m}}{\partial \vec{\theta}},$$

and

$$\frac{\partial A_j}{\partial \mathbf{J}} = \frac{\partial A_j}{\partial q_{0m}} \frac{\partial q_{0m}}{\partial \mathbf{J}}, \qquad (A-16)$$

and straightforward manipulation, we can show that

$$\begin{aligned} S_2 &= \frac{\partial}{\partial J_i}\left[T_{jm} \frac{\partial q_{0j}}{\partial \theta_i} \vec{\Omega} \cdot \frac{\partial q_{0m}}{\partial \theta}\right] \frac{\partial}{\partial \theta_i}\left[T_{jm} \frac{\partial q_{0j}}{\partial J_i} \vec{\Omega} \cdot \frac{\partial q_{0m}}{\partial \vec{\theta}}\right] \\ &= \left[q_{0j}, T_{jm}\, \vec{\Omega} \cdot \frac{\partial q_{0m}}{\partial \vec{\theta}}\right]_{\vec{\theta},\mathbf{J}} = \left[q_{0j}, T_{jm}\, \vec{\Omega} \cdot \frac{\partial q_{0m}}{\partial \theta}\right]_{\mathbf{q}_0,\mathbf{p}_0} \end{aligned} \qquad (A-17)$$

where T_{jm} is the antisymmetric tensor

$$T_{jm} = \frac{\partial A_j}{\partial q_{0m}} - \frac{\partial A_m}{\partial q_{0j}}. \qquad (A-18)$$

Noticing

$$\vec{\Omega} \cdot \frac{\partial q_{0m}}{\partial \vec{\theta}} = \vec{\Omega} \cdot \frac{\partial \mathbf{J}}{\partial p_{0m}} = \frac{\partial H_0}{\partial \mathbf{J}} \cdot \frac{\partial \mathbf{J}}{\partial p_{0m}} = \frac{\partial H_0}{\partial p_{0m}},$$

we expand Eq. (17) to yield

$$S_2 = \frac{\partial q_{0j}}{\partial q_{0\ell}} \frac{\partial}{\partial p_{0\ell}} T_{jm} \frac{\partial H_0}{\partial p_{0m}} = T_{\ell m} \frac{\partial^2 H_0}{\partial p_{0\ell} p_{0m}} = 0 \qquad (A-19)$$

since $\left(\partial^2 H_0/\partial p_{0\ell}\partial p_{-m}\right)$ is symmetric in ℓ and m. Thus we have proved that for our action-angle variable Eq. (11) is valid, which allows us to write the Vlasov equation in an equivalent form

$$\frac{\partial F}{\partial t} + \dot{\vec{\theta}} \cdot \frac{\partial F}{\partial \vec{\theta}} + \dot{\mathbf{J}} \cdot \frac{\partial F}{\partial \mathbf{J}} = \frac{\partial F}{\partial t} + \frac{\partial}{\partial \vec{\theta}} \cdot (\dot{\vec{\theta}} F) + \frac{\partial}{\partial \mathbf{J}} \cdot (\dot{\mathbf{J}} F) = 0, \qquad (A-20)$$

which is clearly more suitable for the explicit demonstration of conservation laws associated with the Vlasov equation. It is obvious that the Vlasov equation conserves particles because

$$\frac{\partial N}{\partial t} = 0 \qquad (A-21)$$

where

$$N = \int d\mathbf{J}\ d\vec{\theta}\ F. \qquad (A-22)$$

References

1. M. Kruskal, J.M.P. **3**, (1962).
2. A. Lenard, Ann. Phy. (N.Y.) **6**, 261-276 (1959).
3. A.N. Kaufman, Phys. Fluids **15**, 1063 (1972).
4. D.A. Hitchcock, R.D. Hazeltine, and S.M. Mahajan, Phys. Fluids **26**, 2603 (1983); R.D. Hazeltine, S.M. Mahajan and D.A. Hitchcock, Phys. Fluids **25**, 350 (1982); S.M. Mahajan, R.D. Hazeltine and D.A. Hitchcock, Phys. Fluids **26**, 700 (1983); R.D. Hazeltine, Phys. Fluids **24**, 1164 (1981).
5. H. Goldstein, *Classical Mechanics* (Second Edition, Addison-Wesley Publishing Co., Inc., 1980) p. 499-506.
6. N.A. Krall and A.W. Trivelpiece, *Principles of Plasma Physics* (McGraw-Hill Book Co., 1973), p. 399.
7. T.H. Dupree, D.J. Tetreault, Phys. Fluids **21**, 425 (1978).

THE UNIVERSITY OF TEXAS AT AUSTIN
INSTITUTE FOR FUSION STUDIES
AUSTIN, TEXAS 78712

June 20, 1985

Professor F.L. Ribe
Editor, The Physics of Fluids
AERL Building, FL-10
University of Washington
Seattle, Washington 98195

Dear Professor Ribe:

Enclosed are three (3) copies of "Plasma Kinetic Theory in Action-Angle Variables", by S.M. Mahajan and C.Y. Chen. We are submitting this paper for publication in The Physics of Fluids.

Thank you for your consideration.

Sincerely,

Swadesh M Mahajan

S.M. Mahajan

/sc
Enc.: 3 copies ms

TECHNICAL PROGRESS REPORT FORM

Title of Paper: Plasma Kinetic Theory in Action-Angle Variables

Author (first): S. M. Mahajan

First Draft: JUN 14 1985

1st Corrections:	JUN 17 1985	11th Corrections:	
2nd Corrections:	JUN 19 1985	12th Corrections:	
3rd Corrections:	JUN 20 1985	13th Corrections:	
4th Corrections:		14th Corrections:	
5th Corrections:		15th Corrections:	
6th Corrections:		16th Corrections:	
7th Corrections:		17th Corrections:	
8th Corrections:		18th Corrections:	
9th Corrections:		19th Corrections:	
10th Corrections:		20th Corrections:	

FINAL VERSION: JUN 20 1985

Journal & Date Sent: Physics of Fluids JUN 20 1985

Date sent to Central Duplicating: JUN 24 1985

Date picked up from Central Duplicating:

Person filling out this form: Suzy

Th

Dr. Allan
Departmen
Universi
Berkeley

Dear Al

Enclose
Angle V

I have
you wo

Thank

Since

Sm

Swa

Enc.:

CENTRAL DUPLICATING WORK ORDER FORM

Institute for Fusion Studies: (Acct. #) 20-7488-0050

Fusion Research Center: (Acct. #) ____________

Title of Paper: Plasma Kinetic Theory . . . .

Report #: 194

First Author: S. M. Mahajan

# of copies needed: 235

# of pages (incl. window and figs.): 23

Single-side window - duplex the rest: ✓

Single side entire manuscript: ____________

Collate: ✓ Staple: ✓

Journal name and date submitted: Phys. Fluids JUN 20 1985

Internal distribution only: ____________

Whose mailing list(s)? Smm

TEXnician: Suzy

Date sent to CD: JUN 24 1985

Date picked up from CD: JUL 05 1985

(Ticket# ____________)

THE UNIVERSITY OF TEXAS AT AUSTIN
CENTRAL DUPLICATING

RLM

DEPARTMENT NAME Fusion Studies

TICKET NO. H № 15713

ACCOUNT NO. 20-7482-0050

OBJECT CODE 157

DATE 6-24 1985

PAPER
☑ Letter ☐ Legal ☐ Ink ☐ Other

OFFSET	Number of Originals	Number of Copies	TOTAL	Price
☑ Simplex cover	23	235	5405	92.00
☑ Duplex rest	30 (SRH)	450	13,500	120.00
☐ COPIES				

		Price
☑ ASSEMBLE		29.80
☑ STAPLE	two on left	68.50
☐ FOLD		
☐ TRIM		
☐ DRILL		
☐ BINDING		
☐ SPECIAL		
☐ ADDRESS		

Distribution Special Instructions

JOB DESCRIPTION
Must be adequate to indicate work is of an official University nature
IFS # 194 + 175

Signature R Hazeltine

☐ WHEN READY NOTIFY Suzy 6122 PAX CTX

DATE PROMISED

☐ DELIVER (if large enough)

RECEIVED BY R Hazeltine

TOTAL CHARGE 370 30

4. FOR DEPARTMENT

DOE/ET-53088-194

IFSR #194

PLASMA KINETIC THEORY IN ACTION-ANGLE VARIABLES

S.M. Mahajan

and

C.Y. Chen
Institute for Fusion Studies
The University of Texas at Austin
Austin, Texas 78712

June 1985

Plasma Kinetic Theory in Action-Angle Variables

S.M. Mahajan

and

C.Y. Chen

Institute for Fusion Studies

The University of Texas at Austin

Austin, Texas 78712

Abstract

An appropriate canonical perturbation theory to correctly deal with general electromagnetic field perturbation is developed, and is used to set up plasma kinetic theory in action-angle variables. A variety of test problems are solved to show the unifying power of the method. Basic linear, quasilinear, and nonlinear equations are derived which can serve as the starting point for a whole range of plasma problems.

I. Introduction

Action and angle variables have often provided a very convenient coordinate system in which the problems of classical physics can be formulated and solved. In celestial mechanics, for example, this approach has been extensively used to determine the effects of small perturbations on the motion of the planets. It is a bit strange that this obviously powerful tool has not found much currency amongst researchers in plasma physics. Apart from a few initial studies,[1-2] the use of action invariants to solve the Vlasov equation in complicated geometry was attempted only recently by Kaufmann,[3] and then by other workers[4] who primarily used and extended Kaufmann's formalism. Unfortunately, the basic formalism is fundamentally flawed, and when applied to deal with time-dependent magnetic field perturbations, it leads to incorrect results even for the linear Vlasov theory! (see Sec. IV for details.) The formalism, however, yields correct results for strictly electrostatic perturbations which defines the domain of validity of the analysis of Refs. 3-4.

It is surprising that a method based on the well-known time-dependent canonical perturbation theory (Kaufmann's) would lead to such inconsistencies. In this paper, we show why a mechanical application of the standard perturbation theory is not possible in the presence of the electromagnetic perturbed fields (EMPF). We also develop a modified perturbation theory which can correctly handle the EMPF. Arguments leading to, and the formalism, are presented in Sec. II.

The remaining sections of this paper are devoted to the formulation of plasma kinetic theory in terms of action-angle variables using our modified perturbation theory. The Vlasov equation in the new phase space is derived in Sec. III, while some examples in the linear theory are worked out in Sec. IV, where we show that our results are identical with standard results, and are different from the results of the earlier theory.[3] After having shown that our theory has corrected the inadequacies of the earlier theory, we go ahead and develop quasilinear and nonlinear aspects of plasma kinetic theory in action-angle variables in Secs. V and VI. Expressions for the quasilinear diffusion coefficient, and the convection coefficient are derived in Sec. V, while Sec. VI is essentially devoted to a formal description of the renormalized nonlinear theory. A brief discussion, which includes some comments on the importance of this work, is given in Sec. VII.

II. Modified Canonical Theory for EMPF

An essential step in setting up the time-dependent or the canonical perturbation theory[5] is to obtain motion invariants of the corresponding unperturbed system described by the Hamiltonian H_0. Provided that the found invariants α's and their conjugates β's constitute a complete set to form a new phase space, a generating function G must exist which can perform a canonical transformation (p and q are the original phase space coordinates)

$$(\mathbf{p}, \mathbf{q}) \stackrel{G(\mathbf{q},\vec{\alpha})}{\longrightarrow} (\vec{\alpha}, \vec{\beta}) \tag{1}$$

with the result that the motion of the total system is simply along a constant line in the new phase space ($\vec{\alpha}$, $\vec{\beta}$), i.e., the Hamiltonian H_0 is independent of the β's.

For the perturbed system, the transformation Eq. (1) is still employed, and then the Hamiltonian H can be expressed as a sum of two parts; the formally unchanged, β independent H_0, and h which is a function of $(\mathbf{p}, \mathbf{q})$ as well as ($\vec{\alpha}$, $\vec{\beta}$). Clearly, the α's change only due to h,

$$\dot{\alpha}_i = -\frac{\partial H}{\partial \beta_i} = -\partial h(\alpha_i, \beta_i, t)/\partial \beta_i. \tag{2}$$

Equation (2) provides a basis for a perturbation theory. The first step in the approximation scheme is to substitute the unperturbed constant values of the $\alpha_i = \alpha_{i0}$ in the right-hand side of Eq. (2) (after taking the β_i derivatives). The equation of motion then could be integrated to obtain the perturbed time dependent α_i's. Notice that the invariant actions $\mathbf{J}$ or J_i are to be identified with α_i, and the conjugate angles $\vec{\theta}$ or θ_i with β_i in the action-angle formalism.

Although this procedure is standard and widely used, it runs into considerable difficulty when applied to the treatment of magnetic field perturbations (EMPF). We must point out that there is nothing theoretically incorrect with the method. In fact, the phase space ($\vec{\alpha}$, $\vec{\beta}$) obtained from transformation Eq. (1) is canonical, and does reduce to its counterpart in the corresponding unperturbed system, which seemingly fulfills the requirements for a perturbation theory. However, an important technical problem arises when magnetic field perturbation is present; the generalized momentum (e, m and $\mathbf{v}$ are the charge, mass and velocity of the particle respectively, and $\mathbf{A}$ is the vector potential)

$$\mathbf{p} = m\mathbf{v} + \frac{e}{c}\mathbf{A} \tag{3}$$

contains the perturbed vector potential explicitly. Consequently, the coordinate system ($\vec{\alpha}$, $\vec{\beta}$) obtained from $(\mathbf{p}, \mathbf{q})$ by a canonical transformation also contains the EMPF (in the rest of this paper, the abbreviation EMPF will be used for the perturbed magnetic field, perturbed vector potential etc.). The use of such a coordinate system to describe the motion of any dynamical system must necessarily suffer from severe intrinsic disadvantages.

1) Generally, the object of a calculation will be to determine the EMPF. Thus the meaning of the coordinates will remain quite obscure till the problem is solved.

2) The coordinate frame is, by definition, not the same as its counterpart in the unperturbed system. This confuses the situation because a cornerstone of the perturbation theory is that we do have the knowledge of the unperturbed system and its phase-space variables α_{i0} and β_{i0}. The confusion between (α_i, β_i) and (α_{i0} and β_{i0}) can lead to serious errors as in the linearized solution of Ref. 3. We discuss this point in detail in the next section.

3) Because the coordinate system has fast variation caused by the varying EMPF, it is no longer practical to solve systems like Vlasov equation by decomposing the perturbation into Fourier harmonics; the superposition is no longer valid, and the resultant Fourier transform of the equation will have convolutions even in the linear analysis.

It is thus strongly indicated that we must look for a different coordinate system in order to exploit the powerful and elegant machinery of canonical perturbation theory. A logical choice will be the coordinate system which contains only the equilibrium electromagnetic fields, given by the vector potential $\mathbf{A}_0$, and the scalar potential Φ_0, i.e., the generalized canonical momentum

$$\mathbf{p}_0 = m\mathbf{v} + \frac{e}{c}\mathbf{A}_0 \tag{4}$$

should replace $\mathbf{p}$. It is from this new set $\mathbf{p}_0$ and $\mathbf{q}_0 \equiv \mathbf{q}$ that we obtain the appropriate $\vec{\alpha}_0$ and $\vec{\beta}_0$ by a generating function G,

$$\mathbf{p}_0, \mathbf{q}_0 \xrightarrow{G(\mathbf{q}_0,\, \vec{\alpha}_0)} (\vec{\alpha}_0,\ \vec{\beta}_0). \tag{5}$$

Notice that the transformation given in Eq. (5) is exactly equivalent to solving the equations of motion of a particle in the equilibrium fields. We, of course, assume that this problem is solved for all cases under consideration. For example, Kaufmann[3] has implicitly solved the equilibrium motion of a guiding center plasma in a torus, while Hazeltine, Mahajan and Hitchcock[4] have obtained explicit expressions for the special case of a high aspect-ratio torus.

We wish to point out here that in the absence of the EMPF ($\mathbf{A} = 0$), $\mathbf{p}_0$ is identically equal to $\mathbf{p}$, and thus the two coordinate systems will be equivalent, and the standard perturbation theory will yield correct results. Clearly, the results of all the previous papers following this approach are correct for purely electrostatic perturbations ($\mathbf{A} = 0$, $\Phi \neq 0$).

These new variables, however, do not form a canonical conjugate pair with respect to the total Hamiltonian H, ($\mathbf{A} \neq 0$) i.e., $\dot{\alpha}_{0i} \neq \partial H/\partial \beta_{0i}$. This fact is of crucial importance, because it is the identification of $\dot{\alpha}_{0i}$ with $\partial H/\partial \beta_{0i}$ that constitutes the principal mistake of Refs. 3-4. For our case, we shall have to find appropriate expressions for the rate of change of actions and angles in the presence of perturbations. We must remark that although $\vec{\alpha}_0$ and $\vec{\beta}_0$ are not canonical (for H), they are obtained as a canonical transformation from $\mathbf{p}_0$, $\mathbf{q}_0$ which label the unperturbed state described by the equilibrium Hamiltonian H_0. Thus $(\vec{\alpha}_0,\ \vec{\beta}_0)$ will retain some of the crucial properties of canonical variables.

The nonrelativistic Hamiltonian for a charged particle in an electromagnetic field is given by

$$H = \frac{1}{2m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}_t\right)^2 + e\Phi_t \tag{6}$$

where the scalar and the vector potentials Φ_t and $\mathbf{A}_t$ include both the unperturbed (Φ_0, $\mathbf{A}_0$) and the perturbed (Φ, $\mathbf{A}$) components, i.e.,

$$\Phi_t = \Phi_0 + \Phi,$$

$$\mathbf{A}_t = \mathbf{A}_0 + \mathbf{A}.$$

By making use of the relation

$$\mathbf{p}_0 = \mathbf{p} - \frac{e}{c}\mathbf{A} \tag{7}$$

which follows from Eqs. (3) and (4), the Hamiltonian H can be written in the form

$$H = H_0 + h$$

where

$$H_0 = \frac{1}{2m}\left(\mathbf{p}_0 - \frac{e}{c}\mathbf{A}_0\right)^2 + e\Phi_0, \tag{8}$$

$$h = e\Phi.$$

It may seem peculiar that the troublesome perturbed vector potential $\mathbf{A}$ has completely disappeared from the scene; it appears explicitly in neither H_0 nor h. The effects of $\mathbf{A}$, however, will reappear when we make the transformation from $(\mathbf{p}, \mathbf{q})$ to $\mathbf{p}_0$, $\mathbf{q}_0$.

As we remarked earlier, the first step in the development of the canonical perturbation theory is the solution of the unperturbed problem. We assume that the unperturbed problem has been completely solved: the invariants $\vec{\alpha}_0$, the conjugate angles $\vec{\beta}_0$, and the generating function $G(\mathbf{q}_0, \vec{\alpha}_0)$ which mediates the canonical transformation of Eq. (5), have all been obtained. Since $\vec{\alpha}_0$ are the invariants of the unperturbed system, the Hamiltonian H_0 is a function of $\vec{\alpha}_0$ alone, i.e.,

$$H_0 = H_0(\vec{\alpha}_0) \tag{9}$$

a property which will be extensively used later, and which is a major source of simplification.

The equation of evaluation of any dynamical quantity Q (for example $\vec{\alpha}_0$ and $\vec{\beta}_0$ labelling the trajectory of the changed particle) is governed by the total Hamiltonian H, and is

$$\dot{Q} = \left.\frac{\partial Q}{\partial t}\right|_{(\mathbf{p},\mathbf{q})} + [Q, H]_{(\mathbf{p},\mathbf{q})} \tag{10a}$$

where

$$[f, g] = \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} \tag{10b}$$

is the Poisson bracket. In Eq. (10b), we have used Einstein's summation convention over repeated indices; this convention will be used throughout the paper.

Our aim is now to transform Eq. (10a) to the coordinate system defined by the equilibrium invariants, and their conjugate angles ($\vec{\alpha}_0$, $\vec{\beta}_0$). To do this, we must first transform the equation to the coordinate system $\mathbf{p}_0$, $\mathbf{q}_0$ because it is only these (and not $(\mathbf{p},\mathbf{q})$) which are related to ($\vec{\alpha}_0$, $\vec{\beta}_0$) by a known canonical transformation [Eq. (5)]. This step is the distinguishing feature of our analysis. The primary source of error in the earlier analysis (Refs. 3-4) is that they failed to distinguish between $(\mathbf{p},\mathbf{q})$ and $\mathbf{p}_0$, $\mathbf{q}_0$.

The details of the transformations $(\mathbf{p},\mathbf{q}) \to \mathbf{p}_0,\ \mathbf{q}_0 \to (\vec{\alpha}_0,\ \vec{\beta}_0)$ are worked out in Appendix A. Here we simply give the trajectories of the changed particle in terms of the action-angle variables, $\vec{\alpha}_0 \equiv \mathbf{J}$, $\vec{\beta}_0 \equiv \vec{\theta}$. The evolution equations are

$$\begin{aligned}\dot{\mathbf{J}} &= -e\frac{\partial \Phi}{\partial \vec{\theta}} - \frac{e}{c}\frac{\partial \mathbf{J}}{\partial p_{0j}}\frac{\partial A_j}{\partial t} - \frac{e}{c}\left[\frac{\partial \mathbf{J}}{\partial p_{0j}}\frac{\partial H_0}{\partial \mathbf{J}}\cdot\frac{\partial}{\partial \vec{\theta}} - \left(\frac{\partial \mathbf{J}}{\partial p_{0j}}\cdot\frac{\partial H_0}{\partial \mathbf{J}}\right)\frac{\partial}{\partial \vec{\theta}}\right]A_j \\ &= -e\frac{\partial \Phi}{\partial \vec{\theta}} - \frac{e}{c}\frac{\partial \mathbf{J}}{\partial p_{0j}}\frac{\partial A_j}{\partial t} - \frac{e}{c}\vec{\Omega}\times\left(\frac{\partial \mathbf{J}}{\partial p_{0j}}\times\frac{\partial A_j}{\partial \vec{\theta}}\right),\end{aligned} \tag{11}$$

$$\begin{aligned}(\dot{\delta\vec{\theta}}) \equiv \dot{\vec{\theta}} - \vec{\Omega} = -e\frac{\partial \Phi}{\partial \mathbf{J}} - \frac{e}{c}\frac{\partial \vec{\theta}}{\partial p_{0j}}\frac{\partial A_j}{\partial t} - \frac{e}{c}\bigg[\frac{\partial \vec{\theta}}{\partial p_{0j}}\vec{\Omega}\cdot\frac{\partial}{\partial \vec{\theta}} \\ + \left(\frac{\partial \mathbf{J}}{\partial p_{0j}}\cdot\vec{\Omega}\right)\frac{\partial}{\partial \mathbf{J}}\bigg]A_j.\end{aligned} \tag{12}$$

where $\Phi = \Phi(\mathbf{J},\vec{\theta},t)$, and $\mathbf{A} = \mathbf{A}(\mathbf{J},\vec{\theta},t)$. The quantities $(\partial\mathbf{J}/\partial p_{0j})$, $(\partial\vec{\theta}/\partial p_{0j})$ and $\partial H_0/\partial\mathbf{J} \equiv \vec{\Omega}(\mathbf{J})$ are known functions of $\mathbf{J}$ and $\vec{\theta}$ from the solution of the unperturbed problem. Since $\mathbf{J}$'s are the invariants in the equilibrium fields, their evolution is due to the perturbed fields only; $\dot{\theta}$, however, has the additional equilibrium frequencies $\vec{\Omega}(\mathbf{J})$. We remind the reader that H_0 is independent of $\vec{\theta}$.

It can be easily recognized that the first two terms on the right-hand side of Eqs. (11) and (12) reflect the effect of the perturbed electric field (both static and inductive), while the last term is due to the Lorentz force. The effects of the unperturbed fields is contained in the equilibrium quantities $(\partial\mathbf{J}/\partial p_{0i})$, etc..

A very important feature of Eqs. (11) and (12) is the linear dependence of $\dot{\mathbf{J}}$ and $\dot{\vec{\theta}}$ on the perturbed fields Φ and $\mathbf{A}$. Notice that this would not be the case if the coefficients of transformation $(\partial \mathbf{J}/\partial p_{0j})$ etc., depended upon the perturbed fields. This feature assures the validity of the superposition rule.

The rest of this paper is devoted to the application of this method to plasma kinetic theory.

III. Vlasov Equation in Action-Angle Variables

In this section, we set up the Vlasov equation in action-angle variables. The treatment of this section is quite general with the added feature that actions $\mathbf{J}$ are the invariants of the equilibrium system. In fact, the invariance (or adiabatic invariance) is the only reason to prefer $(\mathbf{J}, \vec{\theta})$ to any other systems of coordinates.

To put things in perspective, we deal with a guiding-center plasma in an axisymmetric toroidal configuration. Kaufmann has very elegantly solved the equilibrium problem for this system. We request the reader to consult Ref. 3 for details. The essential results are that the guiding-center motion can be described in terms of three adiabatic invariants: $M = (mc^2/e)\mu$ proportional to the magnetic moment μ; $P\varphi$, the canonical angular momentum (this action is an absolute invariant for an axisymmetric system); and J_p which is proportional to the flux enclosed by a drift surface. These actions have associated conjugate angles θ_g, φ and θ respectively.

Throughout this paper, we shall use the condensed notation

$$\mathbf{J} = (M, P_\varphi, J_p)$$

and

$$\vec{\theta} = (\theta_g,\ \varphi,\ \theta) \tag{13}$$

to denote the action-angle variables. It is also important to define the triad of frequencies

$$\vec{\Omega} = \frac{\partial H_0}{\partial \mathbf{J}} \equiv (\omega_g, \omega_\varphi, \omega_\theta) \tag{14}$$

where $H_0 = mv^2/2 + e\Phi_0$ is the equilibrium Hamiltonian. Notice that for a toroidal plasma, $\theta_g(\omega_g)$ is like a gyro angle (average gyro-frequency) $\varphi(\omega_\varphi)$ is a toroidal-like angle (average toroidal frequency), and $\theta(\omega_\theta)$ is a poloidal-like angle (average poloidal frequency). The generating function G which mediates the transformation from $\mathbf{p}_0$, $\mathbf{q}_0$ to $(\mathbf{J}, \vec{\theta})$ is given and discussed in Kaufmann's paper. Having briefly delineated the solution of a typical equilibrium problem, we proceed to deal with the Vlasov Equation.

In the new variable system, the Vlasov equation becomes

$$\frac{dF}{dt}(\mathbf{J}, \vec{\theta}, t) = \frac{\partial F}{\partial t} + \dot{\vec{\theta}} \cdot \frac{\partial F}{\partial \vec{\theta}} + \dot{\mathbf{J}} \cdot \frac{\partial F}{\partial \mathbf{J}} = 0, \tag{15}$$

while the Hamiltonian takes the form

$$H = H_0(\mathbf{J}) + e\Phi(\mathbf{J}, \vec{\theta}, t). \tag{16}$$

The distribution function is to be decomposed into its equilibrium part f_0, and the fluctuating part f

$$F = f_0 + f(\mathbf{J}, \vec{\theta}, t). \tag{17}$$

Clearly, the equilibrium part f_0 satisfies the equation

$$\vec{\Omega} \cdot \frac{\partial f_0}{\partial \vec{\theta}} = 0 \tag{18}$$

implying the simple solution

$$f_0 = f_0(\mathbf{J}) \tag{19}$$

which is clearly a result of our proper choice of the coordinate system. The equation governing the fluctuating or the perturbed distribution function f is obtained from Eqs. (11)-(19),

$$\frac{\partial f}{\partial t} + \vec{\Omega} \cdot \frac{\partial f}{\partial \mathbf{J}} + \dot{\mathbf{J}} \cdot \frac{\partial f_0}{\partial \mathbf{J}} = -\left[(\delta\dot{\vec{\theta}}) \cdot \frac{\partial f}{\partial \vec{\theta}} + \dot{\mathbf{J}} \cdot \frac{\partial f}{\partial \mathbf{J}}\right] \tag{20}$$

where we have not explicitly substituted expressions for $\dot{\mathbf{J}}$ and $(\delta\dot{\vec{\theta}})$. We shall continue developing the formal theory implicitly, and use detailed forms only where we deal with particular cases. In Eq. (20), all the terms on the left-hand side (right-hand side) are linear (nonlinear) in the perturbed quantities f, Φ and $\mathbf{A}$.

Now we exploit another important characteristic of our coordinate system, the cyclic nature of the $\vec{\theta}$'s, to expand all the perturbed quantities as Fourier series in $\vec{\theta}$. A typical perturbed quantity g is decomposed as

$$g(\mathbf{J}, \vec{\theta}, t) = \sum_{\vec{\ell}, \omega} g_{\vec{\ell}, \omega}(\mathbf{J}) \exp(-i\omega t + i\vec{\ell} \cdot \vec{\theta}) \tag{21a}$$

where the Fourier transform $g_{\vec{\ell},\omega}(\mathbf{J})$ is given by

$$g_{\vec{\ell},\omega}(\mathbf{J}) = \frac{1}{(2\pi)^4}\int d\vec{\theta}\, dt \;\; g(\mathbf{J},\vec{\theta},t)\exp(+i\omega t - i\vec{\ell}\cdot\vec{\theta}), \tag{21b}$$

and $\vec{\ell}$ is a triad of integers labelling the Fourier harmonics. Making use of Eqs. (20)-(22), we obtain $\left[g_{\vec{\ell},\omega}(\mathbf{J}) \equiv g_{\vec{\ell}}\right.$, the index ω, and the argument $\mathbf{J}$ will be generally suppressed in the rest of the paper$\left.\right]$,

$$-i(\omega - \vec{\ell}\cdot\vec{\Omega})f_{\vec{\ell}} + \dot{\mathbf{J}}_{\vec{\ell}}\cdot\frac{\partial f_0}{\partial \mathbf{J}} = -N_{\vec{\ell}} \tag{22}$$

where the nonlinear term $N_{\vec{\ell}}$ is the convolution

$$N_{\vec{\ell}} = \sum_{\vec{\ell}'}\left[i(\vec{\ell}-\vec{\ell}')\cdot(\delta\dot{\vec{\theta}})_{\vec{\ell}'}\;\; f_{\vec{\ell}-\vec{\ell}'} + \dot{\mathbf{J}}_{\vec{\ell}'}\cdot\frac{\partial}{\partial \mathbf{J}}f_{\vec{\ell}-\vec{\ell}'}\right]. \tag{23}$$

Notice that both $\dot{\mathbf{J}}_{\vec{\ell}}$ and $(\delta\dot{\vec{\theta}})_{\vec{\ell}}$ are extremely complicated terms obtained by taking the Fourier transforms of Eqs. (11) and (12), and are given by

$$\dot{\mathbf{J}}_{\vec{\ell}} = -ie\left[\vec{\ell}\Phi_{\vec{\ell}} + \sum_{\vec{\ell}'}\mathbf{K}^j(\vec{\ell}',\vec{\ell}-\vec{\ell}')A^j_{\vec{\ell}'}\right], \tag{24}$$

$$(\delta\dot{\vec{\theta}})_{\vec{\ell}} = -ie\left[-i\frac{\partial}{\partial \mathbf{J}}\Phi_{\vec{\ell}} + \sum_{\vec{\ell}'}\mathbf{L}^j(\vec{\ell}',\vec{\ell}-\vec{\ell}')A^j_{\vec{\ell}'}\right], \tag{25}$$

with the vector operators $\mathbf{K}^j$ and $\mathbf{L}^j$ defined by

$$\mathbf{K}^j = \frac{1}{c}\left[-\omega\mathbf{T}^j_{\vec{\ell}-\vec{\ell}'} + \vec{\Omega}\times\left(\mathbf{T}^j_{\vec{\ell}-\vec{\ell}'}\times\vec{\ell}'\right)\right], \tag{26}$$

$$\mathbf{L}^j = \frac{1}{c}\left[-\omega\mathbf{S}^j_{\vec{\ell}-\vec{\ell}'} + (\vec{\ell}'\cdot\vec{\Omega})S^j_{\vec{\ell}-\vec{\ell}'} - i\left(\vec{\Omega}\cdot\mathbf{T}^j_{\vec{\ell}-\vec{\ell}'}\right)\frac{\partial}{\partial \mathbf{J}}\right] \tag{27}$$

where $\mathbf{T}^j_{\vec{\ell}}$ and $\mathbf{S}^j_{\vec{\ell}}$ are the Fourier transforms of the known quantities. $\partial\mathbf{J}/\partial p_{0j}$ and $\partial\vec{\theta}/\partial p_{0j}$ respectively. Equation (22) is the main result of this section, and is an expression of the exact, nonlinear, Fourier transformed Vlasov equation in the presence of fully electromagnetic perturbations in a general magnetic field geometry. The formally simple structure of this equation shows the power and elegance of the action-angle variable approach which allows the Vlasov equation in a complicated geometry to look exactly like the Vlasov equation describing a field-free plasma. This formal equivalence follows from the fact that in the invariant action-angle space, the particle trajectories are always straight lines (as in the field-free case). The result is a unified formalism to deal with a whole class of plasma problems; we do not have to begin with a different looking Vlasov equation every time we change the equilibrium geometry. All formal manipulations can be carried out on Eq. (22), and depending on what $(\mathbf{J}, \vec{\theta})$ we use, it could, for example, describe the infinite homogeneous field-free plasma, or the response of trapped particles in a tokamak. The translation $(\mathbf{J}, \vec{\theta}) \leftrightarrow (\mathbf{p}_0, \mathbf{q}_0)$ is, of course, given by the solution of the equilibrium problem.

IV. Linearized Vlasov Equation

Setting $N_{\vec{\ell}} = 0$ in Eq. (22) leads to the linearized Vlasov equation

$$-i(\omega - \vec{\ell} \cdot \vec{\Omega}) f_{\vec{\ell}} + \dot{\mathbf{J}} \cdot \frac{\partial f_0}{\partial \mathbf{J}} = 0 \tag{27}$$

which is readily solved to obtain

$$f_{\vec{\ell}} = \frac{\dot{\mathbf{J}}_{\vec{\ell}} \cdot \frac{\partial f_0}{\partial \mathbf{J}}}{i(\omega - \vec{\ell} \cdot \vec{\Omega})}. \tag{28}$$

Notice that our clever choice of the coordinate has resulted in making the linear solution $f_{\vec{\ell}}$ independent of $\delta\dot{\vec{\theta}}_{\vec{\ell}}$.

A. Free Field Limit

One of the main reasons for this paper was to correct the errors in the earlier formulation of action-angle variable theories. Let us see how our new linear results compare with known standard results. The simplest check is to obtain the field-free ($\mathbf{A}_0 = 0$, $\Phi_0 = 0$) limit of Eq. (28). This is easily accomplished by the transformations $\mathbf{J} \rightarrow \mathbf{p}$, the ordinary conserved momentum, $\vec{\theta} \rightarrow \mathbf{x}$, $\vec{\ell} \rightarrow \mathbf{k}(\partial/\partial\vec{\theta} \rightarrow \partial/\partial\mathbf{x})$, and $\vec{\Omega} \rightarrow \mathbf{v}$. The results are

$$\mathbf{T}^j_{\vec{\ell}} \rightarrow \hat{e}_i \delta_{i,j} \delta_{\mathbf{k},0} \tag{29}$$

where $\hat{e}_i$ is the unit vector along the direction i, the operator,

$$\mathbf{K}^j \rightarrow \frac{1}{c}\left[-\omega \hat{e}_i \delta_{ij} + \mathbf{v} \times (\hat{e}_i \times \mathbf{k})\delta_{ij}\right]\delta_{\mathbf{k},\mathbf{k}'}, \tag{30}$$

$$\dot{\mathbf{J}}_{\vec{\ell}} \rightarrow -ie\left\{ \mathbf{k}\Phi_{\mathbf{k}} + \frac{1}{c}\left[-\omega \mathbf{A}_{\mathbf{k}} - \mathbf{v} \times (\mathbf{k} \times \mathbf{A}_{\mathbf{k}})\right]\right\}, \tag{31}$$

and finally ($\mathbf{p} = m\mathbf{v}$, $E_{\mathbf{k}}$ and $B_{\mathbf{k}}$ are respectively the perturbed electric and magnetic fields)

$$f_{\mathbf{k}} = \frac{e}{m}\frac{-i\left[\mathbf{k}\Phi_{\mathbf{k}} - \frac{\omega}{c}\mathbf{A}_{\mathbf{k}} - \frac{1}{c}\mathbf{v} \times (\mathbf{k} \times \mathbf{A}_{\mathbf{k}})\right]}{i(\omega - \mathbf{k}\cdot\mathbf{v})} \cdot \frac{\partial f_0}{\partial \mathbf{v}} \tag{32a}$$

$$\equiv \frac{e}{m}\frac{\left(\mathbf{E}_{\mathbf{k}} + \frac{1}{c}\mathbf{v} \times \mathbf{B}_{\mathbf{k}}\right)}{i(\omega - \mathbf{k}\cdot\mathbf{v})} \cdot \frac{\partial f_0}{\partial \mathbf{v}} \tag{32b}$$

which is precisely the required result.[6] By using Faraday's law $\mathbf{B}_k = (c/\omega)(\mathbf{k} \times \mathbf{E}_{\mathbf{k}})$, we can rewrite Eq. (32b) purely in terms of the electric field

$$f_{\mathbf{k}} = \frac{e}{im\omega}\frac{\left[\omega \mathbf{E}_{\mathbf{k}} + \mathbf{v} \times (\mathbf{k} \times \mathbf{E}_{\mathbf{k}})\right]}{\omega - \mathbf{k}\cdot\mathbf{v}} \cdot \frac{\partial f_0}{\partial \mathbf{v}}$$

$$\equiv \frac{e}{im\omega}\left[\mathbf{E}_{\mathbf{k}} + \frac{(\mathbf{v}\cdot\mathbf{E}_{\mathbf{k}})\mathbf{k}}{\omega - \mathbf{k}\cdot\mathbf{v}}\right] \cdot \frac{\partial f_0}{\partial \mathbf{v}}. \tag{32c}$$

We have derived this form of $f_{\mathbf{k}}$ to compare it with the linear result, Eq. (26),

$$\delta f_{\vec{\ell}}(\mathbf{J},\omega) = \frac{\delta H_{\vec{\ell}}(\mathbf{J},\omega)\vec{\ell}\cdot\frac{\partial f_0}{\partial \mathbf{J}}}{\omega - \vec{\ell}\cdot\vec{\Omega}}$$

of Ref. 3. Using the standard prescription to go to the field-free case, we obtain

$$\delta f_{\mathbf{k}} = \frac{e}{im\omega}\frac{(\mathbf{v}\cdot\mathbf{E}_k)\left(\mathbf{k}\cdot\frac{\partial f_0}{\partial \mathbf{v}}\right)}{(\omega - \mathbf{k}\cdot\mathbf{v})} \tag{33}$$

which does not agree with Eq. (32c); Eq. (33) lacks the first term in the square brackets in Eq. (32). The mistake was made in Eq. (25) of Ref. 3 where $\delta\dot{\mathbf{J}} = -\partial H/\partial\vec{\theta} = -\partial(H_0 + \delta H)/\partial\vec{\theta} = -\partial\delta H/\partial\theta$ was used. As pointed out earlier, $(\mathbf{J}, \vec{\theta})$ are not canonical variables for the total Hamiltonian H, and therefore $\delta\dot{\mathbf{J}}$ is not equal to $-\partial\delta H/\partial\theta$. By using the correctly derived expressions for $\dot{\mathbf{J}}$ and $\dot{\theta}$ [Eqs. (11) and (12)], we do indeed reproduce the standard results. After this demonstration, we shall no longer belabor the point that the earlier treatments using action-angle variables were incorrect.

B. Low-Frequency Trapped Particle Response in Tokamaks

We now show how Eq. (28) can be readily used to obtain low-frequency gyro-averaged, bounce-averaged, response for deeply trapped particles in a tokamak. For simplicity, we derive only the electrostatic limit for which [see Eq. (24)]

$$\dot{\mathbf{J}}_{\vec{\ell}} = -ie\vec{\ell}\Phi_{\vec{\ell}} \tag{34}$$

leading to

$$f_{\vec{\ell}} = \frac{-e\Phi_{\vec{\ell}}\,\vec{\ell}\cdot\frac{\partial f_0}{\partial \mathbf{J}}}{\omega - \vec{\ell}\cdot\vec{\Omega}}$$

where $\vec{\ell} = (\ell_g, \ell_\varphi, \ell_\theta)$ with ℓ_g, ℓ_φ, and ℓ_θ the gyro, toroidal and bounce harmonic numbers respectively. The gyro and bounce averaged response is obtained by simply setting $\ell_g = 0 = \ell_\theta$,

$$f_\ell = -e\Phi_\ell \frac{\ell\frac{\partial f_0}{\partial p_\varphi}}{\omega - \ell\Omega_\varphi} \tag{36}$$

where $\ell_\varphi \equiv \ell$. We must remind the reader that these harmonic numbers are in action-angle space and not in real space. Since Maxwell's equations are simple in real space, we will need to convert f_ℓ's to real space harmonics before we can use them to calculate perturbed current and density. This translation mechanism is adequately discussed in Ref. 4.

Before ending this section we would like to point out that tremendous calculational simplification occurs for a very important special class of distribution functions $f_0(\mathbf{J})$ which depend on $\mathbf{J}$ only through the Hamiltonian H_0; the Maxwellian distribution belongs to this class. For $f_0(\mathbf{J}) = f_0[H_0(\mathbf{J})]$,

$$\frac{\partial f_0}{\partial \mathbf{J}} = \frac{\partial f_0}{\partial H_0}\frac{\partial H_0}{\partial \mathbf{J}} = \vec{\Omega}\frac{\partial f_0}{\partial H_0} \tag{37}$$

which results in the simplification

$$\dot{\mathbf{J}}_{\vec{\ell}}\cdot\frac{\partial f_0}{\partial \mathbf{J}} = -ie\vec{\Omega}\cdot\left[\vec{\ell}\Phi_{\vec{\ell}} - \frac{\omega}{c}\sum_{\vec{\ell}'}\mathbf{T}^j_{\vec{\ell}-\vec{\ell}'}A^j_{\vec{\ell}'}\right]\frac{\partial f_0}{\partial H_0}. \tag{38}$$

Further simplification is possible only when one knows $\mathbf{T}^j_{\vec{\ell}-\vec{\ell}'}$ which depends upon the particular problem being investigated.

V. Quasilinear Theory

A. General Theory

A very important application of the action-angle formalism is the development of the quasilinear transport theory for complicated geometries. In fact, Kaufmann's original paper was precisely intended for this purpose.

The principal object of the quasilinear transport theory is to obtain an equation for the slow evolution of the equilibrium (or averaged) distribution function f_0 in response to the perturbing electromagnetic fields. The calculation is carried out in two distinct steps.

1) The equilibrium distribution function f_0 is given a time dependence $f_0 = f_0(\mathbf{J}, t)$ so that Eq. (20) becomes

$$\frac{\partial f_0}{\partial t} + \frac{\partial f}{\partial t} + \vec{\Omega} \cdot \frac{\partial f}{\partial \vec{\theta}} + \dot{\mathbf{J}} \cdot \frac{\partial f_0}{\partial \mathbf{J}} = -\left[\dot{\mathbf{J}} \cdot \frac{\partial f}{\partial \mathbf{J}} + (\dot{\delta \vec{\theta}}) \cdot \frac{\partial f}{\partial \vec{\theta}} \right]. \tag{39}$$

Averaging over the angles $\vec{\theta}$ yields ($\langle f \rangle \equiv f_0$)

$$\begin{aligned} \frac{\partial f_0}{\partial t} &= -\langle \dot{\mathbf{J}} \rangle \cdot \frac{\partial f_0}{\partial \mathbf{J}} - \langle \dot{\mathbf{J}} \cdot \frac{\partial f}{\partial \mathbf{J}} \rangle - \langle (\dot{\delta \vec{\theta}}) \cdot \frac{\partial f}{\partial \vec{\theta}} \rangle \\ &= -\langle \dot{\mathbf{J}} \rangle \cdot \frac{\partial f_0}{\partial \mathbf{J}} - \langle \dot{\mathbf{J}} \cdot \frac{\partial f}{\partial \mathbf{J}} \rangle + \langle f \frac{\partial}{\partial \vec{\theta}} \cdot (\dot{\delta \vec{\theta}} \rangle \end{aligned} \tag{40}$$

because the fast $\vec{\theta}$ variation gives zero average for the terms linear in the fluctuating part f. Notice that although $\dot{\mathbf{J}}$ is linear in the fluctuating fields Φ and $\mathbf{A}$, it has been retained in the equation because the coefficients $\partial \mathbf{J}/\partial p_{0j}$ can also have $\vec{\theta}$ dependence which could give nonzero average $\dot{\mathbf{J}}$. However, the factors $(\partial \mathbf{J}/\partial p_{0j})$ can have only slow equilibrium dependence in a quiescent plasma, and the fluctuations, in general, are characterized by large $\vec{\ell}$, it is quite safe to set $\langle \dot{\mathbf{J}} \rangle = 0$ for most practical problems. In the context of the preceding discussion, we can write the general equation

$$\frac{\partial f_0}{\partial t} = -\langle \dot{\mathbf{J}} \cdot \frac{\partial f}{\partial \mathbf{J}} \rangle + \langle f \frac{\partial}{\partial \vec{\theta}} \cdot (\dot{\delta \vec{\theta}}) \rangle, \tag{41}$$

where f is still the exact fluctuating distribution function, i.e., the solution of Eq. (20).

2) Approximating the total fluctuating f by its linear value f^L is the essential assumption of the quasilinear theory, and leads to

$$\frac{\partial f_0}{\partial t} = -\langle \dot{\mathbf{J}} \cdot \frac{\partial f^L}{\partial \mathbf{J}} \rangle + \langle f^L \frac{\partial}{\partial \vec{\theta}} \cdot (\dot{\delta \vec{\theta}}) \rangle. \tag{42}$$

Using the definition

$$f^L = \sum_{\vec{\ell}} f_{\vec{\ell}} \quad e^{i \vec{\ell} \cdot \vec{\theta}},$$

the definitions of $\dot{\mathbf{J}}_{-\vec{\ell}}$ and $(\delta\dot{\vec{\theta}})_{-\vec{\ell}}$, and Eq. (28), we can write Eq. (42) in the equivalent form

$$\frac{\partial f_0}{\partial t} = \frac{\partial}{\partial \mathbf{J}} \cdot \overleftrightarrow{D} \cdot \frac{\partial f_0}{\partial \mathbf{J}} + \vec{\Lambda} \cdot \frac{\partial f_0}{\partial \mathbf{J}} \tag{43}$$

where

$$\overleftrightarrow{D} = i \sum_{\vec{\ell}} \frac{\dot{\mathbf{J}}_{\vec{\ell}} \dot{\mathbf{J}}_{-\vec{\ell}}}{\omega - \vec{\ell} \cdot \vec{\Omega}} \tag{44a}$$

is the coefficient of quasilinear diffusion, and

$$\vec{\Lambda} = -\sum_{\vec{\ell}} \frac{\left[\left(\frac{\partial}{\partial \mathbf{J}} \cdot \dot{\mathbf{J}}_{-\vec{\ell}}\right) - i\vec{\ell} \cdot (\delta\dot{\vec{\theta}})_{-\vec{\ell}}\right] \dot{\mathbf{J}}_{\vec{\ell}}}{\omega - \vec{\ell} \cdot \vec{\Omega}} \tag{44b}$$

denotes convection in the action-angle space. Notice that the origin of the convective term $\vec{\Lambda}$ is entirely due to the non-canonical nature of $\mathbf{J}$ and $\vec{\theta}$. For canonical variables

$$\frac{\partial}{\partial \mathbf{J}_c} \cdot \dot{\mathbf{J}}_c + \frac{\partial}{\vec{\theta}_c} \cdot \delta\dot{\vec{\theta}}_c = \frac{\partial}{\partial \mathbf{J}_c} \left(-\frac{\partial H}{\partial \vec{\theta}_c}\right) + \frac{\partial}{\partial \vec{\theta}_c} \cdot \left(\frac{\partial H}{\partial \mathbf{J}_c}\right) \equiv 0, \tag{45}$$

or equivalently

$$\frac{\partial}{\partial \mathbf{J}_c} \cdot (\dot{\mathbf{J}}_c)_{-\vec{\ell}} - i\vec{\ell} \cdot (\delta\dot{\vec{\theta}}_c)_{-\vec{\ell}} = 0 \tag{46}$$

which is precisely the term that determines $\vec{\Lambda}$.

If the perturbation were pure electrostatic ($\mathbf{A} = 0$), the action-angle variables will be canonical, $\vec{\Lambda}$ (electrostatic)$\equiv 0$, and

$$\frac{\partial f_0}{\partial t} = \frac{\partial}{\partial \mathbf{J}} \cdot \overleftrightarrow{D}_{es} \cdot \frac{\partial}{\partial \mathbf{J}} f_0 \tag{47}$$

where [use of Eq. (24) for $\dot{\mathbf{J}}_{es}$]

$$\overleftrightarrow{D}_{es} = ie^2 \sum_{\vec{\ell}} \vec{\ell}\, \vec{\ell} \frac{\Phi_{\vec{\ell}}\, \Phi_{-\vec{\ell}}}{\omega - \vec{\ell} \cdot \vec{\Omega}}. \tag{48}$$

Since it is the presence of the magnetic perturbation $\mathbf{A}$ that destroys the canonical nature of $\mathbf{J}$ and $\vec{\theta}$, the convection term $\vec{\Lambda}$ will be a functional of $\mathbf{A}$ alone.

From the analysis in the preceding section, it is quite clear that the quasilinear transport theory becomes considerably more complicated when electromagnetic perturbations are present. The change is not only quantitative, but qualitative also in the sense that the evolution of f_0 is no more purely diffusive in the action-angle space. The relative importance of diffusion, and convection terms would clearly depend upon the details of the situation. Equations (43) and (44), the main results of this section, provide the starting equations to develop, for example, a theory of radial transport in tokamaks caused by electromagnetic fluctuations.

Nonlinear Theory

Equation (22) is the fully nonlinear Fourier transformed equation for the fluctuating distribution function. It can serve as a basis for developing a general renormalized turbulence theory for a Vlasov plasma in arbitrary geometry.

The standard renormalization procedure[7] consists in breaking up the nonlinearity $N_{\vec{\ell}}$ into three terms; the first is proportional to (coherent with) $f_{\vec{\ell}}$, the second is proportional to the perturbed electromagnetic fields, and the third is called the fully incoherent term. The effects of the electromagnetic fields are contained in $(\delta\dot{\vec{\theta}})$ and $\dot{\mathbf{J}}$ in our formulation. Thus we seek the decomposition of the nonlinear term as

$$N_{\vec{\ell}} = D_{\vec{\ell}} f_{\vec{\ell}} + \mathbf{C}_{1\vec{\ell}} \cdot (\delta\dot{\vec{\theta}})_{\vec{\ell}} + \mathbf{C}_{2\vec{\ell}} \cdot \dot{\mathbf{J}}_{\ell} + \tilde{N}_{\vec{\ell}} \tag{49}$$

where $\tilde{N}_{\vec{\ell}}$ is the remaining incoherent part of $N_{\vec{\ell}}$ after the coherent parts have been subtracted. Clearly, the term $D_{\vec{\ell}} f_{\vec{\ell}}$ will renormalize the linear propagator on the left-hand side of Eq. (22).

Detailed expressions for $D_{\vec{\ell}}$, $C_{1\vec{\ell}}$, and $\mathbf{C}_{2\vec{\ell}}$ are obtained by the following simple procedure. We add $D_{\vec{\ell}} f_{\vec{\ell}}$ to both sides of Eq. (22) to obtain the formal solution

$$f_{\vec{\ell}} = g_{\vec{\ell}} \left[+\dot{\mathbf{J}}_{\vec{\ell}} \cdot \frac{\partial f_0}{\partial \mathbf{J}} + N_{\vec{\ell}} - D_{\vec{\ell}} f_{\vec{\ell}} \right] \tag{50}$$

where

$$g_{\vec{\ell}} = \frac{1}{i(\omega - \vec{\ell} \cdot \vec{\Omega} + i D_{\vec{\ell}})} \tag{51}$$

is the renormalized propagator. Substituting Eq. (50) into the nonlinear term Eq. (23) leads to

$$N_{\vec{\ell}} = \sum_{\vec{\ell}'} \left[i(\vec{\ell} - \vec{\ell}') \cdot (\delta\dot{\vec{\theta}})_{\vec{\ell}'} + \dot{\mathbf{J}}_p \cdot \frac{\partial}{\partial \mathbf{J}} \right] g_{\vec{\ell}-\vec{\ell}'} \left[N_{\vec{\ell}-\vec{\ell}'} + \dot{\mathbf{J}}_{\vec{\ell}-\vec{\ell}'} \cdot \frac{\partial f_0}{\partial \mathbf{J}} - D_{\vec{\ell}-\vec{\ell}'} f_{\vec{\ell}-\vec{\ell}'} \right] \tag{52}$$

where $\frac{\partial}{\partial \mathbf{J}}$ operates on all quantities on its right. It is straightforward to see that the terms coherent with $f_{\vec{\ell}}$ and $\mathbf{A}_{\vec{\ell}}$ can come only from $N_{\vec{\ell}-\vec{\ell}'}$, because $\vec{\ell}' = 0$ is not permitted. Denoting this part by $\bar{N}_{\vec{\ell}}$, and substituting Eq. (23) into Eq. (52), we obtain

$$\bar{N}_{\vec{\ell}} = \sum_{\vec{\ell}',\vec{\ell}''} \left[i(\vec{\ell} - \vec{\ell}') \cdot (\delta\dot{\vec{\theta}})_{\vec{\ell}'} + \dot{\mathbf{J}}_{\vec{\ell}'} \cdot \frac{\partial}{\partial \mathbf{J}} \right] g_{\vec{\ell}-\vec{\ell}'}$$
$$\left[i(\vec{\ell} - \vec{\ell}' - \vec{\ell}'') \cdot (\delta\dot{\vec{\theta}})_{\vec{\ell}''} f_{\vec{\ell}-\vec{\ell}'-\vec{\ell}''} + \dot{\mathbf{J}}_{\vec{\ell}''} \cdot \frac{\partial}{\partial \mathbf{J}} f_{\vec{\ell}-\vec{\ell}'-\vec{\ell}''} \right] \tag{53}$$

from which we can readily extract

$$D_{\vec{\ell}} f_{\vec{\ell}} = \sum_{\vec{\ell}'} \left[i(\vec{\ell} - \vec{\ell}') \cdot (\delta\dot{\vec{\theta}})_{\vec{\ell}'} + \dot{\mathbf{J}}_{\vec{\ell}'} \cdot \frac{\partial}{\partial \mathbf{J}} \right] g_{\vec{\ell}-\vec{\ell}'}$$
$$\left[i\vec{\ell} \cdot (\delta\dot{\vec{\theta}})_{-\vec{\ell}'} + \dot{\mathbf{J}}_{-\vec{\ell}'} \cdot \frac{\partial}{\partial \mathbf{J}} \right] f_{\vec{\ell}}\,, \tag{54}$$

$$\mathbf{C}_{1\vec{\ell}} \cdot (\delta\dot{\vec{\theta}})_{\vec{\ell}} = -i \sum_{\vec{\ell}'} \left[i(\vec{\ell} - \vec{\ell}') \cdot (\delta\dot{\vec{\theta}})_{\vec{\ell}'} + \dot{\mathbf{J}}_{\vec{\ell}'} \cdot \frac{\partial}{\partial \mathbf{J}} \right] \left(g_{\vec{\ell}-\vec{\ell}'} f_{-\vec{\ell}'} \vec{\ell}' \cdot (\delta\dot{\vec{\theta}})_{\vec{\ell}} \right) \tag{55}$$

and

$$\mathbf{C}_{2\vec{\ell}} \cdot \dot{\mathbf{J}}_{\vec{\ell}} = \sum_{\vec{\ell}'} \left[i(\vec{\ell} - \vec{\ell}') \cdot (\delta\dot{\vec{\theta}})_{\vec{\ell}'} + \dot{\mathbf{J}}_{\vec{\ell}'} \cdot \frac{\partial}{\partial \mathbf{J}} \right] \left(g_{\vec{\ell}-\vec{\ell}'} \frac{\partial f_{-\vec{\ell}'}}{\partial \mathbf{J}} \cdot \dot{\mathbf{J}}_{\vec{\ell}} \right). \tag{56}$$

By making use of expression for $(\delta\dot{\vec{\theta}})_{\vec{\ell}}$ and $\dot{\mathbf{J}}_{\vec{\ell}}$, one can easily express $\mathbf{C}_{1\vec{\ell}} \cdot (\delta\dot{\vec{\theta}})_{\vec{\ell}} + \mathbf{C}_{2\vec{\ell}} \cdot (\dot{\mathbf{J}})_{\vec{\ell}}$ as $a_{\vec{\ell}}\, \Phi_{\vec{\ell}} + \mathbf{b}_{\vec{\ell}} \cdot \mathbf{A}_{\vec{\ell}}$. Renormalized Eq. (22) reads

$$f_{\vec{\ell}} = g_{\vec{\ell}} \left\{ \left[\left(\frac{\partial f_0}{\partial \mathbf{J}}\right) + \mathbf{C}_{2\vec{\ell}} \right] \cdot \dot{\mathbf{J}}_{\ell} + \mathbf{C}_{1\ell} \cdot (\delta\dot{\vec{\theta}})_{\vec{\ell}} + \tilde{N}_{\ell}. \right\} \tag{57}$$

In the coherent renormalization theories, $\bar{N}_{\vec{\ell}} = 0$, and Eq. (57) serves as a formal nonlinear solution of the Vlasov equation.

VII. Discussion and Conclusions

We have shown that a mechanical application of the canonical perturbation theory can lead to erroneous results when magnetic field perturbations ($\mathbf{A}$) are present. The principal reason is that for $\mathbf{A} \neq 0$, the invariant actions and their associated angles (derived for the unperturbed system) cease to be canonical variables of the total Hamiltonian H, i.e., $\dot{\mathbf{J}} \neq -\partial H/\partial\vec{\theta}$, and $\dot{\vec{\theta}} \neq \partial H/\partial\mathbf{J}$. We have also derived approximate expression for the orbits $\mathbf{J}$ and $\vec{\theta}$ for a changed particle moving in general electromagnetic fields (perturbed and unperturbed) to develop a modified perturbation theory which is used to formulate the kinetic theory of plasmas in the action-angle variables. Simple tests of the correctness of our theory are provided by comparing our results with standard known results. The general formalism is used to derive basic kinetic equation for the study of linear, quasilinear, and renormalized nonlinear theories.

As stated earlier, the great advantage of the action-angle variable approach is to unify the treatment of such diverse plasma problems as the determination of fluctuating distribution function for a field free plasma, and for trapped particles in a tokamak. The entire formal structure is the same, because in the invariant action-angle space, the particle trajectories are always straight lines. We believe that our Eqs. (28), (43)-(44), and (54)-(57) can be sued as starting points for a wide variety of kinetic plasma problems.

The application of this formalism to specific problems of interest will be the subject of a later paper; this paper is intended to be a general delineation of the theory.

Acknowledgment

This work was supported by the U.S. Department of Energy, contract no. DE-FG05-80ET-53088.

Appendix A

We first derive a general perturbation formalism for the evolution of equilibrium invariants of a charged particle. The Hamiltonian H of the particle is given by Eqs. (6) and (8), and the variables $(\mathbf{p}_0, \mathbf{q}_0)$ and $(\vec{\alpha}_0, \vec{\beta}_0)$ are described by Eqs. (4) and (5) and (7). The equation of motion of α_{0i} is [Eq. 10a]

$$\dot{\alpha}_{0i} = \left.\frac{\partial \alpha_{0i}}{\partial t}\right|_{(\mathbf{p},\mathbf{q})} + [\alpha_{0i}, H]_{(\mathbf{p},\mathbf{q})}. \qquad (A-1)$$

From Eq. (7), and $q = q_0$, we derive the transformation coefficients,

$$\frac{\partial p_{0j}}{\partial p_i} = \delta_{ij} \quad , \quad \frac{\partial q_{0j}}{\partial p_i} = 0$$

$$\frac{\partial p_{0j}}{\partial q_{ii}} = -\frac{e}{c}\frac{\partial A_j}{\partial q_i} \quad , \quad \frac{\partial q_{0j}}{\partial q_i} = \delta_{ij} \qquad (A-2)$$

which are used to obtain

$$\left.\frac{\partial \alpha_{0i}}{\partial t}\right|_{(\mathbf{p},\mathbf{q})} = -\frac{e}{c}\frac{\partial \alpha_{0i}}{\partial p_{0j}}\frac{\partial A^j}{\partial t}, \qquad (A-3)$$

$$[\alpha_{0i,H}]_{(\mathbf{p},\mathbf{q})} = [\alpha_{0i,H}]_{\mathbf{p}_0,\mathbf{q}_0} + \langle \alpha_{0i,H} \rangle_{\mathbf{p}_0,\mathbf{q}_0} \qquad (A-4)$$

where the bracket $\langle \quad \rangle$ $\mathbf{p}_0$, $\mathbf{q}_0$ is defined by

$$\langle f, g \rangle_{\mathbf{p}_0,\mathbf{q}_0} \equiv -\frac{e}{c}\left\{ [A^j, g]_{\mathbf{p}_0,\mathbf{q}_0}\frac{\partial f}{\partial p_{0j}} - [A^j, f]_{\mathbf{p}_0,\mathbf{q}_0}\frac{\partial g}{\partial p_{0j}} \right\}. \qquad (A-5)$$

Since both the brackets $[\quad]$ and $\langle \quad \rangle$ are linear in their arguments, Eq. (A4) can be split into four terms

$$[\alpha_{0i}, H]_{(\mathbf{p},\mathbf{q})} = [\alpha_{0i}, H_0]_{\mathbf{p}_0,\mathbf{q}_0} + [\alpha_{0i,h}]_{\mathbf{p}_0,\mathbf{q}_0}$$
$$+ \langle \alpha_{0i}, H_0 \rangle_{\mathbf{p}_0,\mathbf{q}_0} + \langle \alpha_{0i}, h \rangle_{\mathbf{p}_0,\mathbf{q}_0}. \qquad (A-6)$$

Since the perturbed Hamiltonian $h \equiv e\Phi$ is independent of p_{0i}, the fourth term of Eq. (A-6) vanishes

$$\langle \alpha_{0i}, h \rangle_{\mathbf{p}_0,\mathbf{q}_0} = 0.$$

Since $(\vec{\alpha}_0, \vec{\beta}_0)$, though not canonical variables for the total Hamiltonian H, are obtained as a canonical transformation from $\mathbf{p}_0$, $\mathbf{q}_0$, the Poisson bracket remains invariant, i.e.

$$[f, g]_{\mathbf{p}_0,\mathbf{q}_0} = [f, g]_{\vec{\alpha}_0,\vec{\beta}_0}$$

which immediately leads to

$$[\alpha_{0i}, H_0]_{\mathbf{p}_0,\mathbf{q}_0} = [\alpha_{0i}, H_0(\vec{\alpha}_0)]_{\vec{\alpha}_0,\vec{\beta}_0} = 0.$$

In order to apply the invariance of Poisson brackets to other terms, we need probably do a formal treatment of both the electrostatic field Φ and electromagnetic field $\mathbf{A}$. We assume that we have obtained the explicit solutions from the transformation Eq. (5),

$$\begin{aligned} \mathbf{q}_0 &= \mathbf{q}_0(\vec{\alpha}_0, \vec{\beta}_0) \\ \mathbf{p}_0 &= \mathbf{p}_0(\vec{\alpha}_0, \vec{\beta}_0). \end{aligned} \qquad (A-7)$$

Since the α_0's are motion invariants of the particle in the unperturbed situation, we use them as labels to mark the particle. The fields acting on the particle are expressed as

$$\begin{aligned} \Phi_{\vec{\alpha}0}(\vec{\beta}_0, t) &= \Phi(\mathbf{q}_0(\vec{\alpha}_0, \vec{\beta}_0), t) \\ \mathbf{A}_{\vec{\alpha}0}(\vec{\beta}, t) &= \mathbf{A}(\mathbf{p}_0(\vec{\alpha}_0, \vec{\beta}_0), t). \end{aligned} \qquad (A-8)$$

With this understanding, Eq. (A-6) can be further simplified:

$$[\alpha_{0i}, H_0]_{\mathbf{p}_0,\mathbf{q}_0} = -e\frac{\partial \Phi}{\partial \beta_{0i}}$$

$$\begin{aligned} \langle \alpha_{0i}, H_0 \rangle_{\mathbf{p}_0,\mathbf{q}_0} &= -\frac{e}{c}\left\{ \left[A^j, H_0\right]_{\alpha_0,\beta_0} \frac{\partial \alpha_{0i}}{\partial p_{0j}} - \left[A^j, \alpha_{0i}\right]_{\alpha_0,\beta_0} \frac{\partial H_0}{\partial p_{0j}} \right\} \\ &= -\frac{e}{c}\left\{ \frac{\partial A^j}{\partial \beta_{0k}} \cdot \frac{\partial H_0}{\partial \alpha_{0k}} \cdot \frac{\partial \alpha_{0i}}{\partial p_{0j}} - \frac{\partial A^j}{\partial \beta_{0i}} \cdot \frac{\partial H_0}{\partial \alpha_{0k}} \cdot \frac{\partial \alpha_{0k}}{\partial p_{0j}} \right\} \end{aligned}$$

where Einstein's rule is obeyed (and the same in the rest of the paper). Finally, we find

$$\begin{aligned} \dot{\alpha}_{0i} = -e\frac{\partial \Phi}{\partial \beta_{0i}} - \frac{e}{c}\frac{\partial \alpha_{0i}}{\partial p_{0j}} \cdot \frac{\partial A^j}{\partial t} - \frac{e}{c}\left\{ \frac{\partial \alpha_{0i}}{\partial p_{0j}} \cdot \frac{\partial H_0}{\partial \alpha_{0k}} \frac{\partial}{\partial \beta_{0k}} \right. \\ \left. - \frac{\partial \alpha_{0k}}{\partial p_{0j}} \cdot \frac{\partial H_0}{\partial \alpha_{0k}} \cdot \frac{\partial}{\partial \beta_{0i}} \right\} A^j \end{aligned} \qquad (A-9)$$

or equivalently (the subscript 0 if suppressed for α and β)

$$\dot{\vec{\alpha}} = -e\frac{\partial \Phi}{\partial \vec{\beta}} - \frac{e}{c}\frac{\partial \vec{\alpha}}{\partial p_{0j}} \frac{\partial A_j}{\partial t} - \frac{e}{c}\left[\frac{\partial \vec{\alpha}}{\partial p_{0j}} \cdot \left(\frac{\partial H_0}{\partial \vec{\alpha}} \cdot \frac{\partial}{\partial \vec{\beta}} \right) - \left(\frac{\partial \vec{\alpha}}{\partial p_{0j}} \cdot \cdot \frac{\partial H_0}{\partial \vec{\alpha}} \right) \frac{\partial}{\partial \vec{\beta}} \right] A^j.$$

Equation (A-9) is the evolution equation for an invariant under the influence of electrostatic and electromagnetic perturbed field.

Following a similar procedure, an equation for the conjugate β_{0i} is readily obtained

$$\dot{\vec{\beta}} = \frac{\partial H_0}{\partial \vec{\alpha}} - e\frac{\partial \Phi}{\partial \vec{\alpha}} - \frac{e}{c}\frac{\partial \vec{\beta}}{\partial p_{0j}} \cdot \frac{\partial A^j}{\partial t}$$
$$- \frac{e}{c}\left\{ \frac{\partial \vec{\beta}}{\partial p_{0j}} \left(\frac{\partial H_0}{\partial \vec{\alpha}} \cdot \frac{\partial}{\partial \vec{\beta}} \right) + \left(\frac{\partial \vec{\alpha}}{\partial p_{0j}} \cdot \frac{\partial H_0}{\partial \vec{\alpha}} \right) \frac{\partial}{\partial \vec{\alpha}} \right\} A^j. \qquad (A-10)$$

References

1. M. Kruskal, *Atomic Energy Commission Report*, NYO-7903.
2. A. Lenard, Ann. Phy. (N.Y.) **6**, 261-276 (1959).
3. A.N. Kaufmann, Phys. Fluids **15**, 1063 (1972).
4. D.A. Hitchcock, R.D. Hazeltine, and S.M. Mahajan, Phys. Fluids **26**, 2603 (1983); R.D. Hazeltine, S.M. Mahajan and D.A. Hitchcock, Phys. Fluids **25**, 350 (1982); S.M. Mahajan, R.D. Hazeltine and D.A. Hitchcock, Phys. Fluids **26**, 700 (1983); R.D. Hazeltine, Phys. Fluids **24**, 1164 (1981).
5. H. Goldstein, *Classical Mechanics* (Second Edition, Addison-Wesley Publishing Co., Inc., 1980) p. 499-506.
6. N.A. Krall and A.W. Trivelpiece, *Principles of Plasma Physics* (McGraw-Hill Book Co., 1973), p. 399.
7. T.H. Dupree, D.J. Tetreault, Phys. Fluids **21**, 425 (1978).