PSFC/JA-09-37

Kinetic Theory for Distribution Functions of non-Markovian Wave-Particle Interactions in Plasmas

Y. Kominis, a A.K. Ram, and K. Hizanidis a

January 2010

Plasma Science and Fusion Center, Massachusetts Institute of Technology
Cambridge, MA 02139 U.S.A.

a National Technical University of Athens
Association EURATOM-Hellenic Republic
Zografou, Athens 15773. Greece

This work was supported by the U.S. Department of Energy, This work is supported by DoE grants DE-FG02-99ER-54521 and DE-FG02-91ER-54109, and by Association EURATOM, Hellenic Republic. Reproduction, translation, publication, use and disposal, in whole or in part, by or for the United States government is permitted.

The presence of coherent electromagnetic waves and their interaction with charged particles are ubiquitous phenomena in magnetized plasmas that are encountered in space plasmas and in laboratory fusion devices. The presence of waves modifies the distribution functions of the charged particles which, in turn, through Maxwell’s equations, modify the electromagnetic fields. The wave-particle interactions can, for example, saturate the growth of an instability in space plasmas, or change the current profile in a fusion device.

The evolution of the particle distribution function, when interacting with coherent waves, is usually described by the quasilinear theory [1]. The theory leads to a velocity (action) diffusion equation in which the wave-particle interactions are included through a diffusion operator. It is assumed that the electromagnetic waves continuously act on the particles and that their motion is randomized, with respect to the phase of the wave, after one interaction with the wave. This is akin to the Markovian assumption used, for example, in studying Brownian motion. The motion is then characterized by completely uncorrelated particle orbits, phase-mixing, loss of memory, and ergodicity. These statistical properties lead to an important advantage - the long time behavior of particle dynamics is the same as that after one interaction time with the wave. However, there is one significant drawback. The diffusion coefficient is singular, with a Dirac delta function singularity [1], and is not amenable to implementation in numerical codes. More importantly, the Markovian assumption is contrary to the dynamical behavior of particles interacting with coherent waves [2]. The phase space of the particles is a mix of chaotic and coherent motion with islands of coherent motion embedded within chaotic regions. Furthermore, the phase space is bounded with the effect of the wave being limited to particles having a resonant interaction with the waves. Near the boundaries of the bounded phase space, or near islands, particles can get stuck and undergo coherent, correlated motion for times very much longer than the interaction time. Even when the amplitude of the waves is assumed to be impractically large so that the entire phase is chaotic, as in the standard map, the quasilinear theory fails to give an appropriate description of the evolution of the distribution function [3]. The persistence of long time correlations invalidates the Markovian assumption [4–6].

An additional complexity is related to the fact that, in practice, particles do not continuously interact with the same spectrum of waves. For example, in tokamaks, where radio frequency waves are used for heating and current drive, the waves fields are spatially confined. Any given particle, during its toroidal excursion, will interact with the fields over a short fraction of its single transit path length. On its next transit, it will most likely have drifted, due to the inhomogeneity of the magnetic field, away from the location where the previous interaction took place. Thus the interaction of particles with electromagnetic waves encompasses interesting and complex physics, which in most of the realistic cases are not within the domain of validity of the commonly used statistical assumptions.

We derive a hierarchy of functional mappings for the evolution of distribution functions of particles interacting with coherent electromagnetic waves under without resorting to any statistical assumptions or phenomenological arguments. The effect of the electromagnetic waves is assumed to be perturbative. The unperturbed motion being determined by the equilibrium magnetic field confining the plasma [7]. The canonical perturbation theory is utilized so that all the essential features of particle dynamics and the corresponding inhomogeneous phase space are incorporated in the operators defining the functional mappings. Under certain conditions, the mapping equations reduce to the conventional Fokker-Planck action diffusion equation [2].

The Hamiltonian for the particle dynamics

\[ H(\mathbf{J}, \theta, t) = H_0(\mathbf{J}) + \epsilon H_1(\mathbf{J}, \theta, t) \]  

(1)
consists of two parts: the integrable part $H_0(J)$ that is a function of the constants of the motion $J$ of a particle moving in a prescribed equilibrium field [7, 8], and

$$
H_1(J, \theta, t) = \sum_{m \neq 0} A_m(J) e^{i(m \theta - \omega_m t)}
$$

which includes the effect of electromagnetic waves and other perturbations in the equilibrium field. $z = (J, \theta)$ is a vector composed of the action $(J)$-angle $(\theta)$ variables, for $H_0$. In the guiding center approximation for an axisymmetric system (tokamak), the three actions correspond to the magnetic moment, the canonical angular momentum and the toroidal flux enclosed by a drift surface. The respective conjugate angles correspond to the gyrophore, azimuthal and poloidal angle [7]. $\epsilon$ is an ordering parameter indicating that the effect of $H_1$ is perturbative.

The time evolution of any well-behaved function $f(z, t)$ of $z(t)$ and time $t$ from time $t_0$ to time $t$ is given by

$$
f(z(t); t_0, t) = S_H(t; t_0) f(z_0, t_0) \text{ where } S_H(t; t_0) \text{ is the time evolution operator.}
$$

The derivation of $S_H(t; t_0)$ is equivalent to solving the equations of motion. An appropriate way to determine $S_H(t; t_0)$ is to transform to a new set of canonical variables $z' = (J', \theta')$ using an operator $T(z, t)$. The transformation is such that the new Hamiltonian $K(z')$ leads to a time evolution operator $S_K(t; t_0)$ that can be readily determined. A particularly useful transformation is one for which $K$ is a function of the actions $J'$ only. Then $J'$ are constants of the motion and $S_K(t; t_0)$ evolves the angles $\theta'$ so that $f(z'(t; t_0), t) = S_K(t; t_0) f(z_0', t_0) = f(J_0', \theta_0' + \Delta \theta')$ where $\Delta \theta' = \int_{t_0}^{t} \omega_K(J_0', s) ds$ and $\omega_K(J_0', t) = \nabla J_m K (J_0', t)$.

The operator $T(z, t)$ is determined using the Lie transform theory: $T = e^{-L_f}$ where $L_f = [w, f]$. The Poisson bracket is defined as $[a, b] = \nabla_a b - \nabla_b a$. The function $w(z)$ is the Lie generator. The Lie transform theory generates canonical transformations such the operator $T$ commutes with any function of $z$ [9]. This important property implies that the evolution of $f(z, t)$ can be evaluated by transforming to $z'$, applying the time evolution operator $S_K(t; t_0)$ to the transformed function, and then transforming back to $z$. Thus [9],

$$
f(z(t); t_0, t) = T(z_0, t_0) S_K(t; t_0) T^{-1}(z_0, t_0) f(z_0, t_0)
$$

where $T^{-1} = e^{L_f}$ is the inverse operator.

For physical systems of interest described by (2), it is unlikely that $T$ can be completely determined. However, for nearly integrable systems, the Lie transform theory can be applied to determine $T$ perturbatively as a power series in $\epsilon$ [9].

The old Hamiltonian $H$, the new Hamiltonian $K$, the transformation operator $T$, and the Lie generator $w$ are expressed as a power series in $\epsilon$: $X(z, t, \epsilon) = \sum_{n=0}^{\infty} \epsilon^n X_n(z, t)$ where $X$ represents any of the variables $H, K, T, L, w$ [9]. Here $w_0$ is chosen so that $T_0$ is the identity transformation. Second order, the transformations $T$ and $T^{-1}$ are $T_0 = I, T_1 = -L_1, T_2 = -\frac{1}{2} L_2 + \frac{1}{2} L_1^2$ and $T_0^{-1} = I, T_1^{-1} = L_1, T_2^{-1} = \frac{1}{2} L_2 + \frac{1}{2} L_1^2$, respectively. The generating functions are given by

$$
\frac{\partial w_n}{\partial t} + [w_n, H_0] = n(K_n - H_n) - \sum_{m=1}^{n-1} J_{n-m} K_m - m T_{n-m} H_m
$$

The left hand side of Eq. (4) is the total time derivative of $w_n$ along the unperturbed orbits obtained from $H_0$. Then $w_n$ is determined by integrating along these orbits. In order to eliminate the dependence of the new Hamiltonian on $\theta$, we impose the condition that $K_n$'s are either functions of the new actions only or constants. Then,

$$
w_1 = - \sum_{m \neq 0} A_m(J) e^{i(m \theta - \omega_m(J) t)} e^{i\Omega_m(J) t} - e^{i\Omega_m(J) t_0} \frac{\partial \Omega_m(J)}{\partial t}
$$

where we have set $K_1 = 0$ and we have defined $\Omega_m(J) = \mathbf{m} \cdot \omega_0(J) - \omega_m$ with $\omega_0(J) = \nabla J H_0$ being the frequency vector of the unperturbed system. Similarly, we can set $K_2 = 0$ and derive an equation for $w_2$.

The Lie generators in the finite time interval $[t_0, t]$ lead to $w_n(z_0, t_0) = 0$ and, consequently, $T(z_0, t_0) = I$. Since $K_n = 0, (n = 1, 2)$, the evolution operator $S_K$ is the evolution in time along the unperturbed orbits given by $H_0$. Thus, $S_K = S_{K_0} = S_{H_0}$. We can now determine the evolution of $f(z, t)$ in Eq. (3) from $t_1$ to $t_2$.

$$
f(z)_{t_2} = T^{-1}(z_{t_1} + \Delta z, t_2) f(z)_{t_1}
$$

where $f(z)_{t_1} = f(z(t))$ with $\Delta z$ being evaluated along unperturbed orbits. Equation (6) is a functional mapping which maps $f$ at time $t_1 = t_2$ to $f$ at time $t_1$. If we choose $f = z, i.e., f$ is the set composed of the dynamical variables, the mapping (6) gives a near-symplectic mapping for the evolution of $z$ [10]. When $f(z)$ is chosen to be the particle distribution function, Eq. (6) is an approximation to the original Vlasov (Liouville) equation to the same order as the operator $T^{-1}$. Equation (6) is an iterative scheme for the time evolution of $f$ in the same way as symplectic [11], or near-symplectic [10], mappings are for the evolutions of particle orbits. The accuracy of the mapping depends on an effective perturbation strength which is proportional to $\epsilon$ as well as to the time step $\Delta t = t_2 - t_1$. Thus, Eq. (6) applies to any perturbation strength provided that the time step has been chosen sufficiently small to control the accuracy of the mapping.

For a particle distribution function $f(J, \theta, t)$, let us define a function $F(J, \theta) = \int f(J, \theta, t) d\theta$. Then, from (6)

$$
F(J, \theta)_{t_2} = \langle T^{-1}(J, \theta, \Delta \theta, t_1) \rangle_{\theta} F(J, \theta)_{t_1}
$$

where $\langle \cdots \rangle_{\theta}$ denotes averaging over $\theta$. Here the operator $T^{-1}(z, t)$, averaged over $\theta$, acts on a function of $J$ and $\theta$. 


From the second order expansion, and the fact that all functional dependencies on \( \theta \) are periodic with respect to \( \theta \), we obtain

\[
F(\hat{z}, t_2) = \nabla_\hat{z} [D(\hat{z}, t_2) \nabla_\hat{z} F(\hat{z}, t_1)] + C(\hat{z}, t_2) \nabla_\hat{z} F(\hat{z}, t_1)
\]

with \( \hat{z} = (J, \theta) \).

\[
D(\hat{z}, t) = \frac{1}{2} \begin{pmatrix}
\langle (\nabla_\theta w_1)^2 \rangle - \langle (\nabla_\theta w_1)(\nabla_\theta w_1) \rangle \\
- \langle (\nabla_\theta w_1)(\nabla_\theta w_1) \rangle
\end{pmatrix} \frac{\langle (\nabla_3 w_1)^2 \rangle}{\theta}
\]

and

\[
C(\hat{z}, t) = \langle (\nabla_\theta (w_1 + (1/2)w_2)) \theta, \nabla_\theta (w_1 + (1/2)w_2) \theta \rangle
\]

The Lie generating functions \( w_n \) are calculated in the interval \([t_1, t_2]\). If we set \( t_1 = t_0 \) as the initial time and \( t_2 = t \) as the running time, differentiating (8) gives

\[
\frac{\partial F(\hat{z}, t)}{\partial t} = \nabla_\hat{z} [D_t(\hat{z}, t) \nabla_\hat{z} F(\hat{z}, t_0)] + C_t(\hat{z}, t) \nabla_\hat{z} F(\hat{z}, t_0)
\]

where \( D_t = \partial D / \partial t, C_t = \partial C / \partial t \). Since, on the right hand side, \( F \) depends on the initial time \( t_0 \), Eq. (11) is not a Fokker-Planck type of equation.

Substituting \( f = z \) in Eq. (6) gives

\[
C_t = \lim_{\Delta t \to 0} \frac{\langle (\Delta z)_{\theta} \rangle}{\Delta t}, \quad D_t = \lim_{\Delta t \to 0} \frac{\langle (\Delta z)(\Delta z)_{\theta} \rangle}{2\Delta t}
\]

where \( (\Delta z) \) is the variation of \( z \). This form of \( C_t \) and \( D_t \) is similar to the usual quasilinear diffusion coefficients [13].

The Lie transform technique can be carried out to higher orders in \( \epsilon \). However, the higher order corrections to particle dynamics lead to higher order derivatives of \( F \) appearing on the right hand side of Eqs. (8) and (11) [12]. This is analogous to the Kramers-Moyal expansion of the master equation in stochastic processes [13].

Since Lie operators acting on any function of the dynamical variables can be commuted through the function to act directly on the dynamical variables, the evolution of the particle distribution function is directly related to single particle dynamics. This is evident from (8) and (11). Consequently, the Lie generating functions in \( D \) in (9) are directly related to approximate invariants of the particle dynamics. When the infinite time interval is extended to infinity, the solutions to (4) give the approximate invariants of the motion. So all the essential information for the resonant structure of the dynamical phase space is included [2].

The inhomogeneity of the phase space due to the existence of resonant islands is included in the quasilinear tensor \( D \) through \( w_n \). The kinetic equation (11) includes the topology of all phase space.

If we do not average over any of the angles, Eq. (11) is the evolution equation for the full distribution function. The sequential averaging of one angle at a time generates a hierarchy of evolution equations for the appropriately angle-averaged distribution function. As the evolution is averaged over each angle, the dimension of the phase space for the distribution function is correspondingly reduced. While each angle variable varies more rapidly than its canonically conjugate action variable, they may not necessarily evolve faster than the time for wave-particle interactions. For example in a tokamak plasma the particle gyration angle is averaged over since it corresponds to the fastest time scale. However, the poloidal or toroidal angles of the particle vary more slowly and can be included in the hierarchical description [8]. The averaging process does not affect the accuracy of the perturbation theory. The elements of \( D \) in (9) can be analytically evaluated even when we include all the canonical angles. Then the effect of the physics associated with averaging over one or more angles can be directly determined by the change in each element of \( D \).

When all the angles have been averaged over, the resulting evolution equation (11) is for a distribution function that is a function of the actions only. Then \( C = C_t = 0 \) and \( D \) and \( D_t \) are completely determined by \( w_1 \) in Eq. (5). So the second order, in \( \epsilon \), time evolution equation for the action distribution function is determined by first order effects in the particle dynamics. This result is akin to Madey’s theorem for wave-particle interactions in microwave sources [14]. Our procedure provides a relation between the diffusion tensor and the friction vector in the evolution equation for any Hamiltonian system with arbitrary number of degrees of freedom [15]. In action space,

\[
D(J, \Delta t) = \sum_{m \neq 0} mm |A_m(J)|^2 \frac{1 - \cos (\Omega_m(J) \Delta t)}{\Omega_m(J)^2}
\]

and

\[
D_t(J, t) = \sum_{m \neq 0} mm |A_m(J)|^2 \frac{\sin (\Omega_m(J) t)}{\Omega_m(J)}
\]

where \( mm \) is a dyadic. Then Eq. (11), with \( D_t \) in (14), reduces to a Fokker-Planck equation when

\[
\nabla_3 F(J, t) = \nabla_3 F(J, t_0)
\]

which relates the distribution at time \( t_0 \) with that at time \( t \) on both sides of Eq. (11). Physically, this implies that we consider evolution over times that are smaller than the relaxation time \( t_{rel} \) for \( F \). We obtain a Fokker-Planck equation with a time-dependent tensor \( D_t \) (14). In the vicinity of the resonances given by \( \Omega_m = 0 \), the tensor is continuous and non-singular. The resonance width decreases with time. The time-dependent diffusion tensor \( D_t \) is similar to the “running diffusion tensor” discussed by Balescu [16]. However, there is one significant difference. The \( D_t \) obtained above depends on the dynamical actions and includes inhomogeneous resonant structure of...
the phase space. Balescu’s tensor is independent of actions and applies to a Markovian-type chaotic phase space.

The step size in time $t_{\text{per}}$ for integrating Eq. (11) is determined, in first order perturbation theory, by the effective magnitude of $H_1$. Shorter time steps are required near resonances where $\Omega_m = 0$. Longer time steps are possible when the autocorrelation time $t_{\text{ac}}$ for the waves is smaller than the trapping time $t_r$, in the waves [5]. This is the premise for weak turbulence theory where either the amplitudes of the waves are small, or the wave spectrum is broad enough. In the limit $t \to \infty$, Eq. (14) gives the time-independent quasilinear diffusion tensor

$$D_{qi}(J) = \sum_{m \neq 0} \sum_{n} \Omega_m(J)^2 \delta(\Omega_m(J))$$

where $\delta$ is Dirac’s delta function. The long time limit in the evaluation of $w_1$ is justified only for statistically random processes, such as a Markovian process, in which there is phase mixing and rapid decorrelation of the particle orbits [2]. The singular delta function is difficult to implement numerically and, more importantly, excludes the short time transient effects. Furthermore, the asymptotic time limit results in a time-reversible Fokker-Planck equation, while the time-dependent $D_m$ in (14), being an odd-function of time, results in a time-reversible process.

In conclusion, we have derived a time-reversible evolution equation for the distribution function of particles interacting with coherent waves in a plasma. From this master equation, we derive a hierarchy of evolution equations for dimensionally reduced distribution functions. The particle dynamics or the wave spectra are not subject to any Markovian, or statistical, assumptions which form the basis of standard quasilinear theories. For particles interacting with coherent plasma waves, long-time correlations exist even in the chaotic parts of phase space [2]. These correlations and the inhomogeneity of the dynamical phase space are implemented in our evolution equations. The corresponding time evolution operators are non-singular and time dependent, an are capable of describing both transient and asymptotic collective particle behavior. The appropriate time steps for evolution depend on the effective strength of the perturbing wave fields and the structure of the dynamical phase space.

This work was supported under the Contract of Association ERB 5005 CT 99 0100 between the European Atomic Energy Community and the Hellenic Republic. The content of the publication is the sole responsibility of its author (s) and it does not necessarily represent the views of the Commission or its services. This work was also supported by US DOE Grant numbers DE-FG02-99ER-54521, DE-FG02-91ER-54109, and MIT-France Seed Fund.