Quasilinear theory revisited: general kinetic formulation of wave–particle interactions in plasmas

Kyriakos Hizanidis[1]**, Yannis Kominis**[1] **and Abhay K Ram**[2]

[1] National Technical University of Athens, Association EURATOM-Hellenic Republic, Zografou, Athens 15773, Greece
[2] Plasma Science and Fusion Center, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract
In laboratory fusion devices radio frequency electromagnetic waves are routinely used for heating plasmas and for controlling current profiles. The evolution of particle distribution function in the presence of electromagnetic waves is derived from fundamental equations using the action-angle variables of the dynamical Hamiltonian. Unlike conventional quasilinear theories (QLTs), the distribution function is evolved concurrently with the particle motion. Since the particle dynamics is time reversal invariant, the master equation for the evolution of the distribution function is also time reversal invariant. A sequential averaging of the master equation over the angles leads to a hierarchy of diffusion equations. The diffusion operator in the equation obtained after averaging over all angles is time dependent, in direct contrast to time independent diffusion operator in QLTs. The evolution of the distribution function with time-dependent diffusion operator is markedly different from quasilinear evolution and is illustrated for current drive by a spectrum of coherent electrostatic waves. A proper description of wave–particle interactions is important for fusion plasmas since the velocity space gradients of the distribution function decisively affect collisional relaxation and the associated transport processes.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

The interaction of coherent electromagnetic waves with charged particles in plasmas is a universal phenomenon. In present day experimental fusion devices radio frequency (RF) waves are routinely used for heating the plasma and for generating plasma currents. In the International Thermonuclear Experimental Reactor (ITER [1]) RF waves, particularly in the electron cyclotron range of frequencies, will be used for controlling instabilities deleterious to

the confinement of the plasma. RF waves will also play a prominent role as ITER heads toward a steady state operation. The interaction of RF waves with the plasma constituents leads to transfer of energy and momentum from the waves to the particles. Considering the role of RF waves in fusion plasmas it is imperative to have a proper description of the wave–particle interactions. Besides laboratory plasmas, wave–particle interactions manifest themselves in all sorts of plasma environments, including terrestrial, solar and astrophysical plasmas. The waves in the natural plasmas are usually generated by instabilities. A feature common to RF waves and waves in space is that these waves are coherent and not some sort of statistically random fluctuations. So the wave–particle interaction is between coherent plasma waves and charged particles. In this paper we formulate a kinetic theory for the collective behavior of charged particles interacting with coherent electromagnetic waves. The waves modify the particle distribution functions which, in turn, through Maxwell's equations, modify the electromagnetic fields.

The usual formalism for wave–particle interactions is the quasilinear theory (QLT), in which the evolution of the distribution function is through a diffusion operator acting in velocity (action) space [2]. In deriving this operator it is assumed that the electromagnetic waves act continuously on the particles randomizing their motion, with respect to the phase of the wave, after one interaction time. The interaction time is a measure of the time it takes a particle to go through one phase cycle of the wave spectrum. For a sinusoidal plane wave the interaction time is essentially the time over which the phase of the waves changes by 2π. The de-phasing with respect to the phase of the wave is assumed to lead to random motion of the particle akin to Brownian motion. This is the Markovian assumption and is 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.

The Markovian assumption for particle orbits has some significant drawbacks. The corresponding diffusion coefficient is singular, with a Dirac delta function singularity [2], and, consequently, not amenable to implementation in numerical codes. More importantly, the dynamical behavior of particles interacting with coherent waves is not ergodic and does not satisfy the Markovian assumption [3]. The dynamical phase space of particles interacting with coherent waves is inhomogeneous with phase space islands embedded in a chaotic sea. Furthermore, in all wave–particle interactions, the phase space is bounded with the effect of the wave being limited to particles having a resonant interaction with the waves [4]. 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 diffusion operator fails to give an appropriate description of the evolution of the distribution function [5]. The persistence of long time correlations invalidates the Markovian assumption.

In this paper, we formulate a kinetic description for the evolution of a distribution function of particles that is commensurate with the dynamical phase space of particles interacting with coherent electromagnetic waves in plasmas. In the absence of electromagnetic fields the motion, in a prescribed steady state magnetic field, is assumed to be integrable. The Hamiltonian for this unperturbed motion can then be expressed in terms of the action-angle variables corresponding to the constants of motion. In an axisymmetric tokamak, the action variables are the magnetic moment, toroidal flux and the parallel (to the magnetic field) angular momentum. The corresponding, canonically conjugate, angle variables are the gyrophase, the poloidal angle and the toroidal angle. The effect of the electromagnetic wave is assumed to be perturbative—the magnitude of the wave magnetic field being small compared with the ambient

magnetic field. We then make use of the Lie perturbation method to advance the distribution function in time consistent with the motion of particles in the waves. There is no separation of time scales so that the distribution function evolves along with the dynamics of the particles. The resulting master evolution equation is time reversible just as the equation of motion of the particles is time reversible. If we average the master equation sequentially over the angle variables, averaging first over the fastest varying angle, we obtain a hierarchy of diffusion equations. The equation where all the angle variables have been averaged out is the diffusion equation that can be directly compared with the result from the usual quasilinear analysis. In contrast to the quasilinear diffusion operator, the diffusion operator in our kinetic equation is time dependent. In the limit when time approaches infinity, the time-dependent diffusion operator tends to the quasilinear operator. The evolution of the distribution function obtained from our kinetic theory is different from that in the quasilinear case. This is demonstrated by considering diffusion in a continuous spectrum of electrostatic waves propagating along the magnetic field in a plasma.

2. Detailed formalism

Let us consider a general form of the perturbed Hamiltonian system:

$$H(\boldsymbol{J}, \boldsymbol{\theta}, t) = H_0(\boldsymbol{J}) + \varepsilon H_1(\boldsymbol{J}, \boldsymbol{\theta}, t), \tag{1}$$

where $H_0(\boldsymbol{J})$ corresponds to an integrable Hamiltonian depending only on the actions $\boldsymbol{J}$,

$$H_1(\boldsymbol{J}, \boldsymbol{\theta}, t) = \sum_{m \neq 0} A_m(\boldsymbol{J}) e^{i(m \cdot \boldsymbol{\theta} - \omega_m t) + \gamma_m t} + cc \tag{2}$$

is a perturbation that makes the full Hamiltonian non-integrable and ε is an ordering parameter indicating the perturbative effect of H_1. The perturbation H_1 is expressed as a Fourier series in the periodic angle variables $\boldsymbol{\theta}$, with amplitudes A_m, real frequencies ω_m, and a growth or damping rate of the wave given by the magnitude and sign of γ_m. The action-angle variables $\boldsymbol{X} = (\boldsymbol{J}, \boldsymbol{\theta})$ are obtained from H_0 and the time evolution of $\boldsymbol{X}$ from an initial time t_0 to a later time t is governed by Hamilton's equations of motion. The time evolution of any function $f(\boldsymbol{X}, t)$ from t_0 to t is given by

$$f(\boldsymbol{X}, t) = S_H(t_0 \rightarrow t) \circ f(\boldsymbol{X}_0, t_0), \tag{3}$$

where $\boldsymbol{X}_0 = \boldsymbol{X}(t_0)$ are the initial conditions and $S_H(t_0 \rightarrow t)$ is the time evolution operator. The evaluation of $S_H(t_0 \rightarrow t)$, which is equivalent to solving the equations of motion, may not be possible for the original choice of variables. Then make use of the Lie transform theory to map the phase space in $\boldsymbol{X}$ onto a phase space spanned by a new set of variables $\boldsymbol{Y}$. The canonical transformation $T(\boldsymbol{X}, t)$ for this mapping is such that $\boldsymbol{Y} = T(\boldsymbol{X}, t) \cdot \boldsymbol{X}$, where $T(\boldsymbol{X}, t) = \exp[-L(\boldsymbol{X}, t)]$ with $L(\boldsymbol{X}, t)$ being the Lie operator. $L(\boldsymbol{X}, t)$ is obtained from the generating function $w(\boldsymbol{X}, t)$ such that $L \cdot f = [w, f]_{\mathrm{PB}}$, where $[,]_{\mathrm{PB}}$ denotes the Poisson bracket in $\boldsymbol{X}$ phase space. The transformation is chosen in such a way that the new Hamiltonian $K(\boldsymbol{Y}, t)$ with the corresponding time evolution operator $S_K(t_0 \rightarrow t)$ is easier to evaluate. An important and basic property of the Lie transform operator is that it generates canonical transformations and that it commutes with any function of the phase space variables. The latter property implies that the evolution of $f(\boldsymbol{X}_0, t_0)$ can be obtained by transforming to the new variable set $\boldsymbol{Y}_0$, applying the time evolution operator $S_K(t_0 \rightarrow t)$ to the transformed function and then transforming back to the original variables $\boldsymbol{X}$ [6],

$$f(\boldsymbol{X}, t) = T(\boldsymbol{X}_0, t_0) \circ S_K(t_0 \rightarrow t) \circ T^{-1}(\boldsymbol{X}_0, t_0) \circ f(\boldsymbol{X}_0, t_0). \tag{4}$$

For a nearly integrable Hamiltonian system represented by equation (1), with $\varepsilon \ll 1$, a perturbation scheme can be developed in which T is expressed as a power series in ε [7]. According to this scheme, the old Hamiltonian H, the new Hamiltonian K, the transformation operator T and the Lie generator w are expanded in power series of ε. We can set $w_0 = 0$, such that T_0 and T_0^{-1} are both the identity operator I. Up to second order in ε we obtain

$$T \simeq I - \varepsilon L_1 + \frac{\varepsilon^2}{2}(L_1^2 - L_2), \qquad T^{-1} \simeq I + \varepsilon L_1 + \frac{\varepsilon^2}{2}(L_1^2 + L_2). \qquad (5)$$

Then the transformed zero order Hamiltonian is $K_0 = H_0$, while the Lie generators, up to second order in ε are given by

$$\frac{\mathrm{d}w_1}{\mathrm{d}t} = K_1 - H_1, \qquad \frac{\mathrm{d}w_2}{\mathrm{d}t} = 2(K_2 - H_2) - L_1 \cdot (K_1 + H_1). \qquad (6)$$

The time derivative is along the unperturbed orbit given by H_0,

$$\frac{\mathrm{d}}{\mathrm{d}t} = \partial_t + [\,, H_0]_{\mathrm{PB}} \qquad (6a)$$

with ∂_t denotes the partial derivative. The solutions to equations (6) and (6a) are obtained by integrating the right-hand side along known unperturbed orbits. The K_n's ($n \geqslant 1$) are conveniently chosen so that only the slowly varying terms appear. We choose K_n's so as to eliminate the dependence on $\boldsymbol{\theta}$ up to the second order in ε. Thus, $K_n = 0$ for $n = 1, 2$. We can then calculate the evolution of particles that is accurate up to ε^2. The first order generator w_1 is readily obtained from equations (2) and (6) by integration from t_0 to t:

$$w_1(\boldsymbol{J}, \boldsymbol{\theta}, t) = \mathrm{i} \sum_{m \neq 0} A_m(\boldsymbol{J}) \mathrm{e}^{\mathrm{i}m \cdot [\boldsymbol{\theta} - \omega_0(\boldsymbol{J})t]} \frac{\mathrm{e}^{\mathrm{i}\Omega_m(\boldsymbol{J})t + \gamma_m t} - \mathrm{e}^{\mathrm{i}\Omega_m(\boldsymbol{J})t_0 + \gamma_m t_0}}{\Omega_m(\boldsymbol{J}) - \mathrm{i}\gamma_m} + cc, \qquad (7)$$

where $\omega_0(\boldsymbol{J}) = \partial H_0 / \partial \boldsymbol{J}$ is the frequency vector of the unperturbed system and $\Omega_m(\boldsymbol{J}) = \boldsymbol{m} \cdot \omega_0(\boldsymbol{J}) - \omega_m$. If the dynamical system is periodic with respect to all angles $\boldsymbol{\theta}$, $w_2 = 0$ and $L_2 = 0$.

The evolution of $f(\boldsymbol{X}, t)$ over a small time interval $\Delta t = t - t_0$ can be obtained by the same perturbation scheme. Then, to second order in ε,

$$f(\boldsymbol{X}, t) - f(\boldsymbol{X}_0, t_0) = [T^{-1}(\boldsymbol{X}_0, t_0 + \Delta t) - I] \circ f(\boldsymbol{X}_0, t_0) + O(\varepsilon^n; n > 2), \qquad (8)$$

where

$$T^{-1}(\boldsymbol{X}_0, t_0 + \Delta t) = S_K(t_0 \to t_0 + \Delta t) \circ T^{-1}(\boldsymbol{X}_0, t_0) + O(\varepsilon^n; n > 2). \qquad (9)$$

Given that $\Delta t \ll 1$, and using equation (5) a Taylor series expansion of the left-hand side of equation (9) results in

$$T^{-1}(\boldsymbol{X}_0, t_0 + \Delta t) = I + \Delta t \partial_t T^{-1}(\boldsymbol{X}, t) = \Delta t \partial_t(\varepsilon L_1 + \varepsilon^2 L_1^2 / 2). \qquad (10)$$

From equation (8),

$$f(\boldsymbol{X}, t) - f(\boldsymbol{X}_0, t_0) = \Delta t \partial_t(\varepsilon L_1 + \varepsilon^2 L_1^2 / 2) \circ f(\boldsymbol{X}_0, t_0) + O(\varepsilon^n; n > 2). \qquad (11)$$

In the limit $\Delta t \to 0$ this yields

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \varepsilon [\partial_t w_1, f]_{\mathrm{PB}} + \frac{\varepsilon^2}{2}([\partial_t w_1, [w_1, f]_{\mathrm{PB}}]_{\mathrm{PB}} + [w_1, [\partial_t w_1, f]_{\mathrm{PB}}]_{\mathrm{PB}}) + O(\varepsilon^n; n > 2), \qquad (12)$$

where $w_1 = w_1(\boldsymbol{J}, \boldsymbol{\theta}, t)$ is given by equation (7). It is important to note that in this scheme the new Hamiltonian is independent of $\boldsymbol{\theta}$—the $\boldsymbol{\theta}$ dependence having been transformed away to terms which are of higher order than ε^2.

4

Let $f(\boldsymbol{X}, t)$ be the particle distribution function and let us define angle-averaged distribution functions $F_\ell(\boldsymbol{J}, \theta_{m\neq\ell}, \theta_{n\neq\ell}, t)$, $F_{\ell m}(\boldsymbol{J}, \theta_{n\neq\ell,m}, t)$ and $F_{\ell mn}(\boldsymbol{J}, t)$ with $\{\ell, m, n\} = \{1, 2, 3\}$, such that

$$F_\ell(\boldsymbol{J}, \theta_{m\neq\ell}, \theta_{n\neq\ell}t) \equiv \langle f(\boldsymbol{J}, \boldsymbol{\theta}, t)\rangle_\ell \equiv \frac{1}{2\pi} \oint \mathrm{d}\theta_\ell f(\boldsymbol{J}, \boldsymbol{\theta}, t), \tag{13a}$$

$$F_{\ell m}(\boldsymbol{J}, \theta_{n\neq\ell,m}, t) \equiv \langle f(\boldsymbol{J}, \boldsymbol{\theta}, t)\rangle_{\ell m} \equiv \frac{1}{(2\pi)^2} \oint \mathrm{d}\theta_\ell \oint \mathrm{d}\theta_m f(\boldsymbol{J}, \boldsymbol{\theta}, t), \quad (13b)$$

$$F_{\ell mn}(\boldsymbol{J}, t) \equiv \langle f(\boldsymbol{J}, \boldsymbol{\theta}, t)\rangle_{\ell mn} \equiv \frac{1}{(2\pi)^3} \oint \mathrm{d}\theta_1 \oint \mathrm{d}\theta_2 \oint \mathrm{d}\theta_3 f(\boldsymbol{J}, \boldsymbol{\theta}, t). \quad (13c)$$

If the system is periodic with respect to all angles $\boldsymbol{\theta}$, than an averaging equation (12) over one or more angles yields

$$\partial_t F + [F, H]_{\mathrm{PB}} = \nabla_{\boldsymbol{Q}} \cdot (\overset{\leftrightarrow}{\boldsymbol{D}} \cdot \nabla_{\boldsymbol{Q}} F), \tag{14}$$

where F is any one of the three defined in equations (13a)–(13c), $\boldsymbol{Q}$ represents the respective reduced phase spaces $(\boldsymbol{J}, \theta_{m\neq\ell}, \theta_{n\neq\ell})$, $(\boldsymbol{J}, \theta_{n\neq\ell,m})$ and $(\boldsymbol{J})$, and the diffusion tensor $\boldsymbol{D}$ is given by

$$\overset{\leftrightarrow}{\boldsymbol{D}}(\boldsymbol{Q}, t) \equiv \frac{\varepsilon^2}{2} \partial_t \begin{bmatrix} \langle \nabla_\theta w_1 \nabla_\theta w_1 \rangle & -\langle \nabla_\theta w_1 \nabla_{\boldsymbol{J}} w_1 \rangle \\ -\langle \nabla_{\boldsymbol{J}} w_1 \nabla_\theta w_1 \rangle & \langle \nabla_{\boldsymbol{J}} w_1 \nabla_{\boldsymbol{J}} w_1 \rangle \end{bmatrix}. \tag{15}$$

Here $\langle \rangle$ denotes averaging over the angles complementary to $\boldsymbol{Q}$. The Poisson bracket term in the left-hand side of equation (14) exists as long as the distribution F is a function of one or more angles. Equations (12) and (14) form a hierarchy of four Fokker–Planck types of equations whose dimensionality depends upon the number of angles retained in the appropriate description. When all the angles are retained, we obtain the evolution of F in the complete six-dimensional phase space. As each angle is averaged out, we reduce the phase space by projecting F onto lower dimension. When all the angles are averaged out the evolution equation is for F projected onto a three-dimensional phase space corresponding to space spanned by only action variables. When wave–particle interactions do not exist, the right-hand side of equation (14) is zero and we obtain the Vlasov equation for the unperturbed particle motion in reduced phase space.

Corresponding to the canonical transformation, $\boldsymbol{X} \rightarrow \boldsymbol{X}'$, there exists a transformation between the respective subspace, $\boldsymbol{Q} \rightarrow \boldsymbol{Q}'$. Let $M_{\boldsymbol{Q}\rightarrow\boldsymbol{Q}'}$ be the corresponding Jacobian matrix of the subspace transformation and $|M_{\boldsymbol{Q}\rightarrow\boldsymbol{Q}'}|$ be its determinant. Even though the transformation is canonical, the sub-matrix of the transformation is not necessarily unitary. Then, in the new subspace $\boldsymbol{Q}'$, equation (14) becomes

$$\partial_t F' + [F', H']_{\mathrm{PB}} = -F'[\ln|\overset{\leftrightarrow}{\boldsymbol{M}}_{\boldsymbol{Q}\rightarrow\boldsymbol{Q}'}|, H'_0]_{\mathrm{PB}} + |\overset{\leftrightarrow}{\boldsymbol{M}}_{\boldsymbol{Q}\rightarrow\boldsymbol{Q}'}|^{-1}\overset{\leftrightarrow}{\boldsymbol{M}}_{\boldsymbol{Q}\rightarrow\boldsymbol{Q}'} \cdot \nabla_{\boldsymbol{Q}'} \cdot [(\overset{\leftrightarrow}{\boldsymbol{D}} \cdot \overset{\leftrightarrow}{\boldsymbol{M}}_{\boldsymbol{Q}\rightarrow\boldsymbol{Q}'})$$
$$\cdot \nabla_{\boldsymbol{Q}'}|\overset{\leftrightarrow}{\boldsymbol{M}}_{\boldsymbol{Q}\rightarrow\boldsymbol{Q}'}|F']. \tag{16}$$

The first term in the right-hand side can be interpreted as a 'non-inertial' term which is due to the evolution of the transformation itself. This term vanishes if $\boldsymbol{Q} = \boldsymbol{J}$, that is, when all the angular dependence has been averaged out. In this case $|M_{\boldsymbol{Q}\rightarrow\boldsymbol{Q}'}|$ depends only on $\boldsymbol{J}$.

3. Comparison with the QLT

Consider the evolution equation of the angle-averaged distribution function in action space. In equation (14) $\boldsymbol{Q} = \boldsymbol{J}$, and the diffusion tensor in equation (15) takes on the following

dyadic form:

$$\overset{\leftrightarrow}{\boldsymbol{D}}(\boldsymbol{J}, t) = \sum_{m \neq 0} \boldsymbol{mm} \frac{|\varepsilon A_m(\boldsymbol{J})|^2 e^{2\gamma_m t}}{\Omega_m^2(\boldsymbol{J}) + \gamma_m^2} \{\gamma_m (1 - e^{-2\gamma_m t} \cos[\Omega_m(\boldsymbol{J})t])$$

$$+ e^{-\gamma_m t} \Omega_m(\boldsymbol{J}) \sin[\Omega_m(\boldsymbol{J})t]\}. \tag{17}$$

In contrast to the traditional QLT [2], the action-space Fokker–Planck equation given in equations (14), with the diffusion tensor in equation (17), possesses a time-dependent diffusion tensor that treats resonant and non-resonant interactions on equal footing. The Fokker–Planck equation evolves on the same time scale as the averaged action-space distribution function does. In contrast to QLT, our formalism is the same for growing ($\gamma_m > 0$) and damped ($\gamma_m < 0$) waves. Even for $\gamma_m = 0$, $\boldsymbol{D}$ is continuous and non-singular in the vicinity of resonances ($\Omega_m = 0$). In this case, the width of a resonance decreases with time. Although $\boldsymbol{D}$ resembles the 'running diffusion tensor' introduced by Balescu [8], it is fundamentally different. The $\boldsymbol{D}$ given above depends on the action variables and it fully incorporates the inhomogeneous resonant structure of the phase space. The running diffusion tensor by Balescu is for a chaotic, Markovian-type, phase space.

For very long times $t \to \infty$, and with $\gamma_m = 0$, we readily obtain the quasilinear result

$$\lim[\overset{\leftrightarrow}{\boldsymbol{D}}(\boldsymbol{J}, t \to \infty)] = \sum_{m \neq 0} \boldsymbol{mm} |\varepsilon A_m(\boldsymbol{J})|^2 \delta[\Omega_m(\boldsymbol{J})], \tag{18}$$

where δ is the Dirac's delta function. However, this limit is questionable since it assumes statistically random, or Markovian, processes [3]. The limit of extending time to infinity in the canonical perturbation theory, such as in the evaluation of w_1, can only be justified for statistical process of the Markovian type in which there is phase mixing and rapid de-correlation of the particle orbits [5]. The loss of memory in the Fokker–Planck equation with the time and action-space dependent inhomogeneous diffusion tensor of equation (17) is not of a statistical nature. It is rather a loss of memory of the initial phases (angles) over which we have averaged. The singular delta function in QLT completely ignores the short time effects that are important in the evolution the distribution function. While the diffusion operator is evolving so is the distribution function. The particles that determine the diffusion operator are part of the distribution function. Additionally, the delta function requires ad hoc smoothing remedies in numerical implementations.

In order to illustrate the differences between our kinetic theory and the standard QLT we consider the case of electrons interacting with electrostatic fields in a strongly magnetized plasma. The electrostatic fields are assumed to be propagating along the magnetic field so that only motion of electrons along the magnetic field needs to be considered. In laboratory fusion devices this would correspond to the interaction of lower hybrid waves with electrons for driving plasma currents. The one-dimensional unperturbed Hamiltonian for the electrons is $H_0(\boldsymbol{J}) = J^2/2$. The averaged distribution function in action space is initially taken to be a Maxwellian distribution with thermal spread J_0, $F(\boldsymbol{J}, t = 0) = \exp(-J^2/2J_0^2)/(2\pi J_0^2)^{1/2}$. We assume that the electron motion is perturbed by a set of three discrete modes. The evolution of $F(J, t)$ is obtained from equations (15) and (16) and shown in figure 1. In figure 1(*a*) the discrete modes are assumed to be in steady state while in figure 1(*b*) they are assumed to be damped. In the case of the damped modes, $F(J, t)$ approaches a steady state. For discrete modes, the usual QLT is not applicable since smoothing of the delta function in the diffusion coefficient in equation (17) cannot be justified. The electron motion is not chaotic for discrete waves with arbitrarily small amplitudes.
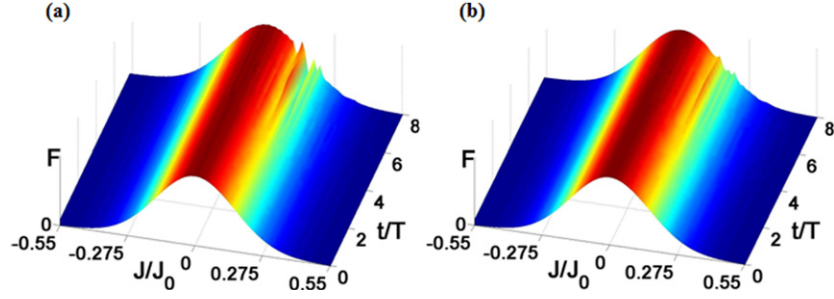
Figure 1. Evolution of $F(J, t)$ for a discrete set of three modes. Here $H_0(J) = J^2/2$, $m = (3; 5; 7)$, $\varepsilon = 6 \times 10^{-4}$, $A_m = \omega_m = 1$, $\gamma_m = 0$ (*a*), $\gamma_m = -0.01$ (*b*) for all m. D is obtained from equation (17). The time is normalized to the wave period $T = 2\pi/\omega_m$, with $\omega_m = 1$ for all m's.
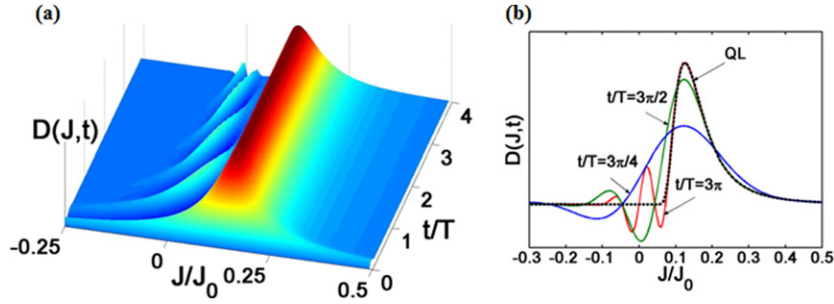

Figure 2. (*a*) $D(J, t)$, from equation (17), as a function of J and t for a continuous spectrum of waves, $A_m = \exp[-(m - m_0)^2/\Delta m]$, $\varepsilon = 10^{-2}$, $m_0 = 3$, $\Delta m = 1.8$, $\omega_m = 1$ and $\gamma_m = 0$; (*b*) $D(J, t)$ as a function of J for $t = 3\pi/4$ (blue), $3\pi/2$ (green), 3π (red). The black dashed curve corresponds to the quasilinear diffusion coefficient. The time is normalized to the wave period $T = 2\pi/\omega_m$, with $\omega_m = 1$ for all m's. (Colour online.)

We next consider a continuous spectrum of undamped ($\gamma_m = 0$) waves of Gaussian form, that is, $A_m = \exp[-(m - m_0)^2/\Delta m]$, where m_0 and Δm are the center and the spectral width, respectively. For such a spectrum the usual QLT is applicable. This spectrum of waves has a unidirectional phase velocity so that, when interacting with electrons, it will generate plasma currents. In figure 2 we plot $D(J, t)$, from equation (17), as a function of J and t. It is evident that as $t \to \infty$, $D(J; t)$ approaches, as expected, the quasilinear form. The difference between our model and the usual QLT becomes clear when we look at the evolution of the angle averaged $F(J, t)$, as determined from equations (15) and (16), and compare it with the results from QLT. In figure 3 the current and the thermal spread from the two formalisms are plotted as a function of time for two different spectra: the larger amplitude spectrum being farther out in the tail of the distribution function and the smaller amplitude spectrum being closer to the bulk.

The differences between the results obtained from our kinetic model and from the QLT are quite evident—the differences increasing with time. The figure for the thermal spreading of the distribution function shows that the QLT underestimates the heating of electrons. In figure 3(*b*) QLT results for the thermal spread, for the two cases considered, evolve very closely and appear almost indistinguishable for the times considered. The differences in the current generated and in the thermal spreading of the electron distribution function are due to the 'tails'

7
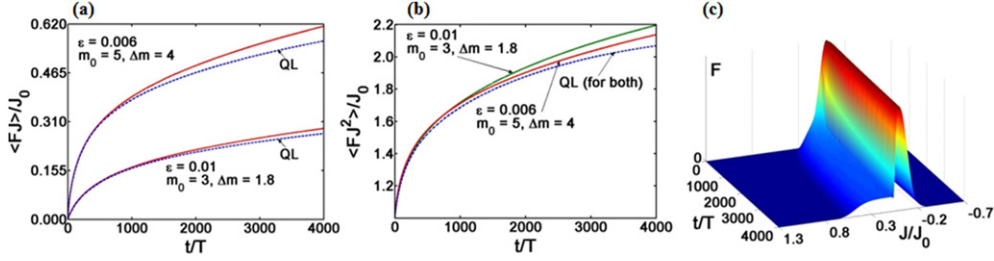
Figure 3. (*a*) Evolution of the normalized (to the thermal velocity) first moment of $F(J, t)$ (current) for the case of continuous spectrum shown in figure 2 (for the lower set of curves) and a second case with $\varepsilon = 6 \times 10^{-3}$, $m_0 = 5$, $\Delta m = 4$ (upper set of curves); (*b*) the associated evolution of the normalized (to the thermal velocity) square root of the second moment of $F(J, t)$ (thermal speed); QL refers to the quasilinear result; (*c*) the respective evolution of $F(J, t)$ as obtained from (14) using (17) for the continuous spectrum with $\varepsilon = 6 \times 10^{-3}$ and $A_m = \exp[-(m - m_0)^2/\Delta m]$, where $m_0 = 5$, $\Delta m = 4$ and $\gamma_m = 0$. The time is normalized to the wave period $T = 2\pi/\omega_m$, with $\omega_m = 1$ for all m's.

in our kinetic diffusion coefficient, shown in figures 2(*a*) and (*b*), that persist as a function of time and are not present in the quasilinear diffusion coefficient.

The implication of these differences is very important. If we take the limit $t \to \infty$ of equation (14), then this limit cannot be commuted through the derivative on the right-hand side. If the limit commuted with the derivative, the current and the thermal spread evaluated from our kinetic model would have been the same as for the QLT for long times. Consequently, the usual QLT is incapable of accounting for diffusive effects at early times which affect the long time behavior of the particle distribution function.

4. Conclusion

In magnetic fusion devices, electromagnetic (RF) waves are used for heating and for generating plasma currents in magnetic confinement devices. The plasma waves are collective particle oscillations which, in turn, interact with the plasma particles. From first principles, we have derived a kinetic theory for the evolution of the particle distribution function when interacting with electromagnetic waves inside plasmas. The interaction of particles with coherent waves is non-Markovian so that the complete structure of the dynamical phase space is included in our formalism. The kinetic theory leads to a time-dependent diffusion operator which evolves on the same time scale as the particle orbits. It is, thus, markedly different from the diffusion operator which one gets in QLTs based on the Markovian assumption. The statistical assumptions in QLTs are in contradiction to the particle dynamics in coherent electromagnetic waves as obtained from the equations of motion. Consequently, based on our kinetic description, the evolution of a distribution function of particles interacting with RF plasma waves leads to considerably different results when compared with the conventional QLTs. This is not only relevant to present day experiments but also to the International Thermonuclear Experimental Reactor [1] in which electron cyclotron waves will be used for local current generation and control of plasma instabilities. An additional complexity is related to the fact that, in practice, particles do not continuously interact with the same spectrum of waves. In tokamaks, the waves' fields used for heating and current drive, for example the electron cyclotron waves, are spatially confined inside the plasma. 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 in QLTs. The kinetic formalism developed in this paper accounts for this complex phase space structure.

In the high temperature fusion plasma the time scale for particle collisions is much longer than any time scale for wave–particle interactions between RF waves and the charged particles. Thus, we can justify the neglect of collisions in the kinetic description of wave–particle interaction. However, when we consider the steady state attained on collisional time scales we do need to have a proper description of the particle distribution function, since the velocity space gradients of the distribution function decisively affect the collisional relaxation and the associated transport processes. This could also be the case in the cooler plasma edge, where coherent electromagnetic waves are generated by instabilities. Then, in the left-hand side of equation (14) an angle-averaged collisional term would have to be added [9]. A perturbation scheme with collisional relaxation in the presence of coherent electromagnetic waves could then be formulated. This is a topic for future research.

In the guiding center approximation for an axisymmetric tokamak, the three actions correspond to the magnetic moment, μ, the canonical toroidal momentum, p_ϕ, and the poloidal canonical momentum, J_p. These three actions can always be defined even when the guiding center approximation is not valid, for example, when frequencies and wave numbers are comparable to the gyro-frequencies and gyro-radii, respectively. The conjugate angles correspond to the gyrophase Θ_g, the toroidal angle Φ and the poloidal angle Θ. Of the three actions, two are velocity-like, namely the magnetic moment and the toroidal momentum. The poloidal canonical momentum is mainly space-like and it is associated with the flux surface near which the motion either the circulating or banana-trapped particles is confined. The canonical toroidal and poloidal momenta can be evaluated from the toroidal flux function ψ and the poloidal flux function ψ_p. In magnetic field coordinates, the toroidal flux is commonly used to express, in covariant form, the equilibrium magnetic field [10], $\boldsymbol{B} = G(\psi)\nabla\Phi + I(\psi)\nabla\Theta + \Delta(\psi, \Theta)\nabla\psi$. Here $G(\psi)$ and $I(\psi)$ are related to the poloidal and the toroidal currents, respectively, while $\Delta(\psi, \Theta)$ is related to the non-orthogonality of the coordinate system. The poloidal flux is determined from the safety factor $q(\psi_\mathrm{p}) = \mathrm{d}\psi/\mathrm{d}\psi_\mathrm{p}$. Averaging over the gyrophase leads to an equation of the type given in equation (14), which determines the evolution of a distribution function for time scales longer than the gyration period. Such a description is suitable for waves or perturbing magnetic fields with frequencies small compared with the gyro-frequencies. The description includes the dynamics of both circulating and banana-trapped particles. This property is lost when we average over the toroidal angle, which is similar to bounce averaging. This averaging provides separate equations for trapped and passing particles. These equations incorporate the spatial dynamics in the poloidal plane and the velocity-like dynamics. If collisions are important at this stage, collisional trapping and de-trapping will couple these equations together. Averaging over the poloidal angle, which is the same as flux surface averaging, leads to a kinetic equation which captures the radial dynamics, across flux surfaces while, at the same time, preserving the velocity space dynamics. This kinetic equation includes the long time behavior of the local distribution functions when RF waves are used for heating and current drive, and for stabilizing the neoclassical tearing mode. By including collisions we can obtain an action-space kinetic equation which can be used for evaluating the transport coefficients in the presence of wave–particle interactions due to coherent electromagnetic waves.

References

[1] ITER Physics Expert Group 2007 *Nucl. Fusion* **47** S128
[2] Vedenov A A, Velikhov E P and Sagdeev R Z 1961 *Nucl. Fusion* **1** 82
    Drummond W E and Pines D 1962 *Nucl. Fusion (Suppl. 3)* 1049
    Kennel C F and Engelmann F 1966 *Phys. Fluids* **9** 2377
    Kaufman A N 1972 *Phys. Fluids* **15** 1063
[3] Cary J R, Escande D F and Verga A D 1990 *Phys. Rev. Lett.* **65** 3132
    Benisti D and Escande D F 1998 *Phys. Rev. Lett.* **80** 4871
[4] Ram A K, Bers A, and Kupfer K 1989 *Phys. Lett.* A **138** 288
[5] Rechester A B and White R B 1980 *Phys. Rev. Lett.* **44** 1586
[6] Kominis Y, Ram A K and Hizanidis K 2008 *Phys. Plasmas* **15** 122501
    Kominis Y, Ram A K and Hizanidis K 2010 *Phys. Rev. Lett.* **104** 235001
[7] Deprit A 1969 *Celest. Mech.* **1** 12
    Cary J R 1981 *Phys. Rep.* **79** 129
[8] Balescu R 2005 *Aspects of Anomalous Transport (Series in Plasmas Physics)* (Bristol: Institute of Physics Publishing)
[9] Cohen R, Hizanidis K, Molvig K and Bernstein I 1984 *Phys. Fluids* **27** 377
[10] D'haeseleer W D, Hitchon W N G, Callen J D and Shohet J L 1991 *Flux Coordinates and Magnetic Field Structure: A Guide to a Fundamental Tool of Plasma Theory* (Berlin: Springer)