Mode conversion in cold low-density plasma with a sheared magnetic field

I. Y. Dodin, D. E. Ruiz, and S. Kubo

Articles you may be interested in

Weak turbulence theory for beam-plasma interaction
Physics of Plasmas **25**, 011603 (2017); 10.1063/1.5017518

One-dimensional particle simulation of wave propagation and generation of second harmonic waves in a composite of plasma and metamaterial
Physics of Plasmas **24**, 122112 (2017); 10.1063/1.5001108

Nonlinear waves and coherent structures in quasi-neutral plasmas excited by external electromagnetic radiation
Physics of Plasmas **24**, 123105 (2017); 10.1063/1.5003322

Particle-in-cell simulation of Buneman instability beyond quasilinear saturation
Physics of Plasmas **24**, 122103 (2017); 10.1063/1.5006463

Self-organized criticality in a cold plasma
Physics of Plasmas **24**, 120701 (2017); 10.1063/1.5005560

Dynamic stabilization of filamentation instability
Physics of Plasmas **25**, 011601 (2017); 10.1063/1.5017517

Mode conversion in cold low-density plasma with a sheared magnetic field

I. Y. Dodin,[1] D. E. Ruiz,[2] and S. Kubo[3]
[1]*Princeton Plasma Physics Laboratory, Princeton, New Jersey 08543, USA*
[2]*Sandia National Laboratories, P.O. Box 5800, Albuquerque, New Mexico 87185, USA*
[3]*National Institute for Fusion Science, National Institutes of National Sciences, 509-5292 Toki, Gifu, Japan*

A theory is proposed that describes mutual conversion of two electromagnetic modes in cold low-density plasma, specifically, in the high-frequency limit where the ion response is negligible. In contrast to the classic (Landau–Zener-type) theory of mode conversion, the region of resonant coupling in low-density plasma is not necessarily narrow, so the coupling matrix cannot be approximated with its first-order Taylor expansion; also, the initial conditions are set up differently. For the case of strong magnetic shear, a simple method is identified for preparing a two-mode wave such that it transforms into a single-mode wave upon entering high-density plasma. The theory can be used for reduced modeling of wave-power input in fusion plasmas. In particular, applications are envisioned in stellarator research, where the mutual conversion of two electromagnetic modes near the plasma edge is a known issue. *Published by AIP Publishing.* https://doi.org/10.1063/1.5003931

I. INTRODUCTION

Mode conversion (MC) is the exchange of action (quanta) between normal modes of a dispersive medium when the parameters of the medium evolve in time or in space.[1–4] Here, we discuss linear MC, which is most efficient when both the frequencies and the wave vectors of interacting modes are close to each other. Regions of such resonant interaction are usually assumed well-localized, so MC theories typically approximate the coupling matrix with its Taylor expansion near the resonance. Then, the general two-wave coupling problem can be reduced,[4–8] at least in the absence of dissipation, to the classic Landau–Zener problem from quantum mechanics.[9,10] This leads to compact asymptotic formulas for the mode amplitudes (which were also rediscovered *ad hoc* in various contexts; e.g., see Refs. 2 and 11–13). However, there are systems where this somewhat universal "Landau–Zener paradigm" is inapplicable. They include inhomogeneous media with degenerate and nearly degenerate wave spectra, such as isotropic or weakly anisotropic dielectrics[13–15] and nonmagnetized or weakly magnetized plasmas as a special case.[16] Although the procedure for deriving the governing equations for such media is known in general,[16,17] calculating the coupling matrices and solving the wave equations explicitly remains an open research area.

Here, we study MC in a specific medium with a nearly degenerate wave spectrum, namely, cold magnetized low-density plasma. As opposed to the standard treatment of the MC in magnetized plasma,[18] which is known as the O-X conversion, our theory allows for a sheared magnetic field. Previous theoretical studies of the O-X conversion in a sheared field[19–28] were either pursued numerically or assumed planar geometry or specific limits (e.g., the high-density limit), or addressed the dynamics of polarization instead of mode amplitudes *per se*. Hence, there is still a lack of a general analytical theory that could explicitly describe the exchange of quanta between the electromagnetic (EM) modes in the low-density case.

The problem of MC in cold three-dimensional low-density plasma with a sheared magnetic field is presently of applied interest in stellarator research.[28–30] Due to a strong magnetic shear and a relatively smooth density profile near the plasma edge in a stellarator, an externally launched single-mode EM wave can lose its quanta to the other EM mode through MC, hence affecting the overall wave-plasma coupling in the device. The MC occurs due to the fact that, in low-density plasma, both EM modes have close-to-vacuum dispersion, i.e., are nearly resonant; then, even a weak inhomogeneity of the magnetic field can couple them easily. To improve the efficiency of the wave-power input into a stellarator, a simple three-dimensional theory of MC in edge plasma with a sheared magnetic field could be beneficial. Here, we construct such theory analytically by considering the plasma density as a small parameter.

The paper is organized as follows. In Sec. II, we introduce a reduced equation for MC in a general context, basically, by restating results from Refs. 8, 16, and 31. Later, this formulation is tailored to weakly anisotropic media, namely, an ordinary differential equation (ODE) is derived for the mode complex envelope along geometrical-optics (GO) rays. The result represents an alternative to the well-known Budden-Kravtsov equations.[2,13] In appropriate variables, the coefficients in our ODE depend on the medium parameters through only one real function. Moreover, if this function remains smooth enough, only its asymptotics matter. Using this, we propose a simple method for predicting how the amplitudes of the two EM modes evolve within a given wave. In Sec. III, we apply these findings to cold magnetized low-density plasma and explain how they can be used for optimizing the wave-power input in fusion plasmas. In Sec. IV, we summarize our results.

II. GENERAL THEORY

A. Basic equations

Let us consider a stationary EM wave governed by

$$\hat{\mathbf{D}}\mathbf{E} = 0. \tag{1}$$

The dispersion operator $\hat{\mathbf{D}}$ that determines the evolution of the electric field $\mathbf{E}$ is obtained by combining Ampere's and Faraday's laws and can be expressed as follows:

$$\hat{\mathbf{D}} = (c/\omega)^2(-\nabla\nabla + \mathbb{1}_3\nabla^2) + \mathbb{1}_3 + \hat{\boldsymbol{\chi}}. \qquad (2)$$

Here, c is the speed of light, ω is the wave frequency, $\mathbb{1}_N$ denotes a $N \times N$ unit matrix, and $\hat{\boldsymbol{\chi}}$ is the medium susceptibility, which is generally an operator. Let us assume that the characteristic wavelength λ of $\mathbf{E}$ is sufficiently small as characterized by the GO parameter

$$\bar{\epsilon} \doteq \lambda/\ell \ll 1. \qquad (3)$$

(Here, the symbol $\doteq$ denotes definitions, and ℓ is the smallest length scale of the inhomogeneities, including the wave-envelope length scales.) Under this assumption, Eq. (1) can be simplified as follows.

Let us consider the phase-space representation of the dispersion operator. Specifically, we apply the Weyl transform, $\hat{\mathbf{D}} \mapsto \mathbf{D}$, where the Weyl image $\mathbf{D}$ is a (3×3)-matrix function of the spatial coordinate $\mathbf{x}$ and the momentum (wave-vector) coordinate $\mathbf{p}$.[32] Then

$$\mathbf{D}(\mathbf{x}, \mathbf{p}) = \mathbf{D}_0(\mathbf{p}) + \boldsymbol{\chi}(\mathbf{x}, \mathbf{p}). \qquad (4)$$

(The dependence on ω is assumed.) Here, $\mathbf{D}_0$ corresponds to the vacuum part of the dispersion operator. In terms of components

$$[\mathbf{D}_0(\mathbf{p})]^\mu{}_\nu = (c/\omega)^2(p^\mu p_\nu - p^2\delta^\mu_\nu) + \delta^\mu_\nu, \qquad (5)$$

where $\mu, \nu = 1, 2, 3$. (In the Euclidean metric, which is henceforth assumed, upper and lower indexes are interchangeable.) Likewise, $\boldsymbol{\chi}$ is the Weyl image of $\hat{\boldsymbol{\chi}}$. For clarity, we assume that there is no dissipation, so $\boldsymbol{\chi}$ is Hermitian. (Weak dissipation can be introduced additively and does not affect our general approach.) Then, the matrix $\mathbf{D}(\mathbf{x}, \mathbf{p})$ is Hermitian too. From the spectral theorem, it has three orthonormal eigenvectors $\mathbf{h}_r(\mathbf{x}, \mathbf{p})$, and the corresponding eigenvalues $\mathcal{E}_r(\mathbf{x}, \mathbf{p})$ are real, which will be used below.

Let us express the electric field as $\mathbf{E} = e^{i\theta}\boldsymbol{\Psi}$, where $\boldsymbol{\Psi}$ is the slow envelope and θ is the rapid phase. We treat the latter as a prescribed function that remains to be specified (see below). The phase θ also determines the wave vector $\mathbf{k} \doteq \nabla\theta$. By Taylor-expanding $\mathbf{D}(\mathbf{x}, \mathbf{p})$ around $\mathbf{p} = \mathbf{k}(\mathbf{x})$ to the first order in $\bar{\epsilon}$, we get

$$\mathbf{D}(\mathbf{x}, \mathbf{p}) \approx \mathbf{D}(\mathbf{x}) + [p_\mu - k_\mu(\mathbf{x})]\mathcal{V}^\mu(\mathbf{x}). \qquad (6)$$

Here, summation over the repeating index μ is assumed. Also, $\mathbf{D}(\mathbf{x}) \doteq \mathbf{D}(\mathbf{x}, \mathbf{k}(\mathbf{x}))$, $\mathcal{V}^\mu(\mathbf{x}) \doteq \mathcal{V}^\mu(\mathbf{x}, \mathbf{k}(\mathbf{x}))$, and $\mathcal{V}^\mu(\mathbf{x}, \mathbf{p}) \doteq \partial\mathbf{D}(\mathbf{x}, \mathbf{p})/\partial p_\mu$. Then, by applying the inverse Weyl transform to Eq. (6) and substituting the result in Eq. (1), we obtain an approximate envelope equation[33]

$$\mathbf{D}(\mathbf{x})\boldsymbol{\Psi} - \frac{i}{2}\left[\partial_\mu \circ \mathcal{V}^\mu(\mathbf{x}) + \mathcal{V}^\mu(\mathbf{x}) \circ \partial_\mu\right]\boldsymbol{\Psi} \approx 0. \qquad (7)$$

Here, $\boldsymbol{\Psi}$ is a shortened notation for $\boldsymbol{\Psi}(\mathbf{x})$, ∂_μ is a shortened notation for $\partial/\partial x^\mu$, and the symbol $\circ$ denotes the composition

of operators. Specifically, $[\partial_\mu \circ \mathcal{V}^\mu(\mathbf{x})]\boldsymbol{\Psi}$ means, by definition, that $\boldsymbol{\Psi}$ first gets multiplied by the matrix $\mathcal{V}^\mu(\mathbf{x})$ and then the *whole product* is differentiated, so the result is $\partial_\mu[\mathcal{V}^\mu(\mathbf{x})\boldsymbol{\Psi}]$. Likewise, $[\mathcal{V}^\mu(\mathbf{x}) \circ \partial_\mu]\boldsymbol{\Psi}$ means that $\boldsymbol{\Psi}$ is first differentiated and then is multiplied by $\mathcal{V}^\mu(\mathbf{x})$, so the result is $\mathcal{V}^\mu(\mathbf{x})(\partial_\mu\boldsymbol{\Psi})$. Hence, the two terms differ by $\partial_\mu[\mathcal{V}^\mu(\mathbf{x})]\boldsymbol{\Psi}$. Similar expansions were also used, e.g., in Ref. 8.

Since $\boldsymbol{\Psi}$ is considered a slow function, the second term in Eq. (7) is $O(\bar{\epsilon})$ so one can see that

$$\mathbf{D}(\mathbf{x})\boldsymbol{\Psi} = O(\bar{\epsilon}). \qquad (8)$$

Then, it is convenient to introduce the representation of $\boldsymbol{\Psi}$ in the basis formed by the eigenvectors of $\mathbf{D}(\mathbf{x})$, namely, $\mathbf{h}_r(\mathbf{x}) \doteq \mathbf{h}_r(\mathbf{x}, \mathbf{k}(\mathbf{x}))$:

$$\boldsymbol{\Psi} = \sum_{r=1}^{3} \mathbf{h}_r(\mathbf{x})\psi_r(\mathbf{x}). \qquad (9)$$

Here, ψ_r are scalar functions and can be understood as follows. Consider multiplying Eq. (8) by $\mathbf{h}_r^\dagger(\mathbf{x})$ from the left. That gives $\mathcal{E}_r(\mathbf{x})\psi_r(\mathbf{x}) = O(\bar{\epsilon})$, where $r = 1, 2, 3$ (no summation over r is assumed), and $\mathcal{E}_r(\mathbf{x}) \doteq \mathcal{E}_r(\mathbf{x}, \mathbf{k}(\mathbf{x}))$. This shows that, for a given r, there are two possibilities: (i) $\psi_r(\mathbf{x})$ is small or (ii) $\mathcal{E}_r(\mathbf{x})$ is small. In case (i), the polarization $\mathbf{h}_r$ does not correspond to a propagating wave mode *per se*; the small nonzero projection of $\boldsymbol{\Psi}$ on $\mathbf{h}_r$ is only due to the fact that the wave field is not strictly sinusoidal. We call such "modes" passive. In case (ii), $(\omega, \mathbf{k})$ are actually close to those of a wave eigenmode that would exist in a homogeneous medium ($\bar{\epsilon} = 0$) with the same parameters. In this case, ψ_r can be understood as the local scalar amplitude of rth mode, so that $\psi_r = O(1)$ is allowed.

Below, we shall consider two active modes, so the third mode ($r = 3$) is automatically passive. In other words, $\psi_{1,2} = O(1)$ and $\psi_3 = O(\bar{\epsilon})$. [Even if an active mode has a zero amplitude initially, it can be excited later through MC from the other active mode, so its amplitude is considered an $O(1)$ quantity.] With this in mind, let us substitute Eq. (9) into Eq. (7) and multiply the resulting equation from the left by $\Pi \doteq \mathbf{h}_1\mathbf{h}_1^\dagger + \mathbf{h}_2\mathbf{h}_2^\dagger$, which is an operator projecting a given vector on the active-mode subspace. Then, one obtains

$$[\boldsymbol{\mathcal{E}} - i\mathbf{v}\cdot\nabla - (i/2)(\nabla\cdot\mathbf{v}) - \mathbf{U}]\boldsymbol{\psi} \approx 0 \qquad (10)$$

up to a term $O(\bar{\epsilon}^2)$, which is neglected.[33] Here

$$\boldsymbol{\psi} \doteq \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \qquad (11)$$

and we introduced the following 2×2 diagonal matrices:

$$\boldsymbol{\mathcal{E}}(\mathbf{x}) \doteq \operatorname{diag}\{\mathcal{E}_1(\mathbf{x}), \mathcal{E}_2(\mathbf{x})\}, \qquad (12)$$

$$\mathbf{v}(\mathbf{x}) \doteq \operatorname{diag}\{\partial_\mathbf{p}\mathcal{E}_1(\mathbf{x}, \mathbf{p}), \partial_\mathbf{p}\mathcal{E}_2(\mathbf{x}, \mathbf{p})\}_{\mathbf{p}=\mathbf{k}(\mathbf{x})}. \qquad (13)$$

Also, $\mathbf{U}(\mathbf{x}) = O(\bar{\epsilon})$ is 2×2 Hermitian matrix given by

$$\mathbf{U} \doteq -(\boldsymbol{\Xi}^\dagger\mathcal{V}^\mu\partial_\mu\boldsymbol{\Xi})_A. \qquad (14)$$

Here, $\Xi(\mathbf{x}) = (\mathbf{h}_1(\mathbf{x}), \mathbf{h}_2(\mathbf{x}))$ is a non-square, 3×2 matrix whose columns are $\mathbf{h}_1$ and $\mathbf{h}_2$ and the index A stands for "the anti-Hermitian part of". Notice that, within the adopted accuracy, ψ_3 does not appear in Eq. (10), even though ψ_3 is generally nonzero. This occurs because the eigenvectors $\mathbf{h}_r$ are orthogonal to each other.

Equation (10) can be reduced to an ODE in two limits. One is the limit of a weakly anisotropic medium, which is discussed below. The other one is the limit of a plane wave propagating in a plane-layered medium, possibly at an angle to the inhomogeneity axis. The latter case will not be considered here but could be approached similarly.

B. Weakly anisotropic medium

Suppose that a medium is only weakly anisotropic, namely, such that $\mathcal{E} = \mathcal{E}_0 \mathbb{1}_2 + O(\bar{\epsilon})$, where $\mathcal{E}_0$ is a scalar. Then, we can approximate $\mathbf{v}$ with a scalar matrix, $\mathbf{v} \approx \mathbf{v}_0 \mathbb{1}_2$, where $\mathbf{v}_0 = [\partial \mathcal{E}_0(\mathbf{x}, \mathbf{p})/\partial \mathbf{p}]_{\mathbf{p}=\mathbf{k}(\mathbf{x})}$ is a vector. [Since $\nabla \psi = O(\bar{\epsilon})$, this approximation introduces an $O(\bar{\epsilon}^2)$ error in Eq. (10), but that is beyond the adopted $O(\bar{\epsilon})$ accuracy of our theory.] Then, $\mathbf{v} \cdot \nabla \approx \mathbb{1}_2(-2d/dl)$, where d/dl is the appropriately normalized spatial derivative along $\mathbf{v}_0$, and the factor -2 is introduced for convenience (see Sec. III). Hence, Eq. (10) can be written as an ODE

$$i\boldsymbol{\psi}' = \mathbf{H}\boldsymbol{\psi}, \tag{15}$$

$$\mathbf{H} \doteq \frac{1}{2}\left[\mathbf{U} - \mathcal{E} - \frac{i}{2}(\nabla \cdot \mathbf{v}_0)\mathbb{1}_2\right], \tag{16}$$

where the prime denotes d/dl. Let us also split $\mathbf{H}$ into its traceless part $\mathcal{H}$ and the remaining scalar part $H_0 \mathbb{1}_2$,

$$\mathcal{H} = \mathbf{H} - H_0 \mathbb{1}_2, \quad H_0 = \operatorname{Tr}\mathbf{H}/2. \tag{17}$$

Then, by using the variable transformation

$$\boldsymbol{\psi} = \exp\left(-i\int H_0 \, dl\right)\mathbf{a}, \tag{18}$$

one can also represent Eq. (15) as

$$i\mathbf{a}' = \mathcal{H}\mathbf{a}, \tag{19}$$

where the coupling matrix $\mathcal{H}$ can be understood as the wave Hamiltonian. As a reminder, this theory describes the field *in the basis formed by* $\mathbf{h}_r$ [Eq. (9)] rather than in the basis determined by the scalar part of $\mathbf{D}$ as in the Budden-Kravtsov theory.[2,13]

Since $\mathcal{H}$ is Hermitian, one can readily notice the following. First, Eq. (19) has a corollary

$$|a_1|^2 + |a_2|^2 = \text{const}, \tag{20}$$

which reflects the conservation of wave quanta. (More specifically, the conserved quantity is the density of the wave action flux, or equivalently, the energy flux density.) Second, $\mathcal{H}$ can be parameterized as follows:

$$\mathcal{H} = \begin{pmatrix} -\alpha & -i\beta \\ i\beta^* & \alpha \end{pmatrix}, \tag{21}$$

where α is real. Let us introduce the new variable

$$\mathbf{q} \doteq \begin{pmatrix} e^{-i\gamma/2} & 0 \\ 0 & e^{i\gamma/2} \end{pmatrix}\mathbf{a}, \tag{22}$$

where $\gamma \doteq \arg \beta$. Let us also replace the independent variable l with $\tau \doteq \int |\beta| \, dl$, henceforth called "time" for brevity. (Replacing l with τ can only be done if β is nonzero in the whole region of interest.) Then, Eq. (19) becomes

$$i\frac{d}{d\tau}\begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} -s(\tau) & -i \\ i & s(\tau) \end{pmatrix}\begin{pmatrix} q_1 \\ q_2 \end{pmatrix}, \tag{23}$$

where

$$s \doteq \frac{\alpha}{|\beta|} - \frac{\dot{\gamma}}{2}. \tag{24}$$

Here the dot denotes a derivative with respect to τ. Note that the medium parameters enter Eq. (23) through only one real function, s. Also note that Eq. (23) can be equivalently written as a second-order ODE for just one of the mode amplitudes. For example, the equation for q_1 is

$$\ddot{q}_1 + (1 + s^2 - i\dot{s})q_1 = 0. \tag{25}$$

C. Landau–Zener model

Equations similar to Eq. (25) emerge also in other MC theories.[4–8] In those works, it is assumed that (i) the wave trajectory starts and ends in regions with $s \gg 1$, and (ii) in the resonance region ($|s| \lesssim 1$), s can be approximated by a linear function of τ, so Eq. (25) becomes the Weber equation. This approximation is justified when the time needed for a wave to traverse the resonance region, which is of order $\dot{s}^{-1}$, is small compared to the characteristic time scale of s, which is of order one in our units; thus, generally speaking, $\dot{s}$ must be large.[34] Under these assumptions, the MC problem becomes identical to the Landau–Zener problem.[9,10] Then, a general asymptotic solution for $\mathbf{q}$ exists

$$\begin{pmatrix} q_1(+\infty) \\ q_2(+\infty) \end{pmatrix} = \begin{pmatrix} \mathcal{T} & -\mathcal{C}^* \\ \mathcal{C} & \mathcal{T} \end{pmatrix}\begin{pmatrix} q_1(-\infty) \\ q_2(-\infty) \end{pmatrix}. \tag{26}$$

Here, $|\mathcal{T}|^2 + |\mathcal{C}|^2 = 1$, or more specifically

$$\mathcal{T} = \exp\left(-\pi|\kappa|^2\right), \quad \mathcal{C} = -\frac{\sqrt{2\pi\kappa}}{\kappa\Gamma(-i|\kappa|^2)}, \tag{27}$$

where Γ is the gamma function, and $\kappa \doteq -i(2\dot{s})^{-1/2}$.

But note that the Landau–Zener model is not entirely universal. For example, when a wave starts in vacuum, the initial wave numbers of the two modes are identical, so the modes are in resonance from the very beginning. This means that $|s| \lesssim 1$ initially. Likewise, the initial $\dot{s}$ can be small, and then it evolves gradually as the wave enters the medium, so the assumption of constant $\dot{s}$ is inapplicable. Hence, in order to describe MC near boundaries, the Landau–Zener model must be replaced with a different approach, which will be discussed below.

D. Spin analogy

Note that Eq. (23) is identical to the generic equation describing a two-level quantum system. For example, one can interpret $q_{1,2}$ as the two components of the wave function describing a spin-1/2 electron. Then, $S_\mu \doteq \mathbf{q}^\dagger \boldsymbol{\sigma}_\mu \mathbf{q}$, where $\boldsymbol{\sigma}_\mu$ are the Pauli matrices,

$$\boldsymbol{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

serves as the expectation value of the μth component of the spin vector in units $\hbar/2$. The resulting vector

$$\mathbf{S} = \begin{pmatrix} S_x \\ S_y \\ S_z \end{pmatrix} \tag{28}$$

is akin to but different from the commonly known Stokes vector[35,36] since q_r are mode amplitudes rather than electric-field components *per se*. Importantly, by assuming the parameterization

$$\mathbf{q} = \bar{q} e^{i\Gamma} \begin{pmatrix} e^{-i\upsilon/2} \cos{(\varphi/2)} \\ e^{i\upsilon/2} \sin{(\varphi/2)} \end{pmatrix}, \tag{29}$$

where $\bar{q} \doteq \sqrt{|q_1|^2 + |q_2|^2}$, $\Gamma \doteq (\arg q_1 + \arg q_2)/2$, and $\upsilon \doteq \arg q_2 - \arg q_1$, one has

$$\mathbf{S} = \bar{q} \begin{pmatrix} \sin{\varphi} \cos{\upsilon} \\ \sin{\varphi} \sin{\upsilon} \\ \cos{\varphi} \end{pmatrix}. \tag{30}$$

Thus, $|\mathbf{S}| = \bar{q}$, and one can infer the values of $q_{1,2}$ from a given $\mathbf{S}$ up to their common phase Γ.

Now that we have introduced Pauli matrices $\boldsymbol{\sigma}_\mu$, let us also use them as a basis to represent the dimensionless Hamiltonian in Eq. (23) as follows:

$$\begin{pmatrix} -s & -i \\ i & s \end{pmatrix} = \frac{1}{2} \boldsymbol{\sigma}_\mu W^\mu. \tag{31}$$

The expansion coefficients W^μ form a three-dimensional vector $\mathbf{W}$ given by

$$\mathbf{W} = \begin{pmatrix} 0 \\ 2 \\ -2s \end{pmatrix}. \tag{32}$$

Then, from Eq. (23), one can deduce that $\mathbf{S}$ is governed by a precession equation[31]

$$\dot{\mathbf{S}} = \mathbf{W} \times \mathbf{S}. \tag{33}$$

Let us assume $\dot{s} \ll 1$. In this case, there is a well-defined local frequency of the complex amplitudes [see Eq. (25)], which also serves as half of the local precession frequency $W = 2\sqrt{1 + s^2}$. Then, the orientation of $\mathbf{S}$, or, more precisely, of the precession plane, follows that of $\mathbf{W}$.[37] This is due to the conservation of the adiabatic invariant associated with the precession, so $\dot{s}$ can be viewed as the adiabaticity parameter. Using it, one can construct a Wentzel–Kramers–Brillouin

(WKB) asymptotic solution of the MC problem, as shown in the Appendix. Below, we show how it can also be understood without detailed calculations, at least in some limits.

1. Nonresonant interaction

Suppose that $s \gg 1$ at all times. Then, $\mathbf{W}$ remains approximately parallel to the z axis. This means that the precession occurs in the (x, y) plane. If the initial state is a pure mode, i.e., the initial $\mathbf{S}$ is parallel to the z axis, then $\mathbf{S}$ remains approximately parallel to the z axis (i.e., the precession trajectory radius remains zero) forever, so no significant MC is possible. This is understood as a nonresonant interaction.

2. Resonant interaction

Now suppose that a wave starts at $s = 0$, so the initial $\mathbf{W}$ is in the y direction. As the wave enters a medium, s begins to grow and $\mathbf{W}$ starts rotating in the (y, z) plane. Suppose that the dispersion curves of the two modes grow apart eventually, so s becomes large ($s \gg 1$). Then, the final direction of $\mathbf{W}$ is along the z axis. If the initial state is a pure mode, i.e., the initial $\mathbf{S}$ is parallel to the z axis, then $\mathbf{S}$ starts precessing in the (x, z) plane first, but eventually the precession plane orients transversely to the z axis [Fig. 1(a)]. This means that, in the final state, the two modes have equal amplitudes. Conversely, in order to obtain a pure state when $s \gg 1$, one needs to start with a mixed-mode state corresponding to $\mathbf{S}$ along the y axis. As shown in Fig. 1(b), as the vector $\mathbf{W}$ starts rotating in the (y, z) plane, the vector $\mathbf{S}$ precesses around $\mathbf{W}$. The final state of $\mathbf{S}$ is aligned to the z axis, which is a pure state mode.

These general arguments can also be reformulated in terms of our original variables $\mathbf{a}$ instead of $\mathbf{q}$. In this case, we define

$$S_\mu \doteq \mathbf{a}^\dagger \boldsymbol{\sigma}_\mu \mathbf{a} \tag{34}$$

and use Eq. (19) to get

$$\mathbf{S}' = \mathbf{W} \times \mathbf{S} \tag{35}$$

(note that the independent variable here is l rather than τ; hence the prime), where the new vector $\mathbf{W}$ is

$$\mathbf{W} = 2 \begin{pmatrix} \mathrm{Im}\,\beta \\ \mathrm{Re}\,\beta \\ -\alpha \end{pmatrix}. \tag{36}$$

The vacuum value of $\mathbf{W}$ is not necessarily along the y axis now. Still, if $\mathbf{W}$ evolves slowly ($\dot{s} \ll 1$), the method for ensuring a single-mode operation in the large-s region is the same as before, namely, $\mathbf{S}$ must be initialized parallel to $\mathbf{W}$.

III. LOW-DENSITY PLASMA

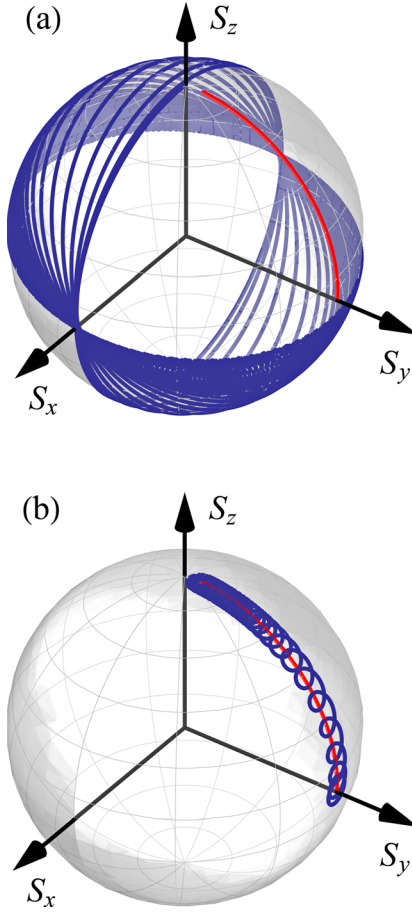In this section, we apply the above general theory to EM waves in cold magnetized low-density plasma.

FIG. 1. Examples of the dynamics of **S** governed by Eq. (33) with a prescribed slowly changing **W**. Shown are the trajectories of **S** (blue) and **W**/W (red) on a unit sphere. (a) Initially, **S** is in the z direction and **W** is in the y direction. Later, **W** rotates and becomes pointed in the z direction. This causes the precession plane to rotate accordingly, so eventually, the vector **S** ends up precessing in the horizontal plane. Hence, the final state is a mixed-mode state. (b) Both **S** and **W** are in the y direction initially and then slowly rotate together towards the z axis. The final state corresponds to $S_z = 1$, which is a pure-mode state.

A. Basic approximations

Suppose a low-density plasma such that χ is comparable to $\bar{\epsilon}$. In this case, we can choose θ such that **k** is constant and satisfies the vacuum dispersion relation

$$k^2 = \omega^2/c^2. \tag{37}$$

(This does *not* mean that the plasma dispersion is neglected. We simply choose to describe it as an effect on **a** rather than as an effect on θ. As long as **a** remains slow, these descriptions are equivalent.) We also choose $\mathbf{D}_0$ for $\mathcal{E}_0 \mathbb{1}_2$ (Sec. II B). Then, $\mathcal{E}_0(\mathbf{x}, \mathbf{p}) = 1 - (pc/\omega)^2$, $\mathbf{v}_0 = -2\mathbf{k}/k^2$, $\nabla \cdot \mathbf{v}_0 = 0$, and

$$\frac{d}{dl} = \frac{k^\mu}{k^2} \frac{\partial}{\partial x^\mu} = \frac{\mathbf{e}_k \cdot \nabla}{k}, \quad \mathbf{e}_k \doteq \frac{\mathbf{k}}{k}. \tag{38}$$

The true eigenvalues of $\mathbf{D}(\mathbf{x}, \mathbf{p})$ in the plasma can be found by considering χ as a small perturbation to $\mathbf{D}_0$ and by using the standard perturbation theory[38]

$$\mathcal{E}_r(\mathbf{x}, \mathbf{p}) \approx \mathcal{E}_0(\mathbf{x}, \mathbf{p}) + (\boldsymbol{\eta}_r^\dagger \chi \boldsymbol{\eta}_r)(\mathbf{x}, \mathbf{p}). \tag{39}$$

Here, $\boldsymbol{\eta}_r$ is the zero-density limit of $\mathbf{h}_r$, and $r = 1, 2$ (henceforth assumed). Since $\mathcal{E}_0(\mathbf{x}, \mathbf{k}) = 0$, this gives

$$\mathcal{E}_r(\mathbf{x}) = (\boldsymbol{\eta}_r^\dagger \chi \boldsymbol{\eta}_r)(\mathbf{x}, \mathbf{k}(\mathbf{x})). \tag{40}$$

In order to calculate **U**, which is already of the first order in $\bar{\epsilon}$, let us use the zeroth-order approximation for $\mathcal{V}^\mu$, namely, $\mathcal{V}^\mu(\mathbf{x}, \mathbf{p}) \approx \partial \mathbf{D}_0/\partial p_\mu$.[39] This gives

$$\mathcal{V}^\mu = \frac{c^2}{\omega^2} \left[(\mathbf{e}^\mu)\mathbf{k}^\dagger + \mathbf{k}(\mathbf{e}^\mu)^\dagger - 2k^\mu \mathbb{1}_3 \right], \tag{41}$$

where $\mathbf{e}^\mu$ is a unit vector along the μ axis. [The adjoints of the real vectors $\mathbf{e}^\mu$ and **k** are the row vectors obtained simply by transposing $\mathbf{e}^\mu$ and **k**, correspondingly. Hence, for example, $(\mathbf{e}^\mu)\mathbf{k}^\dagger$ is a 3×3 matrix but $\mathbf{k}^\dagger(\mathbf{e}^\mu)$ is a scalar.] Let us also use the approximation $\Xi \approx (\boldsymbol{\eta}_1, \boldsymbol{\eta}_2)$. Since $\boldsymbol{\eta}_1$ and $\boldsymbol{\eta}_2$ are *vacuum* polarization vectors, they are orthogonal to **k**, so

$$\mathbf{k}^\dagger \Xi = (\mathbf{k}^\dagger \boldsymbol{\eta}_1, \mathbf{k}^\dagger \boldsymbol{\eta}_2) = 0 \tag{42}$$

and, similarly, $\Xi^\dagger \mathbf{k} = 0$. From this and the fact that **k** is constant, it is readily seen that the first two terms in Eq. (41) do not contribute to **U** in Eq. (14). This leads to

$$\mathbf{U} = 2(\Xi^\dagger \Xi')_A, \tag{43}$$

where we invoked Eq. (37). Then, using the fact that

$$\boldsymbol{\eta}_1^\dagger \boldsymbol{\eta}_1 = \boldsymbol{\eta}_2^\dagger \boldsymbol{\eta}_2 = 1, \quad \boldsymbol{\eta}_1^\dagger \boldsymbol{\eta}_2 = \boldsymbol{\eta}_2^\dagger \boldsymbol{\eta}_1 = 0,$$

one can express **U** as follows:

$$\mathbf{U} = -2i \begin{pmatrix} \boldsymbol{\eta}_1^\dagger \boldsymbol{\eta}_1' & \boldsymbol{\eta}_1^\dagger \boldsymbol{\eta}_2' \\ -(\boldsymbol{\eta}_1^\dagger \boldsymbol{\eta}_2')^* & \boldsymbol{\eta}_2^\dagger \boldsymbol{\eta}_2' \end{pmatrix}. \tag{44}$$

B. Polarization vectors $\eta_{1,2}$

Let us now explicitly calculate $\boldsymbol{\eta}_{1,2}$ assuming the plasma is cold. In order to do so, let us temporarily adopt coordinates such that the local dc magnetic field $\mathbf{B}_0$ is along the z axis, and the y axis is orthogonal to the plane formed by $\mathbf{B}_0$ and **k**, i.e.,

$$\mathbf{b} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{k} = \begin{pmatrix} k\sin\vartheta \\ 0 \\ k\cos\vartheta \end{pmatrix}. \tag{45}$$

(The general result for arbitrary direction of **k** and $\mathbf{B}_0$ is presented at the end of this section.) Here, $\mathbf{b} \doteq \mathbf{B}_0/B_0$ is a unit vector along $\mathbf{B}_0$, and ϑ is the angle between **k** and $\mathbf{B}_0$. Then[40]

$$\chi = \begin{pmatrix} S-1 & -iD & 0 \\ iD & S-1 & 0 \\ 0 & 0 & P-1 \end{pmatrix}. \tag{46}$$

We shall limit our consideration to high-frequency waves, so the ion response can be neglected entirely. Hence,[40]

$$S = 1 - \frac{\omega_p^2}{\omega^2 - \Omega^2}, \quad D = \frac{\Omega}{\omega} \frac{\omega_p^2}{\omega^2 - \Omega^2}, \quad P = 1 - \frac{\omega_p^2}{\omega^2},$$

where $\omega_p \doteq (4\pi n_e e^2 / m)^{1/2}$ is the electron plasma frequency, $\Omega \doteq eB_0/(mc)$ is the electron gyrofrequency, n_e is the electron density, $e < 0$ is the electron charge, and m is the electron mass. Also[40]

$$\mathbf{D}(\mathbf{x}, \mathbf{k}) = \begin{pmatrix} S - N^2 \cos^2 \vartheta & -iD & N^2 \sin \vartheta \cos \vartheta \\ iD & S - N^2 & 0 \\ N^2 \sin \vartheta \cos \vartheta & 0 & P - N^2 \sin^2 \vartheta \end{pmatrix},$$

where $N \doteq k\omega/c$ is the refraction index. From $\mathbf{D}(\mathbf{x}, \mathbf{k})\boldsymbol{\eta}_{1,2} = 0$, we find that each of $\boldsymbol{\eta}_{1,2}$ satisfies

$$\frac{\eta_y}{\eta_x} = -\frac{iD}{S - N^2}, \tag{47}$$

$$\frac{\eta_z}{\eta_x} = -\frac{N^2 \sin \vartheta \cos \vartheta}{P - N^2 \sin^2 \vartheta}. \tag{48}$$

Also, the dispersion relation $\det \mathbf{D} = 0$ gives[40]

$$N_{1,2}^2 = \frac{B \pm \sqrt{B^2 - 4AC}}{2A}, \tag{49}$$

where

$$A = S \sin^2 \vartheta + P \cos^2 \vartheta, \tag{50}$$

$$B = (S^2 - D^2) \sin^2 \vartheta + PS(1 + \cos^2 \vartheta), \tag{51}$$

$$C = P(S^2 - D^2). \tag{52}$$

Then, in the zero-density limit, one obtains[41]

$$\boldsymbol{\eta}_1 = \frac{1}{\sqrt{1 + g_1^2}} \begin{pmatrix} -\cos \vartheta \\ ig_1 \\ \sin \vartheta \end{pmatrix}, \tag{53}$$

$$\boldsymbol{\eta}_2 = -\frac{i\varsigma}{\sqrt{1 + g_2^2}} \begin{pmatrix} -\cos \vartheta \\ ig_2 \\ \sin \vartheta \end{pmatrix}, \tag{54}$$

where we introduced

$$g_{1,2} \doteq u^{-1} \mp \varsigma \sqrt{1 + u^{-2}}, \tag{55}$$

$$u \doteq \frac{2\omega}{\Omega} \csc \vartheta \cot \vartheta, \tag{56}$$

and $\varsigma \doteq \mathrm{sgn}\, u$. The normalization in $\boldsymbol{\eta}_{1,2}$ is chosen such that, at $\vartheta \to \pi/2$, one gets $\boldsymbol{\eta}_1 \to \mathbf{e}_z$ and $\boldsymbol{\eta}_2 \to \mathbf{e}_y$, which corresponds to the O and X waves, correspondingly. (Here, $\mathbf{e}_\mu$ denotes a unit vector along the axis μ. Also remember that our axes are tied to the local $\mathbf{B}_0$.) Accordingly, at $\vartheta \to 0$, one gets $\boldsymbol{\eta}_1 \to -(\mathbf{e}_x + i\mathbf{e}_y \, \mathrm{sgn}\, \Omega)/\sqrt{2}$ and $\boldsymbol{\eta}_2 \to (i\, \mathrm{sgn}\, \Omega)(\mathbf{e}_x - i\mathbf{e}_y \, \mathrm{sgn}\, \Omega)/\sqrt{2}$. Then, $\boldsymbol{\eta}_1$ corresponds to the L wave at $\Omega < 0$ (R wave at $\Omega > 0$), and $\boldsymbol{\eta}_2$ corresponds to the R wave at $\Omega < 0$ (L wave at $\Omega > 0$).

Finally, let us rewrite $\boldsymbol{\eta}_{1,2}$ in the invariant form. To do this, note that $\mathbf{k} = k_x \mathbf{e}_x + k_z \mathbf{e}_z$ and $\mathbf{e}_z = \mathbf{b}$, so

$$\mathbf{e}_x = \frac{\mathbf{k} - \mathbf{b}k_z}{k_x} = \frac{\mathbf{e}_k - \mathbf{b}\cos \vartheta}{\sin \vartheta}. \tag{57}$$

Then, one obtains

$$\boldsymbol{\eta}_1 = \frac{\mathbf{e}_* + ig_1 \mathbf{e}_y}{\sqrt{1 + g_1^2}}, \quad \boldsymbol{\eta}_2 = -\frac{i\varsigma(\mathbf{e}_* + ig_2 \mathbf{e}_y)}{\sqrt{1 + g_2^2}}, \tag{58}$$

where $\mathbf{e}_*$ is introduced, merely to shorten the notation, as a unit vector given by $\mathbf{e}_* \doteq (\mathbf{b} - \mathbf{e}_k \cos \vartheta)/\sin \vartheta$. Fully invariant expressions for $\boldsymbol{\eta}_{1,2}$ can be obtained by using $\mathbf{e}_y = (\mathbf{b} \times \mathbf{e}_k)/\sin \vartheta$ (this equality is seen from the fact that, by definition, $\mathbf{e}_y$ is orthogonal to both $\mathbf{b}$ and $\mathbf{k}$ and has a unit length) and

$$\cos \vartheta = \mathbf{e}_k \cdot \mathbf{b}, \quad \sin \vartheta = |\mathbf{e}_k \times \mathbf{b}|. \tag{59}$$

C. Wave Hamiltonian

Using Eq. (40) along with Eqs. (53) and (54), one readily finds that

$$\mathcal{E}_{1,2} = -\frac{\omega_p^2}{\omega^2 - \Omega^2} \left[1 - \frac{\Omega^2}{2\omega^2} (\mathbf{e}_k \times \mathbf{b})^2 \right.$$
$$\left. \mp \frac{|\Omega|}{\omega} \sqrt{(\mathbf{e}_k \cdot \mathbf{b})^2 + \frac{\Omega^2}{4\omega^2} (\mathbf{e}_k \times \mathbf{b})^4} \right]. \tag{60}$$

Note that $\mathcal{E}_{1,2}$ are linear with respect to the plasma density due to the perturbative approach [Eq. (40)]. By using Eq. (44), one also finds that

$$\mathbf{U} = 2 \begin{pmatrix} \epsilon_z & \epsilon_x - i\epsilon_y \\ \epsilon_x + i\epsilon_y & -\epsilon_z \end{pmatrix}, \tag{61}$$

where $\epsilon_i = O(\bar{\epsilon})$ are scalar functions given by

$$\epsilon_x = -\frac{u'}{2(1 + u^2)} = \frac{\delta'}{2}, \tag{62}$$

$$\epsilon_y = \frac{\zeta}{\sqrt{1 + u^2}} = \zeta \cos \delta, \tag{63}$$

$$\epsilon_z = -\frac{\zeta u}{\sqrt{1 + u^2}} = \zeta \sin \delta. \tag{64}$$

Here, $\delta \doteq -\arctan u$ [where u is given by Eq. (56)] and $\zeta \doteq \mathbf{e}_* \cdot \mathbf{e}_y'$, or, more explicitly,

$$u \doteq \frac{2\omega(\mathbf{e}_k \cdot \mathbf{b})}{\Omega(\mathbf{e}_k \times \mathbf{b})^2}, \quad \zeta \doteq \frac{(\mathbf{e}_k \times \mathbf{b}) \cdot \mathbf{b}'}{(\mathbf{e}_k \times \mathbf{b})^2}, \tag{65}$$

so ζ serves as a measure of the magnetic shear. [The coefficient ϵ_x is the same at the one known for shearless fields, where $\zeta = 0$ (Ref. 21).] Also, $\rho \doteq \mathcal{E}_1 - \mathcal{E}_2$, namely,

$$\rho = \frac{2|\Omega|\omega_p^2}{\omega(\omega^2 - \Omega^2)} \sqrt{(\mathbf{e}_k \cdot \mathbf{b})^2 + \frac{\Omega^2}{4\omega^2} (\mathbf{e}_k \times \mathbf{b})^4}. \tag{66}$$

Hence, the traceless Hamiltonian $\mathcal{H}$ that governs MC [Eq. (17)] is as follows:

$$\mathcal{H} = \begin{pmatrix} \epsilon_z - \rho/2 & \epsilon_x - i\epsilon_y \\ \epsilon_x + i\epsilon_y & -\epsilon_z + \rho/2 \end{pmatrix}. \tag{67}$$

Accordingly,

$$\alpha = \frac{\rho}{2} - \epsilon_z, \quad \beta = \epsilon_y + i\epsilon_x, \quad \gamma = \arctan\left(\frac{\epsilon_x}{\epsilon_y}\right). \tag{68}$$

Also, the corresponding vector $\mathbf{W}$ [Eq. (36)] that governs the precession of $\mathbf{S}$ is

$$\mathbf{W} = \begin{pmatrix} 2\epsilon_x \\ 2\epsilon_y \\ 2\epsilon_z - \rho \end{pmatrix}. \tag{69}$$

D. Propagation parallel to $\mathbf{B}_0$

When a wave propagates parallel to a dc magnetic field (i.e., $\vartheta = 0$), Eq. (60) gives

$$\mathcal{E}_{1,2} = -\frac{\omega_p^2}{\omega(\omega \mp \Omega)}, \tag{70}$$

and Eqs. (61)–(65) give $\mathbf{U} = 0$. In this particular case, the two modes are uncoupled, and Eq. (15) leads to

$$\psi_r = \exp\left(\frac{i}{2}\int \mathcal{E}_r \, dl\right)\psi_{r0}, \tag{71}$$

where the constants ψ_{r0} are determined by the initial conditions. Hence, each $|\psi_r|$ is conserved and $\mathcal{E}_r/2$ serves a correction to the refraction index. (As a reminder, dl is a ray-path element measured in units $k^{-1} = c/\omega$.) The total refraction indexes in this case are

$$N_{1,2} \approx 1 - \frac{\omega_p^2}{2\omega(\omega \mp \Omega)}. \tag{72}$$

This is in agreement with the low-density asymptotics of the known L- and R-wave refraction indexes.[40]

E. Propagation perpendicular to $\mathbf{B}_0$

When a wave propagates perpendicular to a dc magnetic field (i.e., $\vartheta = \pi/2$), Eq. (60) gives

$$\mathcal{E}_1 = -\frac{\omega_p^2}{\omega^2}, \quad \mathcal{E}_2 = -\frac{\omega_p^2}{\omega^2 - \Omega^2}. \tag{73}$$

Then, one obtains

$$\rho = \frac{\omega_p^2 \Omega^2}{\omega^2(\omega^2 - \Omega^2)}. \tag{74}$$

Also, in this case, one has $u = 0$, so $\epsilon_x = \epsilon_z = 0$, and

$$\epsilon_y = \zeta = (\mathbf{e}_k \times \mathbf{b}) \cdot \mathbf{b}'. \tag{75}$$

The corresponding function s that enters Eq. (23) is $s = \rho/(2|\zeta|)$, and $\tau' = |\zeta|$. The adiabaticity parameter in this case is $\dot{s} \sim \rho'/\zeta^2$, where we assumed constant shear (more specifically, $\zeta'/\zeta \ll \rho'/\rho$). We can estimate ρ' as $\rho/(kL_n)$, where L_n is the characteristic scale of the plasma density profile.

Here, ρ can be evaluated at the edge of the adiabaticity domain ($s \sim 1$). This gives $\rho \sim \zeta$. Then, $\dot{s} \sim L_b/L_n$, where $L_b \doteq |k\zeta|^{-1}$ is the characteristic scale of the magnetic-field shear. If $L_n \lesssim L_b$ (weak magnetic shear), then $\dot{s} \gtrsim 1$, so the wave leaves the resonance region before it has time to mode-convert (as long as the GO approximation is satisfied, i.e., $L_n \gg \lambda$). Then, a wave that is initially a pure mode remains such upon entering dense plasma. In contrast, if $L_b \ll L_n$ (strong magnetic shear), then $\dot{s} \ll 1$, so the wave "spin" $\mathbf{S}$ follows $\mathbf{W}$ (Sec. II D). Then, substantial MC is possible. In particular, this explains the results presented in Ref. 28.

Let us consider the case of strong magnetic shear. Then, according to the argument in Sec. II D, a wave that is a pure mode in vacuum eventually transforms into a mixture of the O and X waves with equal amplitudes ($|a_1| = |a_2|$). Conversely, a wave that is composed of two modes in vacuum can asymptotically transform into a single-mode wave upon entering high-density plasma. In the case of the O wave, this requires that the initial amplitudes satisfy Eq. (A10), i.e., the initial vacuum wave must be circularly polarized. Ending up with a pure X wave instead of a pure O wave requires starting with the opposite circular polarization.

For the case of general wave propagation with respect to the magnetic field (arbitrary ϑ), the MC process can be understood similarly except that s is given by a more general formula [Eqs. (24) and (68)].

IV. CONCLUSIONS

In summary, we developed a theory of EM mode conversion in cold low-density plasma, specifically, in the high-frequency limit where the ion response is negligible. In contrast to the classic (Landau–Zener-type) theory of mode conversion, the region of resonant coupling in low-density plasma is not necessarily narrow, so the coupling matrix cannot be approximated with its first-order Taylor expansion; also, the initial conditions are set up differently. For the case of strong magnetic shear, a simple method is identified for preparing a two-mode wave such that it transforms into a single-mode wave upon entering a high-density plasma. The theory can be used for reduced modeling of wave-power input in fusion plasmas. In particular, applications are envisioned in stellarator research, where the mutual conversion of two EM modes near the plasma edge is a known issue.[28–30]

ACKNOWLEDGMENTS

APPENDIX: WKB MODEL

1. Basic equations

Here, we present a formal WKB derivation of the slow "spin-precession" dynamics discussed in Sec. II D. We shall refer to the medium as plasma, and s will be treated as a

measure of the plasma density. We also adopt that $s(\tau = 0) = 0$ corresponds to vacuum. (This is the case for the example considered in Sec. III E.) However, the general idea holds in a broader context too. Also notably, the following model can be understood as a generalization of the "helical-wave" GO discussed in Ref. 22.

Let us start with rewriting Eq. (25) as

$$\ddot{q}_1 + Q(\tau)q_1 = 0, \quad Q \doteq 1 + s^2 - i\dot{s}. \quad (A1)$$

Suppose that

$$\frac{d}{d\tau}\left(\frac{2\pi}{\sqrt{Q}}\right) \ll 1, \quad (A2)$$

a sufficient condition for which is $\dot{s} \ll 1$. Then, the WKB approximation is applicable

$$q_1(\tau) \approx \frac{1}{[Q(\tau)]^{1/4}}\left[C_+ e^{i\phi(\tau)} + C_- e^{-i\phi(\tau)}\right], \quad (A3)$$

where $C_\pm$ are constants determined by the initial conditions at $\tau = 0$, and $\phi \doteq \int_0^\tau \sqrt{Q(\tilde{\tau})}\,d\tilde{\tau}$. Since $\dot{s}$ is assumed small, we Taylor-expand $\sqrt{Q}$ to get $\phi = \phi_{re} + i\phi_{im}$, where

$$\phi_{re} \approx \sigma_s \int_0^\tau \sqrt{1 + s^2(\tilde{\tau})}\,d\tilde{\tau}, \quad (A4)$$

$$\phi_{im} = -\sigma_s \int_0^s \frac{d\tilde{s}}{2\sqrt{1+\tilde{s}^2}} = -\frac{1}{2}\ln\left(|s| + \sqrt{1+s^2}\right). \quad (A5)$$

The sign of ϕ is a matter of convention and depends on the interpretation of $C_\pm$. We introduced a sign factor $\sigma_s \doteq \operatorname{sgn} s(\tau)$ only to ensure that, at $s \gg 1$, one can unambiguously identify the term $\propto e^{i\phi}$ as the first mode (Mode I) and the term $\propto e^{-i\phi}$ as the second mode (Mode II), as seen from Eq. (23). At $s \lesssim 1$, both terms contribute to both modes.

Within this WKB model, if a wave starts *and* ends outside plasma, the mode amplitude is preserved, namely, $\phi_{im}(\infty) = 0$ due to $s(\infty) = 0$. However, note that this
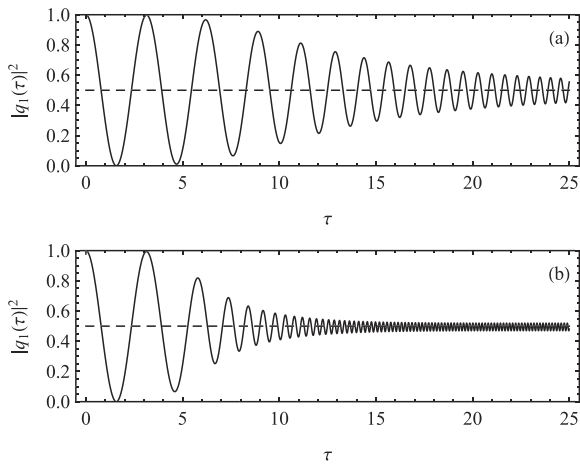

FIG. 2. Numerical solutions of Eq. (A1) for $|q_1|^2$ for two sample profiles: (a) $s(\tau) = 0.02\tau^2$ and (b) $s(\tau) = 40\tanh[(\tau/15)^3]$. The envelope q_1 is measured in units $\bar{q}$ such that $|\bar{q}|^2$ is the total number of quanta, which is conserved [Eq. (20)]. The initial conditions are $q_1(0) = 1$ and $q_2(0) = 0$, i.e., $\dot{q}_1(0) = 0$. Both solutions approach the same asymptote $|q_1|^2 = 1/2$ (dashed), as predicted [Eq. (A7)].

requires the low-density approximation to hold at all τ, which is usually not the case. As a rule, a wave eventually enters a high-density region where rays of the two modes diverge or some dissipation occurs. Thus, even if the radiation escapes plasma later, the original single-mode wave is not quite restored. Hence, the process of rays leaving the plasma will not be considered.

2. Starting with a pure mode

Suppose a wave outside plasma is a pure Mode I, so $q_1(0) = \bar{q}$, $q_2(0) = 0$, and $s(0) = \dot{s}(0) = 0$, so $\dot{q}_1(0) = 0$ and $Q(0) = 1$. Then, $C_+ = C_- = \bar{q}/2$, so Eq. (A3) gives
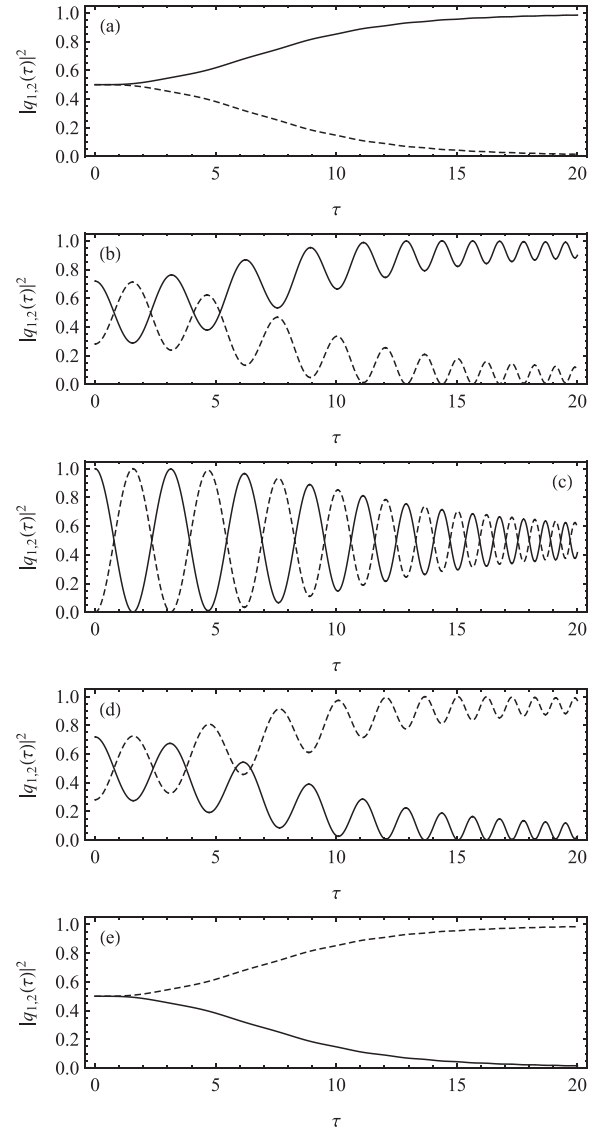

FIG. 3. Numerical solutions of Eq. (A1) for $|q_1|^2$ (solid) and $|q_2|^2$ (dashed) for $s(\tau) = 0.02\tau^2$. The envelopes $q_{1,2}$ are measured in units $\bar{q}$ such that $|\bar{q}|^2$ is the total number of quanta, which is conserved [Eq. (20)]. The initial conditions are $q_1(0) = \sqrt{1 - u^2/2}$ and $\dot{q}_1(0) = -q_2(0) = iu/\sqrt{2}$, so $|q_1(0)|^2 + |q_2(0)|^2 = 1$. Here (a) $u = 1$, (b) $u = 3/4$, (c) $u = 0$, (d) $u = -3/4$, and (e) $u = -1$. In the cases (a) and (e), one starts with a circularly polarized wave in vacuum and ends up with a single-mode wave in dense plasma. In the case (c), where the solid curve is the same as in Fig. 2(a), one starts with a pure Mode I and ends up with a mixture of Modes I and II with equal number of quanta.

$$\frac{q_1(\tau)}{\bar{q}} \approx \frac{\cos\phi(\tau)}{[Q(\tau)]^{1/4}}. \tag{A6}$$

In the high-density limit $(s \gg 1)$, one has $Q \approx s^2$, $2|\cos\phi| \approx \exp(|\phi_{\rm im}|)$, and $2|\phi_{\rm im}| \approx \ln(2|s|)$. Then, $|q_1/\bar{q}|^2$ asymptotically approaches a universal constant

$$\left|\frac{q_1(\infty)}{\bar{q}}\right|^2 \approx \frac{e^{2|\phi_{\rm im}|}}{4|s|} \approx \frac{1}{2}. \tag{A7}$$

Due to Eq. (20), one also gets $|q_2(\infty)/\bar{q}|^2 \approx 1/2$. Thus, during the MC, an initially pure Mode I asymptotically loses half of its action flux (which we also term loosely as "quanta") to Mode II. This conclusion also agrees with numerical calculations (Fig. 2), and similar results apply if the initial conditions of the two modes are interchanged.

3. Ending with a pure mode

Suppose now that a wave *becomes* a pure Mode I asymptotically in the high-density limit $(s \gg 1)$. This means that $C_- = 0$. Also, using Eq. (A5), we obtain

$$q_1 = \frac{C_+}{Q^{1/4}} \, e^{i\phi} \approx C_+ e^{i\phi_{\rm re}} \sqrt{1 + \frac{s}{\sqrt{1+s^2}}}. \tag{A8}$$

In the limit $s \gg 1$, this gives $q_1 \approx C_+\sqrt{2} e^{i\phi_{\rm re}} = \bar{q} e^{i\phi_{\rm re}}$, where $\bar{q} \doteq C_+\sqrt{2}$ can be interpreted as the asymptotic amplitude of Mode I deep inside the plasma. Then

$$q_1(0) \approx \frac{\bar{q}}{\sqrt{2}}, \quad \dot{q}_1(0) \approx \frac{i\sigma_s\bar{q}}{\sqrt{2}}, \tag{A9}$$

where we used $Q = 1$ for vacuum [and thus $\dot{\phi}(0) = 1$]. Since $q_2 = -\dot{q}_1 + isq_1$ [Eq. (23)], one can also rewrite Eqs. (A9) as

$$q_1(0)/q_2(0) \approx i\sigma_s, \tag{A10}$$

which corresponds to $S_y(0) = \pm 1$. These initial conditions ensure that a wave that is initially a mixture of Modes I and II asymptotically converts into the pure Mode-I upon entering high-density plasma. These conclusions also agree with numerical calculations, as illustrated in Fig. 3.

[1]Y. A. Kravtsov and Y. I. Orlov, *Geometrical Optics of Inhomogeneous Media* (Springer-Verlag, New York, 1990).

[2]V. V. Zheleznyakov, V. V. Kocharovskii, and V. V. Kocharovskii, "Linear coupling of electromagnetic waves in inhomogeneous weakly-ionized media," Usp. Fiz. Nauk. **141**, 257 (1983) [Sov. Phys. Usp. **26**, 877 (1983)].

[3]K. Y. Bliokh and S. V. Grinyok, "On the criterion for effectiveness of wave linear transformation in a smoothly inhomogeneous medium and of nonadiabatic transitions during atomic collisions," Zh. Eksp. Teor. Fiz. **120**, 85 (2001) [J. Exp. Theor. Phys. **93**, 71 (2001)].

[4]E. R. Tracy, A. J. Brizard, A. S. Richardson, and A. N. Kaufman, *Ray Tracing and Beyond: Phase Space Methods in Plasma Wave Theory* (Cambridge University Press, New York, 2014).

[5]E. R. Tracy, A. N. Kaufman, and A. J. Brizard, "Ray-based methods in multidimensional linear wave conversion," Phys. Plasmas **10**, 2147 (2003).

[6]E. R. Tracy and A. N. Kaufman, "Metaplectic formulation of linear mode conversion," Phys. Rev. E **48**, 2196 (1993).

[7]L. Friedland, G. Goldner, and A. N. Kaufman, "Four-dimensional eikonal theory of linear mode conversion," Phys. Rev. Lett. **58**, 1392 (1987).

[8]L. Friedland and A. N. Kaufman, "Congruent reduction in geometric optics and mode conversion," Phys. Fluids **30**, 3050 (1987).

[9]L. Landau, "Zur Theorie der Energieübertragung. II," Phys. Z. Sowjetunion **2**, 46 (1932).

[10]C. Zener, "Non-adiabatic crossing of energy levels," Proc. R. Soc. London A **137**, 696 (1932).

[11]V. L. Ginzburg, *The Propagation of Electromagnetic Waves in Plasmas* (Pergamon, Oxford, 1970).

[12]N. S. Erokhin and S. S. Moiseev, *Wave Processes in Inhomogeneous Plasma, in Reviews of Plasma Physics*, edited by M. A. Leontovich (Consultants Bureau, New York, 1979), Vol. 7.

[13]Y. A. Kravtsov, O. N. Naida, and A. A. Fuki, "Waves in weakly anisotropic 3D inhomogeneous media: Quasi-isotropic approximation of geometrical optics," Usp. Fiz. Nauk. **141**, 166 (1996) [Phys. Usp. **39**, 129 (1996)].

[14]K. Y. Bliokh, A. Niv, V. Kleiner, and E. Hasman, "Geometrodynamics of spinning light," Nat. Photonics **2**, 748 (2008).

[15]D. E. Ruiz and I. Y. Dodin, "First-principles variational formulation of polarization effects in geometrical optics," Phys. Rev. A **92**, 043805 (2015).

[16]D. E. Ruiz and I. Y. Dodin, "Extending geometrical optics: A Lagrangian theory for vector waves," Phys. Plasmas **24**, 055704 (2017).

[17]D. E. Ruiz, "Geometric theory of waves and its applications to plasma physics," Ph.D. thesis (Princeton University, 2017).

[18]J. Preinhaelter and V. Kopecky, "Penetration of high-frequency waves into a weakly inhomogeneous magnetized plasma at oblique incidence and their transformation to Bernstein modes," J. Plasma Phys. **10**, 1 (1973).

[19]I. Fidone and G. Granata, "Propagation of electromagnetic waves in a plasma with a sheared magnetic field," Nucl. Fusion **11**, 133 (1971).

[20]D. B. Melrose, "Mode coupling in the solar corona. I. Coupling near the plasma level," Aust. J. Phys. **27**, 31 (1974).

[21]V. V. Zheleznyakov, V. V. Kocharovskii, and V. V. Kocharovskii, "Linear wave interaction in a plasma with an inhomogeneous magnetic field," Zh. Eksp. Teor. Fiz. **77101**, 101 (1979) [Sov. Phys. JETP **50**, 51 (1979)].

[22]V. V. Kocharovskii and V. V. Kocharovskii, "Linear wave conversion in an inhomogeneous plasma with magnetic shear," Fiz. Plazmy **6**, 565 (1980) [Sov. J. Plasma Phys. **6**, 308 (1980)].

[23]M. Brambilla and M. Moresco, "The influence of mode mixing on reflectometry density measurements in a reversed field pinch," Plasma Phys. Controlled Fusion **29**, 381 (1987).

[24]A. Airoldi, A. Orefice, and G. Ramponi, "Polarization and energy evolution of electromagnetic waves in sheared toroidal plasmas," Phys. Fluids B **1**, 2143 (1989).

[25]U. Bellotti and M. Bornatici, "Polarized radiative transfer in plasmas with sheared magnetic field," Plasma Phys. Controlled Fusion **41**, 1277 (1999).

[26]S. E. Segre, "Exact analytic expressions for the evolution of polarization for radiation propagating in a plasma with nonuniformly sheared magnetic field," ENEA Technical Report No. RT/ERG/FUS/2001/01.

[27]A. Yu. Popov, "On O-X mode conversion in 2D inhomogeneous plasma with a sheared magnetic field," Plasma Phys. Controlled Fusion **52**, 035008 (2010).

[28]S. Kubo, H. Igami, T. I. Tsujimura, T. Shimozuma, H. Takahashi, Y. Yoshimura, M. Nishiura, R. Makino, and T. Mutoh, "Plasma interface of the EC waves to the LHD peripheral region," AIP Conf. Proc. **1689**, 090006 (2015).

[29]T. Notake, S. Kubo, T. Shimozuma, H. Idei, Y. Yoshimura, S. Inagaki, K. Ohkubo, S. Kobayashi, Y. Mizuno, S. Ito, Y. Takita, T. Watari, K. Narihara, T. Morisaki, I. Yamada, Y. Nagayama, K. Tanaka, S. Sakakibara, R. Kumazawa, T. Seki, K. Saito, T. Mutoh, A. Shimizu, A. Komori, and the LHD Experiment Group, "Optimization of incident wave polarization for ECRH in LHD," Plasma Phys. Controlled Fusion **47**, 531 (2005).

[30]T. I. Tsujimura, S. Kubo, H. Takahashi, R. Makino, R. Seki, Y. Yoshimura, H. Igami, T. Shimozuma, K. Ida, C. Suzuki, M. Emoto, M. Yokoyama, T. Kobayashi, C. Moon, K. Nagaoka, M. Osakabe, S. Kobayashi, S. Ito, Y. Mizuno, K. Okada, A. Ejiri, and T. Mutoh, and the LHD Experiment Group, "Development and application of a ray-tracing code integrating with 3D equilibrium mapping in LHD ECH experiments," Nucl. Fusion **55**, 123019 (2015).

[31]D. E. Ruiz and I. Y. Dodin, "Lagrangian geometrical optics of nonadiabatic vector waves and spin particles," Phys. Lett. A **379**, 2337 (2015).

[32]For a brief overview of the Weyl transform, see, e.g., Ref. 16.

[33]Details of the calculation presented in Sec. II A will be reported separately in a broader context. For closely related calculations, see Refs. 8, 16, and 31.

[34]Of course, this is only an order-of-magnitude estimate, and $\dot{s} \lesssim 1$ may be tolerable for specific s.

[35]Y. A. Kravtsov, B. Bieg, and K. Y. Bliokh, "Stokes-vector evolution in a weakly anisotropic inhomogeneous medium," J. Opt. Soc. Am. A **24**, 3388 (2007).

[36]M. Born and E. Wolf, *Principles of Optics* (Cambridge University Press, New York, 1999).

[37]S. E. Barnes, "Theory of spinmotive forces in ferromagnetic structures," in *Spin Current*, edited by S. Maekawa, S. O. Valenzuela, E. Saitoh, and T. Kimura (Oxford University Press, Oxford, 2012), Chap. 7.

[38]L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory* (Addison-Wesley, Reading, MA, 1958).

[39]In *cold* plasma, this is an exact equality, since then χ is independent of the wave vector.

[40]T. H. Stix, *Waves in Plasmas* (AIP, New York, 1992).

[41]The calculations were facilitated by *Mathematica* © 1988–2011 Wolfram Research, Inc., version number 8.0.4.0.