Princeton Plasma Physics Laboratory

PPPL-5330

Extending geometrical optics: A Lagrangian theory for vector waves

D.E. Ruiz

December 2016

Prepared for the U.S. Department of Energy under Contract DE-AC02-09CH11466.

Princeton Plasma Physics Laboratory
Report Disclaimers

Extending geometrical optics: A Lagrangian theory for vector waves

D. E. Ruiz[1] and I. Y. Dodin[1,2]

[1]*Department of Astrophysical Sciences, Princeton University, Princeton, New Jersey 08544, USA*
[2]*Princeton Plasma Physics Laboratory, Princeton, New Jersey 08543, USA*

(Dated: December 8, 2016)

Even diffraction aside, the well-known equations of geometrical optics (GO) are not entirely accurate. Traditional GO treats wave rays as classical particles, which are completely described by their coordinates and momenta, but rays have another degree of freedom, namely, polarization. The polarization degree of freedom manifests itself as an effective (classical) spin that can be assigned to rays and can affect the wave dynamics accordingly. A well-known example of associated effects is wave-mode conversion, which can be interpreted as spin precession. However, there are also other, less-known manifestations of the wave spin, such as polarization-driven bending of ray trajectories. This work presents an extension and reformulation of GO as a first-principle Lagrangian theory, whose effective-gauge Hamiltonian governs all the aforementioned polarization phenomena simultaneously. As an example, the theory is applied to describe the polarization-driven divergence of right-hand and left-hand circularly polarized electromagnetic waves in weakly magnetized plasma.

I. INTRODUCTION

A. Motivation

Geometrical optics (GO) is a reduced model of wave dynamics [1, 2] that is widely used in many contexts ranging from quantum dynamics to electromagnetic (EM), acoustic, and gravitational phenomena [3–5]. Mathematically, GO is an asymptotic theory with respect to a small parameter ϵ that is a ratio of the wave relevant characteristic period (temporal or spatial) to the inhomogeneity scale of the underlying medium. Practical applications of GO are traditionally restricted to the lowest-order theory, where each wave is basically approximated with a local eigenmode of the underlying medium at each given spacetime location. Then, the wave dynamics is entirely determined by a single branch of the local dispersion relation. However, this approximation is not entirely accurate, even when diffraction is neglected. If a dispersion relation has more than one branch, i.e., a vector wave with more than one polarization at a given location, then the interaction between these branches can give rise to important polarization effects that are missed in the traditional lowest-order GO.

One of the particularly interesting manifestations of polarization effects is the polarization-driven bending of ray trajectories. At the present moment, it is known primarily in two contexts. One is quantum mechanics, where polarization effects manifest as the Berry phase [6] and the associated Stern-Gerlach force experienced by vector particles, i.e., quantum particles with spin. Another one is optics, where a related effect has been known as the *Hall effect of light* [7–11]. (Namely, even in an isotropic dielectric, rays propagate somewhat differently depending on polarization if the dielectric is inhomogeneous [12].) But the same effect can also be anticipated for waves in plasmas, e.g., radiofrequency (RF) waves in tokamaks. In fact, since ϵ for RF waves in laboratory plasma is typically larger than that for quantum and optical waves, the polarization-driven bending of ray trajectories in this case can be more important and perhaps should be taken into account in practical ray-tracing simulations. But *ad hoc* theories of polarization effects available from optics are inapplicable to plasma waves, which have much more complicated dispersion and thus require more fundamental approaches. Thus, a new general theory is needed that would allow to calculate the polarization-bending of ray trajectories for plasma waves.

Previously, relevant work was done in Refs. [13, 14], where a systematic procedure was proposed to asymptotically diagonalize the wave dispersion operator for general vector waves. Here polarization effects emerge as $\mathcal{O}(\epsilon)$ corrections to the GO dispersion relation. However, the analysis in Refs. [13, 14] is limited to adiabatic dynamics, i.e., mode conversion (linear exchange of quanta between different branches of the local dispersion relation) is excluded. Mode conversion was extensively studied separately, e.g., in Refs. [15–20]. However, these works considered wave modes that are resonant in small, localized regions of the wave phase space. Hence, the nonadiabatic dynamics was formulated as an asymptotic scattering problem between two wave modes, so no polarization-driven bending of ray trajectories was included.

The first general theory that captures the polarization-driven bending of ray trajectories and mode conversion simultaneously was proposed in Ref. [21]. The theory was successfully benchmarked against simpler *ad hoc* theories of the Hall effect of light [22]. However, the theory in Ref. [21] is still limited since it requires that the wave equation be brought to a certain (multisymplectic) form resembling the Dirac equation. Albeit any nondissipative vector wave allows for such representation [21, 23], casting the wave dynamics into the specific framework in Ref. [21] can be complicated. Thus, practical applications require a more flexible formulation that do not rely on this specific framework.

Here we propose such a theory. In addition to generalizing the results of Ref. [21], we also introduce, in a unified context and an instructive manner (hopefully),

B. Extended wave function

As shown in Refs. [21, 23], reduced models of wave propagation are convenient to develop when the action is of the symplectic form; namely,

$$\mathcal{S}_{\text{symplectic}} \doteq \langle \Psi | (\hat{p}_0 \mathbb{I}_{\bar{N}} - \hat{\mathcal{H}}) | \Psi \rangle \,, \tag{6}$$

where $\hat{p}_0 = i\partial_t$ (in the x-representation) and "the wave Hamiltonian" $\hat{\mathcal{H}} = \mathcal{H}(\hat{x}, \hat{\mathbf{p}})$ is some Hermitian operator that is local in time, i.e., commutes with $\hat{t}$. (For extended discussions, see Refs. [23, 31].

In order to cast the general action (4) into the symplectic form (6), let us perform the so-called Feynman reparameterization [32, 33] that lifts the wave dynamics governed by Eq. (4) from $\mathbb{R}^4$ to $\mathbb{R}^5$. Specifically, we let the wave field depend on spacetime and on some parameter τ so that $\Psi(\tau, x) = \langle x | \Psi(\tau, x) \rangle$. Then, we consider the following "extended" action:

$$\mathcal{S}_{\text{X}} \doteq \int d\tau \, L, \tag{7}$$

where $L \doteq L_\tau + L_D$,

$$L_\tau \doteq -(i/2) \left(\langle \Psi | \partial_\tau \Psi \rangle - \text{c.c.} \right), \tag{8a}$$

$$L_D \doteq \langle \Psi | \hat{\mathcal{D}} | \Psi \rangle \,, \tag{8b}$$

and $\partial_\tau \Psi(\tau, x) = \langle x | \partial_\tau \Psi \rangle$. The ELE corresponding to the action $\mathcal{S}_{\text{X}}$ is given by

$$i\partial_\tau \, |\Psi\rangle = \hat{\mathcal{D}} \, |\Psi\rangle \,. \tag{9}$$

Note that $\hat{\mathcal{D}}$ acts as a Hamiltonian operator in the extended variable space. Hence, the dynamics of the original system (5) can be considered as a special case of the dynamics governed by Eq. (9) that corresponds to a steady state with respect to the parameter τ; i.e., $\partial_\tau \Psi = 0$. The advantage of the representation (7) is that the system action has the manifestly symplectic form, so we can proceed as follows.

IV. EIGENMODE REPRESENTATION

A. Variable transformation

We introduce a unitary τ-independent transformation $\hat{\mathcal{T}}$ that maps $|\Psi\rangle$ to some $\bar{N}$-dimensional abstract vector $|\bar{\psi}\rangle$ yet to be defined:

$$|\Psi\rangle = \hat{\mathcal{T}} \, |\bar{\psi}\rangle \,. \tag{10}$$

Inserting Eq. (10) into Eqs. (8) leads to

$$L_\tau = -(i/2) \left(\langle \bar{\psi} | \partial_\tau \bar{\psi} \rangle - \text{c.c.} \right), \tag{11a}$$

$$L_D = \langle \bar{\psi} | \hat{\mathcal{D}}_{\text{eff}} | \bar{\psi} \rangle \,, \tag{11b}$$

where $\hat{\mathcal{D}}_{\text{eff}} \doteq \hat{\mathcal{T}}^\dagger \hat{\mathcal{D}} \hat{\mathcal{T}}$. In what follows, we seek to construct $\hat{\mathcal{T}}$ such that the operator $\hat{\mathcal{D}}_{\text{eff}}$ is simplified in a manner specified below.

B. Weyl representation

Let us consider Eq. (11b) in the Weyl representation. (Readers who are not familiar with the Weyl calculus are encouraged to read Appendix A before continuing further.) In this representation, L_D is written as [28]

$$L_D = \text{Tr} \int \mathrm{d}^4x \, \mathrm{d}^4p \, D_{\text{eff}}(x,p) W(\tau, x, p), \tag{12}$$

where 'Tr' represents the matrix trace. The Wigner tensor $W(\tau, x, p)$ corresponding to $|\bar{\psi}\rangle$ is defined as

$$W^m{}_n(\tau, x, p) \doteq \int \frac{\mathrm{d}^4 s}{(2\pi)^4} \, e^{ip\cdot s} \, \langle x + \frac{s}{2} | \bar{\psi}^m \rangle \, \langle \bar{\psi}_n | x - \frac{s}{2} \rangle \,, \tag{13}$$

and $D_{\text{eff}}(x,p)$ is the Weyl symbol [Eq. (A1)] corresponding to the operator $\hat{\mathcal{D}}_{\text{eff}}$. It can be written explicitly as

$$D_{\text{eff}}(x,p) = [T^\dagger](x,p) \star D(x,p) \star T(x,p), \tag{14}$$

where '$\star$' is the Moyal product [Eq. (A6)] and $D(x,p)$, $T(x,p)$, and $[T^\dagger](x,p)$ are the Weyl symbols corresponding to $\hat{\mathcal{D}}$, $\hat{\mathcal{T}}$, and $\hat{\mathcal{T}}^\dagger$, respectively. Also, the Weyl representation of the unitary condition, $\hat{T}^\dagger \hat{T} = \hat{\mathbb{I}}_{\bar{N}}$, is

$$[T^\dagger](x,p) \star T(x,p) = \mathbb{I}_{\bar{N}}, \tag{15}$$

which will be used below.

C. Eigenmode representation

Let us assume that the symbols D_{eff} and T can be expanded in powers of the GO parameter

$$\epsilon = \max \left\{ \frac{1}{\omega \mathrm{t}}, \frac{1}{|\mathbf{k}| \ell} \right\} \ll 1, \tag{16}$$

where ω and $|\mathbf{k}|$ are understood as the wave frequency and wave number, respectively. Also, t and ℓ are the characteristic time and length scales of the background medium, correspondingly. Hence, we write

$$D_{\text{eff}}(x,p) = \Lambda(x,p) + \epsilon U(x,p) + \mathcal{O}(\epsilon^2), \tag{17a}$$

$$T(x,p) = T_0(x,p) + \epsilon T_1(x,p) + \mathcal{O}(\epsilon^2), \tag{17b}$$

where (Λ, U, T_0, T_1) are $\bar{N} \times \bar{N}$ matrices of order unity.

To the lowest-order in ϵ, the Moyal products in Eqs. (14) and (15) reduce to ordinary products, so

$$\Lambda(x,p) = [T_0^\dagger](x,p) D(x,p) T_0(x,p), \tag{18}$$

$$[T_0^\dagger](x,p) T_0(x,p) = \mathbb{I}_{\bar{N}}. \tag{19}$$

By properties of the Weyl transformation, the fact that $\hat{\mathcal{D}}$ is a Hermitian operator ensures that $D(x,p)$ is a Hermitian matrix. Hence, $D(x,p)$ has $\bar{N}$ orthonormal eigenvectors $\mathbf{e}_q(x,p)$, which correspond to some real eigenvalues will be dropped, and we adopt

$$L_\tau \doteq -(i/2)\int \mathrm{d}^4x\,\left[\psi^\dagger(\partial_\tau\psi)-(\partial_\tau\psi^\dagger)\psi\right], \tag{29a}$$

$$L_D \doteq \mathrm{Tr}\int \mathrm{d}^4x\,\mathrm{d}^4p\,[[\Lambda+\epsilon U]]\,[[W]]. \tag{29b}$$

Here ψ is a complex-valued function with N components, and $[[W]]$ is the $N\times N$ Wigner tensor with elements

$$[[W]]^m{}_n(\tau,x,p)=\int\frac{\mathrm{d}^4s}{(2\pi)^4}\,e^{ip\cdot s}\,\langle x+\frac{s}{2}|\psi^m\rangle\,\langle\psi_n|x-\frac{s}{2}\rangle. \tag{30}$$

Since we consider the coupled dynamics of some N resonant modes, only N columns of T_0 actually contribute to $[[D_{\text{eff}}]]$. For clarity, let us denote the resonant eigenmodes as $\mathbf{e}_q$ with indices $q=1,...,N$. Then, in order to calculate $[[U]]$, one can use Eq. (25). After block-diagonalizing U and introducing the $N\times N$ matrix

$$\Xi(x,p)=[\mathbf{e}_1(x,p),\,\dots,\,\mathbf{e}_N(x,p)], \tag{31}$$

one obtains

$$[[U]]=\frac{i}{4}\{\Xi_0^\dagger,\Xi_0\}\Lambda+\frac{i}{4}\Lambda\{\Xi_0^\dagger,\Xi_0\}+\frac{i}{2}\{\Lambda,\Xi_0\}_{\Xi_0^\dagger}+\frac{i}{2}\{\Xi_0^\dagger,\Lambda\}_{\Xi_0}-\frac{i}{2}\{\Xi_0^\dagger,\Xi_0\}_D, \tag{32}$$

which is a $N\times N$ Hermitian matrix.

Furthermore, it is convenient to split $[[D_{\text{eff}}]]$ as follows:

$$[[D_{\text{eff}}]]=\lambda\mathbb{I}_N+\epsilon\mathcal{U}, \tag{33}$$

where $\lambda\doteq N^{-1}\mathrm{Tr}\,[[D_{\text{eff}}]]$ is the average eigenvalue of the block and $\mathcal{U}\doteq[[D_{\text{eff}}]]-\lambda\mathbb{I}_{\bar{N}}$ is the remaining traceless part of $[[D_{\text{eff}}]]$.

In the special case when all $\lambda^{(q)}$ within the block are identical and $[[U]]$ is traceless, then $\Lambda=\lambda\mathbb{I}_N$, and $\mathcal{U}=[[U]]$. We call such modes degenerate. Then, the expression (32) for $[[U]]$ simplifies, and one obtains

$$\begin{aligned}\mathcal{U}(x,p)&=\frac{i}{4}\{\Xi^\dagger,\Xi\}\lambda+\frac{i}{4}\lambda\{\Xi^\dagger,\Xi\}+\frac{i}{2}\{\lambda,\Xi\}_{\Xi^\dagger}+\frac{i}{2}\{\Xi^\dagger,\lambda\}_{\Xi}-\frac{i}{2}\{\Xi^\dagger,\Xi\}_D\\&=-\frac{1}{2i}\Xi^\dagger\{\lambda,\Xi\}-\frac{1}{2i}\{\Xi^\dagger,\lambda\}\Xi+\frac{1}{2i}\{\Xi^\dagger,\Xi\}_D-\frac{1}{2i}\{\Xi^\dagger,\Xi\}_\lambda\\&=-\frac{1}{2i}\left[\Xi^\dagger\{\lambda,\Xi\}-\text{h.c.}\right]+\frac{1}{2i}\left[(\partial_p\Xi^\dagger)(D-\lambda\mathbb{I}_N)(\partial_x\Xi)-\text{h.c.}\right]\\&=-\left[\Xi^\dagger\{\lambda,\Xi\}\right]_A+\left[(\partial_p\Xi^\dagger)(D-\lambda\mathbb{I}_N)(\partial_x\Xi)\right]_A,\end{aligned} \tag{34}$$

where we used the bracket introduced in Eq. (24) and the subscript 'A' denotes "anti-Hermitian part;" i.e., for any matrix M, then $M_A\doteq(M-M^\dagger)/(2i)$. The expression in Eq. (34) can also be written more explicitly as

$$\mathcal{U}(x,p)=\left(-\frac{\partial\lambda}{\partial p_\mu}\right)\left(\Xi^\dagger\frac{\partial\Xi}{\partial x^\mu}\right)_A+\left(\frac{\partial\lambda}{\partial x^\mu}\right)\left(\Xi^\dagger\frac{\partial\Xi}{\partial p_\mu}\right)_A+\left[\frac{\partial\Xi^\dagger}{\partial p_\mu}(D-\lambda\mathbb{I}_N)\frac{\partial\Xi}{\partial x^\mu}\right]_A. \tag{35}$$

Examples of degenerate systems where these simplified formulas are applicable include spin-1/2 particles [21, 24] and EM waves propagating in isotropic dielectrics [22].

B. Parameterization of the action

In order to derive the corresponding ELEs, let us adopt the following parameterization:

$$\psi(\tau,x)=\sqrt{\mathcal{I}(\tau,x)}\,z(\tau,x)\,e^{i\theta(\tau,x)}. \tag{36}$$

Here $\theta(\tau,x)$ is a real variable that serves as the rapid phase common for all N modes (remember that all modes within the block of interest are approximately resonant to each other). Also, $\mathcal{I}(\tau,x)$ is a real function, and $z(\tau,x)$ is a N-dimensional complex unit vector ($z^\dagger z=1$), whose components describe the amount of quanta in the corresponding modes. (Since we parameterize the N-dimensional complex vector ψ by the N-dimensional complex vector z plus two independent real functions θ and $\mathcal{I}$, not all components of z are truly independent. For an extended discussion, see Ref. [21].)

After substituting the ansatz (36) into Eq. (29a), the Lagrangian L_τ is given by

$$L_\tau=\int\mathrm{d}^4x\,\mathcal{I}\left[\partial_\tau\theta-(i\epsilon/2)(z^\dagger\partial_\tau z-\text{c.c.})\right]. \tag{37}$$

(Here we formally introduce ϵ to denote that z is a slowly-varying quantity; however, this ordering parameter will be removed later.) Now, we calculate the Wigner tensor (30). Substituting Eq. (36) into Eq. (30), we obtain dispersion relation. For an in-depth discussion of these equations, see, e.g., Refs. [1, 4].

B. Point-particle model

The ray equations corresponding to the above field equations can be obtained as the point-particle limit. In this limit, $\mathcal{I}$ can be approximated with a delta function

$$\mathcal{I}(\tau, x) = \mathcal{I}_0 \delta^4(x - X(\tau)). \quad (46)$$

Here $\mathcal{I}_0$ denotes the total action, which is conserved according to Eq. (45a). The value of $\mathcal{I}_0$ is not essential below so we adopt $\mathcal{I}_0 = 1$ for brevity.

In this representation, the wave packet is located at the position $X^\mu(\tau)$ in space-time, and the independent parameter is τ. [This means that at a given τ, the wave packet is located at the spatial point $\mathbf{X}(\tau)$ at time $t(\tau)$.] When inserting Eq. (46) into Eq. (43), the first term in the action gives the following:

$$\begin{aligned}
&\int \mathrm{d}\tau\, \mathrm{d}^4x\, \mathcal{I}\, \partial_\tau \theta \\
&= \int \mathrm{d}\tau\, \mathrm{d}^4x\, \delta^4(x - X(\tau))\, \partial_\tau \theta(\tau, x) \\
&= -\int \mathrm{d}\tau\, \mathrm{d}^4x\, \theta(\tau, x)[\partial_\tau \delta^4(x - X(\tau))] \\
&= \int \mathrm{d}\tau\, \mathrm{d}^4x\, \theta(\tau, x)[\dot{X}^\mu(\tau)\partial_\mu \delta^4(x - X(\tau))] \\
&= -\int \mathrm{d}\tau\, \mathrm{d}^4x\, \partial_\mu \theta(\tau, x)\dot{X}^\mu(\tau)\delta^4(x - X(\tau)) \\
&= \int \mathrm{d}\tau\, P_\mu(\tau)\dot{X}^\mu(\tau), \qquad (47)
\end{aligned}$$

where $P_\mu(\tau) \doteq -\partial_\mu \theta(\tau, X(\tau))$. Similarly,

$$\int \mathrm{d}^4x\, \delta^4(x - X(\tau))\lambda(x, -\partial\theta) = \lambda(X(\tau), P(\tau)). \quad (48)$$

Thus, the point-particle action is expressed as

$$\mathcal{S}_{\mathrm{GO}} = \int \mathrm{d}\tau\, \Big[P(\tau) \cdot \dot{X}(\tau) + \lambda(X, P)\Big]. \quad (49)$$

This is a covariant action, where $X^\mu(\tau)$ and $P_\mu(\tau)$ serve as canonical coordinates and canonical momenta, respectively. Treating X and P as independent variables leads to ELEs matching Hamilton's covariant equations

$$\delta P_\mu : \quad \frac{\mathrm{d}X^\mu}{\mathrm{d}\tau} = -\frac{\partial \lambda}{\partial P_\mu}, \quad (50a)$$

$$\delta X^\mu : \quad \frac{\mathrm{d}P_\mu}{\mathrm{d}\tau} = \frac{\partial \lambda}{\partial X^\mu}. \quad (50b)$$

These are the commonly known ray equations; for instance, see Ref. [1]. They can also be written as

$$\frac{\mathrm{d}X^0}{\mathrm{d}\tau} = -\frac{\partial \lambda}{\partial P_0}, \qquad \frac{\mathrm{d}\mathbf{X}}{\mathrm{d}\tau} = \frac{\partial \lambda}{\partial \mathbf{P}},$$
$$\frac{\mathrm{d}P^0}{\mathrm{d}\tau} = \frac{\partial \lambda}{\partial X^0}, \qquad \frac{\mathrm{d}\mathbf{P}}{\mathrm{d}\tau} = -\frac{\partial \lambda}{\partial \mathbf{X}}.$$

Note that the first term in the integrand in Eq. (50) represents the symplectic part of the canonical phase-space Lagrangian, and the second term represents the Hamiltonian part. Since the Hamiltonian part $\lambda(X, P)$ does not depend explicitly on τ, then $\mathrm{d}\lambda(X, P)/\mathrm{d}\tau = 0$ along the ray trajectories. Thus, the ray dynamics lies on the dispersion manifold defined by

$$\lambda(X, P) = 0. \quad (51)$$

VII. EXTENDED GEOMETRICAL OPTICS

In this section, we explore the polarization effects determined by the $\mathcal{O}(\epsilon)$-accurate Lagrangian (42). For the sake of conciseness, we only discuss the point-particle ray dynamics. For an overview of the continuous-wave model, see Ref. [21].

A. Point-particle model

The ray equations with polarization effects included can be obtained as a point-particle limit of the Lagrangian (42). As in Sec. VI B, we approximate the wave packet to a single point in spacetime [Eq. (46)]. As in Refs. [21, 25], the Lagrangian (42) can be replaced by a point-particle Lagrangian so the action is

$$\begin{aligned}
\mathcal{S}_{\mathrm{XGO}} = \int \mathrm{d}\tau \Big[& P \cdot \dot{X} - (i/2)(Z^\dagger \dot{Z} - \dot{Z}^\dagger Z) \\
& + \lambda(X, P) + Z^\dagger \mathcal{U}(X, P) Z \Big], \qquad (52)
\end{aligned}$$

where $Z(\tau) \doteq z(\tau, X(\tau))$ and we dropped the GO ordering parameter ϵ. In the complex representation, Z and $Z^\dagger$ are canonical conjugate, and

$$Z^\dagger(\tau) Z(\tau) = 1. \quad (53)$$

Even though the components of Z are not independent by definition (Sec. V B), it can be shown [21] that treating them as independent in this point-particle model leads to correct results provided that the initial conditions satisfy Eq. (53). Hence, the independent variables in $\mathcal{S}_{\mathrm{XGO}}$ are $(X, P, Z, Z^\dagger)$, and the corresponding ELEs are

$$\delta P_\mu : \quad \frac{\mathrm{d}X^\mu}{\mathrm{d}\tau} = -\frac{\partial \lambda}{\partial P_\mu} - Z^\dagger \frac{\partial \mathcal{U}}{\partial P_\mu} Z, \quad (54a)$$

$$\delta X^\mu : \quad \frac{\mathrm{d}P_\mu}{\mathrm{d}\tau} = \frac{\partial \lambda}{\partial X^\mu} + Z^\dagger \frac{\partial \mathcal{U}}{\partial X^\mu} Z, \quad (54b)$$

$$\delta Z^\dagger : \quad \frac{\mathrm{d}Z}{\mathrm{d}\tau} = -i\mathcal{U} Z, \quad (54c)$$

$$\delta Z : \quad \frac{\mathrm{d}Z^\dagger}{\mathrm{d}\tau} = i Z^\dagger \mathcal{U}. \quad (54d)$$

Together with Eqs. (31)-(33), Eqs. (54) form a complete set of equations. The first terms on the right-hand side

Let us write Eqs. (62) using the abstract Hilbert space notation. Let $|\mathbf{v}\rangle$ be a state vector representing the velocity field such that $\mathbf{v}(x) = \langle x|\mathbf{v}\rangle$. Likewise, we introduce $|\mathbf{E}\rangle$ and $|\mathbf{B}\rangle$ as the state vectors of $\mathbf{E}(x)$ and $\mathbf{B}(x)$, respectively. Then, Eqs. (62) can be written as follows:

$$\hat{p}_0 |\bar{\mathbf{v}}\rangle = i\hat{\omega}_p |\mathbf{E}\rangle - (\boldsymbol{\alpha}\cdot\hat{\boldsymbol{\Omega}})|\bar{\mathbf{v}}\rangle , \tag{63a}$$

$$\hat{p}_0 |\mathbf{E}\rangle = -i\hat{\omega}_p |\bar{\mathbf{v}}\rangle + ic(\boldsymbol{\alpha}\cdot\hat{\mathbf{p}})|\mathbf{B}\rangle , \tag{63b}$$

$$\hat{p}_0 |\mathbf{B}\rangle = -ic(\boldsymbol{\alpha}\cdot\hat{\mathbf{p}})|\mathbf{E}\rangle , \tag{63c}$$

where $\hat{\omega}_p \doteq \omega_p(\hat{\mathbf{x}})$ and $\hat{\boldsymbol{\Omega}} \doteq \boldsymbol{\Omega}(\hat{\mathbf{x}})$. (As a reminder, $\hat{p}_0 = i\partial_t$ and $\hat{\mathbf{p}} = -i\nabla$ are the components of the four-momentum operator in the x-representation.) Also, $\boldsymbol{\alpha} \doteq (\alpha^1, \alpha^2, \alpha^3)$ are 3×3 Hermitian matrices [39]

$$\alpha^1 \doteq \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \tag{64a}$$

$$\alpha^2 \doteq \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}, \tag{64b}$$

$$\alpha^3 \doteq \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{64c}$$

These matrices serve as generators for the vector product. Namely, for any two column vectors $\mathbf{A}$ and $\mathbf{B}$, one has

$$(\boldsymbol{\alpha}\cdot\mathbf{A})\mathbf{B} = i\mathbf{A}\times\mathbf{B}, \tag{65a}$$

$$\mathbf{A}^{\mathrm{T}}\alpha^j\mathbf{B} = -i(\mathbf{A}\times\mathbf{B})^j, \tag{65b}$$

where the superscript 'T' denotes the matrix transpose.

The next step is to construct a dispersion operator for the electric field state $|\mathbf{E}\rangle$. Starting from Eq. (63a), we solve for the velocity field in terms of the electric field. Hence, we formally obtain the following:

$$\begin{aligned} |\bar{\mathbf{v}}\rangle &= i\hat{\omega}_p(\hat{p}_0\mathbb{I}_3 + \boldsymbol{\alpha}\cdot\hat{\boldsymbol{\Omega}})^{-1}|\mathbf{E}\rangle \\ &= i\hat{\omega}_p\left[\frac{1}{\hat{p}_0} - \frac{\boldsymbol{\alpha}\cdot\hat{\boldsymbol{\Omega}}}{\hat{p}_0^2 - \hat{\Omega}^2} + \frac{(\boldsymbol{\alpha}\cdot\hat{\boldsymbol{\Omega}})^2}{\hat{p}_0(\hat{p}_0^2 - \hat{\Omega}^2)}\right]|\mathbf{E}\rangle , \end{aligned} \tag{66}$$

where $\hat{\Omega} \doteq |\boldsymbol{\Omega}(\hat{\mathbf{x}})|$. Similarly, we obtain $|\mathbf{B}\rangle = -ic(\boldsymbol{\alpha}\cdot\hat{\mathbf{p}})\hat{p}_0^{-1}|\mathbf{E}\rangle$ from Eq. (63c). Substituting these results into Eq. (63b), we obtain

$$\hat{\mathfrak{D}}\,|\mathbf{E}\rangle = 0, \tag{67}$$

where

$$\hat{\mathfrak{D}} \doteq -\hat{p}_0^2 + (\boldsymbol{\alpha}\cdot\hat{\mathbf{p}})^2 + \hat{\omega}_p^2 - \frac{\hat{\omega}_p^2\hat{p}_0(\boldsymbol{\alpha}\cdot\hat{\boldsymbol{\Omega}})}{\hat{p}_0^2 - \hat{\Omega}^2} + \frac{\hat{\omega}_p^2(\boldsymbol{\alpha}\cdot\hat{\boldsymbol{\Omega}})^2}{\hat{p}_0^2 - \hat{\Omega}^2} \tag{68}$$

serves as the dispersion operator for $|\mathbf{E}\rangle$. (For convenience, we let $c = 1$.) Since $\omega_p(\mathbf{x})$ and $\boldsymbol{\Omega}(\mathbf{x})$ are independent of time, then $\hat{p}_0$ commutes with $\hat{\omega}_p$ and $\hat{\boldsymbol{\Omega}}$, so $\hat{\mathfrak{D}}$ is manifestly Hermitian.

Thus, the corresponding action (4) is for the electric field is $\mathcal{S} = \langle\mathbf{E}|\hat{\mathfrak{D}}|\mathbf{E}\rangle$, and the extended action (7) is

$$\mathcal{S}_{\mathrm{X}} \doteq \int \mathrm{d}\tau \left[-\frac{i}{2}\left(\langle\mathbf{E}|\partial_\tau\mathbf{E}\rangle - \mathrm{c.\,c.}\right) + \langle\mathbf{E}|\hat{\mathfrak{D}}|\mathbf{E}\rangle\right]. \tag{69}$$

Note that $\mathbf{E}$ is a three-dimensional vector field, so $\bar{N} = 3$.

B. EM waves in weakly magnetized plasma

We now follow the procedure given in Secs. IV and V to block-diagonalize the dispersion operator. The Weyl symbol of $\hat{\mathfrak{D}}$ is

$$D \doteq -p_0^2 + (\boldsymbol{\alpha}\cdot\mathbf{p})^2 + \omega_p^2 - \frac{\omega_p^2 p_0(\boldsymbol{\alpha}\cdot\boldsymbol{\Omega})}{p_0^2 - \Omega^2} + \frac{\omega_p^2(\boldsymbol{\alpha}\cdot\boldsymbol{\Omega})^2}{p_0^2 - \Omega^2}. \tag{70}$$

For the sake of simplicity, we consider the case of a wave propagating in a weakly magnetized plasma. (The general case will be described in a separate paper.) Thus, supposing that the typical wave frequency is much larger than the gyrofrequency ($\omega \sim p_0 \gg \Omega$), we expand the dispersion symbol (70) in powers of Ω/p_0:

$$D \simeq D_0 + D_1 + \mathcal{O}(\Omega^2/p_0^2), \tag{71}$$

where

$$D_0(\mathbf{x}, p) = -p_0^2 + (\boldsymbol{\alpha}\cdot\mathbf{p})^2 + \omega_p^2(\mathbf{x}), \tag{72a}$$

$$D_1(\mathbf{x}, p_0) = -\omega_p^2(\mathbf{x})(\boldsymbol{\alpha}\cdot\boldsymbol{\Omega})/p_0. \tag{72b}$$

To simplify the following calculation, we assume that $D_1 \sim \mathcal{O}(\Omega/p_0)$ is comparable in magnitude to the GO parameter ϵ, but this is not essential. Hence, we will consider D_1 as a perturbation only.

Following Sec. IV C, the next step is to identify the eigenvalues and eigenmodes of the dispersion symbol $D_0(\mathbf{x}, p)$. The corresponding eigenvalues are

$$\lambda^{(1)}(\mathbf{x}, p) = -p\cdot p + \omega_p^2(\mathbf{x}), \tag{73a}$$

$$\lambda^{(2)}(\mathbf{x}, p) = -p\cdot p + \omega_p^2(\mathbf{x}), \tag{73b}$$

$$\lambda^{(3)}(\mathbf{x}, p_0) = -p_0^2 + \omega_p^2(\mathbf{x}), \tag{73c}$$

where $p\cdot p = p_0^2 - \mathbf{p}^2$. These eigenvalues correspond to the dispersion relations of two transverse EM waves and of longitudinal Langmuir oscillations, respectively. The matrix T_0 defined in Eq. (20) is given by

$$T_0(\mathbf{p}) = [\,\mathbf{e}_1(\mathbf{p}),\, \mathbf{e}_2(\mathbf{p}),\, \mathbf{e}_{\mathbf{p}}(\mathbf{p})\,], \tag{74}$$

where $\mathbf{e}_1(\mathbf{p})$ and $\mathbf{e}_2(\mathbf{p})$ are any two orthonormal vectors in the plane normal to $\mathbf{e}_{\mathbf{p}}(\mathbf{p}) \doteq \mathbf{p}/|\mathbf{p}|$. A right-hand convention is adopted such that $\mathbf{e}_1\times\mathbf{e}_2 = \mathbf{e}_{\mathbf{p}}$. One can easily verify that these vectors are indeed eigenvectors of $D_0(\mathbf{x}, p)$. For example,

$$\begin{aligned} D_0\,\mathbf{e}_1 &= [-p_0^2 + (\boldsymbol{\alpha}\cdot\mathbf{p})(\boldsymbol{\alpha}\cdot\mathbf{p}) + \omega_p^2]\,\mathbf{e}_1 \\ &= (-p_0^2 + \omega_p^2)\,\mathbf{e}_1 - \mathbf{p}\times(\mathbf{p}\times\mathbf{e}_1) \\ &= (-p_0^2 + \mathbf{p}^2 + \omega_p^2)\,\mathbf{e}_1 \\ &= \lambda^{(1)}\,\mathbf{e}_1, \end{aligned} \tag{75}$$

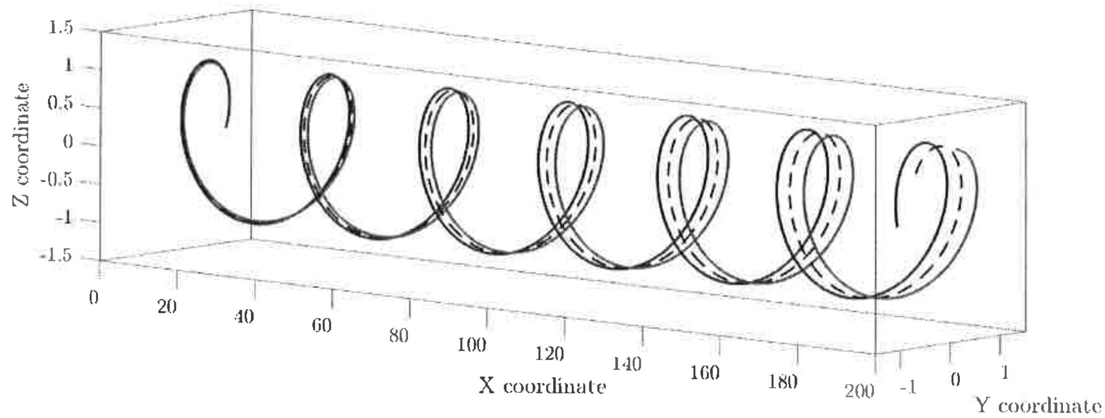

FIG. 1: Comparison between ray trajectories calculated using the equations of traditional GO [Eqs. (101), dashed line] and extended GO [Eqs. (96)]. The blue and red lines represent the ray trajectories for the right-hand and left-hand polarized rays, respectively. For simplicity, nonmanetized plasma is considered, so the Faraday effect is absent. The plasma frequency is given by $\omega_p^2(\mathbf{x}) = y^2 + z^2$. The initial location of the ray trajectories is $\mathbf{X}_0 = (0,1,0)$, and the initial momentum is $\mathbf{P}_0 = (5,0,1)$. (The units are arbitrary, since the figure is a general illustration only.) For this simulation, the GO parameter is roughly $\epsilon \sim 1/\left|\mathbf{P}_0\right| \sim 0.2$. Due to the radial gradient in the plasma frequency, the wave rays follow helical trajectories along the x axis.

where σ_z is another Pauli matrix,

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{90}$$

Here $\Gamma_\pm(\tau)$ represent the wave quanta belonging to the right-hand and left-hand circularly polarized modes, respectively (as defined from the point of view of the source). Also, Γ is normalized such that $\Gamma^\dagger(\tau)\Gamma(\tau) = 1$.

Treating $X(\tau)$, $P(\tau)$, $\Gamma(\tau)$, and $\Gamma^\dagger(\tau)$ as independent variables, we obtain the following ELEs:

$$\delta P_\mu : \quad \frac{\mathrm{d}X^\mu}{\mathrm{d}\tau} = 2P^\mu - \frac{\partial \Sigma}{\partial P_\mu}\Gamma\sigma_z\Gamma, \tag{91a}$$

$$\delta X^\mu : \quad \frac{\mathrm{d}P_\mu}{\mathrm{d}\tau} = \frac{\partial \omega_p^2}{\partial X^\mu} + \frac{\partial \Sigma}{\partial X^\mu}\Gamma\sigma_z\Gamma, \tag{91b}$$

$$\delta \Gamma^\dagger : \quad \frac{\mathrm{d}\Gamma}{\mathrm{d}\tau} = -i\Sigma\sigma_z\Gamma, \tag{91c}$$

$$\delta \Gamma : \quad \frac{\mathrm{d}\Gamma^\dagger}{\mathrm{d}\tau} = i\Gamma\Sigma\sigma_z. \tag{91d}$$

Together with Eq. (85), Eqs. (91) form a complete set of equations. The first terms on the right-hand side of Eqs. (91a) and (91b) describe the ray dynamics in the GO limit. The second terms describe the coupling of the mode polarization and the ray curvature.

D. Restating the Faraday effect

To better understand the polarization equations, let us rewrite Eq. (91c) as an equation in the basis of linearly polarized modes:

$$\dot{Z} = Q\dot{\Gamma} = -i\Sigma Q\sigma_z\Gamma = -i\Sigma(Q\sigma_z Q^{-1})Z = -i\Sigma\sigma_y Z. \tag{92}$$

[This equation could also be obtained if the ray equations were derived directly from the action (84).] Since Σ is a scalar and σ_y is constant, this can be readily integrated, yielding [40]

$$Z(\tau) = \exp(-i\Theta\sigma_y)Z_0 = (\mathbb{I}_2\cos\Theta - i\sigma_y\sin\Theta)Z_0, \tag{93}$$

where $\Theta(\tau) \doteq \int_0^\tau \mathrm{d}\tau'\, \Sigma(\mathbf{X}(\tau'), P(\tau'))$ is the polarization precession angle and $Z_0 \doteq Z(\tau = 0)$. This result can be also be expressed explicitly as follows:

$$Z(\tau) = \begin{pmatrix} \cos\Theta & -\sin\Theta \\ \sin\Theta & \cos\Theta \end{pmatrix} Z_0. \tag{94}$$

It is seen that the polarization of the EM field rotates at the rate $\Sigma(t)$ in the reference frame defined by the basis vectors $(\mathbf{e}_1, \mathbf{e}_2)$. The first term in Eq. (85) is identified as the rate of change of the wave Berry phase [6, 7]. (In optics, the rotation of the polarization plane caused by the Berry phase is also known as the *Rytov rotation* [41, 42].) The second term in Eq. (85) is identified as the rate of change due to Faraday rotation.

E. Dynamics of pure states

If a ray corresponds to a strictly circular polarization such that $\sigma_z\Gamma = \pm\Gamma$, the action (89) can be simplified to $\mathcal{S}_{\mathrm{XGO}} = \int \mathrm{d}\tau\, L_\pm$, where the Lagrangian is given by

$$L_\pm = P\cdot\dot{X} - P\cdot P + \omega_p^2(\mathbf{X}) \pm \Sigma(\mathbf{X}, P). \tag{95}$$

Here the Lagrangian $L_\pm$ governs the propagation of right-hand and left-hand polarization modes, respectively. The

Hence, with the use of the noncanonical coordinates (x, p), the equations of motion no longer depend on the specific choice of $\mathbf{F}(\mathbf{p})$; i.e., they are invariant with respect to the choice (80) of vectors $\mathbf{e}_1$ and $\mathbf{e}_2$. Note that the same equations could be obtained directly from the point-particle limit of Eq. (76), if one substitutes $-\nabla\lambda = \dot{\mathbf{p}}$. For an extended discussion of pure states governed by noncanonical Lagrangians, see Ref. [13].

IX. CONCLUSIONS

Even diffraction aside, the well-known equations of geometrical optics (GO) are not entirely accurate. Traditional GO treats wave rays as classical particles, which are completely described by their coordinates and momenta, but rays have another degree of freedom, namely, polarization. The polarization degree of freedom manifests itself as an effective (classical) spin that can be assigned to rays and affects the wave dynamics accordingly. The goal of this paper is to propose a general first-principle variational formulation that would capture all such effects simultaneously.

We consider general linear nondissipative waves determined by some dispersion operator $\hat{\mathfrak{D}}$. Using the Feynman reparameterization and Weyl calculus, we obtain a reduced Lagrangian for such waves. In contrast with the traditional GO Lagrangian, which has accuracy $\mathcal{O}(\epsilon^0)$ in the GO parameter ϵ, our Lagrangian is $\mathcal{O}(\epsilon)$-accurate, so it captures polarization effects, including both mode conversion and polarization-driven bending of ray trajectories. This effect has been known as the spin-orbital coupling in quantum physics and as the Hall effect of light in optics. Our theory extends its applicability to media with more complicated dispersion, such as plasmas. As an example, we apply the formulation to study the polarization-driven divergence of RF waves propagating in weakly magnetized plasma. The case of strongly magnetized plasma will be discussed in a separate paper.

Importantly, RF waves in laboratory plasmas can have ϵ much larger than quantum matter waves or optical waves. Thus, it stands to reason that polarization effects may be particularly significant in plasma physics. Assessing them quantitatively for plasmas of practical interest will be done in separate publications. Likewise, the method of including dissipation [26] in the above theory will also be described separately.

This work was supported by the U.S. DOE through Contract No. DE-AC02-09CH11466, by the NNSA SSAA Program through DOE Research Grant No. DE-NA0002948, and by the U.S. DOD NDSEG Fellowship through Contract No. 32-CFR-168a.

Appendix A: The Weyl transform

This appendix summarizes our conventions for the Weyl transform. (For more information, see the excellent reviews in Refs. [1, 43–45].) The Weyl symbol $A(x,p)$ of any given operator $\hat{A}$ is defined as

$$A(x,p) \doteq \int \mathrm{d}^4 s\, e^{ip\cdot s} \langle x + s/2|\hat{A}|x - s/2\rangle\,, \qquad \text{(A1)}$$

where $p\cdot s = p_0 s_0 - \mathbf{p}\cdot\mathbf{s}$ and the integrals span over $\mathbb{R}^4$. We shall refer to this description of the operators as a *phase-space representation* since the symbols (A1) are functions of the eight-dimensional phase space. Conversely, the inverse Weyl transformation is given by

$$\hat{A} = \int \frac{\mathrm{d}^4 x\, \mathrm{d}^4 p\, \mathrm{d}^4 s}{(2\pi)^4}\, e^{ip\cdot s/\epsilon} A(x,p)\, |x - s/2\rangle \langle x + s/2|\,. \qquad \text{(A2)}$$

Hence, $A(x,x') = \langle x|\hat{A}|x'\rangle$ can be expressed as

$$A(x,x') = \int \frac{\mathrm{d}^4 p}{(2\pi)^4}\, e^{ip\cdot(x'-x)} A\left(\frac{x+x'}{2}, p\right). \qquad \text{(A3)}$$

In the following, we outline a number of useful properties of the Weyl transform.

- For any operator $\hat{A}$, the trace $\mathrm{Tr}[\hat{A}] \doteq \int \mathrm{d}^4 x\, \langle x|\hat{A}|x\rangle$ can be expressed as

$$\mathrm{Tr}[\hat{A}] = \int \frac{\mathrm{d}^4 x\, \mathrm{d}^4 p}{(2\pi)^4} A(x,p). \qquad \text{(A4)}$$

- If $A(x,p)$ is the Weyl symbol of $\hat{A}$, then $A^\dagger(x,p)$ is the Weyl symbol of $\hat{A}^\dagger$. As a corollary, the Weyl symbol of a Hermitian operator is real.

- For any $\hat{\mathfrak{C}} = \hat{A}\hat{\mathfrak{B}}$, the corresponding Weyl symbols satisfy [46, 47]

$$C(x,p) = A(x,p) \star B(x,p). \qquad \text{(A5)}$$

Here '$\star$' refers to the *Moyal product*, which is given by

$$A(x,p) \star B(x,p) \doteq A(x,p) e^{i\hat{\mathcal{L}}/2} B(x,p), \qquad \text{(A6)}$$

and $\hat{\mathcal{L}}$ is the *Janus operator*

$$\hat{\mathcal{L}} \doteq \overleftarrow{\partial_p} \cdot \overrightarrow{\partial_x} - \overleftarrow{\partial_x} \cdot \overrightarrow{\partial_p} = \{\cdot,\cdot\}. \qquad \text{(A7)}$$

The arrows indicate the direction in which the derivatives act, and $A\hat{\mathcal{L}}B = \{A, B\}$ is the canonical Poisson bracket in the eight-dimensional phase space, namely,

$$\hat{\mathcal{L}} = \frac{\overleftarrow{\partial}}{\partial p^0}\frac{\overrightarrow{\partial}}{\partial x^0} - \frac{\overleftarrow{\partial}}{\partial x^0}\frac{\overrightarrow{\partial}}{\partial p^0} + \frac{\overleftarrow{\partial}}{\partial \mathbf{x}} \cdot \frac{\overrightarrow{\partial}}{\partial \mathbf{p}} - \frac{\overleftarrow{\partial}}{\partial \mathbf{p}} \cdot \frac{\overrightarrow{\partial}}{\partial \mathbf{x}}. \qquad \text{(A8)}$$

Provided that $A\hat{\mathcal{L}}B$ is small, one can use the following asymptotic expansion of the Moyal product:

$$A \star B \simeq AB + \frac{i}{2}\{A, B\}. \qquad \text{(A9)}$$

Princeton Plasma Physics Laboratory
Office of Reports and Publications

Managed by
Princeton University

under contract with the
U.S. Department of Energy
(DE-AC02-09CH11466)

P.O. Box 451, Princeton, NJ 08543
Phone: 609-243-2245
Fax: 609-243-2751

E-mail: publications@pppl.gov

Website: http://www.pppl.gov