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: February 7, 2017)

Even when neglecting diffraction effects, 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 vector-wave rays have another degree of freedom, namely, their polarization. The polarization degree of freedom manifests itself as an effective (classical) “wave spin” that can be assigned to rays and can affect the wave dynamics accordingly. A well-known manifestation of polarization dynamics is mode conversion, which is the linear exchange of quanta between different wave modes and can be interpreted as a rotation of the wave spin. Another, less-known polarization effect is the 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 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 \(\epsilon\) 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 interesting manifestation of such 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; namely, even in an isotropic dielectric, rays propagate somewhat differently depending on polarization if the dielectric is inhomogeneous (see, c.f. Refs. [7–11]). But the same effect can also be anticipated for waves in plasmas, e.g., radiofrequency (RF) waves in tokamaks. In fact, since \(\epsilon\) 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. However, ad hoc theories of polarization effects available from optics are inapplicable to plasma waves, which have more complicated dispersion and thus require more fundamental approaches. Thus, a different theory is needed that would allow the calculation of the polarization-bending of the ray trajectories for plasma waves and, also more broadly, waves in general linear media.

Relevant work was done in Refs. [12, 13], where a systematic procedure was proposed to asymptotically diagonalize the dispersion operator for linear vector waves. Polarization effects emerge as \(\mathcal{O}(\epsilon)\) corrections to the GO dispersion relation. However, this approach excludes mode conversion, i.e., the linear exchange of quanta between different branches of the local dispersion relation. Since the group velocities of the different branches eventually separate, mode conversion is typically followed by ray splitting and, in this particular context, was studied extensively (see, e.g., Refs. [14–20]). However, these works considered wave modes that are resonant in small, localized regions of phase space. Hence, the nonadiabatic dynamics was formulated as an asymptotic scattering problem between two wave modes, so the polarization-driven bending of ray trajectories was not included.

The main message of this work is that mode conversion and the polarization-driven bending of ray trajectories are two sides of the same coin and can be considered simultaneously within a unified theory. The first general theory that captures them both simultaneously was proposed in Ref. [21]. This theory was successfully benchmarked [22] against previous theories describing the Hall effect of light [7–11]. However, the formulation in Ref. [21] is still limited since it requires that the wave equation be brought to a certain (multisymplectic) form resembling the Dirac equation. Although any nondissipative vector wave allows for such representation in prin-
ciple [21, 23], casting the wave dynamics into the specific framework adopted 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, some of the related advances that were made recently in Refs. [24–26]. It is expected that the comprehensive analysis presented in this work will facilitate future practical implementations of the proposed theory, particularly in improving ray-tracing simulations.

B. Outline

We consider general linear nondissipative waves determined by some Hermitian dispersion operator. Using the Feynman reparameterization and the Weyl calculus, we obtain a reduced Lagrangian for such waves. In contrast with the traditional GO Lagrangian, which has an accuracy of $\mathcal{O}(\epsilon^0)$ in the GO parameter $\epsilon$, our Lagrangian is $\mathcal{O}(\epsilon^1)$-accurate, so it captures polarization effects. As an example, we apply the formulation to study polarization effects on the propagation of EM waves propagating in weakly magnetized plasma. (The case of strongly magnetized plasma will be discussed in a separate paper.)

The advantages of our theory are as follows. (i) The theory is derived in a variational form, so the resulting equations are manifestly conservative. (ii) Through the use of the Feynman reparameterization, we can obtain the dynamics of continuous waves and of their rays directly from a variational principle. (iii) The theory assumes no specific wave equation, so quantum spin effects and classical polarization effects can be studied on the same footing. (iv) Moreover, a related formalism [21] is applicable to develop new reduced theories for relativistic spinning particles [21, 24].

This paper is organized as follows. In Sec. II, the basic notation is defined. In Sec. III, the variational formalism used to describe vector waves is presented. In Sec. IV, a general procedure is proposed to block-diagonalize the wave dispersion operator. In Sec. V, we reparametrize the wave action to facilitate asymptotic analysis. In Sec. VI, the leading order GO approximation is discussed. In Sec. VII, the more accurate model that includes polarization effects is discussed. In Sec. VIII, the theory is applied to describe polarization effects on the propagation of EM waves in weakly magnetized plasma. In Sec. IX, our main results are summarized. Finally, Appendix A presents a brief introduction to the Weyl symbol calculus.

II. NOTATION

The following notation is used throughout the paper. The symbol $\dagger$ denotes definitions, “c. c.” denotes “complex conjugate,” and “h. c.” denotes “Hermitian conjugate.” The identity $N \times N$ matrix is denoted by $\mathbb{I}_N$. The Minkowski metric is adopted with signature $(+,−,−,−)$. Greek indices span from 0 to 3 and refer to spacetime coordinates $x^\mu = (x^0, \mathbf{x})$ with $x^0$ corresponding to the time variable $t$. Also, partial derivatives on spacetime will be denoted by $\partial_x$, where the individual components are $\partial_\mu \equiv \partial/\partial x^\mu = (\partial_t, \nabla)$ and $d^4x = dt d^3x$. Latin indices span from 1 to 3 and denote the spatial variables, i.e., $\mathbf{x} = (x^1, x^2, x^3)$ and $\partial_i \equiv \partial/\partial x^i$. Summation over repeated indexes is assumed. In particular, for arbitrary four-vectors $a$ and $b$, we have $a \cdot b = a^\mu b_\mu = a_0 b_0 - a \cdot b$. In Euler-Lagrange equations (ELEs), the notation $\delta a \cdot \Psi$ means that the corresponding equation is obtained by extremizing the action integral with respect to $a$.

III. BASIC EQUATIONS

A. Wave action principle

The dynamics of any nondissipative linear wave can be described by the least action principle, $\delta S = 0$, where the real action $S$ is bilinear in the wave field [23]. We represent a wave field, either classical or quantum, as a complex-valued vector $\Psi(x)$. We allow this vector field to have an arbitrary number of components $N$. In the absence of parametric resonances [27], the action can be written in the form [28]

$$S = \int d^4x \ d^4x' \Psi^\dagger(x) \mathcal{D}(x, x') \Psi(x'),$$

(1)

where $\mathcal{D}$ is a $N \times N$ Hermitian matrix kernel $[\mathcal{D}(x, x') = \mathcal{D}^\dagger(x', x)]$ that describes the underlying medium. Varying $S$ with respect to $\Psi$ (4) leads to

$$\delta \Psi^\dagger : \ 0 = \int d^4x' \mathcal{D}(x, x') \Psi(x').$$

(2)

Similarly, varying with respect to $\Psi$ gives the equation adjoint to Eq. (2), which we do not need to discuss.

It is convenient to describe the wave $\Psi(x)$ as an abstract vector $|\Psi\rangle$ in the Hilbert space of wave states with inner product [23, 29]

$$\langle \Upsilon | \Psi \rangle = \int d^4x \ \Upsilon^\dagger(x) \Psi(x).$$

(3)

In this representation, $\Psi(x) = \langle x | \Psi \rangle$, where $|x\rangle$ are the eigenstates of the coordinate operator $\hat{x}$ such that $\langle x | \hat{x}^\mu | x' \rangle = x^\mu \langle x | x' \rangle = x^\mu \delta^4(x - x')$. We also introduce the momentum (wavevector) operator $\hat{p}$ such that $\langle x | \hat{p}_\mu | x' \rangle = i\partial/\partial x^\mu$ in the $x$-representation [30]. Thus, the action (1) can be rewritten as

$$S = \langle \Psi | \hat{\mathcal{D}} | \Psi \rangle,$$

(4)

where $\hat{\mathcal{D}}$ is the Hermitian dispersion operator such that $\mathcal{D}(x, x') = \langle x | \hat{\mathcal{D}} | x' \rangle$. Treating $\langle \Psi \rangle$ and $|\Psi\rangle$ as independent variables [23] and varying the action (4) gives

$$\delta \langle \Psi \rangle : \ \hat{\mathcal{D}} |\Psi\rangle = 0.$$
(plus the conjugate equation), which is the generalized vector form of Eq. (2). Specifically, Eq. (2) is obtained by projecting Eq. (5) with $\langle x \rangle$ and using the fact that the operator $\int d^4x \ |x\rangle \langle x|$ is 1 is an identity operator.

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,

$$S_{\text{Symplectic}} = \langle \Psi | (\hat{p}\hat{q}_N - \hat{F}) | \Psi \rangle,$$

where $\hat{p}_0 = i \partial_\tau$ (in the $x$-representation) and “the wave Hamiltonian” $\hat{H} = \hat{H}(\hat{x}, \hat{p})$ is some Hermitian operator that is local in time, i.e., commutes with $\hat{L}$. (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 $\tau$ so that $\Psi(\tau, x) = \langle x | \Psi(\tau) \rangle$. Note that $\langle \Psi(\tau) \rangle$ belongs to the same Hilbert space defined in Sec. III A. Thus, the inner product remains the same; i.e., $\langle \Psi(\tau') | \Psi(\tau) \rangle = \int d^4x \bar{\Psi}(\tau', x) \Psi(\tau, x)$. We consider the following “extended” action:

$$S_{\text{X}} = \int d\tau \ L,$$

where $L = L_r + L_d$,

$$L_r = -(i/2) \left[ (\Psi(\tau)) \partial_\tau \Psi(\tau) \right] - c. c., \quad (8a)$$

$$L_d = i \partial_\tau |\Psi(\tau)\rangle = \hat{D} |\Psi(\tau)\rangle,$$

and $\partial_\tau |\Psi(\tau)\rangle = \langle x | \partial_\tau \Psi(\tau) \rangle$. Note that the Lagrangian $L$ is local in the parameter $\tau$; i.e., the abstract vectors are all evaluated at $\tau$. From hereon, all fields will be evaluated at $\tau$, and we will avoid mentioning the dependence of $|\Psi\rangle$ on $\tau$ explicitly. The ELE corresponding to the action $S_X$ is given by

$$i \partial_\tau |\Psi\rangle = \hat{D} |\Psi\rangle.$$

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

IV. EIGENMODE REPRESENTATION

A. Variable transformation

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

$$|\Psi\rangle = \hat{Q} |\psi\rangle.$$

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

$$L_r = -(i/2) \left[ (\bar{\psi} \partial_\tau \psi) - c. c. \right], \quad (11a)$$

$$L_d = i \bar{\psi} \hat{D}_{\text{eff}} |\psi\rangle,$$

where $\hat{D}_{\text{eff}} = \hat{Q} \hat{D} \hat{Q}$. In what follows, we seek to construct $\hat{Q}$ such that the operator $\hat{D}_{\text{eff}}$ is expressed in a block-diagonal form. The procedure used is identical to that given in Refs. [12, 13]. However, in order to account for resonant-mode coupling, $\hat{D}_{\text{eff}}$ will be made only block-diagonal, instead of fully diagonal as in Refs. [12, 13].

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]

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

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

$$W_{mn}(\tau, x, p) = \int \frac{d^4s}{(2\pi)^4} \ e^{ip.s} \langle x + s/2, \bar{\psi}^m \rangle \langle \bar{\psi}^n | x - s/2 \rangle,$$

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

$$D_{\text{eff}}(x, p) = [Q\dagger](x, p) \star D(x, p) \ast Q(x, p), \quad (13)$$

where ‘$\ast$’ is the Moyal product [Eq. (A6)] and $D(x, p)$, $Q(x, p)$, and $[Q\dagger](x, p)$ are the Weyl symbols corresponding to $\hat{D}$, $\hat{Q}$, and $\hat{Q}^\dagger$, respectively. Also, the Weyl representation of the unitary condition, $Q\dagger \hat{Q} = \mathbb{I}_N$, is

$$[Q\dagger](x, p) \ast Q(x, p) = \mathbb{I}_N, \quad (15)$$

which will be used below.

C. Eigenmode representation

Let us assume that the symbols $D_{\text{eff}}$ and $Q$ can be expanded in powers of the GO parameter

$$\epsilon = \max \left\{ \frac{1}{\omega T}, \frac{1}{|k|\ell} \right\} \ll 1,$$

which will be used below.
where \( \omega \) and \( |k| \) are understood as the characteristic wave frequency and wave number, respectively. Also, \( T \) and \( \ell \) are the characteristic time and length scales of the background medium, correspondingly. Hence, we write

\[
D_{\text{eff}}(x,p) = \Lambda(x,p) + \epsilon U(x,p) + \mathcal{O}(\epsilon^2),
\]

\[
Q(x,p) = Q_0(x,p) + \epsilon Q_1(x,p) + \mathcal{O}(\epsilon^2),
\]

where \( (\Lambda, U, Q_0, Q_1) \) are \( \tilde{N} \times \tilde{N} \) matrices of order unity. Hence, \( (\Lambda, U, Q_0, Q_1) \) are \( \tilde{N} \times \tilde{N} \) matrices. Then, Eq. (22) gives

\[
Q_0^\dagger Q_1 + [Q_1^\dagger]Q_0 + (i/2)\{\bar{Q}_0^\dagger, Q_0\} = 0.
\]

Here we assumed that term that involves the Poisson bracket \( \{Q_0^\dagger, Q_0\} \), which arises from the expansion of the Moyal star product [Eq. (A9)], is of the first order in \( \epsilon \). Following Ref. [12], we let \( Q_1 = Q_0(A + iG) \) and \( [Q_1^\dagger] = Q_1^\dagger \), where \( A(x,p) \) and \( G(x,p) \) are \( \tilde{N} \times \tilde{N} \) Hermitian matrices. Then, Eq. (22) gives

\[
A(x,p) = -(i/4)\{\bar{Q}_0^\dagger, Q_0\}.
\]

In order to determine \( G(x,p) \), we write Eq. (14) to the first order in \( \epsilon \). Introducing the bracket

\[
\{A, B\}_C = \frac{\partial A}{\partial p_\mu} C \frac{\partial B}{\partial x^n} - \frac{\partial A}{\partial x^n} C \frac{\partial B}{\partial p_\mu},
\]

and noting that \( DQ_0 = Q_0\Lambda \), we obtain [12]

\[
U(x,p) = Q_0^\dagger DQ_0 + Q_0^\dagger DQ_1 + (i/2)\{Q_0^\dagger D, Q_0\} + (i/2)\{Q_0^\dagger, D\}Q_0
\]

\[
= (A - iG)Q_0^\dagger DQ_0 + Q_0^\dagger DQ_0(A + iG) + (i/2)\{Q_0^\dagger D, Q_0\} + (i/2)\{Q_0^\dagger, D\}Q_0
\]

\[
= \Lambda\Lambda - iG\Lambda + \Lambda A + iA\Lambda + \frac{i}{2}\{Q_0^\dagger D, Q_0\} + (i/2)\{Q_0^\dagger, D\}Q_0 - (i/2)\{Q_0^\dagger, Q_0\}D
\]

\[
= i(\Lambda\Lambda - G\Lambda + \delta U),
\]

where \( \delta U(x,p) \) is a \( \tilde{N} \times \tilde{N} \) matrix given by

\[
\delta U(x,p) = (1/4)\{Q_0^\dagger, Q_0\}\Lambda + (1/4)\Lambda\{Q_0^\dagger, Q_0\} + (1/2)\{\Lambda, Q_0\}Q_0^\dagger + (1/2)\{Q_0^\dagger, \Lambda\}Q_0 - (1/2)\{Q_0^\dagger, Q_0\}D.
\]

Since \( (G\Lambda - GA)^m_n = G^m_n(\lambda^{(m)} - \lambda^{(n)}) \) (no summation is assumed here over the repeating indices), one can diagonalize \( U \) by adopting \( G^m_n = \delta U^{m/n}/(\lambda^{(m)} - \lambda^{(n)}) \), as done in Refs. [12, 13]. However, this method is applicable only when \( |\lambda^{(m)} - \lambda^{(n)}| \gtrsim \mathcal{O}(1) \); otherwise, when \( |\lambda^{(m)} - \lambda^{(n)}| \approx \mathcal{O}(\epsilon) \), \( G \gtrsim \mathcal{O}(\epsilon^{-1}) \) so \( \epsilon Q_1 \gtrsim \mathcal{O}(1) \), which is in violation of the assumed ordering in Eq. (17b). Hence, instead of diagonalizing \( U \), we propose to only block-diagonalize \( U \) as follows. When \( |\lambda^{(m)} - \lambda^{(n)}| \gtrsim \mathcal{O}(1) \), we choose the off-components of \( G^m_n \) so that \( U^{m/n} = 0 \). (We call such modes nonresonant.) When \( |\lambda^{(m)} - \lambda^{(n)}| \approx \mathcal{O}(\epsilon) \), we let \( G^m_n = 0 \). (We call such modes resonant.) By following this prescription and permutating the matrix rows, we obtain \( U \) in the following form:

\[
U(x,p) = \left( \begin{array}{cccc}
U_1(x,p) & 0 & \cdots & 0 \\
0 & U_2(x,p) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & U_J(x,p)
\end{array} \right),
\]

where \( U_j(x,p) \) are \( n_j \times n_j \) Hermitian matrices and \( J \) is the total number of blocks, so \( \sum_{j=1}^{J} n_j = \tilde{N} \). Note that, in the particular case where only nonresonant modes are present, \( U(x,p) \) is diagonal, and one recovers the results obtained in Refs. [12, 13].

Since the matrix \( D_{\text{eff}} \approx \Lambda + \epsilon U \) is made block-diagonal, the Lagrangian (12) is unaffected by the matrix elements...
$W^m_n$ with indices $(m, n)$ such that $U^m_n = 0$. Thus, without loss of generality, we can write

$$L_D = \sum_{j=1}^{J} \text{Tr} \int d^4x \, d^4p \, [[D_{\text{eff}}]]_j \, [[W]]_j,$$  \hspace{1cm} (28)

where $[[D_{\text{eff}}]]_j \simeq [[\Lambda + \epsilon U]]_j$ and $[[W]]_j$ represent the $j$th matrix block of $\Lambda + \epsilon U$ and $W$, respectively. Hence, non-resonant eigenmodes are decoupled while resonant eigenmodes that belong to the same block remain coupled.

V. REDUCED ACTION

A. Basic equations

Now that blocks of mutually nonresonant modes are decoupled, let us focus on the dynamics of modes within a single block of some size $N$. Hence, the block index will be dropped, and we adopt

$$L_T = -(i/2) \int d^4x \left[ \psi^\dagger (\partial_\tau \psi) - (\partial_\tau \psi^\dagger) \psi \right], \hspace{1cm} (29a)$$

$$L_D = \text{Tr} \int d^4x \, d^4p \, [[\Lambda + \epsilon U]] \, [[W]]. \hspace{1cm} (29b)$$

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

$$[[W]]^m_n(x, p) = \int \frac{d^4s}{(2\pi)^4} \, e^{ip \cdot s} \left( x + \frac{s}{2} \right) \, (\psi^m_n)(x - \frac{s}{2}). \hspace{1cm} (30)$$

Since we consider the coupled dynamics of some $N$ resonant modes, only $N$ columns of $Q_0$ actually contribute to $[[D_{\text{eff}}]]$. For clarity, let us denote the resonant eigenmodes as $e_q$ with indices $q = 1, \ldots, N$. Then, in order to calculate $[[U]]$, one can use Eq. (25). After block-diagonalizing $\tilde{U}$ and introducing the $N \times N$ matrix

$$\Xi(x, p) = [e_1(x, p), \ldots, e_N(x, p)], \hspace{1cm} (31)$$

one obtains

$$[[U]] = \frac{i}{4} \left( \Xi^\dagger, \Xi \right) \Lambda + \frac{i}{4} \Lambda \left( \Xi^\dagger, \Xi \right) + \frac{i}{2} \left( \Lambda, \Xi \right) \Xi^\dagger$$

$$+ \frac{i}{2} \left( \Xi, \Lambda \right) \Xi^\dagger - \frac{i}{2} \left( \Xi, \Xi \right)_D, \hspace{1cm} (32)$$

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

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

$$[[D_{\text{eff}}]] = \lambda I_N + dU,$$  \hspace{1cm} (33)

where $\lambda \doteq N^{-1} \text{Tr} [[D_{\text{eff}}]]$ is the average of the eigenvalues of $[[D_{\text{eff}}]]$ and $dU \doteq [[D_{\text{eff}}]] - \lambda I_N$ 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 $\Lambda = \lambda I_N$, and $U = [[U]]$. We call such modes degenerate. Then, the expression (32) for $[[U]]$ simplifies, and one obtains

$$U(x, p) = \frac{i}{4} \left( \Xi^\dagger, \Xi \right) \lambda + \frac{i}{4} \lambda \left( \Xi^\dagger, \Xi \right) + \frac{i}{2} \left( \lambda, \Xi \right) \Xi^\dagger$$

$$+ \frac{i}{2} \left( \Xi^\dagger, \lambda \right) \Xi - \frac{i}{2} \left( \Xi^\dagger, \Xi \right)_D$$

$$= - \frac{1}{2i} \left( \Xi^\dagger, \Xi \right) - \frac{1}{2i} \left( \Xi^\dagger, \lambda \right) \Xi$$

$$+ \frac{1}{2i} \left( \Xi^\dagger, \Xi \right)_D - \frac{1}{2i} \left( \Xi^\dagger, \Xi \right) \lambda$$

$$= - \left[ \Xi^\dagger (\lambda, \Xi) \right] + \left[ (\partial_p \Xi^\dagger)(D - \lambda I_N)(\partial_p \Xi) \right]_A,$$  \hspace{1cm} (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

$$U(x, p) = \left( -\frac{\partial \lambda}{\partial p_{\mu}} \right) \left( \Xi^\dagger \frac{\partial \Xi}{\partial x^\mu} \right)_A + \left( \frac{\partial \lambda}{\partial x^\mu} \right) \left( \Xi^\dagger \frac{\partial \Xi}{\partial p_{\mu}} \right)_A$$

$$+ \left[ \frac{\partial \Xi^\dagger}{\partial p_{\mu}} (D - \lambda I_N) \frac{\partial \Xi}{\partial x^\mu} \right]_A. \hspace{1cm} (35)$$

Examples of physical 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:

$$\psi(\tau, x) = \sqrt{I(\tau, x)} \, z(\tau, x) \, e^{i\theta(\tau, x)}. \hspace{1cm} (36)$$

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, $I(\tau, x)$ is a real function, and $z(\tau, x)$ is a $N$-dimensional complex unit vector $(z^i z = 1)$, whose components describe the amount of quanta in the corresponding modes. (Since we parameterize the $N$-dimensional complex vector $\psi$ by the $N$-dimensional complex vector $z$ plus two independent real functions $\theta$ and $I$, 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 $L_T$ is given by

$$L_T = \int d^4x \, i \frac{\partial}{\partial \tau} \left[ \partial_x \theta - (i/2)(z^i \partial_z z - c.c.) \right]. \hspace{1cm} (37)$$

(Here we formally introduce $\epsilon$ 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
\[
[[W]]_{m}(\tau, x, p) = \int \frac{d^4s}{(2\pi)^4} \sqrt{\mathcal{I}(x + s/2, \tau) z^m(x + s/2, \tau) e^{i\theta(x + s/2, \tau)}} \sqrt{\mathcal{I}(x - s/2, \tau) z^*_m(x - s/2, \tau) e^{-i\theta(x - s/2, \tau)}} e^{ip \cdot s}
\]

where we introduced the four-wavevector \(k\). The lowest-order GO Lagrangian. The terms in the first line on the right-hand side of Eq. (42) represent approximated simply with GO.

\[
\int \frac{d^4s}{(2\pi)^4} \mathcal{I}(\tau, x) z^m(\tau, x) z^*_m(\tau, x) e^{i(p - k) \cdot s}
\]

Summing Eqs. (37) and (39), we obtain the action is the zeroth-order (in \(\epsilon\)) group velocity of the wave. We then introduce the convective derivative \(\frac{d}{d\tau} \approx \frac{\partial}{\partial \tau} + v^\mu(\tau, x) \frac{\partial}{\partial x^\mu}\).

Equation (42), along with the definitions in Eqs. (31), (33), (40), and (41), is the main result of this work. The first line on the right-hand side of Eq. (42) represents the lowest-order GO Lagrangian. The terms in the second line of Eq. (42) are \(\mathcal{O}(\epsilon)\) and introduce polarization effects. [Importantly, diffraction terms would be \(\mathcal{O}(\epsilon^2)\) and thus are safe to neglect in our first-order theory.] In what follows, we discuss the consequences of this theory and provide an example, where we apply the theory to study polarization effects on EM waves in weakly magnetized plasmas.

\[\text{VI. TRADITIONAL GEOMETRICAL OPTICS}\]

A. Continuous wave model

To lowest order in \(\epsilon\), the Lagrangian (42) can be approximated simply with

\[L_{\text{GO}} \approx \int d^4x \mathcal{I} \left[ \partial_\tau \theta + \lambda(x, k) \right],\]

which one may interpret as a Hayes-type representation [34] of the GO wave Lagrangian in the extended \((\tau, x)\) space. This Lagrangian is parameterized by just two functions, the rapid phase \(\theta\) and the total action density \(\mathcal{I}\). Thus, varying the action \(S_{\text{GO}} = \int d\tau L_{\text{GO}}\), we obtain the following ELEs:

Equation (43) gives rise to the photon conservation theorem. Equation (44b) is the local...
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

\[
\mathcal{I}(\tau, x) = \mathcal{I}_0 \delta^4(x - X(\tau)).
\]

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 \( X(\tau) \) in space-time, and the independent parameter is \( \tau \). [This means that at a given \( \tau \), the wave packet is located at the spatial point \( X(\tau) \) at time \( t(\tau) \).] When inserting Eq. (46) into Eq. (43), the first term in the action gives the following:

\[
\int d\tau d^4x \mathcal{I} \partial_\theta \theta = \int d\tau d^4x \delta^4(x - X(\tau)) \partial_\tau \theta(\tau, x)
\]

\[
= \int d\tau d^4x \theta(\tau, x)[\partial_\tau \delta^4(x - X(\tau))]
\]

\[
= \int d\tau d^4x \theta(\tau, x)[\dot{X}^\mu(\tau) \partial_\mu \delta^4(x - X(\tau))]
\]

\[
= -\int d\tau d^4x \partial_\mu \theta(\tau, x) \dot{X}^\mu(\tau) \delta^4(x - X(\tau))
\]

\[
= \int d\tau P_\mu(\tau) \dot{X}^\mu(\tau),
\]

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

\[
\int d^4x \delta^4(x - X(\tau)) \lambda(x, -\partial \theta) = \lambda(X(\tau), P(\tau)).
\]

Thus, the point-particle action is expressed as

\[
\mathcal{S}_{GO} = \int d\tau \left[ P(\tau) \cdot \dot{X}(\tau) + \lambda(X, P) \right].
\]

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

\[
\delta P_\mu : \frac{dX^\mu}{d\tau} = \frac{\partial \lambda}{\partial P_\mu},
\]

\[
\delta X^\mu : \frac{dP_\mu}{d\tau} = \frac{\partial \lambda}{\partial X^\mu}.
\]

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

\[
\frac{dX^0}{d\tau} = -\frac{\partial \lambda}{\partial P_0}, \quad \frac{dX}{d\tau} = \frac{\partial \lambda}{\partial P}, \quad \frac{dP^0}{d\tau} = \frac{\partial \lambda}{\partial X^0}, \quad \frac{dP}{d\tau} = -\frac{\partial \lambda}{\partial X}.
\]

Note that the first term in the integrand in Eq. (49) 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 \( \tau \), then \( d\lambda(X, P)/d\tau = 0 \) along the ray trajectories. Thus, the ray dynamics lies on the dispersion manifold defined by

\[
\lambda(X, P) = 0.
\]

As a reminder, \( \lambda(x, p) \) is defined as the average eigenvalue of the resonant block, i.e., \( \lambda \equiv N^{-1} \text{Tr}[[D_{\text{eff}}]] \). The GO action (49) is only accurate to lowest order in \( \epsilon \); hence, one can approximate \( \lambda(x, p) \approx \lambda^{(n)}(x, p) \), where \( \lambda^{(n)} \) is any particular resonant eigenvalue. This occurs because the resonant eigenvalues differ by \( O(\epsilon) \) and because the polarization coupling is also \( O(\epsilon) \).

VII. EXTENDED GEOMETRICAL OPTICS

In this section, we explore the polarization effects determined by the 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. VIB, we approximate the wave packet to a single point in spacetime [Eq. (46)]. As shown in Refs. [21, 25], the Lagrangian (42) can be replaced by a point-particle Lagrangian so the action is

\[
\mathcal{S}_{XGO} = \int d\tau \left[ P \cdot \dot{X} - \frac{(i/2)(Z^\dagger \dot{Z} - \dot{Z}^\dagger Z)}{0} + \lambda(X, P) + Z^\dagger \mathcal{U}(X, P)Z \right],
\]

where \( Z(\tau) \equiv z(\tau, X(\tau)) \) is the point-particle polarization vector and we dropped the GO ordering parameter \( \epsilon \). In the complex representation, \( Z \) and \( Z^\dagger \) are canonical conjugate, and

\[
Z^\dagger(\tau)Z(\tau) = 1.
\]

Even though the components of \( Z \) are not independent by definition (Sec. VIB), 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 \( \mathcal{S}_{XGO} \) are
Then, we introduce the \((N-1)\)-matrices, with some real coefficients generators \(N_u\). Since \(D\) is traceless part of \([\langle D_{\text{eff}} \rangle\), and \(U\) is \(\langle D_{\text{eff}} \rangle - \lambda I_N\) is the remaining traceless part of \([\langle D_{\text{eff}} \rangle\). As a reminder, \(\lambda(x, p)\) is defined through Eq. (33) as \(\lambda = N^{-1}\text{Tr}[\langle D_{\text{eff}} \rangle]\), and \(U = \langle D_{\text{eff}} \rangle - \lambda I_N\) is \(\lambda\) the remaining traceless part of \([\langle D_{\text{eff}} \rangle\].

As in Sec. VI B, the Hamiltonian part of Eq. (52) is constant along the ray trajectories. As before, the ray dynamics lies on the dispersion manifold defined by setting the Hamiltonian part to zero; i.e.,

\[
\lambda(X, P) + Z^\dagger U(X, P)Z = 0. \tag{55}
\]

Together with Eqs. (31)-(33), Eqs. (54) form a complete set of equations. The first terms on the right-hand side of Eqs. (54a) and (54b) describe the ray dynamics in the GO limit. The second terms describe the coupling to the mode polarization. Equations (54c) and (54d) describe the wave-polarization dynamics.

As in Sec. VI B, the Hamiltonian part of Eq. (52) is constant along the ray trajectories. As before, the ray dynamics lies on the dispersion manifold defined by setting the Hamiltonian part to zero; i.e.,

\[
\lambda(X, P) + Z^\dagger U(X, P)Z = 0. \tag{55}
\]

B. Precession of the wave spin

Let us also describe the rotation of \(Z(\tau)\) as follows. Since \(U(X, P)\) is traceless Hermitian \(N \times N\) matrix, it can be decomposed into a linear combination of \(N^2-1\) generators \(T_u\) of \(SU(N)\), which are traceless Hermitian matrices, with some real coefficients \(-W^u\) [35]:

\[
U = -\sum_{u=1}^{N^2-1} T_u W^u = -\mathcal{T} \cdot W. \tag{56}
\]

Then, we introduce the \((N^2-1)\)-dimensional vector

\[
S(\tau) = Z^\dagger(\tau) T Z(\tau) \tag{57}
\]

so that \(Z^\dagger U Z = -S \cdot W\). The components of \(S(\tau)\) satisfy the following equation:

\[
\begin{align*}
\frac{d}{d\tau} S_w = & Z^\dagger T_u (d_w Z) + (d_z Z^\dagger) T_u Z \\
= & i Z^\dagger U T_u Z - i Z^\dagger T_u U Z \\
= & i Z^\dagger [U, T_u] Z \\
= & -i Z^\dagger [T_u, T_w] Z W^u \\
= & f_{wuv} (Z^\dagger T^v Z) W^u \\
= & f_{wuv} S^v W^u,
\end{align*}
\]

where \(f_{abc}\) are structure constants. They are defined via \([T_a, T_b] = i f_{abc} T^c\) so that the structure constants \(f_{abc}\) are antisymmetric in all indices [35].

For example, consider the case when only two waves are resonant. Then, \(N^2-1 = 3\), \(\mathcal{T}^v\) are the three Pauli matrices divided by two (so \(|S|^2 = 1/2\), and \(f_{wuv}\) is the Levi-Civita symbol, so \(f_{wuv} S^v W^u = (S \times W)_w\). For a Dirac electron, which is a special case, such \(S\) is recognized as the spin vector undergoing the well known precession equation, \(d_\tau S = S \times W\) [21]. In optics, this is an equation for the Stokes vector that was derived earlier to characterize the polarization of transverse EM waves in certain simple media [7, 36, 37].

Hence, it is convenient to extend this quantum terminology also to \(N\) resonant waves. We will call the corresponding \((N^2-1)\)-dimensional vector \(S\) a generalized “wave-spin” vector and express \(f_{wuv} S^v W^u\) symbolically as \((S \times W)_w\), where “\(\times\)” can be viewed as a generalized vector product. Notably, using the concept of spin vector \(S\), one can rewrite Eqs. (54c) and (54d) as follows:

\[
\frac{d}{d\tau} S = S \times W, \tag{59}
\]

which is understood as a generalized precession equation.

In the particular case when \(S\) is conserved (we call such waves “pure states”), then Eqs. (54a), (54b), and (59) form a closed set of equations, and \(\lambda - S \cdot W\) serves as an effective scalar Hamiltonian. The dynamics of \(Z\) and \(Z^\dagger\) does not need to be resolved in this case, so one can rewrite \(S_{\text{XGO}}\) as a functional of \((X, P)\) alone:

\[
S_{\text{XGO}} = \int d\tau [P \cdot \dot{X} + \lambda(X, P) - S \cdot W(X, P)]. \tag{60}
\]

An example of the dynamics described by such action will be discussed in Sec. VIII E.

A more general case is when \(S\) is close to some eigenvector \(w\) of \(W\) that corresponds to some nondegenerate eigenvalue \(\Omega_w\). If \(\Omega_w\) is large enough, then \(S(\tau)\) will remain close to \(w(\tau)\) and will only experience small-amplitude oscillations. These oscillations can be understood as a generalized zitterbewegung effect [38], and they are transient, i.e., vanish when \(W\) becomes zero. In this regime, no mode conversion occurs at \(\tau \to \infty\). In contrast, if \(\Omega_w\) is not large enough, the change of \(S\) governed by Eq. (59) is not necessarily negligible. This corresponds to mode conversion and causes ray splitting at \(\tau \to \infty\) (see, e.g., Refs. [14–20]). This is discussed below.

C. Mode conversion as a form of spin precession

Equation (54c) can also describe mode conversion as it is understood in Refs. [14–20]. This is shown as follows. Let us consider the resonant interaction between two modes as an example; then, \(U\) is a \(2 \times 2\) matrix. From Eq. (33), \(U\) is Hermitian and traceless and can be represented as

\[
U(X, P) = \begin{pmatrix} \Delta \lambda/2 & U_{12} \\ U_{12}^* & -\Delta \lambda/2 \end{pmatrix}, \tag{61}
\]
where $\Delta \lambda(X, P) \equiv [\lambda^{(1)} - \lambda^{(2)}]/2 + (U_{11} - U_{22})/2$ and the coefficient $U_{12}$ determines the mode coupling. Suppose that, absent coupling ($U_{12} = 0$), the dispersion curves of two modes cross at some point $(X, P)$. Suppose also that $\Delta \lambda(\tau) = \Delta \lambda(X(\tau), P(\tau))$ changes along the ray trajectory approximately linearly in $\tau$. Then, $\Delta \lambda \approx \alpha \tau$, where $\alpha$ is some constant coefficient and we chose the origin on the time axis such that $\Delta \lambda(\tau = 0) = 0$ for simplicity. Similarly, $U_{12}(\tau) \equiv U_{12}(X(\tau), P(\tau)) \approx \beta + \gamma \tau$, where $\beta$ and $\gamma$ are some constants. Assuming $\beta$ is sufficiently large, we neglect the term $\gamma \tau$ for it only causes a correction to the dominant effect. Thus, near the mode-conversion region, Eq. (54c) is approximately written as

$$\frac{d}{d\tau} \left( \frac{Z_1}{Z_2} \right) = \left( \frac{\alpha \tau / 2}{\beta^* - \alpha \tau / 2} \right) \left( \frac{Z_1}{Z_2} \right).$$

Equation (62) is the well-known equation for mode conversion that was studied by Zener in Ref. [39]. After eliminating $Z_2$, the governing equation for $Z_1$ is

$$\ddot{Z}_1(\tau) + \left( |\beta|^2 + \frac{i \alpha}{2} + \frac{\alpha^2 \tau^2}{4} \right) Z_1(\tau) = 0. \tag{63}$$

Letting $w \doteq \tau \sqrt{\alpha} e^{i \pi/4}$ and $n \doteq -i|\beta|^2/\alpha$, the equation above can be written as a Weber equation

$$Z''_1(w) + \left( n + \frac{1}{2} - \frac{w^2}{4} \right) Z_1(w) = 0, \tag{64}$$

whose solutions are the parabolic cylinder functions $D_n(w)$. In Refs. [17, 39], the matrix connecting the waves entering and exiting the resonance are obtained by analyzing asymptotics of $D_n(w)$. Specifically,

$$\begin{pmatrix} Z_{1, \text{out}} \\ Z_{2, \text{out}} \end{pmatrix} = \begin{pmatrix} T & -c^* \\ C & T \end{pmatrix} \begin{pmatrix} Z_{1, \text{in}} \\ Z_{2, \text{in}} \end{pmatrix}, \tag{65}$$

where

$$T = \exp(-\pi|\eta|^2), \quad C = -\frac{\sqrt{2\pi|\eta|^2}}{\eta \Gamma(-i|\eta|^2)} \tag{66}$$

where $\Gamma$ is the Gamma function and $\eta \doteq \beta/\sqrt{\alpha}$. The transmission and conversion coefficients for the wave quanta are, correspondingly,

$$|T|^2 = \exp(-2\pi|\beta|^2/|\alpha|), \quad |C|^2 = 1 - |T|^2. \tag{67}$$

(Also see Ref. [16] for a somewhat different approach leading to the same answer.)

This calculation shows that mode conversion, in the way as commonly described in literature [14–20], is nothing but a manifestation of the wave-spin precession described by Eqs. (54c) and (59). Note that the present point-particle model cannot capture ray-splitting because it introduces only one ray for the whole field. However, this theory does predict the transfer of wave quanta, which is a prerequisite for ray-splitting. For a complete analysis on ray-splitting mode conversion, please refer to Refs. [1, 18, 19].

VIII. DISCUSSION: WAVES IN WEAKLY MAGNETIZED PLASMAS

A simplified form of the theory above was applied to describe spin-1/2 particles [21, 24] and waves in isotropic dielectrics [22]. Here we present another example of its application, namely, EM waves in weakly magnetized cold plasmas. (The case of strongly magnetized plasmas will be discussed in a separate paper.) We assume that the plasma response is determined by particles of just one type, e.g., electrons. The generalization to multi-component plasma is straightforward to do.

A. Dispersion operator

The linearized equations of motion are [40]

$$\partial_t \mathbf{v} = \frac{q}{m} \mathbf{E} + \frac{q}{mc} \mathbf{v} \times \mathbf{B}, \tag{69a}$$

$$\partial_t \mathbf{E} = -4\pi q n_0 \mathbf{v} + e \mathbf{\nabla} \times \mathbf{B}, \tag{69b}$$

$$\partial_t \mathbf{B} = -e \mathbf{\nabla} \times \mathbf{E}. \tag{69c}$$

Here $q, m, n_0(x)$, and $\mathbf{v}(t, \mathbf{x})$ are the particle charge, mass, unperturbed background density, and flow velocity, respectively. Also, $\mathbf{E}(t, \mathbf{x})$ denotes the perturbation electric field, $\mathbf{B}_0(t, \mathbf{x})$ is the perturbation magnetic field, $\mathbf{B}_0(\mathbf{x})$ is the background magnetic field, and $c$ is the speed of light. We introduce a re-scaled velocity field $\mathbf{v}(t, \mathbf{x}) \doteq \mathbf{v}(t, \mathbf{x})/[4\pi n_0(\mathbf{x})m^{1/2}]$, so

$$\partial_t \hat{\mathbf{v}} = \omega_p \mathbf{E} + \mathbf{v} \times \mathbf{\Omega}, \tag{70a}$$

$$\partial_t \mathbf{E} = -\omega_p \mathbf{v} + e \mathbf{\nabla} \times \mathbf{B}, \tag{70b}$$

$$\partial_t \mathbf{B} = -e \mathbf{\nabla} \times \mathbf{E}. \tag{70c}$$

where $\omega_p(\mathbf{x}) \doteq [4\pi q^2 n_0(\mathbf{x})/m]^{1/2}$ is the plasma frequency and $\mathbf{\Omega}(\mathbf{x}) \doteq q \mathbf{B}_0(\mathbf{x})/(mc)$ is the gyrofrequency.

Let us write Eqs. (70) using the abstract Hilbert space notation. Let $|\mathbf{v}\rangle$ be a state vector representing the velocity field such that $\mathbf{v}(x) \doteq \langle x | \mathbf{v} \rangle$. 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. (70) can be written as follows:

$$\hat{p}_0 |\mathbf{v}\rangle = i\hat{\omega}_p |\mathbf{E}\rangle - (\mathbf{\alpha} \cdot \hat{\Omega}) |\mathbf{v}\rangle, \tag{71a}$$

$$\hat{p}_0 |\mathbf{E}\rangle = -i\hat{\omega}_p |\mathbf{v}\rangle + ic(\mathbf{\alpha} \cdot \hat{\mathbf{p}}) |\mathbf{B}\rangle, \tag{71b}$$

$$\hat{p}_0 |\mathbf{B}\rangle = -ic(\mathbf{\alpha} \cdot \hat{\mathbf{p}}) |\mathbf{E}\rangle. \tag{71c}$$

where $\hat{\omega}_p \doteq \omega_p(\mathbf{x})$ and $\hat{\Omega} \doteq \Omega(\mathbf{x})$. (As a reminder, $\hat{p}_0 = i\partial_t$ and $\hat{\mathbf{p}} = -i\mathbf{\nabla}$ are the components of the four-momentum operator in the $x$-representation.) Also, $\mathbf{\alpha}$
Here, for any two column vectors $\mathbf{A}$ and $\mathbf{B}$, one has

$$
\mathbf{A} \cdot \mathbf{B} = i \mathbf{A} \times \mathbf{B},
$$

where $\mathbf{A}^T \mathbf{A} = -i (\mathbf{A} \times \mathbf{B})$, where the superscript ‘T’ denotes the matrix transpose.

The next step is to construct a dispersion operator for the electric field state $\langle \mathbf{E} \rangle$. Starting from Eq. (71a), we solve for the velocity field in terms of the electric field. Hence, we formally obtain the following:

$$
\mathbf{D} |\mathbf{E}\rangle = 0,
$$

where

$$
\mathbf{D} \equiv -p_0^2 + (\mathbf{A} \cdot \mathbf{p})^2 + \omega_p^2 - \mathbf{p} \cdot \mathbf{p} + \omega_p^2 (\mathbf{A} \cdot \mathbf{\Omega})^2 + (\mathbf{A} \cdot \mathbf{\Omega})^2.
$$

serves as the dispersion operator for $|\mathbf{E}\rangle$. (For convenience, we let $c = 1$.) Since $\omega_p(x)$ and $\mathbf{\Omega}(x)$ are independent of time, then $p_0$ commutes with $\omega_p$ and $\mathbf{\Omega}$, so $\mathbf{D}$ is manifestly Hermitian. The corresponding action (4) for the electric field is $\mathcal{S} = \langle \mathbf{E} | \mathbf{D} | \mathbf{E}\rangle$, and the extended action (7) is

$$
\mathcal{S}_\chi = \int \mathcal{D} \chi \left[ -\frac{i}{2} (\mathbf{E} | \partial_\tau \mathbf{E} | - \text{c.c.}) + \langle \mathbf{E} | \mathbf{D} | \mathbf{E}\rangle \right].
$$

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

**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 $\mathbf{D}$ is

$$
D \equiv -p_0^2 + (\alpha \cdot \mathbf{p})^2 + \omega_p^2 + \omega_p^2 \frac{\mathbf{A} \cdot \mathbf{\Omega}}{p_0^2 - \Omega^2} + \omega_p^2 \frac{\mathbf{A} \cdot \mathbf{\Omega}}{p_0^2 - \Omega^2}.
$$

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 ($\omega \sim \omega_p$, $\mathbf{p} \gg \mathbf{\Omega}$), we expand the dispersion symbol in powers of $\Omega/p_0$:

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

where

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

$$
D_1(x, p_0) = -\omega_p^2 (\mathbf{A} \cdot \mathbf{\Omega})/p_0.
$$

To simplify the following calculation, we assume that $D_1 \sim \mathcal{O}(\Omega/p_0)$ is comparable in magnitude to the GO parameter $\epsilon$, 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 $D_0(x, p)$. The corresponding eigenvalues are

$$
\lambda^{(1)}(x, p) = -p_0 \cdot \mathbf{p} + \omega_p^2 (x),
$$

$$
\lambda^{(2)}(x, p) = -p_0 \cdot \mathbf{p} + \omega_p^2 (x),
$$

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

where $p \cdot \mathbf{p} = p_0^2 - \mathbf{p}^2$. These eigenvalues correspond to the dispersion relations of two transverse EM waves and of longitudinal Langmuir oscillations, respectively. The matrix $Q_0$ defined in Eq. (20) is given by

$$
Q_0(0) = \left[ e_1(0), e_2(0), e_p(0) \right],
$$

where $e_1(0)$ and $e_2(0)$ are any two orthonormal vectors in the plane normal to $e_p(0) = \mathbf{p}/|\mathbf{p}|$. A right-hand convention is adopted such that $e_1 \times e_2 = e_p$. One can easily verify that these vectors are indeed eigenvectors of $D_0(x, p)$. For example,

$$
D_0 e_1 = [-p_0^2 + (\alpha \cdot \mathbf{p})(\alpha \cdot \mathbf{p}) + \omega_p^2] e_1
$$

$$
= (-p_0^2 + \omega_p^2) e_1 - \mathbf{p} \times (\mathbf{p} \times e_1)
$$

$$
= (-p_0^2 + \omega_p^2) e_1
$$

$$
= \lambda^{(1)} e_1,
$$

where Eq. (73a) was used. Similar calculations follow for the other two eigenmodes $e_2$ and $e_p$.

We now analyze the dynamics of the transverse EM waves. From Sec. V, the eigenvalue is $\lambda(x, p) = -p \cdot \mathbf{p} + \omega_p^2 (x)$, where $\Xi(0) = [e_1(0), e_2(0)]$ is a $3 \times 2$ matrix. Since $\Xi(0)$ only depends on the spatial momentum coordinate, then the polarization-coupling...
Hamiltonian $\mathcal{U}$ in Eq. (35) is given by

$$
\mathcal{U}(x, p) = \frac{\partial \lambda}{\partial x^\mu} \left( \xi^{\mu} \frac{\partial \xi}{\partial p_\mu} \right)_A
- \frac{\partial \lambda}{\partial x} \cdot \left( \xi \frac{\partial \xi}{\partial p} \right)_A
= - \frac{\partial \lambda}{\partial x} \cdot \left[ \left( e^1 \frac{\partial}{\partial p} (e_1, e_2) \right)_A
- \left( e^2 \frac{\partial}{\partial p} e_1 \right)_A \right]
= - \frac{1}{i} \frac{\partial \lambda}{\partial x} \left( \begin{array}{cc}
0 & e^1 \frac{\partial}{\partial p} e_2 \\
0 & e^2 \frac{\partial}{\partial p} e_2
\end{array} \right),
$$

(84)

where $e^q$ is the dual to $e_q$, so $e^q e_r = \delta^q_r$. (Specifically, $e^q$ is a row vector, whose elements are complex-conjugate of those of $e_q$.) Since $\nabla \lambda = \nabla \omega^2$, we can also write Eq. (84) in the form

$$
\mathcal{U}(x, p) = -\sigma_y (\nabla \omega^2) \cdot F,
$$

(85)

where $\sigma_y$ is the $y$-component of the Pauli matrices

$$
\sigma_y = \begin{pmatrix}
0 & -i \\
i & 0
\end{pmatrix}
$$

(86)

and $F(p)$ is a vector with components given by

$$
F(p) = e^1 \frac{\partial}{\partial p} e_2.
$$

(87)

For example, one may choose

$$
e_1(p) = \begin{pmatrix}
\frac{p_x}{\sqrt{p_x^2 + p_y^2}} \\
\frac{p_y}{\sqrt{p_x^2 + p_y^2}} \\
\frac{p_z}{\sqrt{p_x^2 + p_y^2}}
\end{pmatrix},
\quad e_2(p) = \begin{pmatrix}
-\frac{p_y}{\sqrt{p_x^2 + p_y^2}} \\
\frac{p_x}{\sqrt{p_x^2 + p_y^2}} \\
0
\end{pmatrix},
$$

(88)

so that

$$
F(p) = \frac{p_\perp \times p}{|p_\perp| |p|^2}.
$$

(89)

where $p_\perp = (p_x, p_y, 0)^T$; or, more explicitly,

$$
F(p) = \frac{p_\perp \times p}{|p_\perp| |p|^2} \begin{pmatrix}
0 \\
0 \\
p_x
\end{pmatrix}.
$$

(90)

(The specific choice of $e_1$ and $e_2$ does not affect the resulting equations within the accuracy of the present theory. For more details, see Sec. VIII G.)

Returning to the perturbation caused by the background magnetic field, the projection of the eigenmodes on the matrix $D_1(x, p_0)$ is given by

$$
\Xi^\dagger D_1 \Xi
= -\frac{\omega^2}{p_0} \begin{pmatrix}
e^1(\alpha \cdot \Omega) e_1 \\
e^2(\alpha \cdot \Omega) e_1
\end{pmatrix}
-\frac{\omega^2}{p_0} \begin{pmatrix}
e^1(\alpha \cdot \Omega) e_2 \\
e^2(\alpha \cdot \Omega) e_2
\end{pmatrix}
= i \frac{\omega^2}{p_0} \begin{pmatrix}
0 & e_\parallel \cdot \Omega \\
e_\parallel \cdot \Omega & 0
\end{pmatrix}
= -\frac{\omega^2}{p_0} (e_\parallel \cdot \Omega) \sigma_y,
$$

(91)

where we used Eq. (73b).

### C. Ray dynamics

Now, let us discuss the point-particle ray dynamics. Following Sec. VII, we substitute $\lambda(x, p) = -p \cdot p + \omega^2$, Eq. (85), and Eq. (91) into Eq. (52). We then obtain the point-particle action

$$
S_{XGO} = \int d\tau \left[ P \cdot \dot{X} - (i/2)(Z^\dagger \dot{Z} - \dot{Z}^\dagger Z) + P \cdot P + \omega^2 \right]
$$

(92)

where the polarization-coupling matrix is given by

$$
\Sigma(x, p) = -(\nabla \omega^2) \cdot F - \frac{\omega^2}{p_0} (e_\parallel \cdot \Omega)
$$

(93)

and $Z(\tau)$ is a complex-valued vector with two components that describe the degree of polarization along the vectors $e_1$ and $e_2$. It is normalized such that $Z^\dagger Z = 1$.

In the action (92), the two polarization modes are coupled through the Pauli matrix $\sigma_z$. However, these modes can be decoupled when using the basis of circularly polarized modes. We introduce the variable transformation

$$
Z(\tau) = R \Gamma(\tau),
$$

(94)

where

$$
R = \frac{1}{\sqrt{2}} \begin{pmatrix}
1 & 1 \\
1 & -1
\end{pmatrix}
$$

(95)

and $\Gamma(\tau)$ is a new vector with components denoted as

$$
\Gamma(\tau) = \begin{pmatrix}
\Gamma^+ \\
\Gamma^-
\end{pmatrix}.
$$

(96)

Inserting Eq. (94) into the action (92) leads to

$$
S_{XGO} = \int d\tau \left[ P \cdot \dot{X} - (i/2)(\Gamma^+ \dot{\Gamma} - \dot{\Gamma}^\dagger \Gamma) - P \cdot P + \omega^2 \right]
$$

(97)

where $\sigma_z$ is another Pauli matrix,

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

(98)
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, \( \Gamma \) is normalized such that \( \Gamma^\dagger \Gamma = 1 \).

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

\[
\begin{align*}
\delta P_\mu : & \quad \frac{dX^\mu}{d\tau} = 2P^\mu - \frac{\partial \Sigma}{\partial P_\mu} \Gamma \sigma_z \Gamma, \\
\delta X^\mu : & \quad \frac{dP_\mu}{d\tau} = \frac{\partial \omega_\mu^2}{\partial X^\mu} + \frac{\partial \Sigma}{\partial X^\mu} \Gamma \sigma_z \Gamma, \\
\delta \Gamma^\dagger : & \quad \frac{d\Gamma^\dagger}{d\tau} = -i\Sigma \sigma_z \Gamma, \\
\delta \Gamma : & \quad \frac{d\Gamma}{d\tau} = i\Gamma \sigma_z .
\end{align*}
\]

Together with Eq. (93), Eqs. (99) form a complete set of equations. The first terms on the right-hand side of Eqs. (99a) and (99b) 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. (99c) as an equation in the basis of linearly polarized modes:

\[
\dot{Z} = R \dot{\Gamma} = -i\Sigma R \sigma_z \Gamma = -i\Sigma (R \sigma_z R^{-1}) Z = -i\Sigma \sigma_y Z .
\]

[This equation could also be obtained if the ray equations were derived directly from the action (92).] Since \( \Sigma \) is a scalar and \( \sigma_y \) is constant, this can be readily integrated, yielding [42]

\[
Z(\tau) = \exp(-i\Theta \sigma_y) Z_0 = (\bar{v}_2 \cos \Theta - i\sigma_y \sin \Theta) Z_0 ,
\]

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

\[
Z(\tau) = \begin{pmatrix} \cos \Theta & -\sin \Theta \\ \sin \Theta & \cos \Theta \end{pmatrix} Z_0 .
\]

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 \( (e_1, e_2) \). The first term in Eq. (93) is identified as the rate of change of the wave Berry phase [6]. (In optics, the rotation of the polarization plane caused by the Berry phase is also known as the Rytov rotation [37, 43, 44].) The second term in Eq. (93) 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 (97) can be simplified to \( S_{XGO} = \int d\tau L \), where the Lagrangian is given by

\[
L_\pm = P \cdot \dot{X} - P \cdot P + \omega_\mu^2(X) \pm \Sigma(X, P) .
\]

Here the Lagrangian \( L_\pm \) governs the propagation of right-hand and left-hand polarization modes, respectively. The corresponding ELEs are

\[
\begin{align*}
\delta P_\mu : & \quad \frac{dX^\mu}{d\tau} = 2P^\mu - \frac{\partial \Sigma}{\partial P_\mu}, \\
\delta X^\mu : & \quad \frac{dP_\mu}{d\tau} = \frac{\partial \omega_\mu^2}{\partial X^\mu} \pm \frac{\partial \Sigma}{\partial X^\mu} .
\end{align*}
\]
or in terms of spacetime components,
\[
\begin{align*}
\frac{dX^0}{d\tau} &= 2P^0 \mp \frac{\partial \Sigma}{\partial P_0}, \\
\frac{dX}{d\tau} &= 2P \pm \frac{\partial \Sigma}{\partial P}, \\
\frac{dP_0}{d\tau} &= 0, \\
\frac{dP}{d\tau} &= -\frac{\partial \omega_p^2}{\partial X} \mp \frac{\partial \Sigma}{\partial X}.
\end{align*}
\]
The first terms on the right-hand side of Eqs. (104) describe the ray dynamics in the GO limit. The second terms describe the coupling of the mode polarization and the ray curvature. They are also responsible for the polarization-driven bending of ray trajectories.

As shown, \( P_0 \) remains constant because the background medium is time independent. In order to obtain the value of \( P_0 \), we note that the ray Hamiltonian
\[
H_\pm(X, P) = -P \cdot P + \omega_p^2(X) \pm \Sigma(X, P) \tag{105}
\]
is independent of \( \tau \), so one can readily verify that
\[
H_\pm(X(\tau), P(\tau)) = \text{constant}. \tag{106}
\]
Setting the Hamiltonian equal to zero, we use Eq. (105) to determine \( P_0 \). One finds
\[
P_0 \simeq \omega(X, P) \pm \Sigma(X, P_\mu/[2\omega(X, P)]}, \tag{107}
\]
where \( \omega(X, P) \pm (P^2 + \omega_p^2)^{1/2} \) is the wave frequency in the GO limit and \( P_\mu(X, P) \equiv (\omega(X, P), P) \).

**F. Numerical simulations**

To illustrate the polarization-driven divergence of the ray trajectories, Fig. 1 shows the ray trajectories for a right-polarized and left-polarized waves using the Lagrangian (103). For completeness, we also show the calculated ray trajectory as determined by the lowest-order GO ray Lagrangian
\[
L_{GO} = P \cdot \dot{X} - P \cdot P + \omega_p^2(X), \tag{108}
\]
which does not account for polarization effects. As anticipated, the ray trajectories predicted by the Lagrangian (103) differ noticeably from the “spinless” ray trajectory predicted by Eq. (108); namely:
\[
\begin{align*}
\delta P_\mu : \quad & \frac{dX^\mu}{d\tau} = 2P^\mu, \\
\delta X^\mu : \quad & \frac{dP_\mu}{d\tau} = \frac{\partial \omega_p^2}{\partial X^\mu}.
\end{align*}
\]
This divergence along the \( x \)-axis is driven by polarization effects. For EM waves propagating in isotropic non-birefringent dielectrics, this effect is called the Hall effect of light in the optics literature [7].

**G. Noncanonical representation and the Berry connection**

It is possible to obtain an alternative, noncanonical representation of the ray Lagrangian (103) that is invariant with respect to the choice of \( F(p) \) for pure states and explicitly shows the so-called Berry connection. Starting from Eq. (103) and substituting Eq. (93), we can write
\[
L_\pm = P \cdot \dot{X} - P \cdot P + \omega_p^2(X) \\
\mp (\nabla \omega_p^2(X) \cdot F \mp \frac{\omega_p^2(X)}{P_0} [e_p(P) \cdot \Omega(X)]) \\
\simeq P \cdot \dot{X} - P \cdot P + \omega_p^2(X \mp F) \mp \frac{\omega_p^2(X)}{P_0} [e_p(P) \cdot \Omega(X)], \tag{110}
\]
where we assumed that \( \omega_p^2(x) \) is smooth and neglected terms of \( \mathcal{O}(-) \) as usual. Introducing the variables
\[
x^\mu(\tau) = (X^0, X \mp F(P)), \quad p^\mu(\tau) \mp P_\mu \tag{111}
\]
and substituting them into Eq. (110), we obtain
\[
L_\pm \simeq p \cdot \dot{x} - p \cdot p + \omega_p^2(x) \\
\mp p \cdot \dot{F} \mp \frac{\omega_p^2(x)}{P_0} [e_p(p) \cdot \Omega(x)] \\
= p \cdot \dot{x} - p \cdot p + \omega_p^2(x) \\
\pm p \cdot F(p) \mp \frac{\omega_p^2(x)}{P_0} [e_p(p) \cdot \Omega(x)], \tag{112}
\]
where we dropped a perfect time derivative. We also approximated \( X \simeq x \) in the Faraday rotation term in Eq. (112) since it is already \( \mathcal{O}(\epsilon) \). Note that \( |x - X| \) is of the order of the wavelength, i.e., small enough to make \( x \) and \( X \) equally physical as measures of the ray location.

The term \( \dot{p} \cdot F(p) \) is known as the Berry connection term [8]. It is to be noted that adding \( \partial_p \chi(p) \) to \( F(p) \), where \( \chi(p) \) is an arbitrary scalar function, changes \( L_\pm \) by a perfect derivative and does not affect the equations of motion. The ELEs corresponding to the Lagrangian (112) are given by
\[
\begin{align*}
\frac{dx^0}{d\tau} &= 2p_0 \mp \frac{\omega_p^2}{P_0} (e_p \cdot \Omega), \tag{113a} \\
\frac{dx}{d\tau} &= 2p \pm \dot{p} \times (\nabla_p \times F) \mp \frac{\omega_p^2}{P_0} \frac{\partial}{\partial p} (e_p \cdot \Omega), \tag{113b} \\
\frac{dp_0}{d\tau} &= 0, \tag{113c} \\
\frac{dp}{d\tau} &= -\frac{\partial \omega_p^2}{\partial x} \pm \frac{\partial}{\partial x} \left[ \frac{\omega_p^2}{P_0} (e_p \cdot \Omega) \right]. \tag{113d}
\end{align*}
\]
These equations are equivalent to Eqs. (104) within the accuracy of the theory. Substituting Eq. (90), we can also write Eq. (113b) as
\[
\frac{dx}{d\tau} = 2p \pm \frac{\dot{p} \times p}{|p|^3} \mp \frac{\omega_p^2}{P_0} \frac{\partial}{\partial p} (e_p \cdot \Omega). \tag{114}
\]
Hence, with the use of the noncanonical coordinates \((x,p)\), the equations of motion no longer depend on the specific choice of \(F(p)\); i.e., they are invariant with respect to the choice (88) of vectors \(e_1\) and \(e_2\). Note that the same equations could be obtained directly from the point-particle limit of Eq. (84), if one substitutes \(-\nabla \lambda = \dot{p}\). For an extended discussion of pure states governed by noncanonical Lagrangians, see Ref. [12].

### IX. CONCLUSIONS

Even when neglecting diffraction, 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 position and momentum coordinates. However, vector waves have another degree of freedom, namely, their polarization. Polarization dynamics are manifested in two forms: (i) mode conversion, which is the transfer of wave quanta between resonant eigenmodes and can be understood as the precession of the wave spin, and (ii) polarization-driven bending of ray trajectories, which refers to deviations of the GO ray trajectories arising from first-order corrections to the GO dispersion relation. They are easily understood by drawing parallels with quantum mechanics, where similar effects (yet involving \(\hbar\)) are known as spin rotation and spin-orbital coupling.

In this work, we propose a first-principle variational formulation that captures both types of polarization-related effects simultaneously. We consider general linear nondissipative waves, whose dynamics are determined by some dispersion operator \(\hat{D}\). Using the Feynman parameterization and the Weyl calculus, we obtain a reduced Lagrangian model for such general waves. In contrast with the traditional GO Lagrangian, which is \(O(\epsilon^0)\)-accurate in the GO parameter \(\epsilon\), our Lagrangian is \(O(\epsilon)\)-accurate. In our procedure, polarization effects are contained in the \(O(\epsilon)\) corrections to the GO Lagrangian. These corrections may be especially significant for modeling RF waves in laboratory plasmas because such waves can have not-too-small \(\epsilon\) (as opposed, for instance, to quantum particles whose spin effects are typically weak). As an example, we apply the formulation to study the polarization-driven divergence of RF waves propagating in weakly magnetized plasma. Assessing the importance of polarization effects on waves propagating in strongly magnetized plasma will be discussed in a separate paper. 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, 45–47].) The Weyl symbol \(A(x,p)\) of any given operator \(\hat{A}\) is defined as

\[
A(x,p) \doteq \int d^4s \, e^{ip\cdot s} \langle x + s/2 | \hat{A} | x - s/2 \rangle ,
\]

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

\[
\hat{A} = \int d^4x \, d^4p \, d^4s \, e^{ip\cdot s/\epsilon} \langle x - s/2 | \hat{A} | x + s/2 \rangle .
\]

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

\[
A(x,x') = \int \frac{d^4p}{(2\pi)^4} \, e^{ip \cdot (x' - x)} A \left( \frac{x + x'}{2}, p \right).
\]

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

- For any operator \(\hat{A}\), the trace \(\text{Tr}[\hat{A}] = \int d^4x \, \langle x | \hat{A} | x \rangle\) can be expressed as

\[
\text{Tr}[\hat{A}] \doteq \int d^4x \, d^4p \, A(x,p).
\]

- 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 \(\hat{C} = \hat{A} \hat{B}\), the corresponding Weyl symbols satisfy [48, 49]

\[
C(x,p) = A(x,p) * B(x,p).
\]

Here ‘\(*\)’ refers to the Moyal product, which is given by

\[
A(x,p) * B(x,p) \doteq A(x,p) e^{\hat{L}/2} B(x,p),
\]

and \(\hat{L}\) is the Janus operator

\[
\hat{L} \doteq \frac{\partial}{\partial p_0} \cdot \mathbf{\partial}_x - \frac{\partial}{\partial x_0} \cdot \mathbf{\partial}_p = \{ \cdot, \cdot \}.
\]

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

\[
\hat{L} = \frac{\partial}{\partial p^0} \frac{\partial}{\partial x^0} - \frac{\partial}{\partial x^0} \frac{\partial}{\partial p^0} + \frac{\partial}{\partial x} \cdot \frac{\partial}{\partial p} - \frac{\partial}{\partial p} \cdot \frac{\partial}{\partial x}.
\]

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

\[
A * B \simeq AB + \frac{i}{2} \{ A, B \}.
\]
The Moyal product is associative; i.e.,

\[ A \star B \star C = (A \star B) \star C = A \star (B \star C). \]  

(A10)

Now we tabulate some Weyl transforms of various operators. (We use a two-sided arrow to show the correspondence with the Weyl transform.) First of all, the Weyl transforms of the identity, position, and momentum operators are given by

\[ \hat{1} \leftrightarrow 1, \quad \hat{x}^\mu \leftrightarrow x^\mu, \quad \hat{p}_\mu \leftrightarrow p_\mu. \]  

(A11)

For any two functions \( f \) and \( g \), one has

\[ f(\hat{x}) \leftrightarrow f(x), \quad g(\hat{p}) \leftrightarrow g(p). \]  

(A12)

Similarly, using Eq. (A6), one has

\[ \hat{p}_\mu f(\hat{x}) \leftrightarrow p_\mu f(x) + (i/2)\partial_\mu f(x), \]  

(A13)

\[ f(\hat{x})\hat{p}_\mu \leftrightarrow p_\mu f(x) - (i/2)\partial_\mu f(x). \]  

(A14)

[1] 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).

[2] G. B. Whitham, Linear and Nonlinear Waves (Wiley, New York, 2011).

[3] G. B. Whitham, “A general approach to linear and nonlinear dispersive waves using a Lagrangian,” J. Fluid Mech. 22, 273 (1965).

[4] I. Y. Dodin and N. J. Fisch, “Axiomatic geometrical optics, Abraham-Minkowski controversy, and photon properties derived classically,” Phys. Rev. A 86, 053834 (2012).

[5] R. A. Isaacson, “Gravitational radiation in the limit of high frequency. I. The linear approximation and geometrical optics,” Phys. Rev. 166, 1263 (1968).

[6] M. V. Berry, “Quantal phase factors accompanying adiabatic changes,” 392, 45 (1984).

[7] K. Y. Bliokh, A. Niv, V. Kleiner, and E. Hasman, “Geometrodynamics of spinning light,” Nature Photon. 2, 748 (2008).

[8] K. Y. Bliokh, F. J. Rodríguez-Fortuño, F. Nori, and A. V. Zayats, “Spin–orbit interactions of light,” Nature Photon. 9, 796 (2015).

[9] M. Onoda, S. Murakami, and N. Nagaosa, “Hall effect of light,” Phys. Rev. Lett. 93, 083901 (2004).

[10] A. V. Dooghin, N. D. Kundíkova, V. S. Liberman, and B. Y. Zel’ dovich, “Optical Magnus effect,” Phys. Rev. A 45, 8204 (1992).

[11] V. S. Liberman and B. Y. Zel’dovich, “Spin-orbit interaction of a photon in an inhomogeneous medium,” Phys. Rev. A 46, 5199 (1992).

[12] R. G. Littlejohn and W. G. Flynn, “Geometric phases in the asymptotic theory of coupled wave equations,” Phys. Rev. A 44, 5239 (1991).

[13] S. Weigert and R. G. Littlejohn, “Diagonalization of multicomponent wave equations with a Born-Oppenheimer example,” Phys. Rev. A 47, 3506 (1993).

[14] L. Friedland, “Renormalized geometric optics description of mode conversion in weakly inhomogeneous plasmas,” Phys. Fluids 28, 3260 (1985).

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

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

[17] W. G. Flynn and R. G. Littlejohn, “Normal forms for linear mode conversion and Landau-Zener transitions in one dimension,” Ann. Phys. 234, 334 (1994).

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

[19] E. R. Tracy, A. N. Kaufman, and A. Jaun, “Local fields for asymptotic matching in multidimensional mode conversion,” Phys. Plasmas 14, 082102 (2007).

[20] A. S. Richardson and E. R. Tracy, “Quadratic corrections to the metaplectic formulation of resonant mode conversion,” J. Phys. A: Math. Theor. 41, 375501 (2008).

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

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

[23] I. Y. Dodin, “Geometric view on noneikonal waves,” Phys. Lett. A 378, 1598 (2014).

[24] D. E. Ruiz, C. L. Ellison, and I. Y. Dodin, “Relativistic ponderomotive Hamiltonian of a Dirac particle in a vacuum laser field,” Phys. Rev. A 92, 062124 (2015).

[25] D. E. Ruiz and I. Y. Dodin, “On the correspondence between quantum and classical variational principles,” Phys. Lett. A 379, 2623 (2015).

[26] I. Y. Dodin, A. I. Zhmoginov, and D. E. Ruiz, “Variational principles for dissipative (sub)systems, with applications to the theory of dispersion and geometrical optics,” arXiv (2016), 1610.05668v1.

[27] When parameters of the medium oscillate, the wave Lagrangian density generally exhibits additional terms proportional to \( \psi^\dagger \psi \) and \( \psi^\dagger \psi \). Absent parametric resonances, these terms average to zero on large enough time scales and thus have little effect of the time-averaged dynamics [23], except at strong enough modulation.

[28] A. N. Kaufman, H. Ye, and Y. Hui, “Variational formulation of covariant eikonal theory for vector waves,” Phys. Lett. A 120, 327 (1987).

[29] R. G. Littlejohn and R. Winston, “Classifications to classical radiometry,” J. Opt. Soc. Am. A 10, 2024 (2003).

[30] In the \( x \)-representation, we have the following:

\[ \langle x | p_\mu | \psi \rangle = \int d^4 x' \langle x' | p_\mu | x \rangle \langle x | \psi \rangle = \int d^4 x' \partial_\mu \psi(x' - x')/\partial x \psi(x') = i \partial_\mu \psi(x). \]

[31] T. J. Bridges and S. Reich, “Multi-symplectic integrators: numerical schemes for Hamiltonian PDEs that conserve symplecticity,” Phys. Lett. A 284, 184 (2001).

[32] R. P. Feynman, “An operator calculus having applications in quantum electrodynamics,” Phys. Rev. 84, 108
[33] J. P. Aparicio, F. H. Gaioli, and E. T. Garcia Alvarez, “Interpretation of the evolution parameter of the Feynman parametrization of the Dirac equation,” Phys. Lett. A 200, 233 (1995).

[34] W. D. Hayes, “Group velocity and nonlinear dispersive wave propagation,” Proc. R. Soc. Lond. A 332, 199 (1973).

[35] V. Anisovich, M. Kobrinsky, J. Nyiri, Yu. Shabelski, Quark Model and High Energy Collisions, World Scientific, River Edge, 2004. Appendix A.

[36] 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).

[37] K. Y. Bliokh, D. Y. Frolov, and Y. A. Kravtsov, “Non-Abelian evolution of electromagnetic waves in a weakly anisotropic inhomogeneous medium,” Phys. Rev. A 75, 053821 (2007).

[38] E. Schrödinger, Sitzungsber. Preuss. Akad. Wiss. Phys. Math. Kl. 24, 418 (1930); Schrödinger’s derivation is reproduced in A. O. Barut and A. J. Bracken, Phys. Rev. D 23, 2454 (1981).

[39] C. Zener, “Non-adiabatic crossing of energy levels,” Proc. R. Soc. Lond. A 137, 696 (1932).

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

[41] It is to be noted that the $\alpha$ matrices are related to the Gell-Mann matrices, which serve as infinitesimal generators of the special unitary group SU(3).

[42] Here we used the well known Euler formula for Pauli matrices, $e^{i(n \cdot \sigma)} = \hat{I}_2 \cos a + i(n \cdot \sigma) \sin a$.

[43] S. M. Rytov, Dokl. Akad. Nauk SSSR 18, 263 (1938); V. V. Vladimirskii, Dokl. Akad. Nauk SSSR 31, 222 (1941); reprinted in Topological Phases in Quantum Theory, edited by B. Markovski and S. I. Vinitsky (World Scientific, Singapore, 1989).

[44] A. Tomita and R. Y. Chiao, “Observation of Berry’s topological phase by use of an optical fiber,” Phys. Rev. Lett. 57, 937 (1986).

[45] K. Imre, E. Ozizmir, M. Rosenbaum, and P. F. Zweifel, “Wigner method in quantum statistical mechanics,” J. Math. Phys. 8, 1097 (1967).

[46] S. W. McDonald, “Phase-space representations of wave equations with applications to the eikonal approximation for short-wavelength waves,” Phys. Rep. 158, 337 (1988).

[47] G. A. Baker Jr., “Formulation of quantum mechanics based on the quasi-probability distribution induced on phase space,” Phys. Rev. 109, 2198 (1958).

[48] J. E. Moyal, “Quantum mechanics as a statistical theory,” Math. Proc. Cambridge Philos. Soc. 45, 99 (1949).

[49] H. J. Groenewold, “On the principles of elementary quantum mechanics,” Physica 12, 405 (1946).