First-principles variational formulation of polarization effects in geometrical optics

D. E. Ruiz and I. Y. Dodin

First-principle variational formulation of polarization effects in geometrical optics

D. E. Ruiz and I. Y. Dodin

Department of Astrophysical Sciences, Princeton University, Princeton, New Jersey 08544, USA

The propagation of electromagnetic waves in isotropic dielectric media with local dispersion is studied under the assumption of small but nonvanishing λ/ℓ, where λ is the wavelength, and ℓ is the characteristic inhomogeneity scale. It is commonly known that, due to nonzero λ/ℓ, such waves can experience polarization-driven bending of ray trajectories and polarization dynamics that can be interpreted as the precession of the wave "spin". The present work reports how Lagrangians describing these effects can be deduced, rather than guessed, within a strictly classical theory. In addition to the commonly known ray Lagrangian that features the Berry connection, a simple alternative Lagrangian is proposed that naturally has a canonical form. The presented theory captures not only the eigenray dynamics but also the dynamics of continuous wave fields and rays with mixed polarization, or "entangled" waves. The calculation assumes stationary lossless media with isotropic local dispersion, but generalizations to other media are straightforward to do.

I. INTRODUCTION

Electromagnetic (EM) waves propagating in inhomogeneous linear media exhibit a variety of intriguing phenomena associated with wave polarization. Those include precession of the polarization vector, known as the Rytov rotation [1–4], and also the polarization-driven bending of ray trajectories, known as the optical Magnus effect or the optical Hall effect [5–8]. As overviewed recently in Ref. [9], these phenomena can be attributed as manifestations of the Berry phase [10], which is a fundamental concept emerging also in many other areas of physics [11–13]. Hence the subject has been attracting an increased amount of attention, especially with focus on variational formulations, which are particularly elucidating at studying EM polarization effects in the geometrical-optics (GO) limit [7, 14–17]. Yet a straightforward universal theory is still lacking. Within existing models, ray Lagrangians and Hamiltonians either have to be guessed or are derived using additional postulates like quantization, which is, by definition, alien to classical theory. Such *ad hoc* approaches result in a number of limitations; e.g., they are not readily applicable to media with nonlocal and strongly anisotropic dispersion like magnetized plasma. Thus, there is a compelling need for a generalized and simplified description of EM polarization effects from first principles.

The purpose of this paper is to present such a description. More specifically, what we report here is a new application of the general theory that we derived earlier in Ref. [18] for the Dirac electron. In the present paper, we explain how *exactly the same theory* can be applied to GO EM waves and polarization effects in particular. As opposed to Ref. [16], our derivation is entirely classical. It is also straightforward and elementary in the sense that the wave Lagrangian does not need to be guessed, as in Refs. [7, 19], but is rather deduced (basically, using nothing more than matrix multiplication) according to a formalized algorithm. The known results are successfully reproduced and are extended as follows: (i) We show that, in addition to the commonly known ray La-grangian that features the Berry connection [19–21], an alternative ray Lagrangian is possible that naturally has a canonical form and corresponds to a Hamiltonian simpler than that proposed in Ref. [7]. We explain how the two Lagrangians are related and demonstrate their (approximate) equivalence numerically. (ii) As opposed to the result of Ref. [19], our Lagrangians are expressed in terms of physical time and wave vector, so they capture the complete ray dynamics rather than just the ray trajectory. (iii) In addition to eigenray equations, we also derive equations for continuous wave fields and rays with mixed polarization, or "entangled" waves. This description captures both the Rytov rotation and the optical Hall effect simultaneously, and it is also manifestly conservative, since the amplitude equations too are derived from a variational principle.

The calculation presented below applies to arbitrary linear stationary lossless media with isotropic local dispersion, i.e., media whose dielectric and magnetic permittivities are real scalars depending only on spatial coordinates. However, within our theory, generalizations to strongly anisotropic media with nonlocal dispersion, such as magnetized plasma, are straightforward to do. The present paper is intended as an introduction to such calculations, which will be reported separately. Thus, below we intentionally focus on a relatively simple problem to show how our general theory fits existing literature.

This work is organized as follows. In Sec. II, we present a Lagrangian formalism to describe EM wave propagation in isotropic stationary dielectric media with local dispersion. In Sec. III, we obtain a reduced model which captures first-order polarization effects in transverse EM waves. In Sec. IV, we discuss a fluid Lagrangian model, which describes the dynamics of the wave envelope, ray trajectory, and polarization. In Sec. V, we derive ray equations in canonical form. In Sec. VI, we present the noncanonical ray equations, and we compare them with those previously reported. In Sec. VII, we summarize our main results.

II. BASIC EQUATIONS

A. Photon wave function

In this paper, we consider EM wave propagation in isotropic lossless dielectric media with local linear dispersion. In this case, the governing equations for the electric field $\mathbf{E}(t,\mathbf{x})$ and magnetic field $\mathbf{H}(t,\mathbf{x})$ are

$$\partial_t \mathbf{E} = \frac{c}{\varepsilon} \boldsymbol{\nabla} \times \mathbf{H}, \tag{1}$$

$$\partial_t \mathbf{H} = -\frac{c}{\mu} \boldsymbol{\nabla} \times \mathbf{E}, \tag{2}$$

where $\varepsilon = \varepsilon(\mathbf{x})$ is the electric permittivity, and $\mu = \mu(\mathbf{x})$ is the magnetic permeability.

Let us introduce the following normalized fields,

$$\bar{\mathbf{E}} = \sqrt{\varepsilon}\, \mathbf{E}, \tag{3}$$

$$\bar{\mathbf{H}} = \sqrt{\mu}\, \mathbf{H}, \tag{4}$$

so one can rewrite Eqs. (1) and (2) as

$$\partial_t \bar{\mathbf{E}} = \frac{c}{n} \boldsymbol{\nabla} \times \bar{\mathbf{H}} - \frac{c}{n} \boldsymbol{\nabla} (\ln \sqrt{\mu}) \times \bar{\mathbf{H}}, \tag{5}$$

$$\partial_t \bar{\mathbf{H}} = -\frac{c}{n} \boldsymbol{\nabla} \times \bar{\mathbf{E}} + \frac{c}{n} \boldsymbol{\nabla} (\ln \sqrt{\varepsilon}) \times \bar{\mathbf{E}}, \tag{6}$$

where $n(\mathbf{x}) \doteq \sqrt{\varepsilon\mu}$ is the refraction index. (From here one, the symbol "$\doteq$" denotes definitions.) Due to the linearity of these equations, we may formally extend them to complex fields. Then, we can express Eqs. (5) and (6) as a vector Schrödinger equation:

$$i\partial_t \psi = H(\mathbf{x}, -i\boldsymbol{\nabla})\psi, \tag{7}$$

where

$$\psi(t,\mathbf{x}) \doteq \begin{pmatrix} \bar{\mathbf{E}} \\ \bar{\mathbf{H}} \end{pmatrix} \tag{8}$$

is a six-component wave function, which can be interpreted as the photon wave function [22, 23]. The Hamiltonian operator is a 6×6 matrix given by

$$H(\mathbf{x}, \hat{\mathbf{k}}) \doteq \frac{c}{n} \boldsymbol{\lambda} \cdot \hat{\mathbf{k}} + \mathcal{A}, \tag{9}$$

where $\hat{\mathbf{k}} \doteq -i\boldsymbol{\nabla}$ is the wave vector (momentum) operator, the $\boldsymbol{\lambda}$ matrices are 6×6 Hermitian matrices,

$$\boldsymbol{\lambda} \doteq \begin{pmatrix} 0 & i\boldsymbol{\alpha} \\ -i\boldsymbol{\alpha} & 0 \end{pmatrix}, \tag{10}$$

and the matrix $\mathcal{A}(\mathbf{x})$ is defined as follows:

$$\mathcal{A}(\mathbf{x}) \doteq \frac{c}{n} \begin{pmatrix} 0 & -\boldsymbol{\alpha} \cdot \boldsymbol{\nabla}(\ln \sqrt{\mu}) \\ \boldsymbol{\alpha} \cdot \boldsymbol{\nabla}(\ln \sqrt{\varepsilon}) & 0 \end{pmatrix}. \tag{11}$$

Here, the $\boldsymbol{\alpha}$ matrices are 3×3 traceless Hermitian matrices given by [24]

$$\alpha^x \doteq \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \tag{12}$$

$$\alpha^y \doteq \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}, \tag{13}$$

$$\alpha^z \doteq \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{14}$$

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

$$(\boldsymbol{\alpha} \cdot \mathbf{A})\mathbf{B} = i\mathbf{A} \times \mathbf{B}, \tag{15}$$

$$\mathbf{A}^T \alpha^j \mathbf{B} = -i(\mathbf{A} \times \mathbf{B})^j, \tag{16}$$

where T denotes the matrix transpose.

Let us also point out that, since the Hamiltonian is Hermitian, it conserves the wave action:

$$\partial_\mu \left(\psi^\dagger \gamma^\mu \psi \right) = 0, \tag{17}$$

where $\partial_\mu = (\partial_t, \boldsymbol{\nabla})$, $\gamma^\mu \doteq (\mathbb{I}_6, c\boldsymbol{\lambda}/n)$, and $\mathbb{I}_6$ denotes the 6×6 unit matrix. Throughout this work, the Minkowski metric is adopted with signature $(+,-,-,-)$. (Generalizations to curved metrics are straightforward to apply [23].) Greek indices span from 0 to 3 and refer to spacetime coordinates, x^μ, with x^0 corresponding to the time variable, t. Latin indices span from 1 to 3 and denote the spatial variables, x^i (except where specified otherwise). Summation over repeated indices is assumed. Hence, Eq. (17) can be explicitly written as

$$\frac{\partial}{\partial t} \left(\frac{|\bar{\mathbf{E}}|^2 + |\bar{\mathbf{H}}|^2}{16\pi} \right) + \boldsymbol{\nabla} \cdot \left[\frac{c}{n} \frac{\mathrm{Re}(\bar{\mathbf{E}}^* \times \bar{\mathbf{H}})}{8\pi} \right] = 0, \tag{18}$$

where the 8π factor was introduced to emphasize the connection to the well known Poynting theorem [25].

B. Lagrangian density

Equation (7) has a form akin to the so-called multi-symplectic form [26–28] and can be readily given a variational interpretation. Specifically, Eq. (7) can be obtained as the Euler-Lagrange equation, $\delta\Lambda = 0$, for the action integral

$$\Lambda = \int \mathfrak{L} \, \mathrm{d}^4 x, \tag{19}$$

where $\mathrm{d}^4 x \equiv \mathrm{d}t \, \mathrm{d}^3 x$, and the Lagrangian density $\mathfrak{L}$ is given by

$$\mathfrak{L} = \frac{i}{2} \left[\psi^\dagger \gamma^\mu (\partial_\mu \psi) - (\partial_\mu \psi^\dagger) \gamma^\mu \psi \right] + \psi^\dagger \mathcal{M} \psi, \tag{20}$$

Here we adopted natural units, such that $c = 1$. We also introduced

$$\mathcal{M}(\mathbf{x}) = \frac{1}{2n} \begin{pmatrix} 0 & \boldsymbol{\alpha} \\ \boldsymbol{\alpha} & 0 \end{pmatrix} \cdot \boldsymbol{\nabla} \ln Z(\mathbf{x}), \qquad (21)$$

where $Z(\mathbf{x})$ is the impedance of the medium,

$$Z(\mathbf{x}) \doteq \sqrt{\frac{\mu}{\varepsilon}}. \qquad (22)$$

It is to be noted that no approximations have been used in order to obtain this Lagrangian model. Notice also that the Lagrangian density (20) involves only first-order differential operators. In Ref. [18], we showed that such a simple structure is convenient for studying a point-particle model of the spin-orbit coupling for the Dirac electron. Below, we report how the calculation can also be extended to study the effect of polarization of classical waves in dielectrics of the specified type.

III.   REDUCED MODEL

We consider waves such that

$$\psi(t, \mathbf{x}) = \xi e^{i\theta}, \qquad (23)$$

where $\theta(t, \mathbf{x})$ is some rapid real phase, and $\xi(t, \mathbf{x})$ is a vector evolving slowly compared to $\theta(t, \mathbf{x})$. Hence, we assume that the inhomogeneity scale ℓ is large compared to the wavelength; i.e.,

$$\epsilon \doteq \frac{1}{k\ell} \ll 1, \qquad (24)$$

where $k \doteq |\mathbf{k}|$, and

$$\mathbf{k} \doteq \boldsymbol{\nabla}\theta \qquad (25)$$

is the wave vector. Below, we construct a reduced model of EM wave propagation using the smallness of ϵ. Unlike the standard geometrical-optics (GO) theory, which corresponds to a Lagrangian accurate to the zeroth order in ϵ [29, 30], this reduced model will yield a ray Lagrangian accurate to the first order in ϵ.

A.   Eigenmodes in the limit of vanishing ϵ

In general, there exist multiple eigenfrequencies,

$$\omega \doteq -\partial_t \theta, \qquad (26)$$

that correspond to a local wave vector $\mathbf{k}$. At vanishingly small ϵ, these eigenfrequencies are found from the GO limit of Eq. (7), namely,

$$\omega \xi = H_0(\mathbf{x}, \mathbf{k})\xi, \qquad (27)$$

where $H_0(\mathbf{x}, \mathbf{k}) \doteq \boldsymbol{\lambda} \cdot \mathbf{k}/n$. Notice that the matrix $\mathcal{A}$ does not enter H_0, because $\mathcal{A}$ is of the first order in ϵ.

In Eq. (27), there exists in general six independent eigenmodes, since H_0 is Hermitian. These modes can be readily obtained from

$$\omega^2 \xi = H_0^2 \xi = \frac{1}{n^2} \begin{pmatrix} (\boldsymbol{\alpha} \cdot \mathbf{k})^2 & 0 \\ 0 & (\boldsymbol{\alpha} \cdot \mathbf{k})^2 \end{pmatrix} \xi. \qquad (28)$$

Denoting the components of ξ as $\xi^T \doteq \begin{pmatrix} \mathbf{a}_1^T & \mathbf{a}_2^T \end{pmatrix}$, the equation for the first three components is

$$\begin{aligned} \omega^2 \mathbf{a}_1 &= \frac{1}{n^2}(\boldsymbol{\alpha} \cdot \mathbf{k})(\boldsymbol{\alpha} \cdot \mathbf{k})\mathbf{a}_1 \\ &= -\frac{1}{n^2}\mathbf{k} \times (\mathbf{k} \times \mathbf{a}_1) \\ &= \frac{1}{n^2}\left[k^2 \mathbf{a}_1 - (\mathbf{k} \cdot \mathbf{a}_1)\mathbf{k} \right], \end{aligned} \qquad (29)$$

where, in the second line, we have used Eq. (15). As usual, two eigenmodes are related to the propagation of longitudinal ($\mathbf{k} \times \mathbf{a}_1 = 0$) modes with zero frequency, while the other four eigenmodes correspond to the propagation of transversal ($\mathbf{k} \cdot \mathbf{a}_1 = 0$) EM wave modes in the limit of vanishing ϵ.

We will particularly be interested in transversal EM modes with positive phase velocities, such that

$$\omega = k/n. \qquad (30)$$

Corresponding to this frequency, there are two orthonormal eigenvectors h_q, which are given by

$$h_1(\mathbf{k}) = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{pmatrix}, \qquad h_2(\mathbf{k}) = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{e}_2 \\ -\mathbf{e}_1 \end{pmatrix}, \qquad (31)$$

where $\mathbf{e}_1(\mathbf{k})$ and $\mathbf{e}_2(\mathbf{k})$ are any two orthonormal vectors in the plane normal to $\mathbf{e_k} \doteq \mathbf{k}/k$. A right-hand convention is adopted, such that $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e_k}$. Obviously, h_1 and h_2 are eigenvectors of H_0. For example,

$$\begin{aligned} H_0 h_1 &= \frac{1}{n}(\mathbf{k} \cdot \boldsymbol{\lambda}) h_1 \\ &= \frac{i}{n\sqrt{2}} \begin{pmatrix} (\boldsymbol{\alpha} \cdot \mathbf{k})\mathbf{e}_2 \\ -(\boldsymbol{\alpha} \cdot \mathbf{k})\mathbf{e}_1 \end{pmatrix} \\ &= \frac{k}{n\sqrt{2}} \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{pmatrix} \\ &= \omega h_1, \end{aligned} \qquad (32)$$

where in the third line, Eq. (15) was used. A similar calculation follows for h_2.

B.   Eigenmode decomposition

Since h_q form a complete basis, one can write $\xi = h_q \phi^q$, where ϕ^q are scalar functions. We will assume that only transverse modes with positive frequencies are actually excited (we call these modes "active"), whereas others

acquire nonzero amplitudes only through the medium in-homogeneity or the finite width of the EM wave packet (we call these modes "passive"). In this case,

$$\phi^q = \begin{cases} \mathcal{O}(\epsilon^0), & q = 1, 2, \\ \mathcal{O}(\epsilon^1), & q \neq 1, 2. \end{cases} \tag{33}$$

[For future references, note that $\phi^{1,2}(t, \mathbf{x})$ describe the envelopes corresponding to the linearly polarized modes with the electric field $\mathbf{E}$ aligned parallel to the unit vectors $\mathbf{e}_1$ and $\mathbf{e}_2$, respectively.] For given $\phi^{1,2}$, one can always calculate the amplitudes of passive modes using the complete set of Maxwell's equations, but this will not be necessary for our purposes. As shown in Ref. [18], due to the mutual orthogonality of all h_q, the contribution of passive modes to $\mathfrak{L}$ is $o(\epsilon)$, so it can be neglected entirely. In other words, for the purpose of calculating $\mathfrak{L}$, it is sufficient to adopt $\xi \approx h_1\phi^1 + h_2\phi^2$, even though the true ξ may have nonzero projections also on other h_q.

It is convenient to rewrite this decomposition in a matrix form,

$$\xi = \Xi\,\phi, \tag{34}$$

where

$$\phi(t, \mathbf{x}) \doteq \begin{pmatrix} \phi^1 \\ \phi^2 \end{pmatrix}, \tag{35}$$

and Ξ is a 6×2 matrix having h_q as its columns: i.e.,

$$\Xi(\mathbf{k}) \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 \\ \mathbf{e}_2 & -\mathbf{e}_1 \end{pmatrix}. \tag{36}$$

[Below, we also consider Ξ as a function of x^μ in the sense that $\Xi = \Xi(\mathbf{k}(x^\mu))$.] Then, inserting Eq. (34) into Eq. (20), one obtains [18]

$$\mathfrak{L} = \mathcal{K} - \phi^\dagger\left(\mathcal{E} - \mathcal{U}\right)\phi + o(\epsilon), \tag{37}$$

where

$$\mathcal{K} \doteq \frac{i}{2}\left[\phi^\dagger(d_t\phi) - (d_t\phi^\dagger)\phi\right], \tag{38}$$

$$\mathcal{E} \doteq \partial_t\theta + \frac{k}{n}, \tag{39}$$

$$\mathcal{U} \doteq \Xi^\dagger\mathcal{M}\Xi + \frac{i}{2}\left[\Xi^\dagger\gamma^\mu(\partial_\mu\Xi) - (\partial_\mu\Xi^\dagger)\gamma^\mu\Xi\right]. \tag{40}$$

Here, $d_t \doteq \partial_t + \mathbf{v}_0 \cdot \boldsymbol{\nabla}$ is a convective derivative associated to the zeroth order (in ϵ) geometrical optics velocity field,

$$\mathbf{v}_0 \doteq \frac{\partial\omega}{\partial\mathbf{k}} = \frac{1}{n}\mathbf{e_k}. \tag{41}$$

The terms $\mathcal{K}$ and $\mathcal{U}$, which are of order ϵ, represent corrections to the standard, lowest-order GO Lagrangian density. Below, $\mathcal{U}$ will be calculated explicitly, whereas the higher-order terms, $o(\epsilon)$, will be neglected.

C. Stern-Gerlach Hamiltonian

Let us search for a tractable expression for $\mathcal{U}$, which in Refs. [18, 31] was called the "Stern-Gerlach" Hamiltonian. Regarding the first term on the right hand side of Eq. (40), a straightforward calculation gives

$$\begin{aligned} \Xi^\dagger\mathcal{M}\Xi &= \frac{1}{2\sqrt{2}n}\Xi^\dagger\begin{pmatrix} 0 & \boldsymbol{\alpha}\cdot\boldsymbol{\nabla}\ln Z \\ \boldsymbol{\alpha}\cdot\boldsymbol{\nabla}\ln Z & 0 \end{pmatrix}\begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 \\ \mathbf{e}_2 & -\mathbf{e}_1 \end{pmatrix} \\ &= \frac{i}{4nZ}\begin{pmatrix} \mathbf{e}^1 & \mathbf{e}^2 \\ \mathbf{e}^2 & -\mathbf{e}^1 \end{pmatrix}\begin{pmatrix} \boldsymbol{\nabla}Z \times \mathbf{e}_2 & -\boldsymbol{\nabla}Z \times \mathbf{e}_1 \\ \boldsymbol{\nabla}Z \times \mathbf{e}_1 & \boldsymbol{\nabla}Z \times \mathbf{e}_2 \end{pmatrix} \\ &= 0. \end{aligned} \tag{42}$$

Here, we introduced the notation $\mathbf{e}^1 \doteq \mathbf{e}_1^T$ and $\mathbf{e}^2 \doteq \mathbf{e}_2^T$. Notably, these can be understood as adjoint basis vectors; e.g., $\mathbf{e}^2\mathbf{e}_1 \equiv (\mathbf{e}^2)_i(\mathbf{e}_1)^i \equiv \mathbf{e}_2 \cdot \mathbf{e}_1 = 0$. Hence, Eq. (40) is simplified down to

$$\mathcal{U} = \frac{i}{2}\left[\Xi^\dagger\gamma^\mu(\partial_\mu\Xi) - (\partial_\mu\Xi^\dagger)\gamma^\mu\Xi\right]. \tag{43}$$

Using that $\partial_\mu\Xi = (\partial\Xi/\partial k_j)(\partial_\mu k_j)$, one can rewrite $\mathcal{U}$ as follows [18]:

$$\mathcal{U} = \frac{i}{2}\dot{\mathbf{k}} \cdot \left[\Xi^\dagger(\partial_\mathbf{k}\Xi) - (\partial_\mathbf{k}\Xi^\dagger)\Xi\right]. \tag{44}$$

Here, $\dot{\mathbf{k}}(t)$ is given by the lowest-order (in ϵ) ray equation,

$$\dot{\mathbf{k}}(t) = -\partial_\mathbf{x}\omega(\mathbf{x}, \mathbf{k}), \tag{45}$$

and ω is given by Eq. (30). Using Eq. (36), we obtain

$$\begin{aligned} \Xi^\dagger\frac{\partial\Xi}{\partial k^j} - \frac{\partial\Xi^\dagger}{\partial k^j}\Xi = \\ \begin{pmatrix} 0 & \mathbf{e}^1\dfrac{\partial}{\partial k^j}\mathbf{e}_2 - \mathbf{e}^2\dfrac{\partial}{\partial k^j}\mathbf{e}_1 \\ \mathbf{e}^2\dfrac{\partial}{\partial k^j}\mathbf{e}_1 - \mathbf{e}^1\dfrac{\partial}{\partial k^j}\mathbf{e}_2 & 0 \end{pmatrix}. \end{aligned}$$

Since $\mathbf{e}^1\mathbf{e}_2 = 0$, Eq. (44) leads to

$$\mathcal{U} = -\dot{\mathbf{k}} \cdot \mathbf{A}(\mathbf{k})\sigma_y, \tag{46}$$

where σ_y is the y-component of the Pauli matrices,

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \tag{47}$$

and $\mathbf{A}(\mathbf{k})$ is a vector with components given by

$$A_j(\mathbf{k}) \doteq \mathbf{e}^1\frac{\partial}{\partial k^j}\mathbf{e}_2. \tag{48}$$

For example, one may choose [32]

$$\mathbf{e}_1(\mathbf{k}) \doteq \begin{pmatrix} \dfrac{k_x k_z}{k\sqrt{k_x^2 + k_y^2}} \\ \dfrac{k_y k_z}{k\sqrt{k_x^2 + k_y^2}} \\ -\dfrac{\sqrt{k_x^2 + k_y^2}}{k} \end{pmatrix}, \quad \mathbf{e}_2(\mathbf{k}) \doteq \begin{pmatrix} -\dfrac{k_y}{\sqrt{k_x^2 + k_y^2}} \\ \dfrac{k_x}{\sqrt{k_x^2 + k_y^2}} \\ 0 \end{pmatrix}, \tag{49}$$

so that

$$\mathbf{A}(\mathbf{k}) = \frac{\mathbf{e}_\perp \times \mathbf{e}_\mathbf{k}}{k_\perp}, \qquad (50)$$

where $\mathbf{k}_\perp^T \doteq (k_x, k_y, 0)$, and $\mathbf{e}_\perp \doteq \mathbf{k}_\perp/k_\perp$; or, more explicitly,

$$\mathbf{A}(\mathbf{k}) = \frac{k_z}{kk_\perp^2} \begin{pmatrix} k_y \\ -k_x \\ 0 \end{pmatrix}. \qquad (51)$$

D. Lagrangian density: summary

Substituting Eq. (46) into Eq. (37), we can express the Lagrangian density as

$$\mathfrak{L} = -\phi^\dagger \left(\partial_t \theta + \frac{k}{n} \right) \phi$$
$$+ \frac{i}{2} \left[\phi^\dagger (d_t \phi) - (d_t \phi^\dagger) \phi \right] - \phi^\dagger \sigma_y \Sigma \phi, \quad (52)$$

where

$$\Sigma \doteq \dot{\mathbf{k}} \cdot \mathbf{A}(\mathbf{k}). \qquad (53)$$

Let us also use a variable transformation,

$$\phi(t, \mathbf{x}) = Q\, \eta(t, \mathbf{x}), \qquad (54)$$

where

$$Q \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}, \qquad (55)$$

and $\eta(t, \mathbf{x})$ is a new vector with components denoted as

$$\eta(t, \mathbf{x}) \doteq \begin{pmatrix} \eta_+ \\ \eta_- \end{pmatrix}. \qquad (56)$$

Hence, the Lagrangian density (52) can be expressed as

$$\mathfrak{L} = -\eta^\dagger \left(\partial_t \theta + k/n \right) \eta$$
$$+ \frac{i}{2} \left[\eta^\dagger (d_t \eta) - (d_t \eta^\dagger) \eta \right] - \eta^\dagger \sigma_z \Sigma \eta, \quad (57)$$

where σ_z is another Pauli matrix,

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \qquad (58)$$

Here, $\eta_\pm(t, \mathbf{x})$ describe envelopes corresponding to right-hand and left-hand circularly polarized modes, respectively (as defined from the point of view of the source).

The first line of the right hand side of Eq. (57) represents the lowest order GO Lagrangian density for these modes. Additionally, the second line of Eq. (57), which contains $\mathcal{O}(\epsilon)$ terms, introduces polarization effects, as will be explained below.

IV. CONTINUOUS WAVE MODEL

Substituting Eqs. (30) and (45) in Eq. (53), we can rewrite $\mathfrak{L}$ as

$$\mathfrak{L} = -\eta^\dagger \left(\partial_t \theta + \frac{k}{n} \right) \eta$$
$$+ \frac{i}{2} \left[\eta^\dagger (d_t \eta) - (d_t \eta^\dagger) \eta \right] - \eta^\dagger \sigma_z \Sigma(\mathbf{x}, \mathbf{k}) \eta, \quad (59)$$

where

$$\Sigma(\mathbf{x}, \mathbf{k}) = \frac{k}{n^2} \mathbf{A}(\mathbf{k}) \cdot \boldsymbol{\nabla} n. \qquad (60)$$

In this form, the Lagrangian density is analogous to that of a semiclassical Pauli particle and thus can be approached similarly [31]. Let us adopt the representation $\eta = z\sqrt{\mathcal{I}}$, where $\mathcal{I}(t, \mathbf{x}) \doteq \psi^\dagger \psi$ is the action density, and $z(t, \mathbf{x})$ is a unit polarization vector ($z^\dagger z \equiv 1$). Since the common phase of the two components of z can be attributed to θ, we can parameterize z in terms of just two real functions, $\vartheta(t, \mathbf{x})$ and $\zeta(t, \mathbf{x})$:

$$z(\vartheta, \zeta) = \begin{pmatrix} e^{i\vartheta} \cos(\zeta/2) \\ e^{-i\vartheta} \sin(\zeta/2) \end{pmatrix}. $$

Like in the case of the Pauli particle, ζ determines the relative fraction of "spin-up" and "spin-down" quanta, i.e., those corresponding to left-hand and right-hand polarizations. Also, one can understand $\mathbf{S} \doteq z^\dagger \boldsymbol{\sigma} z/2$ as the wave "spin vector" ($\boldsymbol{\sigma}$ denotes the three Pauli matrices, as usual [18, 31]), or, up to a constant factor, as the Stokes vector [33, 34].

Using the above parameterization, the Lagrangian density is rewritten as

$$\mathfrak{L} = -\mathcal{I} \left[\partial_t \theta + \frac{k}{n} + (d_t \vartheta + \Sigma) \cos \zeta \right], \qquad (61)$$

which leads to four Euler-Lagrange equations (ELEs). (For future reference, when an ELE is presented, the denotation "δa :" means that the corresponding equation was obtained by varying the action integral with respect to a.) The first ELE is the action conservation theorem,

$$\delta\theta : \quad \partial_t \mathcal{I} + \boldsymbol{\nabla} \cdot (\mathcal{I}\mathbf{V}) = 0. \qquad (62)$$

The flow velocity is given by $\mathbf{V} = \mathbf{v}_0 + \mathbf{u}$, and

$$\mathbf{u} \doteq \frac{\partial}{\partial \mathbf{k}} \left(\frac{\mathbf{k} \cdot \boldsymbol{\nabla}\vartheta}{nk} + \Sigma \right) \cos \zeta. \qquad (63)$$

Notice that $\mathbf{u}$ represents the polarization-driven deflection of the ray's "center of gravity" predicted in Refs. [5–7]. The second ELE is a Hamilton-Jacobi equation,

$$\delta\mathcal{I} : \quad \partial_t \theta + \frac{k}{n} + (d_t \vartheta + \Sigma) \cos \zeta = 0, \qquad (64)$$

whose gradient yields an equation for $\partial_t \mathbf{k}$, i.e., the momentum equation [31]. The third ELE is

$$\delta\zeta: \quad d_t\vartheta = -\Sigma. \tag{65}$$

As it will become clear below, this describes the rotation of the wave polarization. Finally, the fourth ELE is

$$\delta\vartheta: \quad \partial_t(\mathcal{I}\cos\zeta) + \boldsymbol{\nabla}\cdot(\mathcal{I}\mathbf{v}_0\cos\zeta) = 0. \tag{66}$$

Together, Eqs. (62)-(66) provide a complete "fluid" description of continuous waves. Note that waves are allowed to be nonstationary and "entangled", i.e., contain mixed polarization. In fact, a combination of Eq. (66) with Eq. (62) gives

$$\partial_t(\cos\zeta) + \mathbf{V}\cdot\boldsymbol{\nabla}(\cos\zeta) = \frac{1}{\mathcal{I}}\boldsymbol{\nabla}\cdot(\mathcal{I}\mathbf{u}\cos\zeta), \tag{67}$$

which shows that ζ is generally not conserved; i.e., the numbers of "spin-up" and "spin-down" quanta necessarily oscillate, unless a wave is homogeneous. This can be interpreted as an effective *zitterbewegung* of a classical wave in an inhomogeneous medium.

V. RAY DYNAMICS: CANONICAL REPRESENTATION

A. Basic equations

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

$$\mathcal{I}(t,\mathbf{x}) = \delta(\mathbf{x} - \mathbf{X}(t)), \tag{68}$$

where $\mathbf{X}$ is the location of the wave packet. As in Refs. [18, 31], the Lagrangian density (57) yields a point-particle Lagrangian, $L \doteq \int \mathfrak{L}\, \mathrm{d}^3x$, specifically,

$$L = \mathbf{P}\cdot\dot{\mathbf{X}} - \frac{cP}{n} + \frac{i}{2}\left(Z^\dagger\dot{Z} - \dot{Z}^\dagger Z\right) - \Sigma(\mathbf{X},\mathbf{P})\,Z^\dagger\sigma_z Z. \tag{69}$$

Here $\mathbf{P}(t) \doteq \boldsymbol{\nabla}\theta(t,\mathbf{X}(t))$ is the canonical momentum, $P = |\mathbf{P}|$, and $Z(t) \doteq z(t,\mathbf{X}(t))$ is a two-component complex function. The refraction index and Σ are evaluated at $\mathbf{x} = \mathbf{X}$ and $\mathbf{k} = \mathbf{P}$; e.g., $n = n(\mathbf{X}(t))$. Also, the speed of light constant, c, has been reintroduced for clarity.

Treating $\mathbf{X}(t)$, $\mathbf{P}(t)$, $Z(t)$, and $Z^\dagger(t)$ as independent variables, we obtain the following ELEs:

$$\delta\mathbf{P}: \quad \dot{\mathbf{X}}(t) = \frac{c\mathbf{P}}{nP} + (\partial_{\mathbf{P}}\Sigma)\,Z^\dagger\sigma_z Z, \tag{70}$$

$$\delta\mathbf{X}: \quad \dot{\mathbf{P}}(t) = \frac{cP}{n^2}\boldsymbol{\nabla}n - (\partial_{\mathbf{X}}\Sigma)\,Z^\dagger\sigma_z Z, \tag{71}$$

$$\delta Z^\dagger: \quad \dot{Z}(t) = -i\Sigma\sigma_z Z, \tag{72}$$

$$\delta Z: \quad \dot{Z}^\dagger(t) = i\Sigma Z^\dagger\sigma_z. \tag{73}$$

Together with Eqs. (51) and (60), Eqs. (70)-(73) form a complete set of equations. The first terms on the right hand side of Eqs. (70) and (71) describe the ray dynamics in the GO limit. The second terms describe the coupling of the mode polarization and the ray curvature.

B. Polarization dynamics

To better understand the polarization equations, let us rewrite Eq. (72) as an equation in the basis of linearly polarized modes, i.e., for $\Phi(t) \doteq \phi(t,\mathbf{X}(t))$:

$$\dot{\Phi} = Q\dot{Z} = -i\Sigma Q\sigma_z Z = -i\Sigma(Q\sigma_z Q^{-1})\Phi = -i\Sigma\sigma_y\Phi. \tag{74}$$

[This equation could also be obtained if the ray equations were derived directly from the Lagrangian density (52).] Since Σ is a scalar, and σ_y is constant, this can be readily integrated, yielding [35]

$$\Phi = \exp(-i\Theta\sigma_y)\Phi_0 = (\mathbb{I}_2\cos\Theta - i\sigma_y\sin\Theta)\Phi_0, \tag{75}$$

where $\Theta(t) \doteq \int_0^t \Sigma(t')\,dt'$ can be recognized as the wave Berry phase [10], $\Sigma(t) \doteq \Sigma(\mathbf{X}(t),\mathbf{P}(t))$ and Φ_0 is a vector determined by initial conditions. This result can be also be expressed explicitly as follows:

$$\Phi = \begin{pmatrix} \cos\Theta & -\sin\Theta \\ \sin\Theta & \cos\Theta \end{pmatrix}\Phi_0. \tag{76}$$

It is seen then that the polarization of the EM field rotates at the rate $\Sigma(t)$ in the reference frame defined by the basis vectors $(\mathbf{e}_1,\mathbf{e}_2)$. Such rotation of the polarization plane, also known as *Rytov rotation*, was studied theoretically in Refs. [1, 2, 4] and observed experimentally in Ref. [3]. Clearly, Eq. (65) describes the same effect.

C. Ray dynamics for pure states

If a ray corresponds to a strictly circular polarization, such that $\sigma_z z = \pm z$, the Lagrangian (69) can be simplified down to

$$L_\pm = \mathbf{P}\cdot\dot{\mathbf{X}} - \frac{cP}{n} \mp \Sigma(\mathbf{X},\mathbf{P}), \tag{77}$$

where $L_\pm$ governs the propagation of right and left polarization modes, respectively. This Lagrangian has a canonical form, $L_\pm = \mathbf{P}\cdot\dot{\mathbf{X}} - H_\pm(\mathbf{X},\mathbf{P})$, where $H_\pm$ is a Hamiltonian given by

$$H_\pm(\mathbf{X},\mathbf{P}) = \frac{cP}{n(\mathbf{X})} \pm \frac{cP}{n^2(\mathbf{X})}\mathbf{A}(\mathbf{P})\cdot\boldsymbol{\nabla}n(\mathbf{X}). \tag{78}$$

The variables $\mathbf{X}$ and $\mathbf{P}$ serve as the canonical coordinate and momentum, so they satisfy canonical Hamilton's equations,

$$\delta\mathbf{P}: \quad \dot{\mathbf{X}}(t) = \frac{c\mathbf{P}}{nP} \pm \partial_{\mathbf{P}}\Sigma, \tag{79}$$

$$\delta\mathbf{X}: \quad \dot{\mathbf{P}}(t) = \frac{cP}{n^2}\boldsymbol{\nabla}n \mp \partial_{\mathbf{X}}\Sigma. \tag{80}$$

Since $H_\pm$ is time-independent, one also readily obtains energy (frequency) conservation,

$$H_\pm(\mathbf{X}(t), \mathbf{P}(t)) = \text{const}, \tag{81}$$

along the ray trajectory.

VI. RAY DYNAMICS: NON-CANONICAL REPRESENTATION

A. Ray variables

If the point-particle limit is taken without explicitly invoking Eq. (53), an alternative representation of the ray Lagrangian can be obtained that connects our results to those found in existing literature. For pure states, this alternative Lagrangian is given by

$$L_\pm = \mathbf{p} \cdot \dot{\mathbf{x}} - \frac{cp}{n(\mathbf{x})} \mp \dot{\mathbf{p}} \cdot \mathbf{A}(\mathbf{p}), \tag{82}$$

where $\dot{\mathbf{p}} \cdot \mathbf{A}(\mathbf{p})$ is known as the "Berry connection" term [19]. Notice, in particular, that adding $\partial_\mathbf{k}\chi(\mathbf{k})$ to $\mathbf{A}(\mathbf{k})$, where $\chi(\mathbf{k})$ is an arbitrary scalar function, changes $L_\pm$ merely by a perfect time derivative and thus does not affect the motion equations. Also notice that we have introduced a different notation for the ray variables [$(\mathbf{x}, \mathbf{p})$ instead of $(\mathbf{X}, \mathbf{P})$] for the following reason.

On one hand, as an approximation of the Lagrangian (77), Eq. (82) is expected to yield dynamics similar to that yielded by Eq. (77). On the other hand, notice that Eq. (82) does not have a canonical form. This means that $(\mathbf{x}, \mathbf{p})$ do not satisfy Hamilton's equations and thus, clearly, cannot be the same as $(\mathbf{X}, \mathbf{P})$. To understand the connection between the two sets of variables, let us rewrite Eq. (82) as

$$L_\pm = \mathbf{p} \cdot \dot{\mathbf{x}} - \frac{cp}{n} \mp \frac{d}{dt}(\mathbf{p} \cdot \mathbf{A}) \pm \mathbf{p} \cdot \frac{d\mathbf{A}}{dt}. \tag{83}$$

Dropping the perfect time derivative and introducing

$$\mathbf{q} \doteq \mathbf{x} \pm \mathbf{A}, \tag{84}$$

we obtain a Lagrangian in a canonical form:

$$L_\pm = \mathbf{p} \cdot \dot{\mathbf{q}} - \frac{cp}{n(\mathbf{q} \mp \mathbf{A}(\mathbf{p}))}. \tag{85}$$

The quantity $cp/n(\mathbf{x})$ serves as the canonical energy and is conserved,

$$cp/n(\mathbf{x}) = \text{const}. \tag{86}$$

Also, since $1/n(\mathbf{x})$ is assumed smooth, one can replace this function with its first-order Taylor expansion. Then,

$$L_\pm = \mathbf{p} \cdot \dot{\mathbf{q}} - \frac{cp}{n(\mathbf{q})} \mp \frac{cp}{n^2(\mathbf{q})}\mathbf{A}(\mathbf{p}) \cdot \boldsymbol{\nabla} n(\mathbf{q}), \tag{87}$$

where we omitted terms $\mathcal{O}(\epsilon^2)$, as usual. By comparing this with Eq. (77), we find that

$$\mathbf{x} = \mathbf{X} \mp \mathbf{A}(\mathbf{P}), \qquad \mathbf{p} = \mathbf{P}. \tag{88}$$

Notice that $|\mathbf{x} - \mathbf{X}|$ is of the order of the wavelength, i.e., small enough to make $\mathbf{x}$ and $\mathbf{X}$ equally physical as measures of the ray location.

B. Ray equations

The equations for $\mathbf{x}$ and $\mathbf{p}$ can be obtained by combining Eqs. (79), (80), and (88), or they can be derived directly as ELEs corresponding to the Lagrangian (82):

$$\delta\mathbf{p}: \quad \dot{\mathbf{x}}(t) = \frac{c\mathbf{p}}{np} \pm \dot{\mathbf{p}} \times (\boldsymbol{\nabla}_\mathbf{p} \times \mathbf{A}), \tag{89}$$

$$\delta\mathbf{x}: \quad \dot{\mathbf{p}}(t) = \frac{cp}{n^2}\boldsymbol{\nabla} n. \tag{90}$$

Using Eq. (51), we can also rewrite them as

$$\dot{\mathbf{x}}(t) = \frac{c\mathbf{p}}{np} \pm \frac{\dot{\mathbf{p}} \times \mathbf{p}}{p^3}, \tag{91}$$

$$\dot{\mathbf{p}}(t) = \frac{cp}{n^2}\boldsymbol{\nabla} n. \tag{92}$$

Equations (91) and (92) were reported previously in Refs. [8, 16]. The equations presented in Refs. [14, 15, 19–21] can also be obtained from Eqs. (91) and (92) for the rescaled momentum $\boldsymbol{\rho} \doteq n\mathbf{p}/(cp)$ and rescaled "time" $ds \doteq c\,dt/n(\mathbf{x}(t))$. Specifically, one gets

$$\mathbf{x}' = \frac{\boldsymbol{\rho}}{\rho} \pm \lambda\frac{\boldsymbol{\rho}' \times \boldsymbol{\rho}}{\rho^3}, \quad \boldsymbol{\rho}' = \boldsymbol{\nabla} n, \tag{93}$$

where primes denote derivatives with respect to s, and we used that $\lambda \doteq n(\mathbf{x})/p$ is a constant of motion, as seen from Eq. (86).

C. Comparison of the two models

As we showed explicitly in Sec. VI A (and also by construction), the canonical Lagrangian (77) and the non-canonical Lagrangian (82) differ only by $\mathcal{O}(\epsilon^2)$ and thus are equivalent within their accuracy domain. (The same applies to the earlier theories [8, 14–16, 19–21] too, where the ray Lagrangians are also derived to the first order in ϵ.) This means, in particular, that the effect of the Berry connection term in Eq. (82) can be attributed simply to the choice of coordinates, while the underlying physics can also be described by the canonical Lagrangian (78).

The advantage of the Lagrangian (77) is its manifestly symplectic structure, which is convenient, for instance, for numerical simulations [26]. On the other hand, the Lagrangian (82) leads to "gauge-invariant" equations, in the sense that they are manifestly independent of the coordinate representation for the basis vectors, $\mathbf{e}_1$ and
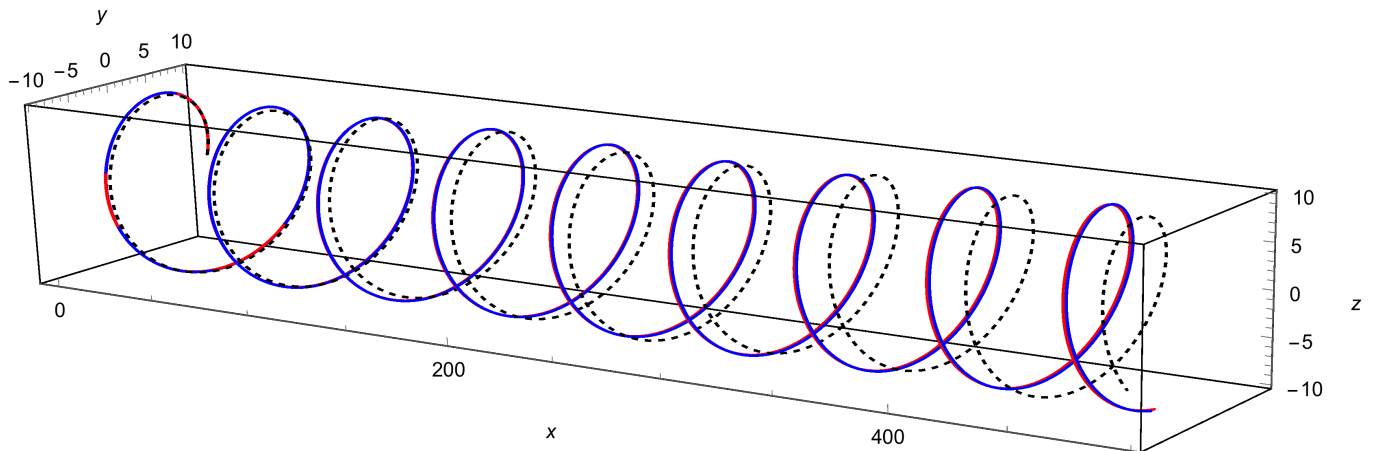
FIG. 1: Comparison between ray trajectories predicted by the canonical Lagrangian (77) and the non-canonical Lagrangian (82). The red line represents $\mathbf{X}(t) - \mathbf{A}(\mathbf{P}(t))$, where $\mathbf{X}(t)$ and $\mathbf{P}(t)$ are found by numerical integration of Eqs. (79) and (80). The blue line shows $\mathbf{x}(t)$, as found by numerical integration of Eqs. (91) and (92). The black dashed line plots the "spinless" ray trajectory governed by the lowest-order GO ray Lagrangian L_0 given by Eq. (94). The refraction index is chosen to be $n(\mathbf{x}) = 1 + \exp[-(y^2 + z^2)/\ell^2]$, the characteristic width is $\ell = 10$, and $\mathbf{A}$ is chosen in the form (51). The length unit is $a \doteq \sqrt{2}/p_0$. The initial location is $\mathbf{x}_0 = (0, 10, 0)$, and the initial momentum is $\mathbf{p}_0 = (1, 0, 1)/a$, so $\epsilon \sim 1/(p_0\ell) \sim 0.07$. It is seen that the canonical Lagrangian and the noncanonical Lagrangian predict results that, for the specified parameters, are essentially indistinguishable from each other yet differ noticeably from those yielded by L_0.

$\mathbf{e}_2$. Also, assuming that $\mathbf{A}$ is chosen in the form (51), using the non-canonical Lagrangian (82) can be advantageous when $p_x^2 + p_y^2 \to 0$, because the right-hand side in Eqs. (91) and (92) remains finite in this case unless $p \to 0$. Therefore, whether the canonical or noncanonical form is more convenient in a given case depends on the specific application.

To illustrate how accurate the agreement is between the two models, we also performed comparative numerical simulations. Figure 1 shows the ray trajectories for a right polarized wave using the canonical (77) and the non-canonical (82) representations. For completeness, we also show the calculated ray trajectory as determined by the lowest-order GO ray Lagrangian,

$$L_0 = \mathbf{p} \cdot \dot{\mathbf{x}} - cp/n(\mathbf{x}), \tag{94}$$

which does not account for polarization effects. As anticipated, the ray trajectories predicted by the Lagrangians (77) and (82) are almost identical and yet differ noticeably from the "spinless" ray trajectory predicted by Eq. (94), namely,

$$\delta\mathbf{p}: \quad \dot{\mathbf{x}}(t) = \frac{c\mathbf{p}}{np}, \tag{95}$$

$$\delta\mathbf{x}: \quad \dot{\mathbf{p}}(t) = \frac{cp}{n^2}\boldsymbol{\nabla}n. \tag{96}$$

VII. CONCLUSIONS

In this paper, we study the propagation of electromagnetic waves in isotropic dielectric media with local dis-

persion under the assumption of small but nonvanishing $\epsilon \doteq \lambda/\ell$, where λ is the wavelength, and ℓ is the characteristic inhomogeneity scale. It is commonly known that, due to nonzero ϵ, such waves can experience polarization-driven bending of ray trajectories and polarization dynamics that can be interpreted as the precession of the wave "spin". Here, we report how Lagrangians describing these effects can be deduced, rather than guessed, within a strictly classical theory. In addition to the commonly known ray Lagrangian that features the Berry connection, a simple alternative Lagrangian is also proposed that naturally has a canonical form. We explain how the two Lagrangians are related and demonstrate their equivalence numerically. The presented theory captures not only the eigenray dynamics but also the dynamics of continuous wave fields and rays with mixed polarization, or "entangled" waves. Our calculation assumes stationary media with isotropic local dispersion, but generalizations to other media are straightforward to do, as will be reported separately.

[1] S. M. Rytov, Dokl. Akad. Nauk SSSR **18**, 263 (1938).
[2] V. V. Vladimirskii, Dokl. Akad. Nauk SSSR **31**, 222 (1941).
[3] A. Tomita and R. Y. Chiao, Phys. Rev. Lett. **57**, 937 (1986).
[4] M. V. Berry, Nature **326**, 277 (1987).
[5] A. V. Dugin, B. Y. Zel'dovich, N. D. Kundikova, and V. S. Liberman, JETP Lett. **53**, 197 (1991).
[6] A. V. Dooghin, N. D. Kundikova, V. S. Liberman, and B. Y. Zel'dovich, Phys. Rev. A **45**, 8204 (1992).
[7] V. S. Liberman and B. Y. Zel'dovich, Phys. Rev. A **46**, 5199 (1992).
[8] M. Onoda, S. Murakami, and N. Nagaosa, Phys. Rev. Lett. **93**, 083901 (2004).
[9] K. Y. Bliokh, F. J. Rodríguez-Fortuño, F. Nori, and A. V. Zayats, arXiv (2015), 1505.02864.
[10] M. V. Berry, Proc. Roy. Soc. Lond. A **392**, 45 (1984).
[11] F. Wilczek and A. Shapere, *Geometric Phases in Physics* (World Scientific, Singapore, 1989).
[12] A. Bohm, A. Mostafazadeh, H. Koizumi, Q. Niu, and J. Zwanziger, *The Geometric Phase in Quantum Systems* (Springer, Berlin, 2003).
[13] S. Murakami, N. Nagaosa, and S.-C. Zhang, Science **301**, 1348 (2003).
[14] K. Y. Bliokh and Y. P. Bliokh, Phys. Rev. E **70**, 026605 (2004).
[15] K. Y. Bliokh and Y. P. Bliokh, Phys. Lett. A **333**, 181 (2004).
[16] M. Onoda, S. Murakami, and N. Nagaosa, Phys. Rev. E **74**, 066610 (2006).
[17] A related calculation was also proposed in H. L. Berk and D. Pfirsch, J. Math. Phys. **21**, 2054 (1980).
[18] D. E. Ruiz and I. Y. Dodin, Phys. Lett. A **379**, 2337 (2015).
[19] K. Y. Bliokh, A. Niv, V. Kleiner, and E. Hasman, Nature Photonics **2**, 748 (2008).
[20] K. Y. Bliokh and V. D. Freilikher, Phys. Rev. B **72**, 035108 (2005).
[21] K. Y. Bliokh, D. Y. Frolov, and Y. A. Kravtsov, Phys. Rev. A **75**, 053821 (2007).
[22] I. Bialynicki-Birula, in *Progress in Optics, Vol. XXXVI* (Elsevier, Amsterdam, 1996), p. 245.
[23] I. Y. Dodin, Phys. Lett. A **378**, 1598 (2014).
[24] It is to be noted that these $\boldsymbol{\alpha}$ matrices are related to the Gell-Mann matrices, which serve as infinitesimal generators of the special unitary group $SU(3)$.
[25] J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1999), 3rd ed.
[26] T. J. Bridges and S. Reich, Physics Letters A **284**, 184 (2001).
[27] T. J. Bridges and S. Reich, J. Phys. A: Math. Gen. **39**, 5287 (2006).
[28] T. J. Bridges, Proc. Roy. Soc. Lond. A **462**, 1531 (2006).
[29] 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).
[30] I. Y. Dodin and N. J. Fisch, Phys. Rev. A **86**, 053834 (2012).
[31] D. E. Ruiz and I. Y. Dodin, Phys. Lett. A **379**, 2623 (2015).
[32] Note that $\mathbf{e}_1$ and $\mathbf{e}_2$ correspond to the spherical unit vectors $\mathbf{e}_\theta$ and $\mathbf{e}_\varphi$ in $\mathbf{k}$-space, respectively.
[33] Y. A. Kravtsov, B. Bieg, and K. Y. Bliokh, J. Opt. Soc. Am. A **24**, 3388 (2007).
[34] Y. A. Kravtsov and B. Bieg, Opt. Appl. **40**, 975 (2010).
[35] Here, we used the well known Euler formula for Pauli matrices, $e^{ia(\mathbf{n}\cdot\boldsymbol{\sigma})} = \mathbb{I}_2 \cos a + i(\mathbf{n}\cdot\boldsymbol{\sigma})\sin a$.