

Gravitational spin Hall effect of light

Marius A. Oancea,^{1,*} Jérémie Joudoux,^{1,†} I. Y. Dodin,^{2,‡} D. E. Ruiz,^{3,§} Claudio F. Paganini,^{4,1,¶} and Lars Andersson^{1,**}

¹*Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Am Mühlenberg 1, D-14476 Potsdam, Germany*

²*Department of Astrophysical Sciences, Princeton University, Princeton, New Jersey 08544, USA*

³*Sandia National Laboratories, P.O. Box 5800, Albuquerque, New Mexico 87185, USA*

⁴*Fakultät für Mathematik, Universität Regensburg, D-93040 Regensburg, Germany*

The propagation of electromagnetic waves in vacuum is often described within the geometrical optics approximation, which predicts that wave rays follow null geodesics. However, this model is valid only in the limit of infinitely high frequencies. At large but finite frequencies, diffraction can still be negligible, but the ray dynamics becomes affected by the evolution of the wave polarization. Hence, rays can deviate from null geodesics, which is known as the gravitational spin Hall effect of light. In the literature, this effect has been calculated *ad hoc* for a number of special cases, but no general description has been proposed. Here, we present a covariant Wentzel-Kramers-Brillouin analysis from first principles for the propagation of light in arbitrary curved spacetimes. We obtain polarization-dependent ray equations describing the gravitational spin Hall effect of light. We also present numerical examples of polarization-dependent ray dynamics in the Schwarzschild spacetime, and the magnitude of the effect is briefly discussed. The analysis reported here is analogous to that of the spin Hall effect of light in inhomogeneous media, which has been experimentally verified.

I. INTRODUCTION

The propagation of electromagnetic waves in curved spacetime is often described within the geometrical optics approximation, which applies in the limit of infinitely high frequencies [1, 2]. In geometrical optics, Maxwell's equations are reduced to a set of ray equations and a set of transport equations along these rays. The ray equations are the null geodesics of the underlying spacetime, and the transport equations govern the evolution of the intensity and the polarization vector. In particular, the geometrical optics approximation predicts that the ray equations determine the evolution of the polarization vector and there is no backreaction from the polarization vector onto the ray equations. However, this model is valid only in the limit of infinitely high frequencies, and there has been interest in calculating the light propagation more accurately. At large but finite frequencies, diffraction can still be negligible but rays can deviate from geodesics. This is known as the gravitational spin Hall effect of light [3].

The mechanism behind the spin Hall effect is the spin-orbit interaction [4], i.e., the coupling of the wave polarization (spin) with the translational (orbital) motion of the ray as a particle, resulting in polarization(spin)-dependent rays. Related phenomena are found in many areas of physics. In condensed matter physics, electrons traveling in certain materials experience a spin Hall ef-

fect, resulting in spin-dependent trajectories, and spin accumulation on the lateral sides of the material [5, 6]. The effect was theoretically predicted by Dyakonov and Perel in 1971 [7, 8], followed by experimental observation in 1984 [9] and 2004 [10]. In optics, the polarization-dependent deflection of light traveling in an inhomogeneous medium is known as the spin Hall effect of light [4, 11]. The effect was predicted by several authors [12–18] and has recently been verified experimentally by Hosten and Kwiat [19] and also by Bliokh *et al.* [20]. The spin Hall effect of light provides corrections to the geometrical optics limit, which scale roughly with the inverse of frequency. This, and several other effects, can be explained in terms of the Berry curvature [4, 20–22].

There are several approaches aiming to describe the dynamics of spinning particles or wave packets in general relativity. Using a multipole expansion of the energy-momentum tensor, the dynamics of massive spinning test particles has been extensively studied in the form of the Mathisson-Papapetrou-Dixon equations [23–27]. A massless limit of these equations was derived by Souriau and Saturnini [28, 29], and particular examples adapted to certain spacetimes have been discussed in Refs. [30–32]. Another commonly used method is the Wentzel-Kramers-Brillouin (WKB) approximation for various field equations on curved spacetimes. For massive fields, this has been done in Refs. [33, 34] by considering a WKB approximation for the Dirac equation. For massless fields, using a WKB approximation for Maxwell's equations on a stationary spacetime, Frolov and Shoom derived polarization-dependent ray equations [35, 36] (see also Refs. [37–40]). With methods less familiar in general relativity, using the Foldy-Wouthuysen transformation for the Bargmann-Wigner equations in a perturbative way, Gosselin *et al.* derived ray equations for photons [41] and electrons [42] traveling in static spacetimes (see also Refs. [43–45]). The gravitational spin Hall effect of

*Electronic address: marius.oancea@aei.mpg.de

†Electronic address: jeremie.joudoux@aei.mpg.de

‡Electronic address: idodin@princeton.edu

§Electronic address: deruiz@sandia.gov

¶Electronic address: claudio.paganini@aei.mpg.de

**Electronic address: lars.andersson@aei.mpg.de

gravitational waves was also considered in Refs. [37, 46]. However, as discussed in Ref. [3], there are inconsistencies between the predictions of these different models, and some of these models only work in particular space-times.

In this work, we are concerned with describing the propagation of electromagnetic waves in curved space-time beyond the traditional geometrical optics approximation. We carry out a covariant WKB analysis of the vacuum Maxwell's equations, closely following the derivation of the spin Hall effect in optics [20, 47, 48], as well as the work of Littlejohn and Flynn [49]. As a result, we derive ray equations that contain polarization-dependent corrections to those of traditional geometrical optics and capture the gravitational spin Hall effect of light. As in optics, these corrections can be interpreted in terms of the Berry curvature. To illustrate the effect, we give some numerical examples of the effective ray trajectories in the Schwarzschild spacetime.

Our paper is organized as follows. In Sec. II, we start by introducing the variational formulation of the vacuum Maxwell's equations. Then, we present the specific form of the WKB ansatz to be used, discuss the role of the Lorenz gauge condition, and state the assumptions that we are considering on the initial conditions. In Sec. III, we present the WKB approximation of the field action and the corresponding Euler-Lagrange equations. After analyzing these equations at each order in the geometrical optics parameter ϵ , we obtain the well-known results of geometrical optics. The dynamics of the polarization vector is expressed in terms of the Berry phase. Finally, we derive an effective Hamilton-Jacobi system that contains $\mathcal{O}(\epsilon)$ corrections over the standard geometrical optics results. In Sec. IV, we use the corrected Hamilton-Jacobi equation to derive the ray equations that account for the gravitational spin Hall effect of light. The gauge invariance of these equations is discussed, and noncanonical coordinates are introduced. In Sec. V, we present some basic examples. For Minkowski spacetime, we analytically show how the effective ray equations reproduce the relativistic Hall effect [50] and the Wigner translations of polarized electromagnetic wave packets [51]. Using numerical computations, we consider the effective ray equations on a Schwarzschild background and compare with the results of Gosselin *et al.* [41]. The magnitude of the effect is also estimated numerically. A summary of the main results, including the effective Hamiltonian and the effective ray equations, can be found in Sec. VI.

Notations and conventions

We consider an arbitrary smooth Lorentzian manifold $(M, g_{\mu\nu})$, where the metric tensor $g_{\mu\nu}$ has signature $-+++$. The absolute value of the metric determinant is denoted as $g = |\det g_{\mu\nu}|$. The phase space is defined as the cotangent bundle T^*M , and phase space points are denoted as (x, p) . The Einstein summation convention

is assumed. Greek indices represent spacetime indices and run from 0 to 3. Latin indices from the beginning of the alphabet, (a, b, c, \dots) , represent tetrad indices and run from 0 to 3. Latin indices from the middle of the alphabet, (i, j, k, \dots) , label the components of 3-vectors and run from 1 to 3. For the curvature, we use the conventions of Hawking and Ellis [52]. Finally, we use the \mathcal{O} notation as follows: a scalar function f depending on a parameter ϵ satisfies $f(\epsilon) = \mathcal{O}(\epsilon^\alpha)$ if there is a constant M such that $|f(\epsilon)| \leq M\epsilon^\alpha$ for small ϵ .

II. MAXWELL'S EQUATIONS AND THE WKB APPROXIMATION

A. Lagrangian formulation of Maxwell's equations

Electromagnetic waves in vacuum can be described by the electromagnetic tensor $\mathcal{F}_{\alpha\beta}$. This is a skew-symmetric real 2-form, which satisfies the vacuum Maxwell's equations [1, Sec. 22.4]

$$\nabla^\alpha \mathcal{F}_{\alpha\beta} = 0, \quad \nabla_{[\alpha} \mathcal{F}_{\beta\gamma]} = 0. \quad (2.1)$$

Solutions to Maxwell's equations can also be represented by introducing the electromagnetic four-potential \mathcal{A}_α , which is a real 1-form. Then, the electromagnetic tensor can be expressed as

$$\mathcal{F}_{\alpha\beta} = 2\nabla_{[\alpha} \mathcal{A}_{\beta]}, \quad (2.2)$$

and Eq. (2.1) becomes [1, Sec. 22.4]

$$\hat{D}_\alpha^\beta \mathcal{A}_\beta = 0, \quad \hat{D}_\alpha^\beta = \nabla^\beta \nabla_\alpha - \delta_\alpha^\beta \nabla^\mu \nabla_\mu. \quad (2.3)$$

This equation can be obtained as the Euler-Lagrange equation of the following action:

$$\begin{aligned} J &= \frac{1}{4} \int_M d^4x \sqrt{g} \mathcal{F}_{\alpha\beta} \mathcal{F}^{\alpha\beta} \\ &= \frac{1}{2} \int_M d^4x \sqrt{g} \mathcal{A}^\alpha \hat{D}_\alpha^\beta \mathcal{A}_\beta, \end{aligned} \quad (2.4)$$

where the last equality is obtained using integration by parts.

B. WKB ansatz

We assume that the vector potential admits a WKB expansion of the form

$$\begin{aligned} \mathcal{A}_\alpha(x) &= \text{Re} \left[A_\alpha(x, k(x), \epsilon) e^{iS(x)/\epsilon} \right], \\ A_\alpha(x, k(x), \epsilon) &= A_{0\alpha}(x, k(x)) + \epsilon A_{1\alpha}(x, k(x)) + \mathcal{O}(\epsilon^2), \end{aligned} \quad (2.5)$$

where S is a real scalar function, A_α is a complex amplitude, and ϵ is a small expansion parameter. The gradient of S is denoted as

$$k_\mu(x) = \nabla_\mu S(x). \quad (2.6)$$

Note that the ansatz (2.5) differs from the classical WKB ansatz since the complex amplitude A_α depends on the phase gradient $k_\alpha(x)$. In other words, we assume that A_α is defined on the Lagrangian submanifold $x \mapsto (x, k(x)) \in T^*M$. Such a dependency can be found in standard textbooks, for example Ref. [53, Sec. 3.3]. Up to an application of the chain rule, our Eq. (2.5) is equivalent to the standard WKB ansatz. In particular, the dependency of A in k appears naturally in the geometrical optics equation (3.16), and we observe in Sec. III D that the polarization vector and the polarization basis naturally depend on k , which is why the k dependence was introduced in Eq. (2.5).

The limit $\epsilon \ll 1$ indicates that the phase of the vector potential rapidly oscillates and its variations are much faster than those corresponding to the amplitude $A_\alpha(x, k, \epsilon)$. The role of the expansion parameter ϵ becomes clear if we consider a timelike observer traveling along the worldline $\lambda \mapsto y^\alpha(\lambda)$ with proper time λ . This observer measures the frequency

$$\omega = -\frac{t^\alpha k_\alpha}{\epsilon}, \quad (2.7)$$

where $t^\alpha = dy^\alpha/d\lambda$ is the velocity vector field of the observer. The phase function S and ϵ are dimensionless quantities. In geometrized units, such that $c = G = 1$ [54, Appendix F], the velocity t^α is dimensionless, and k_α has dimension of inverse length. Hence, ω has the dimension of the inverse length, as expected for frequency. Then, the observer sees the frequency going to infinity as ϵ goes to 0.

C. Lorenz gauge

In this section, we introduce the Lorenz gauge in the context of WKB approximations. We shall impose this gauge condition [cf. (2.21) below] in the rest of the paper, with the exception of Sec. III B, where it is relevant to discuss some aspects of geometrical optics without this condition.

Maxwell's equations in the form (2.3) do not have a well-posed Cauchy problem. In particular, they admit pure gauge solutions. This problem is usually eliminated by introducing a gauge condition. Here we shall focus on the Lorenz gauge condition

$$\nabla_\alpha A^\alpha = 0. \quad (2.8)$$

We reproduce here, in the context of a WKB analysis, the classical argument regarding the gauge fixing for Maxwell's equations (see for instance [55, Lemma 10.2]). Using the identity

$$\hat{D}_\alpha^\beta A_\beta - \nabla_\alpha \nabla^\beta A_\beta = -\nabla^\beta \nabla_\beta A_\alpha + R_{\alpha\beta} A^\beta, \quad (2.9)$$

one observes that, if Maxwell's equations (2.3) and the Lorenz gauge (2.8) are satisfied, then the wave equation

$$-\nabla^\beta \nabla_\beta A_\alpha + R_{\alpha\beta} A^\beta = 0 \quad (2.10)$$

holds. Conversely, by solving Eq. (2.10), with Cauchy data satisfying constraint and gauge conditions, one obtains a solution to Maxwell's equations in the Lorenz gauge.

Note that we consider here approximate solutions to Maxwell's equations

$$\hat{D}_\alpha^\beta A_\beta = \mathcal{O}(\epsilon^0). \quad (2.11)$$

Hence, it is sufficient to consider that the Lorenz gauge is satisfied at the appropriate order:

$$\nabla_\alpha A^\alpha = \mathcal{O}(\epsilon^1). \quad (2.12)$$

We reproduce the standard argument recovering Maxwell's equation in the Lorenz gauge from the wave Eq. (2.10), taking into account that we are considering only approximate solutions. Assume that the wave equation holds:

$$-\nabla^\beta \nabla_\beta A_\alpha + R_{\alpha\beta} A^\beta = \mathcal{O}(\epsilon^0). \quad (2.13)$$

Upon inserting the WKB ansatz, this is equivalent to

$$\begin{aligned} k^\beta k_\beta A_{0\alpha} &= 0, \\ ik^\beta k_\beta A_{1\alpha} + A_{0\alpha} \nabla^\beta k_\beta + 2k^\beta \nabla_\beta A_{0\alpha} &= 0. \end{aligned} \quad (2.14)$$

Furthermore, assume that the initial data for the wave equation (2.13) satisfy

$$\begin{aligned} k_\alpha A_0^\alpha &= 0, \\ \nabla_\alpha A_0^\alpha + ik_\alpha A_1^\alpha &= 0. \end{aligned} \quad (2.15)$$

Equation (2.13) implies that

$$\nabla^\beta \nabla_\beta (\nabla_\alpha A^\alpha) = \mathcal{O}(\epsilon^{-1}). \quad (2.16)$$

The initial data (2.15) for Eq. (2.13) imply that, initially,

$$\nabla_\alpha A^\alpha = \mathcal{O}(\epsilon^1). \quad (2.17)$$

Observe that the condition

$$T^\beta \nabla_\beta (\nabla_\alpha A^\alpha) = \mathcal{O}(\epsilon^0) \quad (2.18)$$

is automatically satisfied, where T^β is a unit future-oriented normal vector to the hypersurface on which initial data are prescribed. Hence, the equation satisfied by the Lorenz gauge source function (2.16) admits initial data as in Eqs. (2.17) and (2.18) vanishing at the appropriate order in ϵ [at $\mathcal{O}(\epsilon^1)$ and $\mathcal{O}(\epsilon^0)$, respectively]. This implies that Maxwell's equations

$$\hat{D}_\alpha^\beta A_\beta = \mathcal{O}(\epsilon^0), \quad (2.19)$$

which can be expanded as

$$\begin{aligned} k^\beta A_{0[\beta} k_{\alpha]} &= 0, \\ 2k^\beta \nabla_\beta A_{0\alpha} - \left(\nabla_\beta A_0^\beta + ik_\beta A_1^\beta \right) k_\alpha - k^\beta \nabla_\alpha A_{0\beta} \\ &\quad - A_0^\beta \nabla_\beta k_\alpha + A_{0\alpha} \nabla_\beta k^\beta + ik^\beta k_\beta A_{1\alpha} = 0, \end{aligned} \quad (2.20)$$

are satisfied in the Lorenz gauge

$$\nabla_\alpha A^\alpha = \mathcal{O}(\epsilon^1) \Leftrightarrow \begin{cases} k_\alpha A_0^\alpha = 0 \\ \nabla_\alpha A_0^\alpha + ik_\alpha A_1^\alpha = 0 \end{cases}. \quad (2.21)$$

D. Assumptions on the initial conditions

We end Sec. II by summarizing the initial conditions that we shall use in the WKB ansatz for Maxwell's equations.

1. The Lorenz gauge (2.21) is satisfied initially. This condition is used to obtain a well-defined solution to the equations of motion, as discussed in Sec. II C.
2. The initial phase gradient k_α is a future-oriented null covector. As will be seen, the condition that k_α is null is a compatibility condition that follows from the Euler-Lagrange equations and the Lorenz gauge condition (2.21) at the lowest order in ϵ ; cf. dispersion relation (3.8) below.
3. Initially, the beam has circular polarization; cf. Eq. (3.47). In Sec. III D we show that the initial state of circular polarization is conserved. In Sec. IV B this assumption ensures a consistent transition transition between the effective dispersion relation and the effective ray equations. Heuristically speaking, due to the spin Hall effect, a localized wave packet that initially has linear polarization can split into two localized wave packets of opposite circular polarization. While this does not represent a problem at the level of Maxwell's equations (which are partial differential equations), the same behavior cannot be captured by the effective ray equations (which are ordinary differential equations) obtained in Sec. IV B.

III. HIGHER-ORDER GEOMETRICAL OPTICS

A. WKB approximation of the field action

We compute the WKB approximation for our field theory by inserting the WKB ansatz (2.5) in the field action (2.4):

$$\begin{aligned} J &= \int_M d^4x \sqrt{g} \operatorname{Re} \left(A^\alpha e^{iS/\epsilon} \right) \hat{D}_\alpha^\beta \operatorname{Re} \left(A_\beta e^{iS/\epsilon} \right) \\ &= \frac{1}{4} \int_M d^4x \sqrt{g} \left[A^{*\alpha} e^{-iS/\epsilon} \hat{D}_\alpha^\beta \left(A_\beta e^{iS/\epsilon} \right) + \text{c.c.} \right] \\ &\quad + \frac{1}{4} \int_M d^4x \sqrt{g} \left[A^\alpha e^{iS/\epsilon} \hat{D}_\alpha^\beta \left(A_\beta e^{iS/\epsilon} \right) + \text{c.c.} \right]. \end{aligned} \quad (3.1)$$

$$D_\alpha^\beta A_\beta - i\epsilon \left(\nabla^\mu D_\alpha^\beta \right) \nabla_\mu A_\beta - \frac{i\epsilon}{2} \left(\nabla_\mu \nabla^\mu D_\alpha^\beta \right) A_\beta = \mathcal{O}(\epsilon^2), \quad (3.5)$$

$$D_\alpha^\beta A^{*\alpha} + i\epsilon \left(\nabla^\mu D_\alpha^\beta \right) \nabla_\mu A^{*\alpha} + \frac{i\epsilon}{2} \left(\nabla_\mu \nabla^\mu D_\alpha^\beta \right) A^{*\alpha} = \mathcal{O}(\epsilon^2), \quad (3.6)$$

$$\nabla_\mu \left[\left(\nabla^\mu D_\alpha^\beta \right) A^{*\alpha} A_\beta - \frac{i\epsilon}{2} \left(\nabla^\mu \nabla^\nu D_\alpha^\beta \right) (A^{*\alpha} \nabla_\nu A_\beta - A_\beta \nabla_\nu A^{*\alpha}) \right] = \mathcal{O}(\epsilon^2). \quad (3.7)$$

If S has a nonvanishing gradient, then $e^{iS/\epsilon}$ is rapidly oscillating. In this case, for f sufficiently regular, the method of stationary phase [56, Sec. 2.3] implies

$$\int_M d^4x \sqrt{g} e^{\pm i2S(x)/\epsilon} f(x) = \mathcal{O}(\epsilon^2). \quad (3.2)$$

Upon expanding the derivative terms in Eq. (3.1), and keeping only terms of the lowest two orders in ϵ , we obtain the following WKB approximation of the field action [for convenience, we are shifting the powers of ϵ , such that the lowest-order term is of $\mathcal{O}(\epsilon^0)$]:

$$\begin{aligned} -\epsilon^2 J &= \int_M d^4x \sqrt{g} \left[D_\alpha^\beta A^{*\alpha} A_\beta \right. \\ &\quad \left. - \frac{i\epsilon}{2} \nabla^\mu D_\alpha^\beta (A^{*\alpha} \nabla_\mu A_\beta - A_\beta \nabla_\mu A^{*\alpha}) \right] + \mathcal{O}(\epsilon^2), \end{aligned} \quad (3.3)$$

where

$$\begin{aligned} D_\alpha^\beta &= \frac{1}{2} k_\mu k^\mu \delta_\alpha^\beta - \frac{1}{2} k_\alpha k^\beta, \\ \nabla^\mu D_\alpha^\beta &= k^\mu \delta_\alpha^\beta - \frac{1}{2} \delta_\alpha^\mu k^\beta - \frac{1}{2} g^{\mu\beta} k_\alpha. \end{aligned} \quad (3.4)$$

Here, D_α^β represents the symbol [57] of the operator \hat{D}_α^β , evaluated at the phase space point $(x, p) = (x, k)$, and we are using the notation $\nabla^\mu D_\alpha^\beta$ for the vertical derivative (Appendix A) of D_α^β , evaluated at the phase space point $(x, p) = (x, k)$.

The action depends on the following fields: $S(x)$, $\nabla_\mu S(x)$, $A_\alpha(x, \nabla S)$, $\nabla_\mu [A_\alpha(x, \nabla S)]$, $A^{*\alpha}(x, \nabla S)$, $\nabla_\mu [A^{*\alpha}(x, \nabla S)]$. Following the calculations in Appendix B, the Euler-Lagrange equations are

In the above equations, the symbol D_α^β and its vertical derivatives are all evaluated at the phase space point (x, k) . Note that the same set of equations can be obtained in a more traditional way, by inserting the WKB ansatz (2.5) directly into the field equation (2.3), or by following the approach presented in Ref. [58]. More generally, a detailed discussion about the variational formulation of the WKB approximation can be found in Ref. [59].

B. Zeroth-order geometrical optics

Starting with Eqs. (3.5)-(3.7), and keeping only terms of $\mathcal{O}(\epsilon^0)$, we obtain

$$D_\alpha^\beta A_{0\beta} = 0, \quad (3.8)$$

$$D_\alpha^\beta A_0^{*\alpha} = 0, \quad (3.9)$$

$$\nabla_\mu \left[\left(\nabla^\mu D_\alpha^\beta \right) A_0^{*\alpha} A_{0\beta} \right] = 0. \quad (3.10)$$

Equation (3.8) can also be written as

$$D_\alpha^\beta A_{0\beta} = \frac{1}{2} (k_\mu k^\mu \delta_\alpha^\beta - k_\alpha k^\beta) A_{0\beta} = 0. \quad (3.11)$$

The matrix D_α^β admits two eigenvalues when k_α is not a null covector. The first eigenvalue is $\frac{1}{2}k_\mu k^\mu$ with eigenspace consisting of covectors perpendicular to k_α . The second eigenvalue is 0 with eigenvector k_α . When k_α is null, the matrix D_α^β is nilpotent. It admits a unique eigenvalue 0 whose eigenspace is the orthogonal to k_α , which contains the covector k_α .

The Lorenz gauge condition (2.21) implies that $A_{0\alpha}$ is orthogonal to k_α . Hence, a necessary condition for Eq. (3.8) to admit a nontrivial solution is that k_α is a null covector. It is also possible to deduce that k_α is a null covector without using the gauge condition. For completeness, we present this argument below.

Equation (3.11) admits nontrivial solutions if and only if $A_{0\beta}$ is an eigenvector of D_α^β with zero eigenvalue. Two cases should be discussed: k_α is a null covector, or k_α is not a null covector.

Assume first that k_α is not a null covector, $k^\mu k_\mu \neq 0$. Then, Eq. (3.11) leads to

$$A_{0\alpha} = \frac{k^\beta A_{0\beta}}{k_\mu k^\mu} k_\alpha. \quad (3.12)$$

This entails that

$$A_{0[\alpha} k_{\beta]} = 0 \quad \text{or} \quad \mathcal{F}_{\alpha\beta} = \nabla_{[\alpha} \mathcal{A}_{\beta]} = \mathcal{O}(\epsilon^0). \quad (3.13)$$

In other words, when k_α is not a null covector, the corresponding solution is, at the lowest order in ϵ , a pure

gauge solution. Since the corresponding electromagnetic field vanishes, we do not consider this case further.

If k_α is null, $k^\mu k_\mu = 0$, Eq. (3.11) implies

$$k^\beta A_{0\beta} = 0. \quad (3.14)$$

This is consistent with the Lorenz gauge condition (2.21) at the lowest order in ϵ . A similar argument can be applied for the complex-conjugate Eq. (3.9), from which we obtain $k_\alpha A_0^{*\alpha} = 0$.

Using Eqs. (3.8)-(3.10), we obtain the well-known system of equations governing the geometrical optics approximation at the lowest order in ϵ :

$$k_\mu k^\mu = 0, \quad (3.15)$$

$$k^\alpha A_{0\alpha} = k_\alpha A_0^{*\alpha} = 0, \quad (3.16)$$

$$\nabla_\mu (k^\mu \mathcal{I}_0) = 0, \quad (3.17)$$

where $\mathcal{I}_0 = A_0^{*\alpha} A_{0\alpha}$ is the lowest-order intensity (more precisely, \mathcal{I}_0 is proportional to the wave action density [59]). Equation (3.17) is obtained from Eq. (3.10) by using the orthogonality condition (3.16). Using Eq. (2.6), we have

$$\nabla_\mu k_\alpha = \nabla_\alpha k_\mu, \quad (3.18)$$

and we can use Eq. (3.15) to derive the geodesic equation for k_μ :

$$k^\nu \nabla_\nu k_\mu = 0. \quad (3.19)$$

C. First-order geometrical optics

Here, we examine Eqs. (3.5) and (3.6) at order ϵ^1 only:

$$D_\alpha^\beta A_{1\beta} - i \left(\nabla^\mu D_\alpha^\beta \right) \nabla_\mu A_{0\beta} - \frac{i}{2} \left(\nabla_\mu \nabla^\mu D_\alpha^\beta \right) A_{0\beta} = 0, \quad (3.20)$$

$$D_\alpha^\beta A_1^{*\alpha} + i \left(\nabla^\mu D_\alpha^\beta \right) \nabla_\mu A_0^{*\alpha} + \frac{i}{2} \left(\nabla_\mu \nabla^\mu D_\alpha^\beta \right) A_0^{*\alpha} = 0. \quad (3.21)$$

Using Eq. (3.4), we can also rewrite Eq. (3.20) as follows:

$$\begin{aligned} k^\mu \nabla_\mu A_{0\alpha} - \frac{1}{2} k_\alpha \nabla_\mu A_0^\mu - \frac{1}{2} k_\beta \nabla_\alpha A_0^\beta \\ - \frac{i}{2} k_\alpha k^\beta A_{1\beta} + \frac{1}{2} A_{0\alpha} \nabla_\mu k^\mu \\ - \frac{1}{4} A_0^\beta \nabla_\beta k_\alpha - \frac{1}{4} A_0^\beta \nabla_\alpha k_\beta = 0. \end{aligned} \quad (3.22)$$

Using Eq. (3.18), we can rewrite the last two terms as

$$-\frac{1}{4} A_0^\beta \nabla_\beta k_\alpha - \frac{1}{4} A_0^\beta \nabla_\alpha k_\beta = -\frac{1}{2} A_0^\beta \nabla_\alpha k_\beta. \quad (3.23)$$

Using Eq. (3.16), we also have

$$0 = \nabla_\alpha (k_\beta A_0^\beta) = k_\beta \nabla_\alpha A_0^\beta + A_0^\beta \nabla_\alpha k_\beta. \quad (3.24)$$

Then, Eq. (3.22) becomes

$$\begin{aligned} k^\mu \nabla_\mu A_{0\alpha} + \frac{1}{2} A_{0\alpha} \nabla_\mu k^\mu \\ - \frac{1}{2} k_\alpha (\nabla_\mu A_0^\mu + i k_\mu A_1^\mu) = 0. \end{aligned} \quad (3.25)$$

The last term can be eliminated by using the Lorenz gauge (2.21). The same steps can be applied to the complex-conjugate Eq. (3.21):

$$\begin{aligned} k^\mu \nabla_\mu A_{0\alpha} + \frac{1}{2} A_{0\alpha} \nabla_\mu k^\mu = 0, \\ k^\mu \nabla_\mu A_0^{*\beta} + \frac{1}{2} A_0^{*\beta} \nabla_\mu k^\mu = 0. \end{aligned} \quad (3.26)$$

Furthermore, using the lowest-order intensity \mathcal{J}_0 , we can write the amplitude in the following way:

$$A_{0\alpha} = \sqrt{\mathcal{J}_0} a_{0\alpha}, \quad A_0^{*\alpha} = \sqrt{\mathcal{J}_0} a_0^{*\alpha}, \quad (3.27)$$

where $a_{0\alpha}$ is a unit complex covector (i.e., $a_0^{*\alpha} a_{0\alpha} = 1$) describing the polarization. Then, from Eq. (3.26), together with Eq. (3.17), we obtain

$$k^\mu \nabla_\mu a_{0\alpha} = k^\mu \nabla_\mu a_0^{*\alpha} = 0. \quad (3.28)$$

The parallel propagation of the complex covector $a_{0\alpha}$ along the integral curve of k^μ is another well-known result of the geometrical optics approximation.

D. The polarization vector in a null tetrad

We observed that the polarization vector satisfies the orthogonality condition

$$k^\alpha a_{0\alpha} = 0. \quad (3.29)$$

Consider a null tetrad [60, Sec. 3] $\{k_\alpha, n_\alpha, m_\alpha, \bar{m}_\alpha\}$ satisfying

$$\begin{aligned} m_\alpha \bar{m}^\alpha &= 1, & k_\alpha n^\alpha &= -1, \\ k_\alpha k^\alpha &= n_\alpha n^\alpha = m_\alpha m^\alpha = \bar{m}_\alpha \bar{m}^\alpha = 0, \\ k_\alpha m^\alpha &= k_\alpha \bar{m}^\alpha = n_\alpha m^\alpha = n_\alpha \bar{m}^\alpha = 0. \end{aligned} \quad (3.30)$$

Note that we use the metric signature opposite to that used in Ref. [60, Sec. 3]. The covectors $n_\alpha, m_\alpha, \bar{m}_\alpha$ are not assumed to be parallel-propagated along the geodesic generated by k^α . It is only k_α that is parallel-propagated along the geodesic generated by k^α , in accordance with Eq. (3.19). Since the null tetrad is adapted to the covector k_α , the orthogonality conditions (3.30) imply that m_α and \bar{m}_α are functions of k_α . The polarization covector $a_{0\alpha}$ is orthogonal to k_α , so we can decompose it as

$$a_{0\alpha}(x, k) = z_1(x) m_\alpha(x, k) + z_2(x) \bar{m}_\alpha(x, k) + z_3(x) k_\alpha(x), \quad (3.31)$$

where z_1, z_2 , and z_3 are complex scalar functions. Since $a_{0\alpha}$ is a unit complex covector, the scalar functions z_1 and z_2 are constrained by

$$z_1^* z_1 + z_2^* z_2 = 1. \quad (3.32)$$

It is important to note that the decomposition (3.31), and more specifically, the choice of m_α , requires choosing a null covector n_α . Fixing n_α is equivalent to choosing a unit timelike covector field t_α that can represent a family of timelike observers. We can always take n_α as

$$t_\alpha = \frac{1}{2\epsilon\omega} k_\alpha + \epsilon\omega n_\alpha. \quad (3.33)$$

Once n_α (or t_α) is fixed, the remaining $\text{SO}(2)$ gauge freedom in the choice of m_α is described by the spin rotation

$$k_\alpha \mapsto k_\alpha, \quad n_\alpha \mapsto n_\alpha, \quad m_\alpha \mapsto e^{i\phi(x)} m_\alpha, \quad (3.34)$$

for $\phi(x) \in \mathbb{R}$. Polarization measurements will always depend on the choice of m_α and \bar{m}_α . However, as shown in Sec. IV B 1, the modified ray equations describing the gravitational spin Hall effect of light do not depend on the particular choice of m_α and \bar{m}_α . Thus, we can work with any smooth choice of m_α and \bar{m}_α that satisfy Eq. (3.30).

Using Eqs. (3.31) and (3.19), the parallel-transport equation for the polarization vector becomes

$$\begin{aligned} 0 &= k^\mu \nabla_\mu a_{0\alpha} \\ &= z_1 k^\mu \nabla_\mu m_\alpha + z_2 k^\mu \nabla_\mu \bar{m}_\alpha + m_\alpha k^\mu \nabla_\mu z_1 \\ &\quad + \bar{m}_\alpha k^\mu \nabla_\mu z_2 + k_\alpha k^\mu \nabla_\mu z_3. \end{aligned} \quad (3.35)$$

Contracting the above equation with \bar{m}^α , m^α , and n^α , we obtain

$$\begin{aligned} k^\mu \nabla_\mu z_1 &= -z_1 \bar{m}^\alpha k^\mu \nabla_\mu m_\alpha, \\ k^\mu \nabla_\mu z_2 &= -z_2 m^\alpha k^\mu \nabla_\mu \bar{m}_\alpha, \\ k^\mu \nabla_\mu z_3 &= -(z_1 m^\alpha + z_2 \bar{m}^\alpha) k^\mu \nabla_\mu n_\alpha. \end{aligned} \quad (3.36)$$

Recall that in the above equations, the covectors m_α and \bar{m}_α are functions of x and $k(x)$. The covariant derivatives are applied as follows:

$$\begin{aligned} k^\mu \nabla_\mu m_\alpha &= k^\mu \nabla_\mu [m_\alpha(x, k)] \\ &= k^\mu \left(\overset{h}{\nabla}_\mu m_\alpha \right) (x, k) \\ &\quad + k^\mu (\nabla_\mu k_\nu) \left(\overset{v}{\nabla}^\nu m_\alpha \right) (x, k) \\ &= k^\mu \overset{h}{\nabla}_\mu m_\alpha, \end{aligned} \quad (3.37)$$

where $\overset{h}{\nabla}_\mu$ is the horizontal derivative (Appendix A). It is convenient to introduce the two-dimensional unit complex vector

$$z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \quad (3.38)$$

which is analogous to the Jones vector in optics [47, 48, 61, 62]. We also use the Hermitian transpose z^\dagger , defined as follows:

$$z^\dagger = (z_1^* \ z_2^*). \quad (3.39)$$

Then, the equations for z_1 and z_2 can be written in a more compact form:

$$k^\mu \nabla_\mu z = ik^\mu B_\mu \sigma_3 z, \quad (3.40)$$

where σ_3 is the third Pauli matrix,

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (3.41)$$

and B_μ is the real 1-form extending to general relativity the Berry connection used in optics [47, 62]:

$$\begin{aligned} B_\mu(x, k) &= \frac{i}{2} \left(\bar{m}^\alpha \overset{h}{\nabla}_\mu m_\alpha - m_\alpha \overset{h}{\nabla}_\mu \bar{m}^\alpha \right) \\ &= i \bar{m}^\alpha \overset{h}{\nabla}_\mu m_\alpha. \end{aligned} \quad (3.42)$$

Furthermore, if we restrict z to an affinely parametrized null geodesic $\tau \mapsto x^\mu(\tau)$, with $\dot{x}^\mu = k^\mu$, we can write

$$\dot{z} = ik^\mu B_\mu \sigma_3 z, \quad (3.43)$$

where $\dot{z} = \dot{x}^\mu \nabla_\mu z$. Integrating along the worldline, we obtain

$$z(\tau) = \begin{pmatrix} e^{i\gamma(\tau)} & 0 \\ 0 & e^{-i\gamma(\tau)} \end{pmatrix} z(0), \quad (3.44)$$

where γ represents the Berry phase [47, 62],

$$\gamma(\tau_1) = \int_{\tau_0}^{\tau_1} d\tau k^\mu B_\mu. \quad (3.45)$$

Using either Eq. (3.36) or Eq. (3.43), we see that the evolution of z_1 and z_2 is decoupled in the circular polarization basis, and the following quantities are conserved along k^μ :

$$\begin{aligned} 1 &= z_1^* z_1 + z_2^* z_2 = z^\dagger z, \\ s &= z_1^* z_1 - z_2^* z_2 = z^\dagger \sigma_3 z. \end{aligned} \quad (3.46)$$

Based on our assumptions on the initial conditions (Sec. II D), we only consider beams which are circularly polarized, i.e. one of the conditions

$$z(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{or} \quad z(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (3.47)$$

holds. Thus, we have $s = \pm 1$, depending on the choice of the initial polarization state.

The results described in this section are similar to the description of the polarization of electromagnetic waves traveling in a medium with an inhomogeneous index of refraction [62].

E. Extended geometrical optics

Now, we take Eqs. (3.5)-(3.7), but without splitting them order by order in ϵ . Our aim is to derive an effective Hamilton-Jacobi system that would give us $\mathcal{O}(\epsilon)$ corrections to the ray equations.

1. Effective dispersion relation

By contracting Eq. (3.5) with $A^{*\alpha}$ and Eq. (3.6) with A_β , and also adding them together, we obtain the following equation:

$$\begin{aligned} D_\alpha^\beta A^{*\alpha} A_\beta \\ - \frac{i\epsilon}{2} \left(\overset{v}{\nabla}^\mu D_\alpha^\beta \right) (A^{*\alpha} \nabla_\mu A_\beta - A_\beta \nabla_\mu A^{*\alpha}) = \mathcal{O}(\epsilon^2). \end{aligned} \quad (3.48)$$

Using Eqs. (3.4) and (3.16), we can rewrite the above equation as follows:

$$\begin{aligned} \frac{1}{2} k_\mu k^\mu (A_0^{*\alpha} A_{0\alpha} + \epsilon A_0^{*\alpha} A_{1\alpha} + \epsilon A_1^{*\alpha} A_{0\alpha}) \\ - \frac{i\epsilon}{2} k^\mu (A_0^{*\alpha} \nabla_\mu A_{0\alpha} - A_{0\alpha} \nabla_\mu A_0^{*\alpha}) \\ + \frac{i\epsilon}{4} k_\alpha (A_0^{*\mu} \nabla_\mu A_0^\alpha - A_0^\mu \nabla_\mu A_0^{*\alpha}) = \mathcal{O}(\epsilon^2). \end{aligned} \quad (3.49)$$

Using Eq. (3.16), we obtain

$$\begin{aligned} 0 &= A_0^{*\mu} \nabla_\mu (k_\alpha A_0^\alpha) \\ &= k_\alpha A_0^{*\mu} \nabla_\mu A_0^\alpha + A_0^{*\mu} A_0^\alpha \nabla_\mu k_\alpha, \end{aligned} \quad (3.50)$$

so we can write

$$\begin{aligned} \frac{i\epsilon}{4} k_\alpha (A_0^{*\mu} \nabla_\mu A_0^\alpha - A_0^\mu \nabla_\mu A_0^{*\alpha}) \\ - \frac{i\epsilon}{2} \nabla_\mu k_\alpha A_0^{*\mu} A_0^\alpha \\ = 0, \end{aligned} \quad (3.51)$$

where the last equality is due to Eq. (3.18). Then, Eq. (3.48) becomes

$$\begin{aligned} \frac{1}{2} k_\mu k^\mu (A_0^{*\alpha} A_{0\alpha} + \epsilon A_0^{*\alpha} A_{1\alpha} + \epsilon A_1^{*\alpha} A_{0\alpha}) \\ - \frac{i\epsilon}{2} k^\mu (A_0^{*\alpha} \nabla_\mu A_{0\alpha} - A_{0\alpha} \nabla_\mu A_0^{*\alpha}) = \mathcal{O}(\epsilon^2). \end{aligned} \quad (3.52)$$

Let us introduce the $\mathcal{O}(\epsilon^1)$ intensity

$$\begin{aligned} \mathcal{J} &= \mathcal{A}^{*\alpha} \mathcal{A}_\alpha \\ &= A_0^{*\alpha} A_{0\alpha} + \epsilon A_0^{*\alpha} A_{1\alpha} + \epsilon A_1^{*\alpha} A_{0\alpha} + \mathcal{O}(\epsilon^2). \end{aligned} \quad (3.53)$$

Then, we can rewrite the amplitude as

$$A_\alpha = \sqrt{\mathcal{J}} a_\alpha = \sqrt{\mathcal{J}} (a_{0\alpha} + \epsilon a_{1\alpha}) + \mathcal{O}(\epsilon^2), \quad (3.54)$$

where a_α is a unit complex covector. Then, from Eq. (3.52) we obtain

$$\frac{1}{2}k_\mu k^\mu - \frac{i\epsilon}{2}k^\mu (a_0^{*\alpha} \nabla_\mu a_{0\alpha} - a_{0\alpha} \nabla_\mu a_0^{*\alpha}) = \mathcal{O}(\epsilon^2). \quad (3.55)$$

This can be viewed as an effective dispersion relation, containing $\mathcal{O}(\epsilon)$ corrections to the geometrical optics equation (3.15). Finally, let us introduce

$$K_\mu = k_\mu - \frac{i\epsilon}{2} (a_0^{*\alpha} \nabla_\mu a_{0\alpha} - a_{0\alpha} \nabla_\mu a_0^{*\alpha}) \quad (3.56)$$

and rewrite the effective dispersion relation as

$$\frac{1}{2}K_\mu K^\mu = \mathcal{O}(\epsilon^2). \quad (3.57)$$

It is worth noting that this equation can also be obtained directly from the effective field action (3.3), specifically by varying the latter with respect to \mathcal{I} .

2. Effective transport equation

Using Eqs. (3.4), (3.15), and (3.16), the effective transport equation (3.7) becomes

$$\begin{aligned} \nabla_\mu \left[k^\mu (A_0^{*\alpha} A_{0\alpha} + \epsilon A_0^{*\alpha} A_{1\alpha} + \epsilon A_1^{*\alpha} A_{0\alpha}) \right. \\ - \frac{i\epsilon}{2} g^{\mu\nu} (A_0^{*\alpha} \nabla_\nu A_{0\alpha} - A_{0\alpha} \nabla_\nu A_0^{*\alpha}) \\ + \frac{i\epsilon}{4} (A_0^{*\alpha} \nabla_\alpha A_0^\mu - A_0^\mu \nabla_\alpha A_0^{*\alpha}) \\ + \frac{i\epsilon}{4} (A_0^{*\mu} \nabla_\alpha A_0^\alpha - A_0^\alpha \nabla_\alpha A_0^{*\mu}) \\ \left. - \frac{\epsilon}{2} k_\alpha (A_0^{*\mu} A_{1\alpha} + A_1^{*\alpha} A_0^\mu) \right] = \mathcal{O}(\epsilon^2). \end{aligned} \quad (3.58)$$

We can perform the following replacements in the above equation:

$$\begin{aligned} A_0^{*\alpha} \nabla_\alpha A_0^\mu &= \nabla_\alpha (A_0^{*\alpha} A_0^\mu) - \nabla_\alpha A_0^{*\alpha} A_0^\mu, \\ \nabla_\alpha A_0^{*\mu} A_0^\alpha &= \nabla_\alpha (A_0^{*\mu} A_0^\alpha) - A_0^{*\mu} \nabla_\alpha A_0^\alpha. \end{aligned} \quad (3.59)$$

After rearranging terms, the effective transport equation becomes

$$\begin{aligned} \nabla_\mu \left[k^\mu (A_0^{*\alpha} A_{0\alpha} + \epsilon A_0^{*\alpha} A_{1\alpha} + \epsilon A_1^{*\alpha} A_{0\alpha}) \right. \\ - \frac{i\epsilon}{2} g^{\mu\nu} (A_0^{*\alpha} \nabla_\nu A_{0\alpha} - A_{0\alpha} \nabla_\nu A_0^{*\alpha}) \\ - \frac{i\epsilon}{2} A_0^\mu (\nabla_\alpha A_0^{*\alpha} - i k_\alpha A_1^{*\alpha}) \\ + \frac{i\epsilon}{2} A_0^{*\mu} (\nabla_\alpha A_0^\alpha + i k_\alpha A_1^\alpha) \\ \left. + \frac{i\epsilon}{4} \nabla_\alpha (A_0^{*\alpha} A_0^\mu) \right] = \mathcal{O}(\epsilon^2). \end{aligned} \quad (3.60)$$

The last term above vanishes due to the symmetry of the Ricci tensor:

$$\begin{aligned} \nabla_\mu \nabla_\alpha (A_0^{*\alpha} A_0^\mu) &= \nabla_{[\mu} \nabla_{\alpha]} (A_0^{*\alpha} A_0^\mu) \\ &= (R_{\alpha\nu\mu}^\nu - R_{\mu\nu\alpha}^\nu) A_0^{*\alpha} A_0^\mu \\ &= (R_{\alpha\mu} - R_{\mu\alpha}) A_0^{*\alpha} A_0^\mu \\ &= 0. \end{aligned} \quad (3.61)$$

Furthermore, after using the Lorenz gauge condition (2.21), we are left with the following form of the effective transport equation:

$$\begin{aligned} \nabla_\mu \left[k^\mu (A_0^{*\alpha} A_{0\alpha} + \epsilon A_0^{*\alpha} A_{1\alpha} + \epsilon A_1^{*\alpha} A_{0\alpha}) \right. \\ \left. - \frac{i\epsilon}{2} g^{\mu\nu} (A_0^{*\alpha} \nabla_\nu A_{0\alpha} - A_{0\alpha} \nabla_\nu A_0^{*\alpha}) \right] = \mathcal{O}(\epsilon^2). \end{aligned} \quad (3.62)$$

Introducing the intensity \mathcal{I} and the vector K^μ , we obtain

$$\begin{aligned} \nabla_\mu \left\{ \mathcal{I} \left[k^\mu - \frac{i\epsilon}{2} g^{\mu\nu} (a_0^{*\alpha} \nabla_\nu a_{0\alpha} - a_{0\alpha} \nabla_\nu a_0^{*\alpha}) \right] \right\} \\ = \nabla_\mu (\mathcal{I} K^\mu) = \mathcal{O}(\epsilon^2). \end{aligned} \quad (3.63)$$

This is an effective transport equation for the intensity \mathcal{I} , which includes $\mathcal{O}(\epsilon)$ corrections to the geometrical optics Eq. (3.17). As discussed in Ref. [59], the direction of K^μ coincides with the direction of the wave action flux.

IV. EFFECTIVE RAY EQUATIONS

A. Hamilton-Jacobi system at the leading order

The lowest-order geometrical optics equations (3.15) and (3.17) can be viewed as a system of coupled partial differential equations:

$$\frac{1}{2} g^{\mu\nu} k_\mu k_\nu = 0, \quad (4.1)$$

$$\nabla_\mu (\mathcal{I}_0 k^\mu) = 0, \quad (4.2)$$

where $k_\mu = \nabla_\mu S$. Equation (4.1) is a Hamilton-Jacobi equation for the phase function S , and Eq. (4.2) is a transport equation for the intensity \mathcal{I}_0 [63]. The Hamilton-Jacobi equation can be solved using the method of characteristics. This is done by defining a Hamiltonian function on T^*M , such that

$$H(x, \nabla S) = \frac{1}{2} g^{\mu\nu} k_\mu k_\nu = 0. \quad (4.3)$$

It is obvious that in this case, the Hamiltonian is

$$H(x, p) = \frac{1}{2} g^{\mu\nu} p_\mu p_\nu. \quad (4.4)$$

Note that in contrast to the dispersion relation (4.3), the Hamiltonian (4.4) is a function on the whole phase space T^*M , with p_μ being an arbitrary covector. Hamilton's equations take the following form:

$$\dot{x}^\mu = \frac{\partial H}{\partial p_\mu} = g^{\mu\nu} p_\nu, \quad (4.5)$$

$$\dot{p}_\mu = -\frac{\partial H}{\partial x^\mu} = -\frac{1}{2} \partial_\mu g^{\alpha\beta} p_\alpha p_\beta. \quad (4.6)$$

Given a solution $\{x^\mu(\tau), p_\mu(\tau)\}$ for Hamilton's equations, we obtain a solution of the Hamilton-Jacobi Eq. (4.3) by taking [64, p. 433]:

$$S(x^\mu(\tau_1), p_\mu(\tau_1)) = \int_{\tau_0}^{\tau_1} d\tau [\dot{x}^\mu p_\mu - H(x, p)] + \text{const.} \quad (4.7)$$

Note that the above equation represents an action, with the corresponding Lagrangian related to the Hamiltonian (4.4) by a Legendre transformation [65, Ex. 3.6.10]. The Euler-Lagrange equation is equivalent to the geodesic equation [65, Th. 3.7.1] and with Hamilton's equations (4.5) and (4.6). Once the Hamilton-Jacobi equation is solved, the transport Eq. (4.2) can also be solved, at least in principle [63]. However, our main interest is in the ray equations governed by the Hamiltonian (4.4). The corresponding Hamilton's equations (4.5) and (4.6) describe null geodesics. These equations can easily be rewritten as

$$\ddot{x}^\mu + \Gamma_{\alpha\beta}^\mu \dot{x}^\alpha \dot{x}^\beta = 0, \quad (4.8)$$

or in the explicitly covariant form:

$$p^\nu \nabla_\nu p^\mu = \dot{x}^\nu \nabla_\nu \dot{x}^\mu = 0. \quad (4.9)$$

B. Effective Hamilton-Jacobi system

The effective dispersion relation (3.57), together with the effective transport equation (3.63) introduce $\mathcal{O}(\epsilon^1)$ corrections over the system discussed above:

$$\frac{1}{2} g^{\mu\nu} k_\mu k_\nu - \frac{i\epsilon}{2} k^\mu (a_0^{*\alpha} \nabla_\mu a_{0\alpha} - a_{0\alpha} \nabla_\mu a_0^{*\alpha}) = \mathcal{O}(\epsilon^2), \quad (4.10)$$

$$\nabla_\mu \left\{ \mathcal{I} \left[k^\mu - \frac{i\epsilon}{2} g^{\mu\nu} (a_0^{*\alpha} \nabla_\nu a_{0\alpha} - a_{0\alpha} \nabla_\nu a_0^{*\alpha}) \right] \right\} = \mathcal{O}(\epsilon^2). \quad (4.11)$$

Using Eq. (3.31), the effective dispersion relation becomes

$$\frac{1}{2} g^{\mu\nu} k_\mu k_\nu - \frac{i\epsilon}{2} k^\mu (z^\dagger \partial_\mu z - \partial_\mu z^\dagger z) - \epsilon s k^\mu B_\mu = \mathcal{O}(\epsilon^2), \quad (4.12)$$

where $B_\mu = B_\mu(x, k)$ is the Berry connection introduced in Eq. (3.42), and $s = \pm 1$, depending on the initial polarization. Using Eq. (3.44), together with the assumption

on the initial polarization, we can write:

$$-\frac{i\epsilon}{2} k^\mu (z^\dagger \partial_\mu z - \partial_\mu z^\dagger z) = \epsilon s k^\mu \partial_\mu \gamma. \quad (4.13)$$

Since the value of s is fixed by the initial conditions, the only unknowns are the phase function S and the Berry phase γ . We can write an effective Hamilton-Jacobi equation for the total phase $\tilde{S} = S + \epsilon s \gamma$:

$$\begin{aligned} H(x, \nabla \tilde{S}) &= \frac{1}{2} g^{\mu\nu} k_\mu k_\nu + \epsilon s k^\mu \partial_\mu \gamma - \epsilon s k^\mu B_\mu + \mathcal{O}(\epsilon^2) \\ &= \frac{1}{2} g^{\mu\nu} \nabla_\mu \tilde{S} \nabla_\nu \tilde{S} - \epsilon s g^{\mu\nu} B_\mu \nabla_\nu \tilde{S} + \mathcal{O}(\epsilon^2). \end{aligned} \quad (4.14)$$

Note that the phase \tilde{S} represents the overall phase factor, up to order $\mathcal{O}(\epsilon^2)$, of a circularly polarized WKB solution, $\mathcal{A}_\alpha = \text{Re}(\sqrt{\mathcal{I}} m_\alpha e^{i\gamma} e^{iS/\epsilon})$ or $\mathcal{A}_\alpha = \text{Re}(\sqrt{\mathcal{I}} \bar{m}_\alpha e^{-i\gamma} e^{iS/\epsilon})$, depending on the state of circular polarization. As discussed in Ref. [15], the Berry phase γ , which comes as a correction to the overall phase of the WKB solution, is responsible for the spin Hall effect of light. The corresponding Hamiltonian function on T^*M is

$$H(x, p) = \frac{1}{2} g^{\mu\nu} p_\mu p_\nu - \epsilon s g^{\mu\nu} p_\mu B_\nu(x, p), \quad (4.15)$$

and we have the following Hamilton's equations:

$$\dot{x}^\mu = g^{\mu\nu} p_\nu - \epsilon s \left(B^\mu + p^\alpha \nabla^\mu B_\alpha \right), \quad (4.16)$$

$$\dot{p}_\mu = -\frac{1}{2} \partial_\mu g^{\alpha\beta} p_\alpha p_\beta + \epsilon s p_\alpha (\partial_\mu g^{\alpha\beta} B_\beta + g^{\alpha\beta} \partial_\mu B_\beta). \quad (4.17)$$

These equations contain polarization-dependent corrections to the null geodesic Eqs. (4.5) and (4.6), representing the gravitational spin Hall effect of light. For $\epsilon = 0$, one recovers the standard geodesic equation in canonical coordinates.

We can also write these ray equations in a more compact form

$$\begin{pmatrix} \dot{x}^\mu \\ \dot{p}_\mu \end{pmatrix} = \begin{pmatrix} 0 & \delta_\nu^\mu \\ -\delta_\mu^\nu & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial H}{\partial x^\nu} \\ \frac{\partial H}{\partial p_\nu} \end{pmatrix}, \quad (4.18)$$

where the constant matrix on the right-hand side is the inverse of the symplectic 2-form, or the Poisson tensor [66].

1. Noncanonical coordinates

The Hamiltonian (4.15) contains the Berry connection B_μ , which is gauge dependent. The latter means that B_μ depends on the choice of m_α and \bar{m}_α ; for example, the transformation $m_\alpha \mapsto m_\alpha e^{i\phi}$ causes the following transformation of the Berry connection:

$$B_\mu \mapsto B_\mu - \nabla_\mu \phi. \quad (4.19)$$

This kind of gauge dependence was considered by Littlejohn and Flynn in Ref. [49], where they also proposed how to make the Hamiltonian and the equations of motion gauge invariant. The main idea is to introduce noncanonical coordinates such that the Berry connection is removed from the Hamiltonian and the symplectic form acquires the corresponding Berry curvature, which is gauge invariant. This is similar to the description of a charged particle in an electromagnetic field in terms of either the canonical or the kinetic momentum of the particle. The Berry connection and Berry curvature play a similar role as the electromagnetic vector potential and the electromagnetic tensor [67].

We start by rewriting the Hamiltonian (4.15) as

$$H(x, p) = H_0(x, p) - \epsilon s g^{\mu\nu} p_\mu B_\nu(x, p), \quad (4.20)$$

where $H_0 = \frac{1}{2} g^{\mu\nu} p_\mu p_\nu$. Following Ref. [49], the Berry connection can be written in the following way, by using the definition of the horizontal derivative:

$$\begin{aligned} p^\mu B_\mu(x, p) &= ip^\mu \bar{m}^\alpha \overset{h}{\nabla}_\mu m_\alpha \\ &= ip^\mu \bar{m}^\alpha \nabla_\mu m_\alpha + ip^\mu p_\sigma \Gamma_{\mu\rho}^\sigma \bar{m}^\alpha \overset{v}{\nabla}^\rho m_\alpha \quad (4.21) \\ &= i \frac{\partial H_0}{\partial p_\mu} \bar{m}^\alpha \nabla_\mu m_\alpha - i \frac{\partial H_0}{\partial x^\mu} \bar{m}^\alpha \overset{v}{\nabla}^\mu m_\alpha. \end{aligned}$$

The Berry connection can be eliminated formally from the Hamiltonian (4.15) by considering the following substitution on T^*M :

$$X^\mu = x^\mu + i\epsilon s \bar{m}^\alpha \overset{v}{\nabla}^\mu m_\alpha, \quad (4.22)$$

$$P_\mu = p_\mu - i\epsilon s \bar{m}^\alpha \nabla_\mu m_\alpha. \quad (4.23)$$

It is possible to obtain this substitution as the linearization of a change of coordinates. For more details, see Appendix D.

Since the symplectic form transforms nontrivially under this substitution, (X, P) are noncanonical coordinates. The Hamiltonian (4.15) is a scalar, so we obtain

$$\begin{aligned} H'(X, P) &= H(x, p) \\ &= H \left(X^\mu - i\epsilon s \bar{m}^\alpha \overset{v}{\nabla}^\mu m_\alpha, P_\mu + i\epsilon s \bar{m}^\alpha \nabla_\mu m_\alpha \right) \\ &= H(X, P) - i\epsilon s \frac{\partial H_0}{\partial x^\mu} \bar{m}^\alpha \overset{v}{\nabla}^\mu m_\alpha + i\epsilon s \frac{\partial H_0}{\partial p_\mu} \bar{m}^\alpha \nabla_\mu m_\alpha \\ &= H_0(X, P). \quad (4.24) \end{aligned}$$

In the new coordinate system (X, P) , we obtain the following Hamiltonian:

$$H'(X, P) = \frac{1}{2} g^{\mu\nu}(X) P_\mu P_\nu. \quad (4.25)$$

The corresponding Hamilton's equations can be written in a matrix form as

$$\begin{pmatrix} \dot{X}^\mu \\ \dot{P}_\mu \end{pmatrix} = T' \begin{pmatrix} \frac{\partial H'}{\partial X^\nu} \\ \frac{\partial H'}{\partial P_\nu} \end{pmatrix}, \quad (4.26)$$

where T' is the Poisson tensor in the new variables. Following Marsden and Ratiu [66, p. 343], we obtain

$$T' = \begin{pmatrix} \epsilon s (F_{pp})^{\nu\mu} & \delta_\nu^\mu + \epsilon s (F_{xp})_\nu^\mu \\ -\delta_\mu^\nu - \epsilon s (F_{xp})_\mu^\nu & -\epsilon s (F_{xx})_{\nu\mu} \end{pmatrix}, \quad (4.27)$$

where we have the following Berry curvature terms:

$$\begin{aligned} (F_{pp})^{\nu\mu} &= i \left(\overset{v}{\nabla}^\mu \bar{m}^\alpha \overset{v}{\nabla}^\nu m_\alpha - \overset{v}{\nabla}^\nu \bar{m}^\alpha \overset{v}{\nabla}^\mu m_\alpha \right. \\ &\quad \left. + \bar{m}^\alpha \overset{v}{\nabla}^{[\mu} \overset{v}{\nabla}^{\nu]} m_\alpha - m_\alpha \overset{v}{\nabla}^{[\mu} \overset{v}{\nabla}^{\nu]} \bar{m}^\alpha \right), \\ (F_{xx})_{\nu\mu} &= i \left(\nabla_\mu \bar{m}^\alpha \nabla_\nu m_\alpha - \nabla_\nu \bar{m}^\alpha \nabla_\mu m_\alpha \right. \\ &\quad \left. + \bar{m}^\alpha \nabla_{[\mu} \nabla_{\nu]} m_\alpha - m_\alpha \nabla_{[\mu} \nabla_{\nu]} \bar{m}^\alpha \right), \\ (F_{px})_\nu^\mu &= - (F_{xp})_\nu^\mu \\ &= i \left(\overset{v}{\nabla}^\mu \bar{m}^\alpha \nabla_\nu m_\alpha - \nabla_\nu \bar{m}^\alpha \overset{v}{\nabla}^\mu m_\alpha \right). \end{aligned} \quad (4.28)$$

The Poisson tensor in noncanonical coordinates T' automatically satisfies the Jacobi identity, since it is a covariant quantity obtained from the Poisson tensor in canonical coordinates T through a change of variables on the cotangent bundle.

Simplified expressions for the Berry curvature terms can be found in Appendix C. Now we can write Hamilton's equations in the new variables:

$$\begin{aligned} \dot{X}^\mu &= P^\mu + \epsilon s P^\nu (F_{px})_\nu^\mu \\ &\quad + \epsilon s \Gamma_{\beta\nu}^\alpha P_\alpha P^\beta (F_{pp})^{\nu\mu}, \end{aligned} \quad (4.29)$$

$$\begin{aligned} \dot{P}_\mu &= \Gamma_{\beta\mu}^\alpha P_\alpha P^\beta - \epsilon s P^\nu (F_{xx})_{\nu\mu} \\ &\quad - \epsilon s \Gamma_{\beta\nu}^\alpha P_\alpha P^\beta (F_{xp})_\mu^\nu. \end{aligned} \quad (4.30)$$

The last term on the right-hand side of Eq. (4.29) is the covariant analogue of the spin Hall effect correction obtained in optics, $(\mathbf{p} \times \mathbf{p})/|\mathbf{p}|^3$, due to the Berry curvature in momentum space [4, 47]. This term is also the source of the gravitational spin Hall effect in the work of Gosselin *et al.* [41]. In Eq. (4.30), the second term on the right-hand side contains the Riemann tensor and resembles the curvature term obtained in the Mathisson-Papapetrou-Dixon equations [27].

Given a null covector P_μ , the class of Lorentz transformations leaving P_μ invariant define the little group, which is isomorphic to $SE(2)$, the symmetry group of the two-dimensional Euclidean plane [51]. In terms of a null tetrad $\{P, n, m, \bar{m}\}$, the action of the little group can be split into the following types of transformations [68, p. 53]:

$$\begin{aligned} \text{Type 1: } & P \mapsto P, \quad n \mapsto n, \\ & m \mapsto m e^{i\phi}, \quad \bar{m} \mapsto \bar{m} e^{-i\phi}, \\ \text{Type 2: } & P \mapsto P, \quad n \mapsto n + \bar{a}m + a\bar{m} + a\bar{a}P, \\ & m \mapsto m + aP, \quad \bar{m} \mapsto \bar{m} + \bar{a}P, \end{aligned} \quad (4.31)$$

where ϕ is a real scalar function and a is a complex scalar function. The transformations of Type 1 are the spin rotations mentioned in Sec. III D, while the transformations of Type 2 can be considered as a change of observer t_μ , based on Eq. (3.33). It can easily be checked that the Berry curvature terms in Eq. (4.28) are invariant under Type 1 transformations. However, the Berry curvature terms are not invariant under Type 2 transformations. As a consequence, the ray equations (4.29) and (4.30) depend on the choice of observer. It is shown in the following section how this observer dependence is related to the problem of localizing massless spinning particles [50, 51].

V. EXAMPLES

In this section, we apply the modified ray equations describing the gravitational spin Hall effect of light to two concrete examples. The first example, concerning the relativistic Hall effect and Wigner translations, is treated analytically, while the second example, describing the propagation of polarized light rays close to a Schwarzschild black hole, is treated numerically.

When working with the modified ray equations, in either the canonical form given in Eqs. (4.16) and (4.17) or the noncanonical form given in Eqs. (4.29) and (4.30), one needs to specify the background metric $g_{\mu\nu}$, and the choice of polarization vectors m^α and \bar{m}^α . The polarization vectors are needed in order to compute the Berry connection and the Berry curvature. A particular choice of polarization vectors can easily be constructed by introducing an orthonormal tetrad $(e_a)^\mu$, with $(e_0)^\mu = t^\mu$ representing our choice of family of time-like observers. Adapting the polarization vectors used in optics [47], we can write $p^\mu = P^a(e_a)^\mu$, $v^\mu = V^a(e_a)^\mu$, and $w^\mu = W^a(e_a)^\mu$, where the components of these vectors are given by

$$P^a = \begin{pmatrix} P^0 \\ P^1 \\ P^2 \\ P^3 \end{pmatrix}, \quad V^a = \frac{1}{P_p} \begin{pmatrix} 0 \\ -P^2 \\ P^1 \\ 0 \end{pmatrix}, \quad (5.1)$$

$$W^a = \frac{1}{P_p P_s} \begin{pmatrix} 0 \\ P^1 P^3 \\ P^2 P^3 \\ -(P_p)^2 \end{pmatrix},$$

where

$$P_p = \sqrt{(P^1)^2 + (P^2)^2}, \quad (5.2)$$

$$P_s = \sqrt{(P^1)^2 + (P^2)^2 + (P^3)^2}.$$

The vectors v^μ and w^μ are real unit spacelike vectors that represent a linear polarization basis satisfying Eq. (C2). They are related to the circular polarization vectors m^α and \bar{m}^α by Eq. (C1). Using this particular choice of polarization vectors, the Berry connection and the Berry

curvature terms can be computed, and the modified ray equations can be integrated, either analytically or numerically.

A. Relativistic Hall effect and Wigner translations

The relativistic Hall effect [50] is a special relativistic effect that occurs when Lorentz transformations are applied to objects carrying angular momentum. In particular, consider a localized wave packet carrying intrinsic angular momentum and propagating in the z direction in Minkowski spacetime. If a Lorentz boost is applied in the x direction, then the location of the Lorentz-transformed energy density centroid is shifted in the y direction, depending on the orientation of the angular momentum. This shift corresponds to the Wigner translation [51, 69, 70].

The following example shows that an effect analogous to the Wigner translation discussed in Ref. [51] appears in the effective ray equations (4.29) and (4.30). We consider the Minkowski spacetime in Cartesian coordinates (t, x, y, z) , with

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2, \quad (5.3)$$

and we want to compare the effective rays obtained from Eqs. (4.29) and (4.30) with two different choices of observer. In the first case, we consider the standard orthonormal tetrad

$$e_0 = \partial_t, \quad e_1 = \partial_x, \quad e_2 = \partial_y, \quad e_3 = \partial_z, \quad (5.4)$$

where $(e_0)^\mu$ is our first choice of observer. With this orthonormal tetrad, the polarization vectors are defined as in Eq. (5.1), and the Berry curvature terms can be computed. The ray equations reduce to the geodesic equations

$$\dot{X}^\mu = P^\mu, \quad \dot{P}_\mu = 0. \quad (5.5)$$

In order to describe light rays traveling in the z direction, we impose initial conditions $X^\mu(0) = (0, 0, 0, 0)$ and $P_\mu(0) = (-1, 0, 0, 1)$, and we obtain

$$X^\mu(\tau) = (\tau, 0, 0, \tau), \quad (5.6)$$

$$P_\mu(\tau) = (-1, 0, 0, 1).$$

As a second case, we apply a time-dependent boost in the x direction to the standard orthonormal tetrad in Eq. (5.3). We obtain

$$e'_0 = \cosh t \partial_t - \sinh t \partial_x, \quad e'_2 = \partial_y$$

$$e'_1 = -\sinh t \partial_t + \cosh t \partial_x, \quad e'_3 = \partial_z, \quad (5.7)$$

where $(e'_0)^\mu$ is our second choice of observer. Note that $(e'_0)^\mu$ represents a family of observers boosted in the x direction, with the rapidity of the boost represented by the time coordinate t . The polarization vectors are chosen as

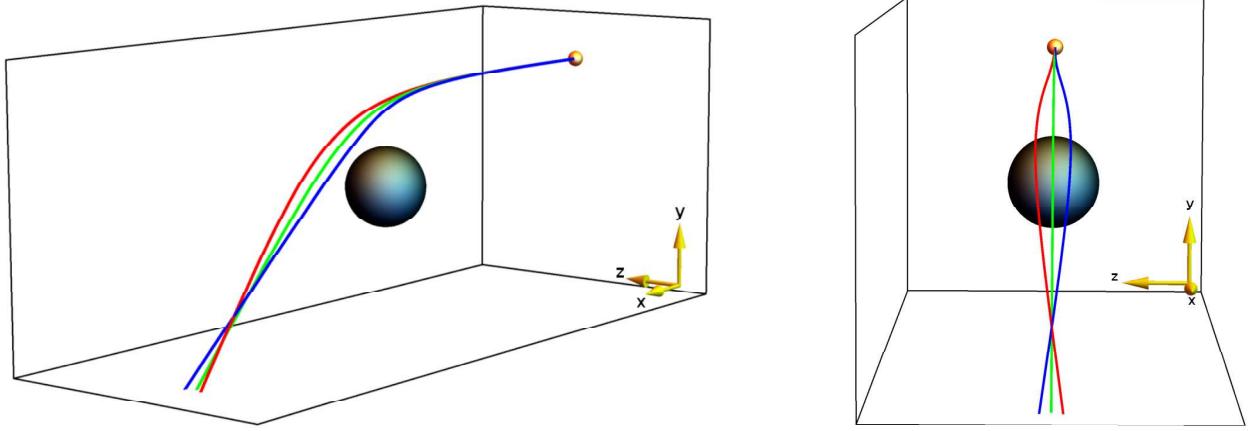


Figure 1: Results of numerical simulations illustrating the gravitational spin Hall effect of light around a Schwarzschild black hole. The effect is exaggerated for visualization purposes. The two figures present the same rays from different viewing angles. The central sphere represents the Schwarzschild black hole, and the small orange sphere represents a source of light. The blue and the red trajectories correspond to rays of opposite circular polarization, $s = \pm 1$, while the green trajectory represents a null geodesic. We take $r_s = 1$, and we start with the initial position $x^i(0) = (-50r_s, 15r_s, 0)$, and initial normalized momentum $p_i = (1, 0, 0)$. The wavelength λ is set to a sufficiently large value to make the effect visible on this plot.

in Eq. (5.1), but this time with respect to the orthonormal tetrad in Eq. (5.7). The Berry curvature terms in Eqs. (4.29) and (4.30) can be explicitly computed, and we obtain

$$\dot{X}^\mu = P^\mu + \epsilon s P^\nu (F_{px})_\nu^\mu, \quad (5.8)$$

$$\dot{P}_\mu = 0, \quad (5.9)$$

where

$$P^\nu (F_{px})_\nu^\mu = \frac{P_t}{[(e'_0)^\mu P_\mu]^2} \begin{pmatrix} 0 \\ 0 \\ P_z \\ -P_y \end{pmatrix}. \quad (5.10)$$

We impose the same initial conditions as in the previous case: $\dot{X}^\mu(0) = (0, 0, 0, 0)$ and $\dot{P}_\mu(0) = (-1, 0, 0, 1)$. Since the frequency is defined as $\omega = -(e'_0)^\mu P_\mu/\epsilon$, the small parameter ϵ can be identified with the wavelength of the initial light ray, as measured by the observer $(e'_0)^\mu$ at the spacetime point $x^\mu = X^\mu(0)$. Then, the ray equations can be analytically integrated, and we obtain

$$\begin{aligned} X^\mu(\tau) &= (\tau, 0, -s\epsilon \tanh \tau, \tau), \\ P_\mu(\tau) &= (-1, 0, 0, 1). \end{aligned} \quad (5.11)$$

Thus, given a circularly polarized light ray traveling in the z direction and two families of observers $(e_0)^\mu$ and $(e'_0)^\mu$, which are related by boosts in the x direction, we obtained the polarization-dependent Wigner translation in the y direction, $\Delta y = s\epsilon \tanh \tau$, in agreement with [51, Eq. 28]. Note that the Wigner translation is always smaller than one wavelength.

Recovering the results of Ref. [51] suggests that a worldline $X^\mu(\tau)$ representing a solution of Eqs. (4.29) and (4.30) could be interpreted as the location of the

energy density centroid of a localized wave packed with definite circular polarization, as measured by the chosen family of observers.

B. The gravitational spin Hall effect on a Schwarzschild background

To illustrate how the polarization-dependent correction terms modify the ray trajectories on a Schwarzschild background, let us provide some numerical examples. For convenience, we perform the numerical computations using canonical coordinates (x, p) and treat x^0 as a parameter along the rays. Hence, Eqs. (4.16) and (4.17) become

$$\dot{x}^0 = 1, \quad (5.12)$$

$$\dot{x}^i = \frac{g^{i\nu} p_\nu - \epsilon s \left(B^i + p^\alpha \nabla^i B_\alpha \right)}{g^{0\nu} p_\nu - \epsilon s \left(B^0 + p^\alpha \nabla^0 B_\alpha \right)}, \quad (5.13)$$

$$\dot{p}_i = \frac{-\frac{1}{2} \partial_i g^{\alpha\beta} p_\alpha p_\beta + \epsilon s p_\alpha (\partial_i g^{\alpha\beta} B_\beta + g^{\alpha\beta} \partial_i B_\beta)}{g^{0\nu} p_\nu - \epsilon s \left(B^0 + p^\alpha \nabla^0 B_\alpha \right)}, \quad (5.14)$$

and p_0 is calculated from

$$\frac{1}{2} g^{\mu\nu} p_\mu p_\nu - \epsilon s g^{\mu\nu} p_\mu B_\nu(x, p) = 0. \quad (5.15)$$

This equation can be solved explicitly, using the fact that the velocity \dot{x}^α is future oriented:

$$p_0 = \frac{1}{g^{00}} \left[- (g^{0i} p_i - \epsilon s g^{0\mu} B_\mu) + \sqrt{(g^{0i} p_i - \epsilon s g^{0\mu} B_\mu)^2 - g^{00} (g^{ij} p_i p_j - 2\epsilon s p_i g^{i\mu} B_\mu)} \right]. \quad (5.16)$$

Note that in general B_μ depends on p_0 . However, since this is an $\mathcal{O}(\epsilon^1)$ term, we can replace the $\mathcal{O}(\epsilon^0)$ expression for p_0 in B_μ .

In order to compare with the results of Gosselin *et al.* [41], we consider a Schwarzschild spacetime in Cartesian isotropic coordinates (t, x, y, z) :

$$ds^2 = - \left(\frac{1 - \frac{r_s}{4R}}{1 + \frac{r_s}{4R}} \right)^2 dt^2 + \left(1 + \frac{r_s}{4R} \right)^4 (dx^2 + dy^2 + dz^2), \quad (5.17)$$

where $r_s = 2GM/c^2$ is the Schwarzschild radius and $R = \sqrt{x^2 + y^2 + z^2}$. We also define the following orthonormal tetrad:

$$\begin{aligned} e_0 &= \frac{1 + \frac{r_s}{4R}}{1 - \frac{r_s}{4R}} \partial_t, & e_1 &= \left(1 + \frac{r_s}{4R} \right)^{-2} \partial_x, \\ e_2 &= \left(1 + \frac{r_s}{4R} \right)^{-2} \partial_y, & e_3 &= \left(1 + \frac{r_s}{4R} \right)^{-2} \partial_z, \end{aligned} \quad (5.18)$$

where $t^\mu = (e_0)^\mu$ is our choice of observer.

The Berry connection B_μ can be explicitly computed by introducing a particular choice of polarization vectors, using Eq. (5.1) and the orthonormal tetrad (5.18). We now have all the elements required for the numerical integration of Eqs. (5.12)-(5.14). For this purpose, we used the NDSOLVE function of *Mathematica* [71]. For these examples, we used the default settings for integration method, precision and accuracy.

After obtaining a numerical solution $(x(t), p(t))$ to Eqs. (5.12)-(5.14), in order to ensure the gauge invariance of our results, we have to evaluate the gauge-invariant noncanonical quantities $(X(t), P(t))$, as given in Eqs. (4.22) and (4.23). These are the quantities used to represent the trajectories in Figs. 1 and 2. A comparative discussion between the use of canonical and non-canonical ray equations in optics, together with numerical examples, can be found in Ref. [47].

As the first step, we numerically compare our ray Eqs. (5.12)-(5.14) with those predicted by Gosselin *et al.* [41]. This is done by numerically integrating Eqs. (5.12)-(5.14), as well as Eq. (23) from Ref. [41]. Up to numerical errors, we obtain the same ray trajectories with both sets of equations. However, while the equations obtained by Gosselin *et al.* only apply to static spacetimes, Eqs. (5.12)-(5.14) do not have this limitation.

The results of our numerical simulations are shown in Fig. 1, which illustrates the general behavior of the gravitational spin Hall effect of light around a Schwarzschild

black hole. [The actual effect is small, so the figure is obtained by numerical integration of Eqs. (5.12)-(5.14) for unrealistic parameters.] Here, we consider rays of opposite circular polarization ($s = \pm 1$) passing close to a Schwarzschild black hole, together with a reference null geodesic ($s = 0$). Except for the value of s , we are considering the same initial conditions, $(x^i(0), p_i(0))$, for these rays. Unlike the null geodesic, for which the motion is planar, the circularly polarized rays are not confined to a plane.

As another example, we used initial conditions $(x^i(0), p_i(0))$ such that the rays are initialized as radially ingoing or outgoing. In this case (not illustrated, since it is trivial), the gravitational spin Hall effect vanishes, and the circularly polarized rays coincide with the radial null geodesic.

Using these numerical methods, we can also estimate the magnitude of the gravitational spin Hall effect. As a particular example, we consider a similar situation to the one presented in Fig. 1, where the black hole is replaced with the Sun. More precisely, we model this situation by considering a Schwarzschild black hole with $r_s \approx 3$ km. We consider the deflection of circularly polarized rays coming from a light source far away, passing close to the surface of the Sun, and then observed on the Earth. This situation is illustrated in Fig. 2. The numerical results are based on the initial data presented in the caption of Fig. 2. When reaching the Earth, the separation distance between the rays of opposite circular polarization depends on the wavelength. For example, taking wavelengths of the order $\lambda \approx 10^{-9}$ m results in a separation distance of the order $d \approx 10^{-15}$ m, while for wavelengths of the order of $\lambda \approx 1$ m we obtain a separation distance of the order $d \approx 10^{-6}$ m. Although the ray separation is small (about six orders of magnitude smaller than the wavelength), what really matters is that the rays are scattered by a finite angle. Therefore, the ray separation grows linearly with distance after the reintersection point. This means that the effect should be robustly observable if one measures it sufficiently far from the Sun. Furthermore, massive compact astronomical objects, such as black holes or neutron stars, are expected to produce a larger gravitational spin Hall effect.

As a consistency check, we also performed the numerical computations using different coordinates, such as the standard Schwarzschild spherical coordinates and Gullstrand–Painlevé coordinates. The results are independent of the choice of coordinates. However, the polarized rays are not invariant under a change of observer. This is due to an effect analogous to the Wigner translations discussed in Sec. V A. For example, instead of the static observer introduced in Eq. (5.18), one could consider a free-falling observer. In this case, the ray trajectories presented in Figs. 1 and 2 are slightly modified, due to the Wigner translations, and preliminary investigations indicate that these modifications are smaller than one wavelength, as in the case discussed in Sec. V A. It is not

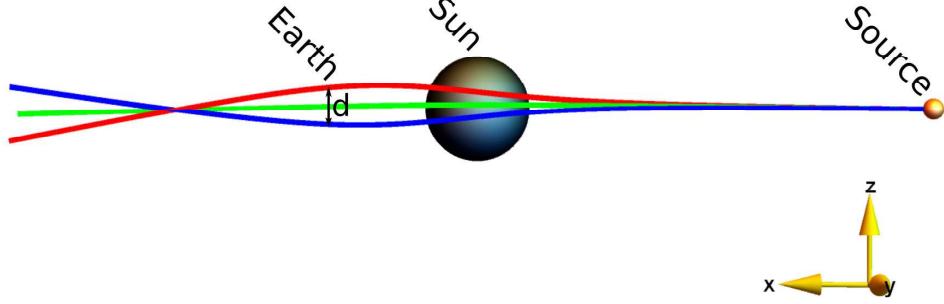


Figure 2: Results of numerical simulations illustrating the gravitational spin Hall effect of light around the Sun. The effect is exaggerated for visualization purposes. The separation distance d is observed from the Earth. The blue and the red trajectories correspond to rays of opposite circular polarization, $s = \pm 1$, while the green trajectory represents a null geodesic. We take $r_s = 3$ km, and we start with the initial position $x^i(0) = (-10^7 r_s, 3 \times 10^5 r_s, 0)$, and initial normalized momentum $p_i = (1, 0, 0)$.

clear how to separate the purely gravitational effect from the observer-dependent Wigner translations. However, this is not a problem. The Wigner translation represents the observer-dependent ambiguity in defining the location of the ray on a single-wavelength scale and remains bounded. In contrast, the purely gravitational effect can affect the angle of light scattering off a gravitating object and thus the ray displacement associated with this effect accumulates linearly with the distance. This means that the latter effect dominates at large distances. A more detailed analysis of the modified ray equations, at both the analytical and the numerical levels, will be carried out in future work.

VI. CONCLUSIONS

In summary, we have presented a first comprehensive theory of the gravitational spin Hall effect that occurs due to the coupling of the polarization with the translational dynamics of the light rays. The ray dynamics is governed by the corrected Hamiltonian

$$H(x, p) = \frac{1}{2} g^{\mu\nu} p_\mu p_\nu - \epsilon s g^{\mu\nu} p_\mu B_\nu(x, p). \quad (6.1)$$

Here, the first term represents the geometrical optics Hamiltonian, and the second term represents a correction of $\mathcal{O}(\epsilon^1)$ that is due to the Berry connection, which is given by

$$\begin{aligned} B_\mu(x, p) &= i\bar{m}^\alpha \left(\frac{\partial}{\partial x^\mu} m_\alpha - \Gamma_{\alpha\mu}^\sigma m_\sigma + \Gamma_{\mu\rho}^\sigma p_\sigma \frac{\partial}{\partial p_\rho} m_\alpha \right) \\ &= i\bar{m}^\alpha \nabla_\mu^h m_\alpha. \end{aligned} \quad (6.2)$$

Assuming the noncanonical coordinates (4.22), the corresponding ray equations are

$$\begin{aligned} \dot{X}^\mu &= P^\mu + \epsilon s P^\nu (F_{px})_\nu^\mu \\ &\quad + \frac{2i\epsilon s}{(t^\alpha P_\alpha)^2} \Gamma_{\beta\nu}^\alpha P_\alpha P^\beta m^{[\nu} \bar{m}^{\mu]}, \end{aligned} \quad (6.3)$$

$$\begin{aligned} \dot{P}_\mu &= \Gamma_{\beta\mu}^\alpha P_\alpha P^\beta + \epsilon s P^\nu \left[iR_{\alpha\beta\mu\nu} m^\alpha \bar{m}^\beta + (\tilde{F}_{xx})_{\nu\mu} \right] \\ &\quad + \epsilon s \Gamma_{\beta\nu}^\alpha P_\alpha P^\beta (F_{px})_\mu^\nu, \end{aligned} \quad (6.4)$$

where the terms F_{px} and \tilde{F}_{xx} and the timelike vector t^α are given in Appendix C. The last term on the right-hand side of Eq. (6.3) is the covariant analogue of the spin Hall correction term usually encountered in optics [20, 41], while the Riemann curvature term in Eq. (6.4) is reminiscent of a similar term appearing in the Mathisson-Papapetrou-Dixon equations [27]. In Minkowski spacetime, the F_{px} term is responsible for the relativistic Hall effect [50] and Wigner translations [51].

The resulting deviation of the ray trajectories from those predicted by geometrical optics is weak but not unobservable. First of all, even small angular deviations are observable at large enough distances. Second, as shown in Ref. [19], weak quantum measurement techniques can be used to detect the spin Hall effect of light, even when the spatial separation between the left-polarized and the right-polarized beams of light is smaller than the wavelength.

Potentially, this work can be naturally extended in two directions. First, the corrected ray equations are yet to be studied more thoroughly, both analytically and numerically. Rigorous numerical investigations are needed to obtain a precise prediction of the effect, in particular for Kerr black holes. Second, Maxwell's equations are a proxy to linearized gravity. It is expected that a similar approach can be carried out to obtain an effective pointwise description of a gravitational wave packet, extending the results of Ref. [46].

As discussed in Ref. [4], the spin Hall effect of light is directly related to the conservation of total angular momentum. For the discussion presented so far, the considered rays carry extrinsic orbital angular momentum, associated with the ray trajectory, and intrinsic spin angular momentum, associated with the polarization. However, it is well-known that light can also carry intrinsic orbital angular momentum [72–74] (see also Ref. [75] and references therein). In principle, the magnitude of the spin Hall effect can be increased by considering optical beams carrying intrinsic orbital angular momentum [76]. The method and ansatz that we have adopted are insufficient to describe this effect. A more realistic and more precise approach involving wave packets, such as Laguerre-Gaussian beams, should be considered. It may be possible to do so using the machinery developed in Ref. [58].

A formulation of the special-relativistic dynamics of massless spinning particles and wave packets beyond the geometrical optics limit has been previously reported by Duval and collaborators (cf. Ref. [77] for the spin-1/2 case; see also Ref. [78]). This analysis relates the modified dynamics to the approach of Souriau [79], making use of so-called spin enslaving. This has been extended to general helicity by Andrzejewski *et al.* [80]. We expect that the Hamiltonian formulation presented here corresponds to a general relativistic version of the models considered in the mentioned papers. This will be considered in a future work.

Acknowledgements

We are grateful to Pedro Cunha for helpful discussions. A significant part of this work was done while one of the authors, L.A., was in residence at Institut Mittag-Leffler in Djursholm, Sweden during the fall of 2019, supported by the Swedish Research Council under Grant no. 2016-06596. I.Y.D. acknowledges support from the U.S. National Science Foundation under Grant No. PHY 1903130. M.A.O. is supported by the International Max Planck Research School for Mathematical and Physical Aspects of Gravitation, Cosmology and Quantum Field Theory. C.F.P. was partially supported by the Australian Research Council Grant no. DP170100630 and partially funded by the SNSF Grant no. P2SKP2 178198.

Sandia National Laboratories is a multimission laboratory managed and operated by the National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy (DOE) National Nuclear Security Administration under Contract No. DE-NA0003525. This paper describes the objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. DOE or the United States Government.

Appendix A: HORIZONTAL AND VERTICAL DERIVATIVES ON T^*M

Let (x^μ, p_μ) be canonical coordinates on T^*M . Considering fields defined on T^*M , such as $u_\alpha(x, p)$ and

$v^\alpha(x, p)$, the horizontal and vertical derivatives are defined as follows [81, Sec. 3.5]:

$$\overset{v}{\nabla}{}^\mu u_\alpha = \frac{\partial}{\partial p_\mu} u_\alpha, \quad (\text{A1a})$$

$$\overset{h}{\nabla}{}_\mu u_\alpha = \frac{\partial}{\partial x^\mu} u_\alpha - \Gamma_{\alpha\mu}^\sigma u_\sigma + \Gamma_{\mu\rho}^\sigma p_\sigma \frac{\partial}{\partial p_\rho} u_\alpha, \quad (\text{A1b})$$

$$\overset{v}{\nabla}{}^\mu v^\alpha = \frac{\partial}{\partial p_\mu} v^\alpha, \quad (\text{A2a})$$

$$\overset{h}{\nabla}{}_\mu v^\alpha = \frac{\partial}{\partial x^\mu} v^\alpha + \Gamma_{\sigma\mu}^\alpha v^\sigma + \Gamma_{\mu\rho}^\sigma p_\sigma \frac{\partial}{\partial p_\rho} v^\alpha. \quad (\text{A2b})$$

The extension to general tensor fields on T^*M is straightforward. Note that in contrast to Ref. [81, Sec. 3.5], we have the opposite sign for the last term in the definition of the horizontal derivative. This is because our fields, $u_\alpha(x, p)$ and $v^\alpha(x, p)$, are defined on T^*M , and not on TM , as is the case in the reference mentioned before. We can make use of the following properties:

$$\begin{aligned} \overset{h}{[\nabla}_\mu, \overset{v}{\nabla}{}^\nu] &= 0, & \overset{v}{[\nabla}{}^\mu, \overset{v}{\nabla}{}^\nu] &= 0, \\ \overset{h}{\nabla}_\mu p_\alpha &= \overset{h}{\nabla}_\mu g_{\alpha\beta} = \overset{v}{\nabla}{}^\mu g_{\alpha\beta} = 0. \end{aligned} \quad (\text{A3})$$

Appendix B: VARIATION OF THE ACTION

Here, we derive the Euler-Lagrange equations that correspond to the action

$$J = \int_M d^4x \sqrt{g} \mathcal{L}, \quad (\text{B1})$$

where the Lagrangian density is of the following form:

$$\begin{aligned} \mathcal{L} = \mathcal{L} \Big(S(x), \nabla_\mu S(x), \\ A_\alpha[x, \nabla S(x)], \nabla_\mu \{A_\alpha[x, \nabla S(x)]\}, \\ A^{*\alpha}[x, \nabla S(x)], \nabla_\mu \{A^{*\alpha}[x, \nabla S(x)]\} \Big). \end{aligned} \quad (\text{B2})$$

Here, $S(x)$ is an independent field, while A_α and $A^{*\alpha}$ cannot be considered independent, since they depend on $\nabla_\mu S$. Following Hawking and Ellis [52, p. 65], we define the variation of a field Ψ_i as a one-parameter family of fields $\Psi_i(u, x)$, with $u \in (-\varepsilon, \varepsilon)$ and $x \in M$. We use the following notation:

$$\left. \frac{\partial \Psi_i(u, x)}{\partial u} \right|_{u=0} = \Delta \Psi_i. \quad (\text{B3})$$

Note that the derivative with respect to the parameter u commutes with the covariant derivative, so we have:

$$\frac{d}{du} \nabla_\mu S(u, x) = \nabla_\mu \left(\frac{\partial S}{\partial u} \right), \quad (\text{B4})$$

$$\frac{d}{du} A_\alpha(u, x, \nabla S(u, x)) = \frac{\partial A_\alpha}{\partial u} + \frac{\partial A_\alpha}{\partial \nabla_\nu S} \nabla_\nu \left(\frac{\partial S}{\partial u} \right), \quad (\text{B5})$$

$$\frac{d}{du} \nabla_\mu [A_\alpha(u, x, \nabla S(u, x))] = \nabla_\mu \left[\frac{d}{du} A_\alpha(u, x, \nabla S(u, x)) \right] = \nabla_\mu \left[\frac{\partial A_\alpha}{\partial u} + \frac{\partial A_\alpha}{\partial \nabla_\nu S} \nabla_\nu \left(\frac{\partial S}{\partial u} \right) \right]. \quad (\text{B6})$$

We consider the variation of the action, taking special care when applying the chain rule:

$$0 = \frac{dJ}{du} \Big|_{u=0} = \int_M d^4x \sqrt{g} \left\{ \frac{\partial \mathcal{L}}{\partial S} \Delta S + \frac{\partial \mathcal{L}}{\partial \nabla_\mu S} \Delta (\nabla_\mu S) + \frac{\partial \mathcal{L}}{\partial A_\alpha} \left[\Delta A_\alpha + \frac{\partial A_\alpha}{\partial \nabla_\mu S} \nabla_\mu (\Delta S) \right] + \frac{\partial \mathcal{L}}{\partial \nabla_\mu A_\alpha} \nabla_\mu \left[\Delta A_\alpha + \frac{\partial A_\alpha}{\partial \nabla_\nu S} \nabla_\nu (\Delta S) \right] + \frac{\partial \mathcal{L}}{\partial A^{*\alpha}} \left[\Delta A^{*\alpha} + \frac{\partial A^{*\alpha}}{\partial \nabla_\mu S} \nabla_\mu (\Delta S) \right] + \frac{\partial \mathcal{L}}{\partial \nabla_\mu A^{*\alpha}} \nabla_\mu \left[\Delta A^{*\alpha} + \frac{\partial A^{*\alpha}}{\partial \nabla_\nu S} \nabla_\nu (\Delta S) \right] \right\}. \quad (\text{B7})$$

Integrating by parts and assuming the boundary terms vanish, we obtain:

$$0 = \frac{dJ}{du} \Big|_{u=0} = \int_M d^4x \sqrt{g} \left\{ \left(\frac{\partial \mathcal{L}}{\partial A_\alpha} - \nabla_\mu \frac{\partial \mathcal{L}}{\partial \nabla_\mu A_\alpha} \right) \Delta A_\alpha + \left(\frac{\partial \mathcal{L}}{\partial A^{*\alpha}} - \nabla_\mu \frac{\partial \mathcal{L}}{\partial \nabla_\mu A^{*\alpha}} \right) \Delta A^{*\alpha} + \frac{\partial \mathcal{L}}{\partial S} \Delta S - \nabla_\mu \left[\frac{\partial \mathcal{L}}{\partial \nabla_\mu S} + \frac{\partial A_\alpha}{\partial \nabla_\mu S} \left(\frac{\partial \mathcal{L}}{\partial A_\alpha} - \nabla_\nu \frac{\partial \mathcal{L}}{\partial \nabla_\nu A_\alpha} \right) + \frac{\partial A^{*\alpha}}{\partial \nabla_\mu S} \left(\frac{\partial \mathcal{L}}{\partial A^{*\alpha}} - \nabla_\nu \frac{\partial \mathcal{L}}{\partial \nabla_\nu A^{*\alpha}} \right) \right] \Delta S \right\}. \quad (\text{B8})$$

Since the above equation must be satisfied for all variations ΔS , ΔA_α , and $\Delta A^{*\alpha}$, we obtain the following Euler-Lagrange equations:

$$\frac{\partial \mathcal{L}}{\partial A^{*\alpha}} - \nabla_\mu \frac{\partial \mathcal{L}}{\partial \nabla_\mu A^{*\alpha}} = \mathcal{O}(\epsilon^2), \quad (\text{B9})$$

$$\frac{\partial \mathcal{L}}{\partial A_\alpha} - \nabla_\mu \frac{\partial \mathcal{L}}{\partial \nabla_\mu A_\alpha} = \mathcal{O}(\epsilon^2), \quad (\text{B10})$$

$$\frac{\partial \mathcal{L}}{\partial S} - \nabla_\mu \left[\frac{\partial \mathcal{L}}{\partial \nabla_\mu S} + \frac{\partial A_\alpha}{\partial \nabla_\mu S} \left(\frac{\partial \mathcal{L}}{\partial A_\alpha} - \nabla_\nu \frac{\partial \mathcal{L}}{\partial \nabla_\nu A_\alpha} \right) + \frac{\partial A^{*\alpha}}{\partial \nabla_\mu S} \left(\frac{\partial \mathcal{L}}{\partial A^{*\alpha}} - \nabla_\nu \frac{\partial \mathcal{L}}{\partial \nabla_\nu A^{*\alpha}} \right) \right] = \mathcal{O}(\epsilon^2). \quad (\text{B11})$$

Furthermore, Eq. (B11) can be simplified by using Eqs. (B9) and (B10). Thus, as a final result, we have the following set of Euler-Lagrange equations:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial A^{*\alpha}} - \nabla_\mu \frac{\partial \mathcal{L}}{\partial \nabla_\mu A^{*\alpha}} &= \mathcal{O}(\epsilon^2), \\ \frac{\partial \mathcal{L}}{\partial A_\alpha} - \nabla_\mu \frac{\partial \mathcal{L}}{\partial \nabla_\mu A_\alpha} &= \mathcal{O}(\epsilon^2), \\ \frac{\partial \mathcal{L}}{\partial S} - \nabla_\mu \frac{\partial \mathcal{L}}{\partial \nabla_\mu S} &= \mathcal{O}(\epsilon^2). \end{aligned} \quad (\text{B12})$$

Appendix C: BERRY CURVATURE

In order to calculate the Berry curvature terms (4.28), it is enough to use a tetrad $\{t^\alpha, p^\alpha, v^\alpha, w^\alpha\}$, where t^α is a

future-oriented timelike vector field representing a family of observers and p^α is a generic vector, not necessarily null, representing the momentum of a point particle (ray). The vectors v^α and w^α are real spacelike vectors related to m^α and \bar{m}^α by the following relations:

$$m^\alpha = \frac{1}{\sqrt{2}} (v^\alpha + iw^\alpha), \quad \bar{m}^\alpha = \frac{1}{\sqrt{2}} (v^\alpha - iw^\alpha). \quad (\text{C1})$$

The elements of the tetrad $\{t^\alpha, p^\alpha, v^\alpha, w^\alpha\}$ satisfy the following relations:

$$\begin{aligned} t_\alpha t^\alpha &= -1, & p_\alpha p^\alpha &= \kappa, & t_\alpha p^\alpha &= -\epsilon\omega, \\ v_\alpha v^\alpha &= w_\alpha w^\alpha = 1, & & & & \\ t_\alpha v^\alpha &= t_\alpha w^\alpha = p_\alpha v^\alpha = p_\alpha w^\alpha = v_\alpha w^\alpha = 0. & & & & \end{aligned} \quad (\text{C2})$$

Note that the vectors v^α and w^α depend of p^μ through the orthogonality condition, while t^α is independent of

p^μ . We start by computing the vertical derivatives of the vectors v^α and w^α . Using the tetrad, we can write:

$$\overset{v}{\nabla}{}^\mu v^\alpha = \frac{\partial v^\alpha}{\partial p_\mu} = c_1{}^\mu t^\alpha + c_2{}^\mu p^\alpha + c_3{}^\mu v^\alpha + c_4{}^\mu w^\alpha, \quad (\text{C3})$$

$$\overset{v}{\nabla}{}^\mu w^\alpha = \frac{\partial w^\alpha}{\partial p_\mu} = d_1{}^\mu t^\alpha + d_2{}^\mu p^\alpha + d_3{}^\mu v^\alpha + d_4{}^\mu w^\alpha, \quad (\text{C4})$$

where $c_i{}^\mu$ and $d_i{}^\mu$ are unknown vector fields that need to be determined. Using the properties from Eq. (C2), we obtain

$$\begin{aligned} \overset{v}{\nabla}{}^\mu v^\alpha &= \frac{\epsilon\omega}{\epsilon^2\omega^2 + \kappa} v^\mu t^\alpha - \frac{1}{\epsilon^2\omega^2 + \kappa} v^\mu p^\alpha + c_4{}^\mu w^\alpha, \\ \overset{v}{\nabla}{}^\mu w^\alpha &= \frac{\epsilon\omega}{\epsilon^2\omega^2 + \kappa} w^\mu t^\alpha - \frac{1}{\epsilon^2\omega^2 + \kappa} w^\mu p^\alpha + d_3{}^\mu v^\alpha. \end{aligned} \quad (\text{C5})$$

Applying the same arguments to the terms $\nabla_\mu v_\alpha$ and $\nabla_\mu w_\alpha$, we also obtain

$$\begin{aligned} \nabla_\mu v_\alpha &= -\frac{1}{\epsilon^2\omega^2 + \kappa} (\epsilon\omega p_\sigma \nabla_\mu v^\sigma + \kappa t_\sigma \nabla_\mu v^\sigma) t_\alpha \\ &\quad + \frac{1}{\epsilon^2\omega^2 + \kappa} (p_\sigma \nabla_\mu v^\sigma - \epsilon\omega t_\sigma \nabla_\mu v^\sigma) p_\alpha + f_{4\mu} w_\alpha, \\ \nabla_\mu w_\alpha &= -\frac{1}{\epsilon^2\omega^2 + \kappa} (\epsilon\omega p_\sigma \nabla_\mu w^\sigma + \kappa t_\sigma \nabla_\mu w^\sigma) t_\alpha \\ &\quad + \frac{1}{\epsilon^2\omega^2 + \kappa} (p_\sigma \nabla_\mu w^\sigma - \epsilon\omega t_\sigma \nabla_\mu w^\sigma) p_\alpha + g_{3\mu} v_\alpha. \end{aligned} \quad (\text{C6})$$

Note that the fields $c_{4\mu}$, $d_{3\mu}$, $f_{4\mu}$, and $g_{3\mu}$ are undetermined within this approach, but this is not a problem, because they do not affect the Berry curvature.

1. \mathbf{F}_{pp}

We compute $(F_{pp})^{\nu\mu}$ by using Eq. (C5) and setting $\kappa = 0$. Since vertical derivatives commute (see Eq. (A3)), we can write

$$\begin{aligned} (F_{pp})^{\nu\mu} &= i \left(\overset{v}{\nabla}{}^\mu \bar{m}^\alpha \overset{v}{\nabla}{}^\nu m_\alpha - \overset{v}{\nabla}{}^\nu \bar{m}^\alpha \overset{v}{\nabla}{}^\mu m_\alpha \right) \\ &= \overset{v}{\nabla}{}^\nu v^\alpha \overset{v}{\nabla}{}^\mu w_\alpha - \overset{v}{\nabla}{}^\mu v^\alpha \overset{v}{\nabla}{}^\nu w_\alpha \\ &= \frac{2}{\epsilon^2\omega^2} v^{[\nu} w^{\mu]} \\ &= \frac{2i}{\epsilon^2\omega^2} m^{[\nu} \bar{m}^{\mu]}. \end{aligned} \quad (\text{C7})$$

2. \mathbf{F}_{xx}

We have

$$\begin{aligned} (F_{xx})_{\nu\mu} &= i \left(\nabla_\mu \bar{m}^\alpha \nabla_\nu m_\alpha - \nabla_\nu \bar{m}^\alpha \nabla_\mu m_\alpha \right. \\ &\quad \left. + \bar{m}^\alpha \nabla_{[\mu} \nabla_{\nu]} m_\alpha - m_\alpha \nabla_{[\mu} \nabla_{\nu]} \bar{m}^\alpha \right). \end{aligned} \quad (\text{C8})$$

The last two terms can be expressed in terms of the Riemann tensor:

$$i \left(\bar{m}^\alpha \nabla_{[\mu} \nabla_{\nu]} m_\alpha - m_\alpha \nabla_{[\mu} \nabla_{\nu]} \bar{m}^\alpha \right) = -i R_{\alpha\beta\mu\nu} m^\alpha \bar{m}^\beta. \quad (\text{C9})$$

The first two terms can be computed using Eq. (C6) and $\kappa = 0$:

$$\begin{aligned} (\tilde{F}_{xx})_{\nu\mu} &= i \left(\nabla_\mu \bar{m}^\alpha \nabla_\nu m_\alpha - \nabla_\nu \bar{m}^\alpha \nabla_\mu m_\alpha \right) \\ &= \nabla_\nu v^\alpha \nabla_\mu w_\alpha - \nabla_\mu v^\alpha \nabla_\nu w_\alpha \\ &= \frac{1}{\epsilon^2\omega^2} \left(p_\sigma \nabla_\mu v^\sigma p_\rho \nabla_\nu w^\rho - p_\sigma \nabla_\nu v^\sigma p_\rho \nabla_\mu w^\rho \right. \\ &\quad - \epsilon\omega p_\sigma \nabla_\mu v^\sigma t_\rho \nabla_\nu w^\rho + \epsilon\omega p_\sigma \nabla_\nu v^\sigma t_\rho \nabla_\mu w^\rho \\ &\quad \left. - \epsilon\omega t_\sigma \nabla_\mu v^\sigma p_\rho \nabla_\nu w^\rho + \epsilon\omega t_\sigma \nabla_\nu v^\sigma p_\rho \nabla_\mu w^\rho \right) \\ &= \frac{1}{\epsilon^2\omega^2} \left(p_\sigma \nabla_\mu m^\sigma p_\rho \nabla_\nu \bar{m}^\rho - p_\sigma \nabla_\nu m^\sigma p_\rho \nabla_\mu \bar{m}^\rho \right. \\ &\quad - \epsilon\omega p_\sigma \nabla_\mu m^\sigma t_\rho \nabla_\nu \bar{m}^\rho + \epsilon\omega p_\sigma \nabla_\nu m^\sigma t_\rho \nabla_\mu \bar{m}^\rho \\ &\quad \left. - \epsilon\omega t_\sigma \nabla_\mu m^\sigma p_\rho \nabla_\nu \bar{m}^\rho + \epsilon\omega t_\sigma \nabla_\nu m^\sigma p_\rho \nabla_\mu \bar{m}^\rho \right). \end{aligned} \quad (\text{C10})$$

3. \mathbf{F}_{px} and \mathbf{F}_{xp}

Since $(F_{px})_\nu{}^\mu = -(F_{xp})_\nu{}^\mu$, it is enough to compute only one term. Using Eqs. (C5) and (C6), and setting $\kappa = 0$, we obtain

$$\begin{aligned} (F_{px})_\nu{}^\mu &= i \left(\overset{v}{\nabla}{}^\mu \bar{m}^\alpha \nabla_\nu m_\alpha - \nabla_\nu \bar{m}^\alpha \overset{v}{\nabla}{}^\mu m_\alpha \right) \\ &= \nabla_\nu v^\alpha \overset{v}{\nabla}{}^\mu w_\alpha - \overset{v}{\nabla}{}^\mu v^\alpha \nabla_\nu w_\alpha \\ &= \frac{1}{\epsilon^2\omega^2} \left[(p_\sigma \nabla_\nu w^\sigma - \epsilon\omega t_\sigma \nabla_\nu w^\sigma) v^\mu \right. \\ &\quad \left. - (p_\sigma \nabla_\nu v^\sigma - \epsilon\omega t_\sigma \nabla_\nu v^\sigma) w^\mu \right] \\ &= \frac{i}{\epsilon^2\omega^2} \left[(p_\sigma \nabla_\nu \bar{m}^\sigma - \epsilon\omega t_\sigma \nabla_\nu \bar{m}^\sigma) m^\mu \right. \\ &\quad \left. - (p_\sigma \nabla_\nu m^\sigma - \epsilon\omega t_\sigma \nabla_\nu m^\sigma) \bar{m}^\mu \right]. \end{aligned} \quad (\text{C11})$$

Appendix D: COORDINATE TRANSFORMATION

The substitution from Eqs. (4.22) and (4.23) can be obtained, up to terms of order ϵ^2 , as a linearization of the following composition of changes of coordinates on the cotangent bundle T^*M . Consider the family of diffeomorphisms (Φ_ϵ) generated by the vector field on M

$$Y = is \bar{m}^\alpha \overset{v}{\nabla}{}^\mu m_\alpha \partial_{x^\mu}, \quad (\text{D1})$$

that is to say

$$\frac{d}{d\epsilon} \Phi_\epsilon(x) = Y(\Phi_\epsilon(x)) \text{ with } \Phi_0(x) = x. \quad (\text{D2})$$

By construction, the Taylor expansion in a coordinate chart of Φ_ϵ at order ϵ^1 leads to Eq. (4.22). Φ_ϵ naturally lifts to the cotangent bundle using the pullback Φ_ϵ^* :

$$\Phi_\epsilon^* : (x, p) \mapsto (\Phi_\epsilon(x), p \circ d\Phi_\epsilon^{-1}|_{\Phi_\epsilon(x)}). \quad (\text{D3})$$

Note that the choice of the lift is not unique. The mapping Φ_ϵ^* is, at order one in ϵ , in coordinates,

$$(x^\mu, p_\mu) \mapsto (x^\mu + i\epsilon \bar{m}^\alpha \nabla^\mu m_\alpha, p_\mu - i\epsilon s p_\beta \partial_{x^\mu} (\bar{m}^\alpha \nabla^\beta m_\alpha)). \quad (\text{D4})$$

Consider next the translation of the momentum variable defined by

$$\Psi_\epsilon : (x, p) \mapsto (x, p - \epsilon\sigma), \quad (\text{D5})$$

where $\sigma = i s (\bar{m}^\alpha \nabla_\mu m_\alpha + p_\beta \partial_{x^\mu} (\bar{m}^\alpha \nabla^\beta m_\alpha)) dx^\mu$. The linearization in ϵ of the diffeomorphism $\Psi_\epsilon \circ \Phi_\epsilon^*$ provides by construction the change of variables in Eqs. (4.22) and (4.23).

[1] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (W. H. Freeman San Francisco, 1973) pp. xxvi, 1279 p. .

[2] P. Schneider, J. Ehlers, and E. E. Falco, *Gravitational Lenses* (Springer, 1992).

[3] M. A. Oancea, C. F. Paganini, J. Joudoux, and L. Andersson, An overview of the gravitational spin Hall effect, arXiv:1904.09963 (2019), arXiv:1904.09963 .

[4] K. Y. Bliokh, F. J. Rodríguez-Fortuño, F. Nori, and A. V. Zayats, Spin-orbit interactions of light, *Nat. Photonics* **9**, 796 (2015).

[5] M. I. Dyakonov and A. V. Khaetskii, Spin Hall Effect, in *Spin Physics in Semiconductors*, edited by M. I. Dyakonov (Springer Berlin Heidelberg, 2008) pp. 211–243.

[6] J. Sinova, S. O. Valenzuela, J. Wunderlich, C. H. Back, and T. Jungwirth, Spin Hall effects, *Rev. Mod. Phys.* **87**, 1213 (2015).

[7] M. I. Dyakonov and V. I. Perel, Possibility of orienting electron spins with current, *JETP Lett.* **13**, 467 (1971).

[8] M. I. Dyakonov and V. I. Perel, Current-induced spin orientation of electrons in semiconductors, *Phys. Lett. A* **35**, 459 (1971).

[9] A. A. Bakun, B. P. Zakharchenya, A. A. Rogachev, M. N. Tkachuk, and V. G. Fleisher, Observation of a surface photocurrent caused by optical orientation of electrons in a semiconductor, *JETP Lett.* **40**, 1293 (1984).

[10] Y. K. Kato, R. C. Myers, A. C. Gossard, and D. D. Awschalom, Observation of the spin Hall effect in semiconductors, *Science* **306**, 1910 (2004).

[11] X. Ling, X. Zhou, K. Huang, Y. Liu, C.-W. Qiu, H. Luo, and S. Wen, Recent advances in the spin Hall effect of light, *Rep. Prog. Phys.* **80**, 066401 (2017).

[12] A. V. Dooghin, N. D. Kundikova, V. S. Liberman, and B. Y. Zel'dovich, Optical Magnus effect, *Phys. Rev. A* **45**, 8204 (1992).

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

[14] M. Onoda, S. Murakami, and N. Nagaosa, Hall Effect of Light, *Phys. Rev. Lett.* **93**, 083901 (2004).

[15] K. Y. Bliokh and Y. P. Bliokh, Modified geometrical optics of a smoothly inhomogeneous isotropic medium: The anisotropy, Berry phase, and the optical Magnus effect, *Phys. Rev. E* **70**, 026605 (2004).

[16] K. Y. Bliokh and Y. P. Bliokh, Topological spin transport of photons: The optical Magnus effect and Berry phase, *Phys. Lett. A* **333**, 181 (2004).

[17] C. Duval, Z. Horváth, and P. A. Horváthy, Fermat principle for spinning light, *Phys. Rev. D* **74**, 021701(R) (2006).

[18] C. Duval, Z. Horváth, and P. A. Horváthy, Geometrical spinoptics and the optical Hall effect, *J. Geom. Phys.* **57**, 925 (2007).

[19] O. Hosten and P. Kwiat, Observation of the spin Hall effect of light via weak measurements, *Science* **319**, 787 (2008).

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

[21] G. Sundaram and Q. Niu, Wave-packet dynamics in slowly perturbed crystals: Gradient corrections and Berry-phase effects, *Phys. Rev. B* **59**, 14915 (1999).

[22] A. Bérard and H. Mohrbach, Spin Hall effect and Berry phase of spinning particles, *Phys. Lett. A* **352**, 190 (2006).

[23] M. Mathisson, Republication of: New mechanics of material systems, *Gen. Relativ. Gravit.* **42**, 1011 (2010).

[24] A. Papapetrou, Spinning test-particles in general relativity. I, *Proc. R. Soc. A* **209**, 248 (1951).

[25] W. Tulczyjew, Motion of multipole particles in general relativity theory, *Acta Phys. Pol.* **18**, 393 (1959).

[26] W. G. Dixon, A covariant multipole formalism for extended test bodies in general relativity, *Il Nuovo Cimento (1955-1965)* **34**, 317 (1964).

[27] W. G. Dixon, The New Mechanics of Myron Mathisson and Its Subsequent Development, in *Equations of Motion in Relativistic Gravity*, edited by D. Puetzfeld, C. Lämmerzahl, and B. Schutz (Springer International Publishing, Cham, 2015) pp. 1–66.

[28] J.-M. Souriau, Modèle de particule à spin dans le champ électromagnétique et gravitationnel, *Ann. Inst. Henri Poincaré A* **20**, 315 (1974).

[29] P. Saturnini, *Un modèle de particule à spin de masse nulle dans le champ de gravitation*, Thesis, Université de Provence (1976).

[30] C. Duval and T. Schücker, Gravitational birefringence of light in Robertson-Walker cosmologies, *Phys. Rev. D* **96**, 043517 (2017).

[31] C. Duval, L. Marsot, and T. Schücker, Gravitational birefringence of light in Schwarzschild spacetime, *Phys. Rev. D* **99**, 124037 (2019).

[32] L. Marsot, How does the photon's spin affect gravitational wave measurements?, *Phys. Rev. D* **100**, 064050 (2019).

[33] J. Audretsch, Trajectories and spin motion of massive spin- $\frac{1}{2}$ particles in gravitational fields, *J. Phys. A* **14**,

411 (1981).

[34] R. Rüdiger, The Dirac equation and spinning particles in general relativity, *Proc. R. Soc. A* **377**, 417 (1981).

[35] V. P. Frolov and A. A. Shoom, Spinoptics in a stationary spacetime, *Phys. Rev. D* **84**, 044026 (2011).

[36] V. P. Frolov and A. A. Shoom, Scattering of circularly polarized light by a rotating black hole, *Phys. Rev. D* **86**, 024010 (2012).

[37] C.-M. Yoo, Notes on spinoptics in a stationary spacetime, *Phys. Rev. D* **86**, 084005 (2012).

[38] S. R. Dolan, Higher-order geometrical optics for circularly-polarized electromagnetic waves, arXiv:1801.02273 (2018), arXiv:1801.02273 .

[39] S. R. Dolan, Geometrical optics for scalar, electromagnetic and gravitational waves on curved spacetime, *Int. J. Mod. Phys. D* **27**, 1843010 (2018).

[40] A. I. Harte, Gravitational lensing beyond geometric optics: I. Formalism and observables, *Gen. Relativ. Gravit.* **51**, 14 (2019).

[41] P. Gosselin, A. Bérard, and H. Mohrbach, Spin Hall effect of photons in a static gravitational field, *Phys. Rev. D* **75**, 084035 (2007).

[42] P. Gosselin, A. Bérard, and H. Mohrbach, Semiclassical dynamics of Dirac particles interacting with a static gravitational field, *Phys. Lett. A* **368**, 356 (2007).

[43] A. J. Silenko and O. V. Teryaev, Semiclassical limit for Dirac particles interacting with a gravitational field, *Phys. Rev. D* **71**, 064016 (2005).

[44] A. J. Silenko, Foldy-Wouthuysen transformation and semiclassical limit for relativistic particles in strong external fields, *Phys. Rev. A* **77**, 012116 (2008).

[45] Y. N. Obukhov, A. J. Silenko, and O. V. Teryaev, General treatment of quantum and classical spinning particles in external fields, *Phys. Rev. D* **96**, 105005 (2017).

[46] N. Yamamoto, Spin Hall effect of gravitational waves, *Phys. Rev. D* **98**, 061701(R) (2018).

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

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

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

[50] K. Y. Bliokh and F. Nori, Relativistic Hall Effect, *Phys. Rev. Lett.* **108**, 120403 (2012).

[51] M. Stone, V. Dwivedi, and T. Zhou, Wigner Translations and the Observer Dependence of the Position of Massless Spinning Particles, *Phys. Rev. Lett.* **114**, 210402 (2015).

[52] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, 1973).

[53] S. Bates and A. Weinstein, *Lectures on the Geometry of Quantization*, Berkeley Mathematics Lecture Notes, Vol. 8 (American Mathematical Society, Providence, RI; Berkeley Center for Pure and Applied Mathematics, Berkeley, CA, 1997) pp. vi+137.

[54] R. M. Wald, *General Relativity* (University of Chicago Press, Chicago, IL, 1984) pp. xiii+491.

[55] Y. Choquet-Bruhat, General Relativity and the Einstein Equations, Oxford Mathematical Monographs (Oxford University Press, Oxford, 2009) pp. xxvi+785.

[56] J. J. Duistermaat, *Fourier Integral Operators*, Modern Birkhäuser Classics (Birkhäuser/Springer, New York, 2011).

[57] S. A. Fulling, Pseudodifferential operators, covariant quantization, the inescapable van Vleck–Morette determinant, and the $\frac{R}{6}$ controversy, *Int. J. Mod. Phys. D* **05**, 597 (1996).

[58] I. Y. Dodin, D. E. Ruiz, K. Yanagihara, Y. Zhou, and S. Kubo, Quasioptical modeling of wave beams with and without mode conversion. I. Basic theory, *Phys. Plasmas* **26**, 072110 (2019).

[59] 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, 2014).

[60] R. Penrose and W. Rindler, *Spinors and Space-Time: Two-Spinor Calculus and Relativistic Fields*, Vol. 1 (Cambridge University Press, 1987).

[61] G. R. Fowles, *Introduction to Modern Optics* (Courier Corporation, 1989).

[62] K. Y. Bliokh, Geometrodynamics of polarized light: Berry phase and spin Hall effect in a gradient-index medium, *J. Opt. A: Pure Appl. Opt.* **11**, 094009 (2009).

[63] B. B. Moussa and G. T. Kossioris, On the System of Hamilton–Jacobi and Transport Equations Arising in Geometrical Optics, *Commun. Part. Diff. Eq.* **28**, 1085 (2003).

[64] H. Goldstein, C. P. Poole, and J. L. Safko, *Classical Mechanics* (Addison Wesley, 2002).

[65] R. Abraham and J. E. Marsden, *Foundations of Mechanics* (Benjamin/Cummings Publishing Co., Inc., Advanced Book Program, Reading, Mass., 1978) pp. xxii+m–xvi+806, second edition.

[66] J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems*, Vol. 17 (Springer Science & Business Media, 2013).

[67] D. Chruciński and A. Jamiołkowski, *Geometric Phases in Classical and Quantum Mechanics*, Vol. 36 (Springer Science & Business Media, 2012).

[68] S. Chandrasekhar, *The Mathematical Theory of Black Holes*, Oxford Classic Texts in the Physical Sciences (The Clarendon Press, Oxford University Press, New York, 1998) pp. xxii+646, reprint of the 1992 edition.

[69] C. Duval, M. Elbistan, P. Horváthy, and P.-M. Zhang, Wigner–souriau translations and lorentz symmetry of chiral fermions, *Physics Letters B* **742**, 322 (2015).

[70] K. Bolonek-Lasoń, P. Kosiński, and P. Maślanka, Lorentz transformations, sideways shift and massless spinning particles, *Physics Letters B* **769**, 117 (2017).

[71] W. R. Inc., *Mathematica*, Version 12.0, Champaign, IL, 2019.

[72] L. Allen, M. W. Beijersbergen, R. J. C. Spreeuw, and J. P. Woerdman, Orbital angular momentum of light and the transformation of Laguerre–Gaussian laser modes, *Phys. Rev. A* **45**, 8185 (1992).

[73] D. L. Andrews and M. Babiker, *The Angular Momentum of Light* (Cambridge University Press, 2012).

[74] M. Krenn and A. Zeilinger, On small beams with large topological charge: II. Photons, electrons and gravitational waves, *New J. Phys.* **20**, 063006 (2018).

[75] S. Aghapour, L. Andersson, and R. Bhattacharyya, Helicity and spin conservation in Maxwell theory and Linearized Gravity, arXiv:1812.03292 (2018), arXiv:1812.03292 .

- [76] K. Y. Bliokh, Geometrical optics of beams with vortices: Berry phase and orbital angular momentum Hall effect, *Phys. Rev. Lett.* **97**, 043901 (2006).
- [77] C. Duval, M. Elbistan, P. A. Horváthy, and P. M. Zhang, Wigner-Souriau translations and Lorentz symmetry of chiral fermions, *Phys. Lett. B* **742**, 322 (2015).
- [78] M. Stone, Berry phase and anomalous velocity of Weyl fermions and Maxwell photons, *Int. J. Mod. Phys. B* **30**, 1550249 (2016).
- [79] J.-M. Souriau, *Structure of Dynamical Systems*, Progress in Mathematics, Vol. 149 (Birkhäuser Boston, Inc., Boston, MA, 1997).
- [80] K. Andrzejewski, A. Kijanka-Dec, P. Kosiński, and P. Maślanka, Chiral fermions, massless particles and Poincare covariance, *Phys. Lett. B* **746**, 417 (2015).
- [81] V. A. Sharafutdinov, *Integral Geometry of Tensor Fields*, Inverse and Ill-posed Problems Series (VSP, Utrecht, 1994) p. 271.