Effects of nonlinear vacuum electrodynamics on the polarization plane of light

Volker Perlick[*]
ZARM, University of Bremen, Am Fallturm, 28359 Bremen, Germany

Claus Lämmerzahl[†]
ZARM, University of Bremen, Am Fallturm, 28359 Bremen, Germany and Institute of Physics, University of Oldenburg, 26111 Oldenburg, Germany

Alfredo Macías[‡]
Departamento de Física, Universidad Autónoma Metropolitana-Iztapalapa, A.P. 55-534, Mexico D.F. 09340, Mexico

(Received 24 September 2018; published 28 November 2018)

We consider the Plebański class of nonlinear theories of vacuum electrodynamics, i.e., Lagrangian theories that are Lorentz invariant and gauge invariant. Our main goal is to derive the transport law of the polarization plane in such a theory, on an unspecified general-relativistic spacetime and with an unspecified electromagnetic background field. To that end we start out from an approximate-plane-harmonic-wave ansatz that takes the generation of higher harmonics into account. By this ansatz, the electromagnetic field is written as an asymptotic series with respect to a parameter α, where the limit $\alpha \to 0$ corresponds to sending the frequency to infinity. We demonstrate that by solving the generalized Maxwell equations to zeroth and first order with respect to α one gets a unique transport law for the polarization plane along each light ray. We exemplify the general results with the Born-Infeld theory.

DOI: 10.1103/PhysRevD.98.105014

I. INTRODUCTION

In the standard Maxwell theory of vacuum electrodynamics, the field strength tensor F_{ab} (which comprises the vector fields $\vec{E}$ and $\vec{B}$) is related to the excitation tensor (which comprises the vector fields $\vec{D}$ and $\vec{H}$) by a linear constitutive law. However, there are good reasons to assume that this law has to be replaced by a nonlinear relation for very strong electromagnetic fields. In the course of history, several such nonlinear modifications of the vacuum Maxwell theory have been suggested.

One of the best known examples is the theory of Born and Infeld [1] from 1934. Its introduction was motivated by the observation that in the standard Maxwell vacuum theory the field energy in an arbitrarily small ball around a point charge is infinite which leads to an infinite self-force, and that this infinity might be overcome if one modifies the constitutive law of the vacuum in a nonlinear fashion. The Born-Infeld theory introduces a new hypothetical constant of Nature, b_0, with the dimension of a (magnetic) field strength. In the limit $b_0 \to \infty$ the theory approaches the standard Maxwell theory, i.e., the fact that the latter is in good agreement with experiments can be understood if one assumes that b_0 is very large. On the basis of the Born-Infeld theory one would have to expect measurable deviations from the vacuum Maxwell theory in electromagnetic fields that are of a similar order of magnitude as b_0. Although not exactly in the main stream of physics, the Born-Infeld theory was always taken seriously by many scientists. In the late 1990s this theory got an additional strong push when Tseytlin [2] realized that it can be derived, as an effective theory, from some kind of string theories.

Another very well known nonlinear modification of the vacuum Maxwell theory is the Heisenberg-Euler theory [3] from 1936. It is a classical field theory which comes about, as an effective theory, if one-loop corrections from quantum electrodynamics are taken into account. In contrast to the Born-Infeld theory, it does not involve any new hypothetical constant of Nature, i.e., it numerically predicts how strong an electromagnetic field has to be in order to produce measurable deviations from the standard vacuum Maxwell theory. Since a few years (magnetic) fields of this strength can be produced in the laboratory.

The Born-Infeld theory and the Heisenberg-Euler theory are Lorentz invariant, they are gauge invariant, and they

[*]volker.perlick@zarm.uni-bremen.de
[†]claus.laemmerzahl@zarm.uni-bremen.de
[‡]amac@xanum.uam.mx

derive from a Lagrangian. The entire class of theories that share these properties was systematically studied by Plebański [4], with important early contributions by Boillat [5]. We refer to it as to the *Plebański class* of electromagnetic theories. The Born-Infeld theory and the Heisenberg-Euler theory are the best known examples in this class, but there are many more. In particular, there are theories of the Plebański class that allow for regular black-hole solutions if they are coupled to Einstein's field equations. The first two examples were found by Ayón-Beato and García [6,7].

It is a general feature of nonlinear theories that the superposition principle is no longer satisfied. As a consequence, the propagation of light is influenced by electromagnetic background fields. This effect is known as "light-by-light scattering" and it has been observed in 1997, see Burke *et al.* [8], in good agreement with the prediction by the Heisenberg-Euler theory. Another effect predicted by most theories of the Plebański class, with the notable exception of the Born-Infeld theory, is birefringence in vacuo. This means that, according to these theories, a light beam that enters into a region with a sufficiently strong electromagnetic background field would split into two beams, as in the case of a light beam entering into a crystal according to ordinary optics, with the two different beams corresponding to two different polarization states. Such a birefringence in vacuo is predicted, in particular, by the Heisenberg-Euler theory. Experimentalists are trying to observe this effect since several years and there is the general expectation that these attempts will be successful soon, see in particular the most recent status report on the so-called PVLAS experiment by Della Valle *et al.* [9]. In this experiment not only the birefringence in vacuo but also the dichroism of the Heisenberg-Euler theory is tried to be measured. The latter means the effect that there are different absorption coefficients for the two different polarization states which results in an apparent rotation of the polarization plane. Finally, we mention that there are also attempts to verify effects from nonlinear electrodynamics with astrophysical observations. A particularly promising idea is to observe the birefringence in vacuo if light passes through a very strong magnetic field, such as in the neighborhood of a magnetar. A first observation that might indicate such an effect was already made, see Mignani *et al.* [10].

We emphasize that experiments searching for birefringence in vacuo cannot be used as tests for the Born-Infeld theory because in the latter there is no such effect. As an alternative, the Born-Infeld theory may be tested with the help of Michelson interferometry. Such an experiment was discussed for the Heisenberg-Euler theory by Boer and van Holten [11], Döbrich and Gies [12], Zavattini and Calloni [13] and Grote [14], for the Heisenberg-Euler and the Born-Infeld theories by Denisov, Krivchenkov and Kravtsov [15], and in detail for a general theory of the Plebański class by Schellstede *et al.* [16]. Moreover, there are suggestions to test the Born-Infeld theory with wave-guides, see Ferraro [17], or with fluid motions in a magnetic background field, see Dereli and Tucker [18]. As of now, none of these experiments has been actually carried through.

In this paper we want to study, for a general theory of the Plebański class, the effect of a background field on the transport law of the polarization plane along a light ray. This will give us a new way of testing these theories, in particular the Born-Infeld theory, experimentally. We emphasize that this is to be distinguished from all the experiments mentioned above. In particular, it is not to be confused with the planned observation of dichroism by the PVLAS experiment: The latter is an effect on the absorption of light, depending on the polarization state. Here we want to investigate the direct effect of an electromagnetic background field on the polarization plane.

To that end we start out from an approximate-plane-harmonic-wave ansatz, taking the generation of higher harmonics into account. By this ansatz the electromagnetic field is written as an asymptotic series with respect to a parameter α. Sending α to zero corresponds to sending the frequency to infinity. We will see that we have to consider the generalized Maxwell equations to zeroth order and to first order with respect to α in order to determine the transport law for the polarization plane. Earlier studies of the high-frequency limit in nonlinear theories were restricted to the derivation of the eikonal equation from the zeroth order of the generalized Maxwell equations. It is well known that, as a result, one finds that the light rays are the null geodesics of two optical metrics; this was first shown by Novello *et al.* [19] and later, in different representations, by Obukhov and Rubilar [20] and by Schellstede *et al.* [16]. To the best of our knowledge, the transport law of the polarization plane was not yet considered for an arbitrary theory of the Plebański class.

We will not specify the nonlinear electromagnetic theory, apart from the fact that we require it to be of the Plebański class. However, we mention that not all theories of this type are to be considered as physically meaningful: Some of them violate causality in the sense that the light cones of the optical metrics are not inside the light cone of the spacetime metric, see Schellstede *et al.* [21]. Also, not all of them give rise to a well-posed initial-value problem, see Abalos *et al.* [22].

The paper is organized as follows. In Sec. II we briefly review the basic features of theories of the Plebański class. In Sec. III we introduce our approximate-plane-wave ansatz on an arbitraty general-relativistic spacetime and for an arbitrary electromagnetic background field. In Sec. IV we evaluate the generalized Maxwell equations to zeroth order which gives us the eikonal equation and an algebraic condition on the polarization plane. In Sec. V we consider the generalized Maxwell equations to first order and discuss the additional conditions they give us on the polarization plane. In Sec. VI we exemplify the results with the Born-Infeld theory.

II. THE PLEBAŃSKI CLASS OF NONLINEAR ELECTRODYNAMICAL THEORIES

We consider a general-relativistic spacetime, i.e., an oriented 4-dimensional manifold with a metric tensor g_{ab} of Lorentzian signature. The covariant derivative associated with the Levi-Civita connection of the metric will be denoted ∇_a. Latin indices take values 0, 1, 2, 3 and are lowered with g_{ab} and raised with its inverse g^{bc}.

The Plebański class [4] consists of all nonlinear electrodynamical theories that derive from an action of the form

$$S[A_c] = \frac{1}{4\pi c}\int_M \left(\mathcal{L}(F,G) + \frac{4\pi}{c} j^a A_a\right)\sqrt{|\det(g_{bc})|}\, d^4x. \tag{1}$$

Here M is a domain of the spacetime, $d^4x = dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3$, j^a is a *given* current density, A_a is the electromagnetic potential,

$$F_{ab} = \nabla_a A_b - \nabla_b A_a \tag{2}$$

is the electromagnetic field strength and $\mathcal{L}$ is the Lagrangian for the electromagnetic field. It is assumed that the latter depends only on the two invariants

$$F = \frac{1}{2}F_{ab}F^{ab} \quad \text{and} \quad G = -\frac{1}{4}F_{ab}{}^{\star}F^{ab}. \tag{3}$$

Here and in the following, * denotes the Hodge star operator, i.e.,

$${}^{\star}F_{ab} = \frac{1}{2}\varepsilon_{abcd}F^{cd} \tag{4}$$

where ε_{abcd} is the totally antisymmetric Levi-Civita tensor field (volume form) associated with the spacetime metric.

By (2) the homogeneous Maxwell equation is automatically satisfied,

$$\varepsilon^{abcd}\nabla_b F_{cd} = 0. \tag{5}$$

Requiring that the variational derivative of the action (1) with respect to the potential A_c vanishes, for all compact domains M and all variations that keep A_c fixed on the boundary of M, leads to the inhomogeneous Maxwell equation,

$$\nabla_a H^{ab} = -\frac{4\pi}{c} j^b, \tag{6}$$

where

$$H^{ab} = -\frac{\partial \mathcal{L}}{\partial F_{ab}} = -2\mathcal{L}_F F^{ab} + \mathcal{L}_G{}^{\star}F^{ab} \tag{7}$$

is the electromagnetic excitation. For the sake of brevity, we write

$$\mathcal{L}_F = \frac{\partial \mathcal{L}}{\partial F}, \qquad \mathcal{L}_G = \frac{\partial \mathcal{L}}{\partial G} \tag{8}$$

and

$$\mathcal{L}_{FF} = \frac{\partial^2 \mathcal{L}}{\partial F^2}, \qquad \mathcal{L}_{GG} = \frac{\partial^2 \mathcal{L}}{\partial G^2}, \qquad \mathcal{L}_{FG} = \frac{\partial^2 \mathcal{L}}{\partial F \partial G}. \tag{9}$$

It is the constitutive law (7) that distinguishes different theories, while the Maxwell equations (5) and (6) are always the same.

Each particular theory of the Plebański class is characterized by a particular Lagrangian and, thereby, by a particular constitutive law. Let us mention the two most important examples: For the Born-Infeld theory [1], the Lagrangian reads

$$\mathcal{L} = b_0^2 - b_0^2\sqrt{1 + \frac{F}{b_0^2} - \frac{G^2}{b_0^4}} \tag{10}$$

where b_0 is a hypothetical constant of Nature with the dimension of a magnetic field strength. For $b_0 \to \infty$ the Born-Infeld theory reproduces the standard Maxwell vacuum theory. For the Heisenberg-Euler theory [3],

$$\mathcal{L} = E_0^2\left(-\frac{1}{2}\frac{F}{E_0^2} + \Lambda\left(\frac{F^2}{E_0^4} + 7\frac{G^2}{E_0^4}\right) + \ldots\right) \tag{11}$$

where $E_0 = m^2c^4/e^3$ and $\Lambda = \hbar c/(90\pi e^2)$. Here m is the electron mass, e is the electron charge, $\hbar$ is the reduced Planck constant and the ellipses in (11) stand for terms of third and higher order in F and G.

III. APPROXIMATE-PLANE-HARMONIC-WAVE ANSATZ

An approximate-plane-harmonic wave is a one-parameter family F^{α}_{cd} of field strength tensors, depending on a real parameter α, of the form

$$F^{\alpha}_{cd} = F_{cd} + \alpha F^{(1)}_{cd} + \sum_{K=2}^{\infty}\alpha^K F^{(K)}_{cd} \tag{12}$$

where

$$F^{(1)}_{cd} = \mathrm{Re}\{e^{iS/\alpha} f^{(11)}_{cd}\} \tag{13}$$

and

$$F_{cd}^{(K)} = \sum_{\tilde{K}=0}^{K} \mathrm{Re}\{e^{i\tilde{K}S/\alpha} f_{cd}^{(K\tilde{K})}\} \quad \text{for } K \geq 2. \tag{14}$$

Here F_{cd} is a given electromagnetic background field that is independent of α, S is a real-valued function and $f_{cd}^{(K\tilde{K})}$ is a complex-valued antisymmetric tensor field for each pair of integers K, $\tilde{K}$ that occurs. We assume that, on the spacetime region considered, the tensor fields $\nabla_a S$ and $f_{cd}^{(11)}$ have no zeros. The series is to be understood as an asymptotic series, not as a convergent series.

The function S is called the *eikonal function*. On a sufficiently small neighborhood, the field $F_{ab}^{(1)}$ is approximately a plane harmonic wave: The surfaces $S = \text{constant}$ are the wave-fronts and the gradient of S divided by alpha defines the wave four-covector. Correspondingly, the frequency measured by an observer with four-velocity U^a is $\omega = U^a \nabla_a S/\alpha$. The limit $\alpha \to 0$ corresponds to sending the frequency to infinity. The idea is to feed the ansatz (12) into Maxwell's equations, to solve these equations iteratively order by order in α and, in this way, to asymptotically approach a one-parameter family of exact solutions.

Our ansatz (12) is a generalization of the standard approximate-plane-harmonic-wave ansatz. The latter goes back to Ralph Luneburg and is detailed, for wave propagation in linear and isotropic media, e.g., in the textbook by Kline and Kay [23]. Our ansatz is more general in two respects: First, we take a nonzero background field into account. In a linear theory, it suffices to consider the case with zero background field because, by the superposition principle, the propagation of the approximate-plane-harmonic wave is independent of a background field. In a nonlinear theory, however, the propagation is influenced by a background field. Second, the higher-order fields, $F_{cd}^{(K)}$ for $K \geq 2$, come not only with the same frequency as the first-order field $F_{cd}^{(1)}$ (the terms with $\tilde{K} = 1$) but also with integer multiples of this frequency (the terms with $\tilde{K} \neq 1$). This reflects the *generation of higher harmonics* which is well-known from optics in nonlinear media. It should not come as a surprise that it has to be taken into account also in the nonlinear vacuum theories of the Plebański class. Higher harmonics play no role if one considers Maxwell's equations only to the lowest order (i.e., α^0). This is the reason why it was not necessary to take them into account in [24] where the eikonal equation was derived for Maxwell's equations with a local but otherwise arbitrary constitutive law. In the present paper, however, we want to derive the transport law for the polarization plane which requires considering Maxwell's equations also to the next order (i.e., α^1). We will see that these equations cannot in general be solved if we set all terms $f_{cd}^{(K\tilde{K})}$ with $\tilde{K} \neq 1$ equal to zero.

For our purpose we need the series (12) up to second order,

$$F_{cd}^{\alpha} = F_{cd} + \alpha \mathrm{Re}\{e^{iS/\alpha} f_{cd}^{(11)}\} + \alpha^2 \mathrm{Re}\{f_{cd}^{(20)} + e^{iS/\alpha} f_{cd}^{(21)} + e^{2iS/\alpha} f_{cd}^{(22)}\} + \dots \tag{15}$$

which includes *frequency doubling* ($\tilde{K} = 2$) and the generation of a nonoscillatory mode, known from nonlinear media as *optical rectification* ($\tilde{K} = 0$). The homogeneous Maxwell equation (5) is automatically satisfied for all α if we assume that (15) derives from a potential,

$$F_{cd}^{\alpha} = \nabla_c A_d^{\alpha} - \nabla_d A_c^{\alpha}. \tag{16}$$

It is easy to see that such a potential (up to an arbitrary gradient term) must be of the form

$$A_d^{\alpha} = A_d + \alpha^2 \mathrm{Re}\{a_d^{(10)} + e^{iS/\alpha} a_d^{(11)}\} + \alpha^3 \mathrm{Re}\{a_d^{(20)} + e^{iS/\alpha} a_d^{(21)} + e^{2iS/\alpha} a_d^{(22)}\} + \dots \tag{17}$$

Then (16) holds to zeroth order in α with

$$F_{cd} = \nabla_c A_d - \nabla_d A_c, \tag{18}$$

to first order with

$$f_{cd}^{(11)} = i(\nabla_c S a_d^{(11)} - \nabla_d S a_c^{(11)}), \tag{19}$$

and to second order with

$$f_{cd}^{(20)} = \nabla_c a_d^{(10)} - \nabla_d a_d^{(10)}, \tag{20}$$

$$f_{cd}^{(21)} = \nabla_c a_d^{(11)} - \nabla_d a_d^{(11)} + i(\nabla_c S a_d^{(21)} - \nabla_d S a_c^{(21)}), \tag{21}$$

$$f_{cd}^{(22)} = 2i(\nabla_c S a_d^{(22)} - \nabla_d S a_c^{(22)}). \tag{22}$$

Here we have used our assumption that the gradient of S has no zeros which implies that $S \neq 0$ almost everywhere and that, accordingly, the functions 1, $\sin(S(x)/\alpha)$, $\cos((S(x)/\alpha))$, $\sin(2S(x)/\alpha))$, and $\cos(2S(x)/\alpha)$ are linearly independent.

Feeding the approximate-plane-harmonic wave (12) into the constitutive law (7) gives, after a rather long but straightforward calculation, an excitation of the form

$$H_{ab}^{\alpha} = H_{ab} + \alpha \mathrm{Re}\{e^{iS/\alpha} h_{ab}^{(11)}\} + \alpha^2 \mathrm{Re}\{h_{ab}^{(20)} + e^{iS/\alpha} h_{ab}^{(21)} + e^{i2S/\alpha} h_{ab}^{(22)}\} + \dots \tag{23}$$

The zeroth order term in (23) is just the excitation of the background field,

$$H_{ab} = -2\mathcal{L}_F F_{ab} + \mathcal{L}_G {}^{\star}F_{ab}, \tag{24}$$

the first-order amplitude is

$$h^{(11)}_{ab} = \frac{1}{2}\chi_{ab}{}^{cd} f^{(11)}_{cd}, \tag{25}$$

and the second-order amplitudes are

$$h^{(20)}_{ab} = \frac{1}{2}\chi_{ab}{}^{cd} f^{(20)}_{cd} + \frac{1}{2}\psi_{ab}{}^{cdef} f^{(11)}_{cd}\bar{f}^{(11)}_{ef}, \tag{26}$$

$$h^{(21)}_{ab} = \frac{1}{2}\chi_{ab}{}^{cd} f^{(21)}_{ef}, \tag{27}$$

$$h^{(22)}_{ab} = \frac{1}{2}\chi_{ab}{}^{cd} f^{(22)}_{cd} + \frac{1}{2}\psi_{ab}{}^{cdef} f^{(11)}_{cd} f^{(11)}_{ef}, \tag{28}$$

with

$$\chi_{ab}{}^{cd} = \mathcal{L}_G \varepsilon_{ab}{}^{cd} - 2\mathcal{L}_F(\delta^c_a\delta^d_b - \delta^d_a\delta^c_b) - 4\mathcal{L}_{FF}F_{ab}F^{cd} + 2\mathcal{L}_{FG}(F_{ab}{}^\star F^{cd} + {}^\star F_{ab}F^{cd}) - \mathcal{L}_{GG}{}^\star F_{ab}{}^\star F^{cd} \tag{29}$$

and

$$\begin{aligned}\psi_{ab}{}^{cdef} = {}& \frac{1}{4}(-2\mathcal{L}_{FF}F_{ab} + \mathcal{L}_{FG}{}^\star F_{ab})(g^{ce}g^{df} - g^{de}g^{cf}) + \frac{1}{2}(\delta^c_a\delta^d_b - \delta^c_b\delta^d_a)(-2\mathcal{L}_{FF}F^{ef} + \mathcal{L}_{FG}{}^\star F^{ef}) \\ & - \frac{1}{8}(-2\mathcal{L}_{FG}F_{ab} + \mathcal{L}_{GG}{}^\star F_{ab})\varepsilon^{cdef} - \frac{1}{4}\varepsilon_{ab}{}^{cd}(-2\mathcal{L}_{FG}F^{ef} + \mathcal{L}_{GG}{}^\star F^{ef}) + \frac{1}{2}(-2\mathcal{L}_{FFF}F_{ab} + \mathcal{L}_{FFG}{}^\star F_{ab})F^{cd}F^{ef} \\ & - \frac{1}{2}(-2\mathcal{L}_{FFG}F_{ab} + \mathcal{L}_{FGG}{}^\star F_{ab})F^{cd\star}F^{ef} + \frac{1}{8}(-2\mathcal{L}_{FGG}F_{ab} + \mathcal{L}_{GGG}{}^\star F_{ab}){}^\star F^{cd\star}F^{ef}.\end{aligned} \tag{30}$$

We see that the first- order constitutive law (25) is of the same form as the constitutive law of a linear medium, but now with a constitutive tensor $\chi_{ab}{}^{cd}$ that depends on the invariants F and G of the background field. Quite generally, such a constitutive tensor can be decomposed into principal part, skewon part, and axion part (see Hehl and Obukhov [25]). In (29), the first term is the axion part, the rest is the principal part, and the skewon part is zero. It is known [25] that the skewon part is always vanishing if the theory derives from a variational principle.

At the second order, we get for each of the three amplitudes $h^{(2\tilde{K})}_{ab}$ a linear law with the same constitutive tensor $\chi_{ab}{}^{cd}$ as for the first order, but for $\tilde{K}=0$ and $\tilde{K}=2$ additional quadratic terms with a second-order constitutive tensor $\psi_{ab}{}^{cdef}$ which looks rather complicated.

We will now evaluate the Maxwell equations. The homogeneous Maxwell equation is satisfied if we express the amplitudes $f^{(K\tilde{K})}_{cd}$ in terms of the potential according to (19), (20), (21), and (22). Feeding the excitation (23) into the inhomogeneous Maxwell equation requires at zeroth order

$$-\frac{4\pi}{c}j_b = \nabla^a H_{ab}, \tag{31}$$

$$0 = \nabla^a S h^{(11)}_{ab}, \tag{32}$$

and at first order

$$0 = \nabla^a h^{(11)}_{ab} + i\nabla^a S h^{(21)}_{ab}, \tag{33}$$

$$0 = \nabla^a S h^{(22)}_{ab}. \tag{34}$$

Here we have assumed that the current j_b is independent of α, i.e., that only the background field may have a source whereas our approximate-plane-harmonic wave is source-free. Moreover, we have again used our assumption that the gradient of S has no zeros which implies that the functions 1, $\sin(S(x)/\alpha)$, $\cos((S(x)/\alpha))$, $\sin(2S(x)/\alpha))$, and $\cos(2S(x)/\alpha)$ are linearly independent.

At zeroth order we get one equation, (32), that has to be satisfied. With (19) and (25) this equation reads

$$0 = \nabla^a S\chi_{ab}{}^{cd}\nabla_c S a^{(11)}_d. \tag{35}$$

We will evaluate this equation in the next section. We will see that it gives us the *eikonal equation* for S and an algebraic condition on $a^{(11)}_d$ which is known as the zeroth order *polarization condition*. Note that $a^{(11)}_d$ is not gauge-invariant: As can be read from (19), the field strength $f^{(11)}_{cd}$ is unchanged if a multiple of $\nabla_d S$ is added to $a^{(11)}_d$. We will see that the zeroth order polarization condition is actually a condition on the (gauge-invariant) plane spanned by $a^{(11)}_d$ and $\nabla_d S$. We refer to this plane as to the *polarization plane*.

At first order we get two equations, (33) and (34), that have to be satisfied. With (19), (20), (21), (25), (27), and (28) these equations read

$$\begin{aligned}0 = {}& \nabla^a(\chi_{ab}{}^{cd}\nabla_c S a^{(11)}_d) + \nabla^a S\chi_{ab}{}^{cd}\nabla_c a^{(11)}_d \\ & + i\nabla^a S\chi_{ab}{}^{cd}\nabla_c S a^{(21)}_d,\end{aligned} \tag{36}$$

$$0 = \nabla^a S\chi_{ab}{}^{cd}\nabla_c S a^{(22)}_d - \psi_{ab}{}^{cdef}\nabla^a S\nabla_c S a^{(11)}_d a^{(11)}_f. \tag{37}$$

We will evaluate these two equations, as far as necessary for our purpose, in Sec. V below. They will give us a differential equation for $a^{(11)}_d$ which is known as the first-order *transport equation* and algebraic conditions on $a^{(21)}_d$ and

$a_d^{(22)}$ which are the first-order polarization conditions. These equations are not in general satisfied if $a_d^{(22)} = 0$, i.e., frequency doubling has to be taken into account if Maxwell's equations are to be solved to first order.

If one wants to go beyond the first order, one can do this step by step. At the Kth level one gets transport equations for the amplitudes $a_d^{(K\tilde{K})}$ and polarization conditions on the amplitudes $a_d^{((K+1)\tilde{K})}$.

IV. EVALUATION OF THE ZEROTH-ORDER FIELD EQUATION

In Sec. IV A we will derive the eikonal equation from the zeroth-order field equation (35), in Sec. IV B we will determine the Hamiltonian for the rays and in Sec. IV C we will evaluate the zeroth-order polarization condition. The main results of Secs. IV A and IV B are not new. In particular, it is known that for any theory of the Plebański class the rays are the null geodesics of two optical metrics. This was first demonstrated by Novello *et al.* [19]. The same result was rederived, using a different representation, by Obukhov and Rubilar [20] who also showed that the optical metrics have Lorentzian signature if they are nondegenerate. Still another form of the optical metrics was derived by Schellstede *et al.* [16]. However, we have to rederive these known results here because in doing so we will also establish a number of new relations that will be needed later. We will use the same representation as in [16].

A. Derivation of the eikonal equation

In the following we write

$$p_a = \nabla_a S, \quad u_a = F_{ab} \nabla^b S, \quad v_a = {}^*F_{ab} \nabla^b S \tag{38}$$

which implies

$$p_a u^a = p_a v^a = 0. \tag{39}$$

Then the zeroth-order field equation (35) can be rewritten as

$$M_b{}^d a_d^{(11)} = 0 \tag{40}$$

where

$$M_b{}^d = \chi_{ab}{}^{cd} p^a p_c = -2\mathcal{L}_F p_a p^a \delta_b^d + 2\mathcal{L}_F p_b p^d - 4\mathcal{L}_{FF} u_b u^d + 2\mathcal{L}_{FG}(u_b v^d + v_b u^d) - \mathcal{L}_{GG} v_b v^d. \tag{41}$$

Note that $M_b{}^d$ is self-adjoint with respect to the spacetime metric, i.e., $M_{ab} = M_{ba}$. This is a consequence of the above-mentioned fact that the skewon part of the constitutive tensor vanishes. Also note that the axion part gives no contribution to (41) which is a general result [25,26].

From (41) we read that p_d is in the kernel of $M_b{}^d$, so (40) is satisfied by $a_d^{(11)} = \psi p_d$ with any scalar factor ψ. However, by (19) such a potential gives a trivial first-order field strength. As we require $f_{cd}^{(11)} \neq 0$, we need a solution $a_d^{(11)}$ of (40) that is linearly independent of p_d, i.e., the kernel of $M_d{}^b$ has to be at least two-dimensional. This is the case if and only if the *adjugate* $A_d{}^b$ of $M_d{}^b$ (also known as the *classical adjoint*) vanishes, cf. Itin [27]. A straightforward (though tedious) calculation shows that the adjugate is given by

$$A_b{}^a = -8\mathcal{L}_F(M(p_c p^c)^2 + N p_c p^c u_d u^d + P(u_d u^d)^2) p_b p^a \tag{42}$$

where

$$M = \mathcal{L}_F^2 + 2\mathcal{L}_F \mathcal{L}_{FG} G - \frac{1}{2}\mathcal{L}_F \mathcal{L}_{GG} F - PG^2, \tag{43}$$

$$N = 2\mathcal{L}_F \mathcal{L}_{FF} + \frac{1}{2}\mathcal{L}_F \mathcal{L}_{GG} - PF, \tag{44}$$

$$P = \mathcal{L}_{FF}\mathcal{L}_{GG} - \mathcal{L}_{FG}^2. \tag{45}$$

Here we have used the well-known [4] identities

$${}^*F_{ac} F^{bc} = -G\delta_a^b, \qquad F_{ac}F^{bc} - {}^*F_{ac}{}^*F^{bc} = F\delta_a^b \tag{46}$$

which imply

$$u_c v^c = -G p_c p^c, \qquad u_c u^c - v_c v^c = F p_c p^c. \tag{47}$$

By (42), the zeroth-order field equation (40) admits a solution $a_d^{(11)}$ giving a nontrivial field strength if and only if

$$0 = \mathcal{L}_F(M(p_c p^c)^2 + N p_c p^c u_d u^d + P(u_d u^d)^2). \tag{48}$$

This is the *eikonal equation*. It is a first-order partial differential equation for the function S. Each solution to this equation determines a family of light rays, in the same way as in Hamiltonian mechanics each solution to the Hamilton-Jacobi equation determines a family of trajectories, see the next subsection. If viewed as an algebraic condition on the covector p_a, (48) is known as the *dispersion relation*, as the *characteristic equation* or as the *Fresnel equation*.

From now on we require $\mathcal{L}_F \neq 0$ because otherwise the eikonal equation is an identity, so there is no well-defined notion of rays. If in addition $M \neq 0$, (48) factorizes according to

$$(\tilde{g}_+^{bc} p_b p_c)(\tilde{g}_-^{de} p_d p_e) = 0 \tag{49}$$

where

$$\tilde{g}^{bc}_{\pm} = g^{bc} + \sigma_{\pm} F^{bd} F^{c}{}_{d} = (1 + \sigma_{\pm} F) g^{bc} + \sigma_{\pm} {}^{*}F^{bd*}F^{c}{}_{d} \tag{50}$$

and

$$\sigma_{\pm} = \frac{N}{2M} \pm \sqrt{\frac{N^2}{4M^2} - \frac{P}{M}}. \tag{51}$$

$\tilde{g}^{bc}_{+}$ and $\tilde{g}^{bc}_{-}$ are known as the *optical metrics*. Note that $\sigma_{\pm}$ is always real because $N^2 - 4MP$ can be rewritten as the sum of two squares,

$$N^2 - 4MP = (\mathcal{L}_F \mathcal{L}_{GG} - N)^2 + 4(\mathcal{L}_F \mathcal{L}_{FG} - PG)^2. \tag{52}$$

The determinant of $\tilde{g}^{cd}_{\pm}$ is

$$\det(\tilde{g}^{cd}_{\pm}) = (1 + \sigma_{\pm} F - \sigma_{\pm}^2 G^2)^2 \det(g^{cd}). \tag{53}$$

As (g^{cd}) is of Lorentzian signature, the right-hand side of (53) is either zero or negative. This demonstrates that the optical metrics are either degenerate or Lorentzian (i.e., of signature $(-+++)$ or $(---+)$), as was already observed by Obukhov and Rubilar [20]. If the determinant is nonzero, the covariant components of the optical metrics are

$$(\tilde{g}^{-1})^{\pm}_{cd} = \frac{g_{cd} - \sigma_{\pm} {}^{*}F_{c}{}^{b*}F_{db}}{1 + \sigma_{\pm} F - \sigma_{\pm}^2 G^2} = \frac{(1 + \sigma_{\pm} F) g_{cd} - \sigma_{\pm} F_{c}{}^{b} F_{db}}{1 + \sigma_{\pm} F - \sigma_{\pm}^2 G^2}. \tag{54}$$

Indeed, with the help of the identities (46) it is easy to check that (50) and (54) imply $(\tilde{g}^{-1})^{\pm}_{ac} \tilde{g}^{cb}_{\pm} = \delta^b_a$.

If $M = 0$, the eikonal equation factorizes as well, but we will not consider this case because it shows some pathologies, see [21]. We restrict for the rest of the paper to background fields for which $\mathcal{L}_F \neq 0$, $M \neq 0$ and $(1 + \sigma_{\pm} F - \sigma_{\pm}^2 G^2) \neq 0$ so that we have two optical metrics of Lorentzian signature. Then the eikonal equation is of the form (49), i.e., it requires $p_a = \nabla_a S$ to be a null covector of at least one of the two optical metrics. This is true if and only if

$$p_a p^a + \sigma_{\pm} u_a u^a = 0 \tag{55}$$

holds with at least one of the two signs where $\sigma_{\pm}$ is given by (51). We refer to the two equations (55) with $p_a = \nabla_a S$ as to the two *partial eikonal equations*.

We end this section with two useful results.

Proposition IV.1. Let σ be one of the two solutions, $\sigma = \sigma_+$ or $\sigma = \sigma_-$, to (51). Then the following conditions are mutually equivalent:

(a) $N^2 = 4MP$, i.e., the two optical metrics coincide, $\sigma = \frac{N}{2M}$.
(b) $\mathcal{L}_F \mathcal{L}_{GG} = N$ and $\mathcal{L}_F \mathcal{L}_{FG} = PG$.
(c) $DM = \mathcal{L}_F^2$, $DN = 2\mathcal{L}_F^2 \sigma$ and $DP = \mathcal{L}_F^2 \sigma^2$.
(d) $2D\mathcal{L}_{FF} = \mathcal{L}_F \sigma (1 + F\sigma)$, $D\mathcal{L}_{GG} = 2\mathcal{L}_F \sigma$ and $D\mathcal{L}_{FG} = \mathcal{L}_F G \sigma^2$.

In (c) and (d), $D = 1 + F\sigma - G^2\sigma^2$.

Proof. (a) ⇔ (b) is obvious from (52). We now assume that one, and thus also the other, of these conditions is true. Then we find from (a) that

$$N = 2M\sigma, \qquad P = M\sigma^2 \tag{56}$$

and from inserting (b) into (43) that

$$M = \mathcal{L}_F^2 - PG^2 - \frac{NF}{2} \tag{57}$$

(56) and (57) demonstrate that then (c) is true. Conversely, (c) obviously implies (a), so we have proven that (a), (b), and (c) are mutually equivalent. Finally, we observe that (a) and (c) together with (44) imply (d) and that (d), if inserted into (43), (44), and (45), implies (a), so all four conditions are indeed mutually equivalent. □

Proposition IV.2. Assume that $p_a = \nabla_a S$ is a solution to the eikonal equation $p_a p^a + \sigma u_a u^a = 0$ with $\sigma = \sigma_+$ or $\sigma = \sigma_-$. Then the eigenvalues of the matrix $(M_b{}^d)$ are $\lambda_1 = \lambda_2 = 0$ and

$$\lambda_3 = 2\mathcal{L}_F \sigma u_a u^a, \tag{58}$$

$$\lambda_4 = (4\mathcal{L}_F \sigma - 4\mathcal{L}_{FF} + 4\mathcal{L}_{FG} G\sigma - \mathcal{L}_{GG}(1 + F\sigma)) u_a u^a \tag{59}$$

Proof. By assumption, zero is a double-eigenvalue of the matrix (41). Then the remaining two eigenvalues λ_3 and λ_4 can be determined in the following way. The formulas for the trace of a matrix and for the trace of the square of a matrix in terms of its eigenvalues yield

$$M_b{}^b = \lambda_3 + \lambda_4, \tag{60}$$

$$M_b{}^d M_d{}^b = \lambda_3^2 + \lambda_4^2. \tag{61}$$

Upon calculating the traces with the help of (39), solving (60) and (61) for the eigenvalues results in the given expressions for λ_3 and λ_4. □

B. Hamiltonian for rays and transport vector fields

We say that S is a solution to the eikonal equation of multiplicity two if $p_a = \nabla_a S$ satisfies the Eq. (55) with both signs, and we say that it is a solution of multiplicity one if (55) holds with one sign but not with the other. The multiplicity may change from point to point.

Each of the two partial eikonal equations has the form of the Hamilton-Jacobi equation, $H(x, \nabla S) = 0$, with the Hamiltonian

$$H_{\pm}(x,p) = \frac{1}{2}\tilde{g}^{bc}_{\pm}(x)p_b p_c. \tag{62}$$

The solutions to Hamilton's equations

$$\dot{x}^a = \frac{\partial H_{\pm}(x,p)}{\partial p_a}, \quad \dot{p}_a = -\frac{\partial H_{\pm}(x,p)}{\partial x^a}, \quad H(x,p) = 0 \tag{63}$$

are known as the *bicharacteristic curves* or as the *rays*. They are the null geodesics of the optical metric. Every solution S to the eikonal equation is associated with a congruence of rays whose tangent vector field is given by

$$K^b_{\pm}(x) = \frac{\partial H_{\pm}(x,p)}{\partial p_b}\bigg|_{p=\nabla S(x)} = \tilde{g}^{bc}_{\pm}(x)\nabla_c S(x), \tag{64}$$

i.e.,

$$K^b_{\pm} = p^b - \sigma_{\pm}F^{bc}u_c. \tag{65}$$

This vector field is known as the *transport vector field* associated with the solution S of the eikonal equation. For solutions of multiplicity two, we have two transport vector fields K^b_+ and K^b_-. However, they are always proportional to each other so that the rays (as unparametrized curves) are uniquely determined. We will prove this in the next section. Note that the nondegeneracy of the optical metric implies that the transport vector field cannot have zeros if we assume that $p_a = \nabla_a S$ has no zeros (as required for an eikonal function of an approximately plane wave), i.e, that "rays cannot stand still."

The following proposition establishes a property of the transport vector field that will be crucial for the next section.

Proposition IV.3. Assume that $p_a = \nabla_a S$ satisfies the eikonal equation $p_a p^a + \sigma u_a u^a = 0$ where $\sigma = \sigma_+$ or $\sigma = \sigma_-$. Let $\tilde{g}^{ab} = g^{ab} + \sigma F^{ac}F^b{}_c$ and $K^a = \tilde{g}^{ab}p_b$. Then

$$\tilde{g}^{cd}p_c u_d = \tilde{g}^{cd}p_c v_d = 0, \qquad \tilde{g}^{cd}u_c v_d = 0, \tag{66}$$

$$\tilde{g}^{cd}u_c u_d = \tilde{g}^{cd}v_c v_d = u^c u_c(1 + \sigma F - \sigma^2 G^2). \tag{67}$$

As a consequence, the transport vector field satisfies

$$K^a p_a = K^a u_a = K^a v_a = 0. \tag{68}$$

Proof. This can be verified in a straightforward manner with the help of the identities (46). □

C. Polarization condition

If we fix a solution $p_a = \nabla_a S$ to the eikonal equation $p_a p^a + \sigma u_a u^a = 0$ with $\sigma = \sigma_+$ or $\sigma = \sigma_-$, the zeroth-order field equation (40) gives an algebraic restriction on $a^{(11)}_b$. This is the zeroth-order polarization condition. In this section we investigate to what extent the polarization condition fixes the allowed values for $a^{(11)}_b$ and, thereby, for the lowest-order field-strength amplitude $f^{(11)}_{cd}$.

Thereby we have to distinguish solutions of multiplicity two from solutions of multiplicity one. Clearly, if the two optical metrics coincide, $\sigma_+ = \sigma_-$, every solution is of multiplicity two. In a background field with $\sigma_+ \neq \sigma_-$, a solution is of multiplicity two if and only if $u_a u^a = 0$. In this case p_a is a *principal null covector*, i.e., a covector with $p_a p^a = 0$ for which u_a and v_a are multiples of p_a. In the following proposition we determine the general form of the matrix $M_b{}^d$ for this special case. For more details on principal null solutions to the eikonal equation we refer to Abalos *et al.* [22] where also pictures of the cones of the optical metrics can be found.

Proposition IV.4. Assume that $p_a = \nabla_a S$ satisfies $p_a p^a = 0$ and $u_a u^a = 0$. Then p_a is a solution of multiplicity two to the eikonal equation. The covectors u_a and v_a are multiples of p_a,

$$u_c = F_c{}^a p_a = \mu p_c, \qquad v_c = {}^*F_c{}^a p_a = \nu p_c, \tag{69}$$

where the coefficients μ and ν satisfy

$$\mu^2 = -\frac{F}{2} + \sqrt{\frac{F^2}{4} + G^2},$$
$$\nu^2 = \frac{F}{2} + \sqrt{\frac{F^2}{4} + G^2}, \qquad \mu\nu = -G. \tag{70}$$

The transport vector fields are proportional to p^a,

$$K^a_{\pm} = \xi_{\pm}p^a, \tag{71}$$

where

$$\xi_{\pm} = 1 - \sigma_{\pm}\mu^2. \tag{72}$$

The matrix $M_b{}^d$ reduces to

$$M_d{}^b = (2\mathcal{L}_F - 4\mathcal{L}_{FF}\mu^2 + 4\mathcal{L}_{FG}\mu\nu - \mathcal{L}_{GG}\nu^2)p_b p^d. \tag{73}$$

Proof. If $p_a p^a = 0$ and $u_a u^a = 0$, (55) is trivially satisfied with both signs, i.e., the covector p_a is lightlike with respect to both optical metrics. Moreover, we read from (66) and (67) that with respect to either of the two optical metrics the covectors u_a and v_a are orthogonal to p_a and lightlike. As two lightlike vectors are orthogonal with respect to a Lorentzian metric if and only if they are linearly dependent, this proves that (69) has to hold with some coefficients μ and ν. Then (70) follows from (46). Inserting (69) into (65) and (41), respectively, yields (71) and (73). □

Recall that the eikonal equation requires the kernel of ${M_b}^d$ to be at least two-dimensional. Proposition (IV.4) implies that the kernel is even three-dimensional if $u_a u^a = 0$. We will now consider the case $u_a u^a \neq 0$.

Proposition IV.5. Assume that $p_a = \nabla_a S$ is a solution to one of the two eikonal equations, $p_a p^a + \sigma u_a u^a = 0$ where σ stands for σ_+ or for σ_-. Let $\tilde{g}^{ab} = g^{ab} + \sigma F^{ac} {F^b}_c$ be the corresponding optical metric and $K^a = \tilde{g}^{ab} p_b$ be the corresponding transport vector field. If $u_a u^a \neq 0$, the three covectors p_a, u_a and v_a are linearly independent. They span the orthocomplement of p_a with respect to $\tilde{g}^{ab}$. The kernel of the matrix ${M_b}^d$ consists of all covectors

$$a_b^{(11)} = \alpha u_b + \beta v_b + \gamma p_b \tag{74}$$

where γ is arbitrary and α and β satisfy

$$\begin{pmatrix} {m_1}^1 & {m_1}^2 \\ {m_2}^1 & {m_2}^2 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{75}$$

where

$$\begin{pmatrix} {m_1}^1 & {m_1}^2 \\ {m_2}^1 & {m_2}^2 \end{pmatrix} = 2\mathcal{L}_F \sigma \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} -4\mathcal{L}_{FF} & 2\mathcal{L}_{FG} \\ 2\mathcal{L}_{FG} & -\mathcal{L}_{GG} \end{pmatrix} \times \begin{pmatrix} 1 & G\sigma \\ G\sigma & 1+F\sigma \end{pmatrix}. \tag{76}$$

The kernel is three-dimensional if and only if $p_a = \nabla_a S$ is a solution of multiplicity two. The kernel is then spanned by p_d, u_d, and v_d, i.e., it coincides with the orthocomplement of p_b with respect to the optical metric.

Proof. Our assumption that $u_a u^a \neq 0$ implies, by (66) and (67) together with $\tilde{g}^{ab} p_a p_b = 0$, that p_a, u_a, and v_a are linearly independent and that they span the $\tilde{g}^{ab}$-orthocomplement of p_a. After normalizing u_a and v_a with the help of (67) we may complement these three covectors to a Newman-Penrose tetrad by choosing a covector w_a with

$$\tilde{g}^{ab} w_a p_b = 1, \qquad \tilde{g}^{ab} w_a w_b = 0,$$
$$\tilde{g}^{ab} w_a u_b = 0, \qquad \tilde{g}^{ab} w_a v_b = 0. \tag{77}$$

From (41) we calculate with the help of (47)

$${M_b}^d w_d = 2\mathcal{L}_F \sigma u_a u^a w_b + 2\mathcal{L}_F (1 + \sigma F^{fg} {F^e}_g w_f p_e) p_b, \tag{78}$$

$${M_b}^d p_d = 0, \tag{79}$$

$${M_b}^d u_d = u_a u^a (2\mathcal{L}_F \sigma - 4\mathcal{L}_{FF} + 2\mathcal{L}_{FG} G\sigma) u_b + u_a u^a (2\mathcal{L}_{FG} - \mathcal{L}_{GG} G\sigma) v_b, \tag{80}$$

$${M_b}^d v_d = u_a u^a (2\mathcal{L}_{FG}(1 + F\sigma) - 4\mathcal{L}_{GG} G\sigma) u_b + u_a u^a (2\mathcal{L}_F \sigma - \mathcal{L}_{GG}(1 + F\sigma) + 2\mathcal{L}_{FG} G\sigma) v_b. \tag{81}$$

The first two equations (78) and (79) demonstrate that ${M_b}^d$ leaves the two-space spanned by w_d and p_d invariant and that it has a one-dimensional kernel on this two-space. The last statement follows from the fact that w_d is not in the kernel: It is mapped onto a covector ${M_b}^d w_d$ that is nonzero if $\sigma = 0$ (because then it is a nonzero multiple of p_b) and also if $\sigma \neq 0$ (because then it has a nonzero component in the direction of w_b). The other two equations (80) and (81) demonstrate that the two-space spanned by u_d and v_d is left invariant as well. On this two-space the matrix ${M_b}^d$ must have a one-dimensional or two-dimensional kernel because the eikonal equation requires that the kernel of the full matrix ${M_b}^d$ is at least two-dimensional. By (80) and (81), a covector $\alpha u_b + \beta v_b$ is in the kernel if and only if (75) holds with (76). The determinant of the matrix (76) vanishes as a consequence of the eikonal equation. Clearly, a (2×2)-matrix has a two-dimensional kernel if and only if it is the zero matrix. The matrix (76) is the zero matrix if and only if the symmetric matrix

$$\frac{1}{u_a u^a D} \begin{pmatrix} {m_1}^1 & {m_1}^2 \\ {m_2}^1 & {m_2}^2 \end{pmatrix} \begin{pmatrix} 1+F\sigma & -G\sigma \\ -G\sigma & 1 \end{pmatrix} = \frac{2\mathcal{L}_F \sigma}{D} \begin{pmatrix} 1+F\sigma & -G\sigma \\ -G\sigma & 1 \end{pmatrix} + \begin{pmatrix} -4\mathcal{L}_{FF} & 2\mathcal{L}_{FG} \\ 2\mathcal{L}_{FG} & -\mathcal{L}_{GG} \end{pmatrix} \tag{82}$$

is the zero matrix, where $D = 1 + F\sigma - G^2\sigma^2$. By comparison with part (d) of Proposition IV.1 we see that this is the case if and only if the two optical metrics coincide. As we assume that $u_a u^a \neq 0$ this is true if and only if $p_a = \nabla_a S$ is a solution of multiplicity two. □

With these results at hand it is now easy to evaluate the polarization condition. We do this first for solutions of multiplicity two.

Proposition IV.6. Let $p_a = \nabla_a S$ be a solution of multiplicity two to the eikonal equation, i.e., $p_a p^a + \sigma_+ u_a u^a = 0$ and $p_a p^a + \sigma_- u_a u^a = 0$. Then the two transport vector fields $K_+^a = p^a - \sigma_+ F^{ab} u_b$ and $K_-^a = p^a - \sigma_- F^{ab} u_b$ are linearly dependent. The polarization condition ${M_b}^d a_d^{(11)} = 0$ is equivalent to $K_\pm^d a_d^{(11)} = 0$ (which holds with one sign if and only if it holds with the other sign), i.e., it restricts $a_d^{(11)}$ to a three-dimensional subspace which contains p_a.

Proof. If $u_a u^a = 0$, this follows from Proposition IV.4. If $u_a u^a \neq 0$ it follows from Proposition IV.5. □

We now prove the analogous statement for solutions of multiplicity one.

Proposition IV.7. Let $p_a = \nabla_a S$ be a solution of multiplicity one to the eikonal equation, i.e., $p_a p^a + \sigma u_a u^a = 0$ with $\sigma = \sigma_+$ or $\sigma = \sigma_-$ but not with both.

Then the polarization condition $M_b{}^d a_d^{(11)} = 0$ is true if and only if $a_b^{(11)} = \alpha u_b + \beta v_b + \gamma p_b$ where α and β satisfy (75) with (76). This condition restricts $a_b^{(11)}$ to a two-dimensional subspace that contains p_b.

Proof. This is an immediate consequence of Proposition IV.5. □

We summarize the results of this section in the following way. For every solution $p_a = \nabla_a S$ to the eikonal equation the polarization condition requires that $a_b^{(11)}$ satisfies $K^b a_b^{(11)} = 0$ where K^b is the corresponding transport vector field. This may be interpreted as a transversality condition. For a solution of multiplicity two there is no additional restriction, i.e., $a_b^{(11)}$ is confined to a three-dimensional subspace that contains p_b. By contrast, for a solution of multiplicity one the polarization condition restricts $a_b^{(11)}$ to a two-dimensional space that contains p_b, i.e., it fixes the polarization plane (the plane spanned by $a_b^{(11)}$ and p_b) uniquely.

V. EVALUATION OF THE FIRST-ORDER FIELD EQUATION

We now turn to the first-order field equation which gives us the two conditions (36) and (37). We can write them, in a slightly more compact form, as

$$0 = \nabla^a(\chi_{ab}{}^{cd} p_c a_d^{(11)}) + p^a \chi_{ab}{}^{cd} \nabla_c a_d^{(11)} + i M_b{}^d a_d^{(21)}, \quad (83)$$

$$0 = M_b{}^d a_d^{(22)} - \psi_{ab}{}^{cdef} p^a p_c a_d^{(11)} a_f^{(11)}. \quad (84)$$

We want to determine what kind of information these equations give us on the polarization plane spanned by $a_b^{(11)}$ and p_b.

We know from the preceding section that for a solution of multiplicity one this plane is already uniquely fixed at the zeroth-order level, so the first-order equations cannot give us any additional information on this plane. One just has to check for consistency, i.e., one has to verify that the sum of the first two terms in (83) is in the image space of $M_b{}^d$ and that the second term in (84) is in the image space of $\psi_{ab}{}^{cdef}$. Then (83) and (84) give us polarization conditions on $a_d^{(21)}$ and $a_d^{(22)}$. We have already emphasized that (84) is not in general satisfied by $a_d^{(22)} = 0$, i.e., that frequency doubling has to be taken into account if the field equation should hold at first order, and that at the next order in general also a nonzero $a^{(20)}$ is needed.

As in this paper we will be satisfied with determining the potential up to first order, there is nothing else to be done for solutions of multiplicity one. Therefore, in the following we will restrict ourselves to solutions of multiplicity two. We know from the preceding section that then $a_d^{(11)}$ is restricted at the zeroth-order level only by the condition $K^d a_d^{(11)} = 0$. This condition restricts the polarization plane to a three-dimensional space, i.e., it still allows the polarization plane to arbitrarily rotate along a ray. We will now demonstrate that the first-order equation (83) gives us a transport law which uniquely determines the polarization plane along a ray if it is given at one point of this ray. We will consider first solutions of multiplicity two with $u_a u^a = 0$ and then with $u_a u^a \neq 0$.

A. Transport equation in the case $u_a u^a = 0$

For a solution of multiplicity two with $u_a u^a = 0$ the rays are lightlike geodesics not only with respect to each of the two optical metrics but also with respect to the spacetime metric. (The affine parametrizations are in general different.) For such a solution we have u_a and v_a parallel to p_a and the matrix $M_b{}^d$ projects onto the line spanned by p_b, recall Proposition IV.4. As a consequence, (83) reduces to

$$\begin{aligned} &4\mathcal{L}_F p^a \nabla_a a_b^{(11)} + 2\nabla_a(\mathcal{L}_F p^a) a_b^{(11)} \\ &\quad - \nabla^a \mathcal{L}_G \varepsilon_{ab}{}^{cd} p_c a_d^{(11)} = \psi p_b \end{aligned} \quad (85)$$

where ψ is an undetermined scalar function. Recall from Proposition IV.4 that in the case at hand the two transport vector fields K_+^a and K_-^a are multiples of p^a, i.e., that p^a is tangent to the rays. Therefore, (85) gives us a first-order ordinary differential equation for $a_b^{(11)}$ along each ray. As ψ is arbitrary, for each initial condition this differential equation has a solution that is unique up to a multiple of p_b. In other words, (85) gives us a unique transport law for the polarization plane.

If $\nabla^a \mathcal{L}_G \varepsilon_{ab}{}^{cd} p_c a_d^{(11)} = 0$, we read from (85) that the polarization plane is parallel with respect to the transport law defined by the Levi-Civita derivative of the spacetime metric, as it is in the standard Maxwell vacuum theory; in general, however, in a theory of the Plebański class a background field with nonconstant $\mathcal{L}_G$ produces a rotation of the polarization plane. This gives us a new experimental test of this type of theories in situations where the rays behave as in the standard vacuum Maxwell theory but the polarization plane does not. We will exemplify this with the Born-Infeld theory in the next section.

B. Transport equation in the case $u_a u^a \neq 0$

We now consider a solution of multiplicity two with $u_a u^a \neq 0$. For such solutions we know from Proposition IV.5 that the matrix $M_b{}^d$ has a three-dimensional kernel spanned by p_b, u_b, and v_b, i.e., that $a_d^{(11)}$ is of the form

$$a_d^{(11)} = \alpha u_d + \beta v_d + \gamma p_d. \quad (86)$$

By the same token, as the matrix $M_b{}^d$ is self-adjoint with respect to the spacetime metric, (83) is true with some $a_d^{(21)}$ if and only if the equation

$$0 = z^b\{\nabla^a(\chi_{ab}{}^{cd}p_c a_d^{(11)}) + p^a\chi_{ab}{}^{cd}\nabla_c a_d^{(11)}\} \quad (87)$$

is true for $z^b = p^b$, $z^b = u^b$, and $z^b = v^b$. It is easy to check that for $z^b = p^b$ the equation is identically satisfied, for all $a_d^{(11)}$ of the form (86). Therefore, we only have to consider it for $z^b = u^b$ and $z^b = v^b$.

To that end, we recall that a solution of multiplicity two with $u_a u^a \neq 0$ exists only if the two optical metrics coincide. In the following we write σ for $\sigma_+ = \sigma_-$ and K^a for $K_+^a = K_-^a$. If we express $\mathcal{L}_{FF}$, $\mathcal{L}_{FG}$, and $\mathcal{L}_{GG}$ with the help of part (d) of Proposition IV.1, we see that $\chi_{ab}{}^{cd}$ can be written as

$$\begin{aligned}\chi_{ab}{}^{cd} &= \mathcal{L}_G\varepsilon_{ab}{}^{cd} - 2\mathcal{L}_F(\delta_a^c\delta_b^d - \delta_a^d\delta_b^c)\\ &\quad - \frac{2\mathcal{L}_F\sigma}{D}F_{ab}((1+\sigma F)F^{cd} - \sigma G{}^*F^{cd})\\ &\quad + \frac{2\mathcal{L}_F\sigma}{D}{}^*F_{ab}(\sigma G F^{cd} - {}^*F^{cd}))\end{aligned} \quad (88)$$

where $D = 1 + \sigma F - \sigma^2 G^2$.

If we insert this expression and (86) into (87) with $x^b = u^b$ and with $x^b = v^b$, we get after some lengthy algebra the two equations

$$4\mathcal{L}_F u_b u^b K^a\nabla_a\alpha = a\alpha + b\beta, \quad (89)$$

$$4\mathcal{L}_F u_b u^b K^a\nabla_a\beta = a\beta - b\alpha, \quad (90)$$

where

$$a = -2\nabla_a(\mathcal{L}_F u_c u^c K^a), \quad (91)$$

$$\begin{aligned}b &= \nabla^a\mathcal{L}_G\varepsilon_{abcd}p^b v^c u^d\\ &\quad + 2\mathcal{L}_F p^b(p^a p^x - p_e p^e g^{ac})({}^*F_a{}^d\nabla_c F_{db} + F_a{}^d\nabla_c{}^*F_{db}).\end{aligned} \quad (92)$$

These equations determine the change of α and β and, thus, of the polarization plane, along each ray. In particular, b determines the rotation of the polarization plane with respect to the basis covectors u_b and v_b which are orthogonal to each other, but not parallely transported, with respect to the optical metric.

VI. EXAMPLE: BORN-INFELD THEORY

As an example, we consider the transport law for the polarization plane in the Born-Infeld theory [1] where the Lagrangian is given by (10). In this case, for any background field, the two optical metrics coincide,

$$\sigma_+ = \sigma_- = \frac{-1}{b_0^2 + F}, \quad (93)$$

$$\tilde{g}_+^{ab} = \tilde{g}_-^{ab} = g^{ab} - \frac{F^{ac}F^b{}_c}{b_0^2 + F} \quad (94)$$

so there is no birefringence and every solution to the eikonal equation is a solution of multiplicity two.

We assume that the underlying spacetime is the Minkowski spacetime with standard inertial coordinates $(x^0 = ct, x^1, x^2, x^3)$, i.e., $g_{ab} = \eta_{ab}$ where $(\eta_{ab}) = \mathrm{diag}(-1,1,1,1)$. As the background field we choose the superposition of a time-dependent electric field and a constant magnetic field, both in x^3 direction; so the only nonvanishing components of the field strength tensor are

$$F_{03} = -F_{30} = \frac{E(t)}{c}, \qquad F_{12} = -F_{21} = B_0. \quad (95)$$

Note that the homogeneous Maxwell equation is indeed satisfied, $\varepsilon^{abcd}\nabla_b F_{cd} = 0$. For this electromagnetic field,

$$p_a = \nabla_a S, \quad S(x^0,x^1,x^2,x^3) = \frac{1}{c}(x^3 - x^0) = \frac{x^3}{c} - t \quad (96)$$

is a principal null covector field, $p_b p^b = 0$ and

$$u_a = F_{ab}p^b = -E(t)p_a, \quad (97)$$

$$v_a = {}^*F_{ab}p^b = -cB_0 p_a, \quad (98)$$

so the eikonal equation is satisfied with $u_a u^a = 0$. The transport vector field K^a is proportional to p^a, i.e., the rays are straight lightlike lines in x^3 direction. Note that they are lightlike not only with respect to the spacetime metric; they are lightlike geodesics also with respect to the optical metric. Whereas K^a is adapted to an affine parameter with respect to the optical metric, p^a is adapted to the parametrization with the time coordinate t which is an affine parameter with respect to the spacetime metric.

The amplitude $a_b^{(11)}$ must be orthogonal to p_c with respect to the spacetime metric, so we may write it in the form

$$a_b^{(11)} = \zeta(\delta_b^1\cos\varphi + \delta_b^2\sin\varphi) + \gamma p_b \quad (99)$$

with scalar coefficients ζ and γ and an angle φ which gives, at each point along the ray, the rotation of the polarization plane with respect to the (x^1, x^2) basis vectors which are parallel with respect to the Levi-Civita derivative of the metric along the ray.

For the electromagnetic field (95), the partial derivatives of the Lagrangian are

$$\mathcal{L}_F = \frac{-1}{2\sqrt{1 + \frac{B_0^2}{b_0^2}}\sqrt{1 - \frac{E(t)^2}{c^2 b_0^2}}}, \quad (100)$$

$$\mathcal{L}_G = \frac{-B_0 E(t)}{cb_0^2\sqrt{1 + \frac{B_0^2}{b_0^2}}\sqrt{1 - \frac{E(t)^2}{c^2 b_0^2}}}. \quad (101)$$

With these results, inserting (99) into (85) yields

$$\dot{\varphi} = \frac{-B_0 \frac{dE(t)}{dt}}{2cb_0^2(1 - \frac{E(t)^2}{c^2 b_0^2})} \tag{102}$$

where the overdot means derivative with respect to t along the ray, $p^a \nabla_a \varphi = \dot{\varphi}$. We see that the time-dependence of $\mathcal{L}_G$ produces a rotation of the polarization plane. Integrating (102) from $t_1 = 0$ with $E(t_1) = 0$ to a time t_2 with $E(t_2) = E_0$ gives

$$\Delta\varphi = \frac{B_0}{2b_0} \mathrm{arctanh}\left(\frac{E_0}{cb_0}\right) = \frac{B_0 E_0}{2cb_0^2}\left(1 + O\left(\frac{E_0^2}{c^2 b_0^2}\right)\right). \tag{103}$$

In principle, this can be utilized for a new laboratory test of the Born-Infeld theory. If in the considered constellation with strong fields E_0 and B_0 no rotation of the polarization plane is detected, this gives a lower bound on b_0. However, the effect on the polarization plane is so small that with present-day technology this test of the Born-Infeld theory is not yet competitive with other tests. E.g., in Ref. [16] we have seen that with interferometric methods one could find a bound on b_0 in the order of $b_0 \gtrsim 7 \times 10^7$ T which corresponds to $b_0 \gtrsim 7 \times 10^{11} \sqrt{\mathrm{g}}/(\sqrt{\mathrm{cm}}\,\mathrm{s})$ in Gaussian units. Assuming that a rotation of the polarization plane by one arcminute could be measured, $\Delta\varphi \approx 3 \times 10^{-4}$ rad, we would need electric and magnetic fields of

$$\left| B_0 \frac{E_0}{c} \right| \approx 3 \times 10^{12}\ \mathrm{T}^2 \tag{104}$$

for being competitive with the interferometric test. This is not reachable in a laboratory experiment in the near future.

VII. CONCLUSIONS

It was the main purpose of this paper to derive the transport law of the polarization plane in nonlinear vacuum electrodynnamics. We have done this, on an unspecified general-relativistic spacetime, for a theory of the Plebański class and an electromagnetic background field which were arbitrary except for some nondegeneracy conditions. To that end we have utilized an approximate-plane-harmonic-wave ansatz which takes the generation of higher harmonics and frequency rectification into account. According to this ansatz, the electromagnetic field is written as an asymptotic series with respect to a parameter α where the limit $\alpha \to 0$ refers to sending the frequency to infinity. We have seen that the generalized Maxwell equations have to be solved to zeroth and to first order with respect to α for determining the transport law of the polarization plane in lowest nontrivial order.

When considering the generalized Maxwell equations to zeroth order, we have rederived the known result that, for every theory of the Plebański class and every background field that satisfy the assumed nondegeneracy conditions, there are two optical metrics which are both of Lorentzian signature. For a solution of the zeroth-order equations one needs a scalar function, the so-called eikonal function S, whose gradient $p_b = \nabla_b S$ is lightlike with respect to at least one of the two optical metrics, and a covector field, $a_b^{(11)}$, which has to satisfy an algebraic equation known as the zeroth-order polarization condition.

We have seen that two cases have to be distinguished. The first case is that of a solution of multiplicity one, i.e., the case that $p_b = \nabla_b S$ is lightlike with respect to only one of the optical metrics. Then the polarization plane (i.e., the plane spanned by $a_b^{(11)}$ and p_b) is uniquely determined by the zeroth-order polarization condition. The first-order equations give no additional information on the polarization plane and have to be checked only for consistency. The second case is that of a solution of multiplicity two, i.e., the case that $p_b = \nabla_b S$ is lightlike with respect to both optical metrics. Then the zeroth-order polarization condition allows an arbitrary rotation of the polarization plane along each ray. However, the generalized Maxwell equations at first order give us a transport equation which determines the polarization plane uniquely along a ray if it is given at one point of this ray. This transport law has a fairly simple form in the case that p_b is a principal null covector of the electromagnetic background field, $F_a{}^b p_b \sim p_a$, see Sec. V A. It is much more awkward if this is not the case, see Sec. V B.

We have exemplified the general results with the Born-Infeld theory. In this theory the two optical metrics coincide, i.e., all solutions of the eikonal equation are of multiplicity two. We have considered a particular solution where p_b is a principal null covector of a background field on Minkowski spacetime. In this example, the rays are straight lightlike lines, i.e., the light propagation is the same as in the standard Maxwell vacuum theory. However, the behavior of the polarization plane is different: Whereas in the standard Maxwell theory it is parallely transported along each ray, here it rotates with respect to a parallely transported plane by an angle $\Delta\varphi$. This is a feature that could be observed in a laboratory experiment, sometimes in the future, when sufficiently strong fields are available.

ACKNOWLEDGMENTS

We thank Gerold Schellstede for helpful discussions. V. P. is grateful to Deutsche Forschungsgemeinschaft for financial support under Grant No. LA 905/14-1. Moreover, C. L. and V. P. gratefully acknowledge support from the Deutsche Forschungsgemeinschaft within the Research Training Group 1620 "Models of Gravity." A. M. acknowledges support from DAAD.

[1] M. Born and L. Infeld, Proc. R. Soc. A **144**, 425 (1934).
[2] A. A. Tseytlin, in *The Many Faces of the Superworld: Yuri Golfand Memorial Volume*, edited by M. Shifman (World Scientific, Singapore, 1999), p. 417.
[3] W. Heisenberg and H. Euler, Z. Phys. **98**, 714 (1936).
[4] J. Plebański, *Lectures on Non-linear Electrodynamics* (Nordita, Copenhagen, 1970).
[5] G. Boillat, J. Math. Phys. (N.Y.) **11**, 941 (1970).
[6] E. Ayón-Beato and A. García, Phys. Rev. Lett. **80**, 5056 (1998).
[7] E. Ayón-Beato and A. García, Phys. Lett. B **464**, 25 (1999).
[8] D. L. Burke, R. C. Field, G. Horton-Smith, J. E. Spencer, D. Walz, S. C. Berridge, W. M. Bugg, K. Shmakov, A. W. Weidemann, C. Bula *et al.*, Phys. Rev. Lett. **79**, 1626 (1997).
[9] F. Della Valle, A. Ejlli, U. Gastaldi, G. Messineo, E. Milotti, R. Pengo, G. Ruoso, and G. Zavattini, Eur. Phys. J. C **76**, 24 (2016).
[10] R. P. Mignani, D. Testa, V. González Caniulef, R. Taverna, R. Turolla, S. Zane, and K. Wu, Mon. Not. R. Astron. Soc. **465**, 492 (2017).
[11] D. Boer and J.-W. van Holten, arXiv:hep-ph/0204207.
[12] B. Döbrich and H. Gies, Europhys. Lett. **87**, 21002 (2009).
[13] G. Zavattini and E. Calloni, Eur. Phys. J. C **62**, 459 (2009).
[14] H. Grote, Phys. Rev. D **91**, 022002 (2015).
[15] V. I. Denisov, I. V. Krivchenkov, and N. V. Kravtsov, Phys. Rev. D **69**, 066008 (2004).
[16] G. O. Schellstede, V. Perlick, and C. Lämmerzahl, Phys. Rev. D **92**, 025039 (2015).
[17] R. Ferraro, Phys. Rev. Lett. **99**, 230401 (2007).
[18] T. Dereli and R. Tucker, Europhys. Lett. **89**, 20009 (2010).
[19] M. Novello, V. A. De Lorenci, J. M. Salim, and R. Klippert, Phys. Rev. D **61**, 045001 (2000).
[20] Y. N. Obukhov and G. F. Rubilar, Phys. Rev. D **66**, 024042 (2002).
[21] G. O. Schellstede, V. Perlick, and C. Lämmerzahl, Ann. Phys. (Berlin) **528**, 738 (2016).
[22] F. Abalos, F. Carrasco, É. Goulart, and O. Reula, Phys. Rev. D **92**, 084024 (2015).
[23] M. Kline and I. W. Kay, *Electromagnetic Theory and Geometrical Optics* (Interscience, New York, 1965).
[24] V. Perlick, J. Math. Phys. (N.Y.) **52**, 042903 (2011).
[25] F. W. Hehl and Y. Obukhov, *Foundations of Classical Electrodynamics* (Birkhäuser, Basel, 2003).
[26] Y. Itin, Phys. Rev. D **76**, 087505 (2007).
[27] Y. Itin, J. Phys. A **42**, 475402 (2009).