Optical Kerr effect in vacuum

Scott Robertson [*]

LAL, Univ. Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, Orsay, France

From an effective field theory of electromagnetism in vacuum including all lowest-order nonlinear terms consistent with Lorentz invariance and locality of photon-photon interactions, we derive an effective-medium description of strong background fields as regards their influence on a weak probe. We mainly consider as background a pump beam with well-defined wave vector and polarization. This leads us to define a nonlinear index of vacuum which, in the Euler-Heisenberg model derived from QED, has an optimal value of 1.555×10^{-33} cm^2/W for a linearly polarized pump as seen by a counterpropagating, orthogonally polarized probe. We further generalize the model to include coupling to an axion field. In the limit where the axion mass is much smaller than the typical photon energy, this yields dispersive corrections, and the axionic signature is found to be greatly enhanced for a circularly polarized pump as compared to a linearly polarized one. The formalism here presented points to a simplification of the DeLLight experiment [Sarazin *et al.*, Eur. Phys. J. D **70**, 13 (2016)] aiming to measure the deflection of a probe by a tightly focused laser pulse.

I. INTRODUCTION

In media, the dependence of optical properties on the electric or magnetic field has been known since the time of Faraday in the mid 19th century [1], though it has gained particular prominence in the last sixty years with the availability of high-intensity lasers and the subsequent development of nonlinear optics [2–6]. Such field-dependent behavior arises from the nontrivial response of bound charges and currents within the medium, leading to a range of nonlinear effects which have been extensively studied in the literature. Since at optical frequencies the intensity associated with a single photon is very low, reaching the nonlinear regime requires an intense field comprising so many photons that it can be treated classically. From the point of view of a weak probe, a system comprising a strong background field and a medium can be treated as a single "dressed" medium, with the background field contributing to the total refractive index [5,6]. In typical dielectric media whose molecules possess inversion symmetry, the refractive index change is proportional to the square of the electric field, a phenomenon usually referred to as the *Kerr effect* after its discoverer John Kerr [7]. When the index change is engendered by intense light (rather than static fields), this *optical Kerr effect* allows the assignment of a nonlinear index n_2 to the medium [5,6] such that the total refractive index includes a term proportional to the intensity I of the wave:

$$n(I) = n_0 + n_2 I, \tag{1}$$

where n_0 is the "bare" index of the medium in the absence of strong fields. Though n_2 depends on the precise configuration (such as the relative polarization between the probe and the background), values typically range from 10^{-16} to 10^{-14} cm^2/W (see, e.g., Table 4.1.2 of [5]).

In vacuum, by contrast, classical electrodynamics is a linear theory, where n_2 is exactly zero. By analogy with the situation in media, we may ask whether this apparent linearity is only a low-field approximation, i.e., whether the field equations might become nonlinear when the fields are strong enough. Indeed, the standard model already answers in the affirmative; in particular, quantum electrodynamics (QED) allows photon-photon scattering mediated by virtual electron-positron pairs, which play a role analogous to that of bound charges in media. In the long-wavelength limit, this yields a nonlinear effective field theory for the electromagnetic field, the Lagrangian of which was derived by Euler, Kockel, and Heisenberg [8–10]. In principle, however, this is but one way in which nonlinearities could be generated: there may well be as-yet-unidentified particles (such as axions [11–14]) coupling to photons and thereby contributing to the effective nonlinear response; alternatively, there may be higher-order corrections to the classical electromagnetic sector of the Lagrangian (such as proposed by Born and Infeld [15]).

It is thus of interest to experimentally probe nonlinear electrodynamics (NLED) in vacuum, in order to test our current predictions and potentially rule out alternative models [16,17]. While there have been several experiments in the high-energy photon regime which also tend to involve charged particles of some kind (a recent example is provided by the heavy ion collisions observed at the LHC [18]; see Ref. [16] for many others), the direct elastic scattering of real photons has not yet been observed, and the low-energy photon regime—where, as in media, the strong background field is classical and nonlinearities can be treated as field-dependent contributions to the total refractive index—remains relatively unexplored. The most sensitive tests of low-energy NLED to date are those of the BMV [19,20] and PVLAS [21,22] experiments, which aim to detect the birefringence induced by a strong magnetic field perpendicular to the direction of the probe wave. These have not yet reached the sensitivity

────────
[*]robertson@lal.in2p3.fr

required to test effects on the order of those predicted by QED. An alternative, complementary approach is to directly exploit the intensity dependence of the refractive index by using strong background fields to deflect the trajectory of a probe wave. This was attempted by Jones in 1960 [23,24] using a static magnetic field, while the recently proposed DeLLight experiment [25] aims to observe such a deflection using the much greater intensities within a tightly focused laser pulse.

Inspired by such proposals, we develop in this paper a theoretical framework for describing the lowest-order nonlinear interactions between electromagnetic waves in vacuum. We are particularly concerned with the generality of the approach, which covers two aspects. First, rather than restricting ourselves to the Euler-Heisenberg model describing the nonlinearities induced by QED, we consider instead a generalized model (in the spirit of Plebański [26] and Boillat [27]) which is consistent with Lorentz invariance and locality of effective photon-photon interactions. As far as we are aware, this generalized Lagrangian was previously adopted only in [28] (for the particular case of static background fields), though the approach used in [29] is reminiscent of it. Moreover, being mindful of proposals for the detection of axions [30–34], we also allow a relaxation of the assumption of local photon-photon interactions through coupling to an axionlike field. This allows an explicit demonstration that, while the inclusion of the axion coupling would lead to a straightforward renormalization of the model parameters if the axion mass were large enough, the breakdown of locality of photon-photon interactions when this is not the case yields dispersive corrections to the behavior of a probe wave. It also, in the case of an elliptically polarized background, induces some elliptical birefringence that is completely absent in the purely local theory.

The second aspect in which this paper generalizes the standard treatment of NLED concerns its explicitization of the effective medium engendered by an arbitrary configuration of strong background fields. That is, assuming an explicit decomposition of the total fields into those of a strong background and a weak probe, the field equations can be linearized in those of the latter, in which case the dressed vacuum comprising the vacuum and strong background fields together behaves as an optical medium in its own right. To this effective medium is assigned a set of well-defined susceptibility tensors. This description is more intuitive than the manifestly covariant formalism commonly used in the literature (as in, e.g., [35]). The general expression for the susceptibility tensors encompasses any background field configuration, describing both static and propagating fields; in the latter case, elliptical polarization is included in a relatively straightforward manner. The behavior of a probe wave is completely determined by the optical properties encoded in the effective susceptibilities. In particular, they allow a straightforward determination of the nonlinear index n_2 of vacuum by analogy with Eq. (1). While a few previous works have given expressions for n_2 of vacuum [36,37], they have worked with the Euler-Heisenberg model[1] rather than the generalized one adopted here, and

explicit values given [5,38] have only been estimates that lie somewhat below the optimized value[2].

The paper is organized as follows. In Sec. II we lay the theoretical foundations by specifying the field normalization and Lagrangian we shall use, the latter being subject to the restrictions of Lorentz invariance and local interactions mentioned above. We also indicate how particularly important models of NLED fit into this generalized framework. In Sec. III, we develop the effective-medium description by explicitly separating the total fields into a strong background and a weak probe, then linearizing the wave equations in the fields of the latter. We pay particular attention to plane probe waves and their eigenstates within the effective medium, i.e., their refractive indices and polarizations, and we use this formalism to derive some known results in the case of static background fields. We consider an intense pump wave as a background in Sec. IV, showing how elliptical polarization of the pump can be taken into account and deriving the nonlinear index of vacuum by analogy with its definition in standard optical media. In Sec. V, we generalize the Lagrangian to include coupling to an axion field of arbitrary mass, yielding an effective theory of NLED which is nonlocal and thus characterized by dispersion. The analysis is carried through as before, and the key differences are emphasized. We summarize our findings and conclude in Sec. VI.

II. PRELIMINARIES

We begin by establishing some theoretical foundations. First, we introduce a convenient normalization for the electromagnetic fields which simplifies the writing of many equations. We then state and discuss the most general form of the Lagrangian for the fields given a set of reasonable constraints. This Lagrangian has three free parameters, and we finish this section by identifying the subsets of these parameters corresponding to two particularly important models of NLED: those of Euler-Heisenberg (EH) and Born-Infeld (BI).

A. Field normalization and equations

We work in Minkowski (flat) space, so that a three-vector description of the electromagnetic fields may be straightforwardly applied. To avoid overuse of the fundamental constants ϵ_0 and μ_0 (respectively, the permittivity and permeability of free space), it is convenient to use the following rescaled definitions for the electric and magnetic fields:

$$\mathbf{E} = \sqrt{\epsilon_0}\,\mathbf{E}_{\mathrm{SI}}, \quad \mathbf{D} = \frac{\mathbf{D}_{\mathrm{SI}}}{\sqrt{\epsilon_0}}, \quad \mathbf{B} = \frac{\mathbf{B}_{\mathrm{SI}}}{\sqrt{\mu_0}}, \quad \mathbf{H} = \sqrt{\mu_0}\,\mathbf{H}_{\mathrm{SI}},$$

$$(2)$$

angle between the propagation directions of pump and probe, and even elliptical polarization, though how arbitrary polarization states enter is rather implicit.

[2]Equation (13.8.9) of [5] gives only an estimate for n_2 as it neglects the tensorial nature of the response and is derived from a nonlinear permittivity that does not apply to weak probe waves on a strong background. An order of magnitude is extracted from numerical simulations in [38], but not in the optimized scenario with the probe and background orthogonally polarized and counterpropagating.

---

[1]In the Euler-Heisenberg context, Eq. (2.21) of [37] gives a general result for propagating background fields that includes an arbitrary

where the subscript "SI" indicates the corresponding fields expressed in SI units. With these definitions, each of the fields $\mathbf{E}$, $\mathbf{D}$, $\mathbf{B}$, and $\mathbf{H}$ has exactly the same units (the square root of an energy density), and the Maxwell equations in the absence of free charges and currents take the following form:

$$\nabla \cdot \mathbf{B} = 0, \quad \nabla \times \mathbf{E} + \partial_{ct}\mathbf{B} = 0, \tag{3a}$$

$$\nabla \cdot \mathbf{D} = 0, \quad \nabla \times \mathbf{H} - \partial_{ct}\mathbf{D} = 0, \tag{3b}$$

where $c = 1/\sqrt{\epsilon_0\mu_0}$ is the speed of light in vacuum. Equations (3a) are automatically satisfied when $\mathbf{E}$ and $\mathbf{B}$ are defined in the standard covariant formulation as components of the antisymmetric tensor $F_{\mu\nu}$, itself defined as the exterior derivative of the four-potential A_μ [39]; in effect they are consistency conditions that allow such a writing to take place. By contrast, Eqs. (3b) are a convenient writing of the Euler-Lagrange equations found by extremizing the action with respect to variations of A_μ, where $\mathbf{D}$ and $\mathbf{H}$ are defined as

$$\mathbf{D} \doteq \frac{\partial \mathcal{L}}{\partial \mathbf{E}}, \quad \mathbf{H} \doteq -\frac{\partial \mathcal{L}}{\partial \mathbf{B}}. \tag{4}$$

While wave equations (3) can only be fully solved once the constitutive equations relating $\mathbf{D}$ and $\mathbf{H}$ to $\mathbf{E}$ and $\mathbf{B}$ are specified, Eqs. (4) indicate that these relations are fully determined once the Lagrangian $\mathcal{L}(\mathbf{E}, \mathbf{B})$ is.

B. Parametrized Lagrangian for NLED

There are only two scalar quantities invariant under proper orthochronous Lorentz transformations (i.e., those continuously connected to the identity, requiring no spatial reflection or time reversal) which can be constructed from the electromagnetic fields alone [39]:

$$\mathcal{F} \doteq -\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} = \tfrac{1}{2}(E^2 - B^2), \quad \mathcal{G} \doteq -\tfrac{1}{4}F_{\mu\nu}\widetilde{F}^{\mu\nu} = \mathbf{E} \cdot \mathbf{B}, \tag{5}$$

where $\widetilde{F}^{\mu\nu} = \tfrac{1}{2}\epsilon^{\mu\nu\alpha\beta}F_{\alpha\beta}$ is the Hodge dual of $F_{\mu\nu}$, $\epsilon^{\mu\nu\alpha\beta}$ being the completely antisymmetric Levi-Civita symbol with $\epsilon^{0123} = 1$. Therefore, a Lorentz invariant Lagrangian containing the electromagnetic fields alone must depend only on $\mathcal{F}$ and $\mathcal{G}$, and if we further assume that the Lagrangian is spatially and temporally *local* (i.e., interactions are purely of the "contact" type), then its value at a given point of spacetime is straightforwardly a function of $\mathcal{F}$ and $\mathcal{G}$ at the same point: $\mathcal{L}(x) = \mathcal{L}(\mathcal{F}(x), \mathcal{G}(x))$. Such a Lagrangian for NLED is said to be of the Plebański class [26]. We recognize $\mathcal{F}$ itself as the standard Lagrangian for electrodynamics in vacuum; it yields the trivial constitutive relations $\mathbf{D} = \mathbf{E}$ and $\mathbf{H} = \mathbf{B}$ when plugged into Eqs. (4), whence we recover the usual Maxwell equations in vacuum when inserted into Eqs. (3). Moreover, if $\mathcal{G}$ is added to the Lagrangian with some constant coefficient, it is straightforward to show that it has no effect on the field equations,[3] and we are thus free to exclude the occurrence of such a term. We thereby conclude that, at lowest order, the Lagrangian is simply $\mathcal{F}$.

In the weakly nonlinear regime, the Lagrangian can be expanded in powers of $\mathcal{F}$ and $\mathcal{G}$ [16,17]. The first nonlinearities will be due to terms quadratic in $\mathcal{F}$ and $\mathcal{G}$ (i.e., quartic in the fields), and we parametrize their contribution as follows:

$$\mathcal{L} = \mathcal{F} + \delta_1 \tfrac{1}{2}\mathcal{F}^2 + \delta_2 \tfrac{1}{2}\mathcal{G}^2 + \delta_3 \mathcal{F}\mathcal{G} + \cdots. \tag{6}$$

The "post-Maxwellian" parameters [29] δ_1, δ_2, and δ_3 have units of inverse energy density. They are not completely arbitrary: we will later see that they must satisfy certain inequalities in order for causality to be respected. The term proportional to δ_3 is often neglected as it breaks invariance under spatial reflection ($\mathcal{P}$) or time reversal ($\mathcal{T}$) transformations, which preserve the sign of $\mathcal{F}$ while causing $\mathcal{G}$ to flip sign. Although QED is invariant under $\mathcal{P}$ and $\mathcal{T}$, the full standard model is not, $\mathcal{P}$-invariance being broken by weak interactions [40,41]. Therefore, if we wish to include possible deviations from QED in our description, there is no fundamental reason why δ_3 should vanish, and we keep it here for the sake of completeness.

Substituting into Eqs. (4), we find the nonlinear constitutive relations

$$\mathbf{D} = \frac{\partial \mathcal{L}}{\partial \mathcal{F}}\mathbf{E} + \frac{\partial \mathcal{L}}{\partial \mathcal{G}}\mathbf{B}, \quad \mathbf{H} = \frac{\partial \mathcal{L}}{\partial \mathcal{F}}\mathbf{B} - \frac{\partial \mathcal{L}}{\partial \mathcal{G}}\mathbf{E}, \tag{7}$$

where

$$\frac{\partial \mathcal{L}}{\partial \mathcal{F}} = 1 + \delta_1\mathcal{F} + \delta_3\mathcal{G} + \cdots, \quad \frac{\partial \mathcal{L}}{\partial \mathcal{G}} = \delta_2\mathcal{G} + \delta_3\mathcal{F} + \cdots. \tag{8}$$

The first nonlinear terms in $\mathbf{D}$ and $\mathbf{H}$ are thus of third order in the fields, and can be thought of as defining the third-order susceptibilities of vacuum by analogy with media. There is a key difference, of course: nonlinearities in media tend to be dominated by a single third-order susceptibility, governing the contribution to $\mathbf{D}$ which is cubic in the electric field.[4] In vacuum, however, the requirement of Lorentz invariance restricts the possible nonlinearities to those described by Eqs. (7) and (8), and thus does not allow this term to occur on its own. An important consequence of this difference is the following: since plane waves satisfy $\mathcal{F} = \mathcal{G} = 0$, they behave in vacuum NLED just as in the linear theory of Maxwell, with the same dispersion relation, $\omega = ck$. In the language of nonlinear optics, we may say that vacuum does not exhibit *self-phase modulation* (SPM), but only *cross-phase modulation* (XPM): interactions are induced only between different plane waves [38]. In media, on the other hand, SPM and XPM both occur, with different (though closely related) nonlinear indices [as defined in Eq. (1)] [6].

C. Euler-Heisenberg and Born-Infeld models

Particular models yield particular values of and/or relations between the coefficients δ_i. We here consider two specific cases.

---

[3]This is related to the fact that, when written in terms of A_μ, $\mathcal{G}$ turns out to be a total derivative [39].

[4]In standard terminology, it is only this term that is referred to as the "Kerr effect."

The EH effective Lagrangian [9,10] is derived from QED by summing over all Feynman diagrams containing a single electron-positron loop. In performing the summation, it is assumed that the electromagnetic fields themselves are constant over the loop, which leads to a local effective theory.[5] Therefore, when expanded to quartic order in the fields [8], the EH Lagrangian takes the form (6), with $\delta_3 = 0$ due to the $\mathcal{P}$- and $\mathcal{T}$-invariance of QED, and with the particular values

$$\delta_1^{(\mathrm{EH})} = \frac{16}{45}\alpha^2 \frac{\hbar_e^3}{m_e c^2} \approx 13.3 \times 10^{-12} \ \mu\mathrm{m}^3/\mathrm{J},$$

$$\delta_2^{(\mathrm{EH})} = \frac{7}{4}\delta_1^{(\mathrm{EH})} \approx 23.3 \times 10^{-12} \ \mu\mathrm{m}^3/\mathrm{J}. \quad (9)$$

Here, $\alpha \approx 1/137$ is the fine-structure constant, m_e is the mass of the electron, and $\hbar_e = \hbar/m_e c$ is the reduced Compton wavelength of the electron.

The BI model [15] is derived from the postulate that there exists a fundamental upper limit on the field strength, thus regularizing the self-energy of charged point particles. This model is also $\mathcal{P}$- and $\mathcal{T}$-invariant so that $\delta_3 = 0$, but it predicts $\delta_1 = \delta_2 \doteq \delta^{(\mathrm{BI})}$, in strict disagreement with the EH result given above. It thus contains one free parameter, usually written as the maximum absolute field strength b, where $\delta^{(\mathrm{BI})} = 1/b^2$. No precise value is predicted, though Born and Infeld considered that the absolute field strength should be approximately that produced by an electron at its own classical radius, and using this prescription one finds

$$\delta^{(\mathrm{BI})} \sim 4\pi\alpha^3 \frac{\hbar_e^3}{m_e c^2} \approx 3.43 \times 10^{-12} \ \mu\mathrm{m}^3/\mathrm{J}, \quad (10)$$

about a factor of 4 smaller than $\delta_1^{(\mathrm{EH})}$ (or a factor of 7 smaller than $\delta_2^{(\mathrm{EH})}$).

III. EFFECTIVE-MEDIUM DESCRIPTION

In this section, we develop the analogy between the dressed vacuum (including strong electromagnetic fields) and an optical medium. The Lagrangian and wave equations are explicitly decomposed into a background term describing the strong fields alone and the lowest-order correction due to the presence of the probe. A general equation for the probe wave eigenstates is derived, and some known results for the case of static background fields are reproduced.

A. Decomposition into background and probe fields

Much like in gravity, where we consider test particles assumed light enough not to have any significant effect on the gravitational field and whose motion is thus entirely

determined by the space-time metric already present, we wish here to consider probe waves propagating in a vacuum whose optical properties have been altered by the presence of strong fields, the probe waves being too weak to contribute to this alteration themselves. To this end, we decompose the total field into a sum of two terms: a background field, much the stronger of the two, entirely responsible for the alteration of the optical properties of the vacuum, and a significantly weaker probe field whose propagation through the altered vacuum we wish to solve for. The dressed vacuum (i.e., the combination of vacuum plus background fields) can be considered as a medium in its own right. The insensitivity of the properties of this effective medium to the presence of the probe implies that the wave equations for the probe fields will be linear, or equivalently that the part of the Lagrangian relevant to the probe will be quadratic in those same fields.

Explicitly, let us write the total fields as $\mathbf{E} = \mathbf{E}_0 + \mathbf{e}$ and $\mathbf{B} = \mathbf{B}_0 + \mathbf{b}$, where $\mathbf{E}_0$ and $\mathbf{B}_0$ represent the background fields while $\mathbf{e}$ and $\mathbf{b}$ are the probe fields. The Lagrangian is written as a Taylor series in the latter:

$$\mathcal{L} = \mathcal{L}_0 + \frac{\partial \mathcal{L}}{\partial E_i}\bigg|_0 e_i + \frac{\partial \mathcal{L}}{\partial B_i}\bigg|_0 b_i + \frac{1}{2}\frac{\partial^2 \mathcal{L}}{\partial E_i \partial E_j}\bigg|_0 e_i e_j$$

$$+ \frac{1}{2}\frac{\partial^2 \mathcal{L}}{\partial B_i \partial B_j}\bigg|_0 b_i b_j + \frac{\partial^2 \mathcal{L}}{\partial E_i \partial B_j}\bigg|_0 e_i b_j + \cdots. \quad (11)$$

The subscript "0" indicates that the quantity in question is to be evaluated for the background fields $\mathbf{E}_0$ and $\mathbf{B}_0$, which are taken to be solutions of the full nonlinear wave equations, extremizing the action by definition. Therefore, the terms linear in $\mathbf{e}$ and $\mathbf{b}$ in Eq. (11), representing the first-order variation of $\mathcal{L}_0$, give zero contribution to the action and can be removed. The first nontrivial terms involving the probe are those quadratic in the $\mathbf{e}$ and $\mathbf{b}$ fields, and we define this quadratic part as the effective Lagrangian for the probe:

$$\mathcal{L}_{\mathrm{probe}} \doteq \tfrac{1}{2}(e^2 - b^2) + \tfrac{1}{2}\mathbf{e}^{\mathrm{T}}\delta\mathcal{L}_{\mathrm{ee}}\mathbf{e} + \tfrac{1}{2}\mathbf{b}^{\mathrm{T}}\delta\mathcal{L}_{\mathrm{bb}}\mathbf{b} + \mathbf{e}^{\mathrm{T}}\delta\mathcal{L}_{\mathrm{eb}}\mathbf{b}, \quad (12)$$

where the superscript "T" indicates the transpose, and where we have introduced the matrices

$$[\delta\mathcal{L}_{\mathrm{ee}}]_{ij} \doteq \frac{\partial^2(\mathcal{L} - \mathcal{F})}{\partial E_i \partial E_j} = \frac{\partial(D_i - E_i)}{\partial E_j} = \frac{\partial(D_j - E_j)}{\partial E_i},$$

$$[\delta\mathcal{L}_{\mathrm{bb}}]_{ij} \doteq \frac{\partial^2(\mathcal{L} - \mathcal{F})}{\partial B_i \partial B_j} = \frac{\partial(B_i - H_i)}{\partial B_j} = \frac{\partial(B_j - H_j)}{\partial B_i},$$

$$[\delta\mathcal{L}_{\mathrm{eb}}]_{ij} \doteq \frac{\partial^2(\mathcal{L} - \mathcal{F})}{\partial E_i \partial B_j} = \frac{\partial(D_i - E_i)}{\partial B_j} = \frac{\partial(B_j - H_j)}{\partial E_i}. \quad (13)$$

These matrices clearly vanish when the full Lagrangian takes the standard Maxwell form (i.e., when $\mathcal{L} = \mathcal{F}$), in which case $\mathcal{L}_{\mathrm{probe}}$ is simply the Maxwell Lagrangian for the probe fields. Differentiating Eqs. (7) as prescribed by the definitions in Eqs. (13), we may write explicit expressions for the matrices $\delta\mathcal{L}_{\mathrm{ee}}$,

---

[5]This amounts to assuming that the typical photon wavelengths are much larger than the Compton wavelength of the electron, which gives the characteristic size of the electron-positron loop. The loop can then be considered as a pointlike vertex. Since $\lambda_e \sim 10^{-12}$ m, this is a good approximation at optical wavelengths $\gtrsim 10^{-7}$ m.

$\delta\mathcal{L}_{bb}$, and $\delta\mathcal{L}_{eb}$:

$$[\delta\mathcal{L}_{ee}]_{ij} = (\delta_1\mathcal{F}_0 + \delta_3\mathcal{G}_0)\delta_{ij} + \delta_1 E_{0,i}E_{0,j} + \delta_2 B_{0,i}B_{0,j} + \delta_3(E_{0,i}B_{0,j} + B_{0,i}E_{0,j}),$$
$$[\delta\mathcal{L}_{bb}]_{ij} = -(\delta_1\mathcal{F}_0 + \delta_3\mathcal{G}_0)\delta_{ij} + \delta_1 B_{0,i}B_{0,j} + \delta_2 E_{0,i}E_{0,j} - \delta_3(E_{0,i}B_{0,j} + B_{0,i}E_{0,j}),$$
$$[\delta\mathcal{L}_{eb}]_{ij} = (\delta_2\mathcal{G}_0 + \delta_3\mathcal{F}_0)\delta_{ij} - \delta_1 E_{0,i}B_{0,j} + \delta_2 B_{0,i}E_{0,j} + \delta_3(E_{0,i}E_{0,j} - B_{0,i}B_{0,j}), \tag{14}$$

where δ_{ij} is the Kronecker delta. Equivalently, using vector and matrix notation, we have

$$\delta\mathcal{L}_{ee} = (\delta_1\mathcal{F}_0 + \delta_3\mathcal{G}_0)\mathbb{1} + \delta_1\mathbf{E}_0\mathbf{E}_0^T + \delta_2\mathbf{B}_0\mathbf{B}_0^T + \delta_3(\mathbf{E}_0\mathbf{B}_0^T + \mathbf{B}_0\mathbf{E}_0^T),$$
$$\delta\mathcal{L}_{bb} = -(\delta_1\mathcal{F}_0 + \delta_3\mathcal{G}_0)\mathbb{1} + \delta_1\mathbf{B}_0\mathbf{B}_0^T + \delta_2\mathbf{E}_0\mathbf{E}_0^T - \delta_3(\mathbf{E}_0\mathbf{B}_0^T + \mathbf{B}_0\mathbf{E}_0^T),$$
$$\delta\mathcal{L}_{eb} = (\delta_2\mathcal{G}_0 + \delta_3\mathcal{F}_0)\mathbb{1} - \delta_1\mathbf{E}_0\mathbf{B}_0^T + \delta_2\mathbf{B}_0\mathbf{E}_0^T + \delta_3(\mathbf{E}_0\mathbf{E}_0^T - \mathbf{B}_0\mathbf{B}_0^T), \tag{15}$$

where $\mathbb{1}$ is the 3×3 identity matrix, and where the vector biproducts of the form $\mathbf{uv}^T$ are "outer products" yielding matrices rather than scalars. Importantly, the $\delta\mathcal{L}$ matrices can be nonzero even when $\mathcal{F}_0$ and $\mathcal{G}_0$ vanish. Therefore, a single plane wave, though in some sense a "linear" solution of the wave equations, will nonetheless generate an effective medium as it will affect a probe wave whose propagation direction is not equal to its own.[6]

The probe Lagrangian (12) allows us to treat $\mathbf{e}$ and $\mathbf{b}$ as the "full" electromagnetic fields, the background fields no longer being treated dynamically but rather having been subsumed into the definition of the effective medium. That is, Eqs. (3) may be applied to the probe fields alone, and the associated constitutive relations are found by inserting $\mathcal{L}_{probe}$ into Eqs. (4):

$$\mathbf{d} = (\mathbb{1} + \overline{\delta\mathcal{L}_{ee}})\mathbf{e} + \overline{\delta\mathcal{L}_{eb}}\mathbf{b}, \quad \mathbf{h} = (\mathbb{1} - \overline{\delta\mathcal{L}_{bb}})\mathbf{b} - \overline{\delta\mathcal{L}_{eb}}^T\mathbf{e}. \tag{16}$$

Here, we have introduced an overbar on the $\delta\mathcal{L}$ matrices to indicate a space-time average over the wavelength and period of the probe; equivalently, the overbar selects their "slowly varying" component with respect to the oscillations of the probe. This ensures that $\mathbf{d}$ and $\mathbf{h}$ inherit the same carrier wave as $\mathbf{e}$ and $\mathbf{b}$, differing only in the form of their slowly varying envelope. In general (and particularly when they are provided by a propagating wave) the background fields are highly oscillatory, and the $\delta\mathcal{L}$ matrices defined in Eqs. (14) and (15) will inherit some of this oscillatory behavior. However, since the $\delta\mathcal{L}$ matrices influence the probe via an accumulated phase, the highly oscillatory terms can (to a good approximation) usually be neglected.[7]

---

[6]In this respect, the original DeLLight proposal [25] is overcomplicated as it suggests using two counterpropagating pump beams to engender a nontrivial refractive index profile as seen by a probe. One of the key points of this paper (examined further in Sec. IV) is that a single pump beam is sufficient for this purpose.

[7]This "rotating wave approximation" is a standard procedure in nonlinear optics; see, e.g., Ref. [6]. The rapidly oscillating terms in $\delta\mathcal{L}$ become significant only when *phase matching* occurs, i.e., when a certain combination of the wave vectors and frequencies involved generates another carrier wave which is itself "on shell," with frequency and wave vector approximately satisfying the dispersion relation $\omega = ck$. For the quartic nonlinear Lagrangian considered here, there will be a total of four such waves in any given combination,

Rearranging Eqs. (16), and neglecting products of the $\delta\mathcal{L}$ matrices [to be consistent with our neglect of higher-order terms in the Lagrangian (6)], we find

$$\mathbf{d} = (\mathbb{1} + \overline{\delta\mathcal{L}_{ee}})\mathbf{e} + \overline{\delta\mathcal{L}_{eb}}\mathbf{h}, \quad \mathbf{b} = (\mathbb{1} + \overline{\delta\mathcal{L}_{bb}})\mathbf{h} + \overline{\delta\mathcal{L}_{eb}}^T\mathbf{e}. \tag{17}$$

Equations (17) are in the standard form with respect to which the susceptibilities of an optical medium are defined. We may thus identify the effective electric, magnetic, and magneto-electric susceptibilities of the dressed vacuum:

$$\chi^e = \overline{\delta\mathcal{L}_{ee}}, \quad \chi^m = \overline{\delta\mathcal{L}_{bb}}, \quad \alpha = \overline{\delta\mathcal{L}_{eb}}. \tag{18}$$

B. Plane probe waves in the effective medium

For definiteness, and without loss of generality, we take the probe to be propagating in the $-z$ direction. The convenience of this choice stems from the fact that, when using a right-handed coordinate system, projections onto the xy plane (with the z axis pointing out of the page) intuitively represent what is "seen" by the probe during its propagation. Using the fact that the averaged $\overline{\delta\mathcal{L}}$ matrices are (by definition) slowly varying with respect to the wavelength and period of the probe, we may locally decompose its electric field into a slowly varying envelope and a carrier wave:

$$\mathbf{e}(z, t) = \text{Re}\{\mathbf{e}^{(0)}e^{-ikz-i\omega t}\} = \tfrac{1}{2}\mathbf{e}^{(0)}e^{-ikz-i\omega t} + \text{c.c.}, \tag{19}$$

where $k > 0$ and $\omega > 0$. Analogous expressions hold for $\mathbf{b}$, $\mathbf{d}$, and $\mathbf{h}$. The vector $\mathbf{e}^{(0)}$ determines the amplitude and polarization of the electric field, and likewise for $\mathbf{d}^{(0)}$, etc., while the ratio of ω to k gives the phase velocity of the wave: $\omega/k = c/n$, where n is the refractive index. In the weakly nonlinear regime we are considering, n will remain very close to 1, in which case it is more convenient to express this relation in the form

$$1 + \delta n = \frac{ck}{\omega}. \tag{20}$$

We wish to determine the refractive index variation δn, which satisfies $0 \leqslant \delta n \ll 1$,[8] and which will (to lowest order) be quadratic in the background fields $\mathbf{E}_0$ and $\mathbf{B}_0$.

---

and these processes are typically referred to as *four-wave mixing*. See [37,42] for four-wave mixing in vacuum NLED.

[8]The positivity of δn stems from the requirement of causality in special relativity, i.e., that signals cannot propagate faster than the speed of light in vacuum, c. Generally speaking, this applies not to

Even before accounting for the constitutive equations (16) relating **e** and **b** to **d** and **h**, the probe must satisfy the independent set of Eqs. (3). For a plane wave, these become

$$\mathbf{k} \cdot \mathbf{b}^{(0)} = 0, \quad c\mathbf{k} \times \mathbf{e}^{(0)} - \omega\mathbf{b}^{(0)} = 0, \quad (21a)$$

$$\mathbf{k} \cdot \mathbf{d}^{(0)} = 0, \quad c\mathbf{k} \times \mathbf{h}^{(0)} + \omega\mathbf{d}^{(0)} = 0, \quad (21b)$$

where in the present case we have $\mathbf{k} = -k\hat{z}$. In each line of Eqs. (21), the second equation implies the first, so that these give only two independent equations rather than four. Considering therefore only the second equation of each line, and using the definition of δn given in Eq. (20), Eqs. (21) may be written as

$$\mathbf{b}^{(0)} = -(1 + \delta n)\Omega_z\mathbf{e}^{(0)}, \quad (22a)$$

$$\mathbf{d}^{(0)} = (1 + \delta n)\Omega_z\mathbf{h}^{(0)}, \quad (22b)$$

where we have defined the 3×3 matrix

$$\Omega_z = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (23)$$

such that, when acting on a three-dimensional vector **v**, we have $\Omega_z\mathbf{v} = \hat{z} \times \mathbf{v}$. We also note that, since $\Omega_z^{\mathrm{T}} = -\Omega_z$, the ordering is faithfully represented when acting on a transposed vector: $\mathbf{v}^{\mathrm{T}}\Omega_z = -(\Omega_z\mathbf{v})^{\mathrm{T}} = -(\hat{z} \times \mathbf{v})^{\mathrm{T}} = (\mathbf{v} \times \hat{z})^{\mathrm{T}}$.

It now remains to impose the constitutive relations (16) for the probe. Using Eq. (22a) so that both $\mathbf{d}^{(0)}$ and $\mathbf{h}^{(0)}$ can be written directly in terms of $\mathbf{e}^{(0)}$ (with no reference to $\mathbf{b}^{(0)}$), and neglecting products of small quantities, we have

$$\mathbf{d}^{(0)} = [\mathbb{1} + \overline{\delta\mathcal{L}_{\mathrm{ee}}} - \overline{\delta\mathcal{L}_{\mathrm{eb}}}\Omega_z]\mathbf{e}^{(0)}, \quad (24a)$$

$$\mathbf{h}^{(0)} = \left[-(1 + \delta n)\Omega_z + \overline{\delta\mathcal{L}_{\mathrm{bb}}}\Omega_z - \overline{\delta\mathcal{L}_{\mathrm{eb}}}^{\mathrm{T}} \right]\mathbf{e}^{(0)}. \quad (24b)$$

Equations (24) are related to each other using Eq. (22b), thus yielding a single homogeneous equation involving the vector $\mathbf{e}^{(0)}$:

$$\begin{aligned} &\left[\mathbb{1} + \Omega_z^2 + 2\delta n\Omega_z^2 + \overline{\delta\mathcal{L}_{\mathrm{ee}}} - \Omega_z\overline{\delta\mathcal{L}_{\mathrm{bb}}}\Omega_z - \overline{\delta\mathcal{L}_{\mathrm{eb}}}\Omega_z \right. \\ &\left. - (\overline{\delta\mathcal{L}_{\mathrm{eb}}}\Omega_z)^{\mathrm{T}} \right]\mathbf{e}^{(0)} = 0. \end{aligned} \quad (25)$$

The value of δn is found by requiring the determinant of the operator in square brackets in Eq. (25) to vanish. Note that, as $\Omega_z^2 = \mathrm{diag}\{-1, -1, 0\}$, δn will appear only at quadratic order in the determinant (as could have been expected, since the two solutions correspond to the two possible polarizations of the probe wave). Moreover, since $\mathbb{1} + \Omega_z^2 = \mathrm{diag}\{0, 0, 1\}$ and all other terms are of first order in δn and the $\delta\mathcal{L}$ matrices, then to lowest nontrivial order only the xy projection of Eq. (25) need

---

the phase velocity ω/k but to the group velocity $d\omega/dk$, for which we may define a "group index" $n_g = n + \omega\,dn/d\omega$ such that the group velocity is c/n_g. Then the causality condition is simply $n_g \geqslant 1$. In the present case, due to the local nature of the electromagnetic self-interaction encoded in the Lagrangian (6), n will be independent of ω, so the phase and group velocities are identical and this condition reduces to $\delta n \geqslant 0$. In Sec. V, we shall examine a dispersive case where δn can be negative, yet the positivity of δn_g is still respected.

be considered.[9] Using expressions (15) for the $\delta\mathcal{L}$ matrices and the above-mentioned identification of Ω_z with the cross product operator $\hat{z}\times$, Eq. (25) reduces to the following 2×2 eigenvalue problem:

$$\tfrac{1}{2}[\delta_1\overline{\mathcal{E}\mathcal{E}^{\mathrm{T}}} + \delta_2\overline{\mathcal{B}\mathcal{B}^{\mathrm{T}}} + \delta_3(\overline{\mathcal{E}\mathcal{B}^{\mathrm{T}}} + \overline{\mathcal{B}\mathcal{E}^{\mathrm{T}}})]\mathbf{e}_{\perp}^{(0)} = \delta n\mathbf{e}_{\perp}^{(0)}, \quad (26)$$

where the subscript "$\perp$" indicates the projection onto the xy plane, and where we have defined [in very similar fashion to Eq. (20) of [28]]

$$\mathcal{E} = \mathbf{E}_{0,\perp} - \hat{z} \times \mathbf{B}_{0,\perp} = -\hat{z} \times (\hat{z} \times \mathbf{E}_0) - \hat{z} \times \mathbf{B}_0,$$

$$\mathcal{B} = \mathbf{B}_{0,\perp} + \hat{z} \times \mathbf{E}_{0,\perp} = -\hat{z} \times (\hat{z} \times \mathbf{B}_0) + \hat{z} \times \mathbf{E}_0. \quad (27)$$

It is clear that $\mathcal{E}$ and $\mathcal{B}$ lie in the xy plane, and from their definition it immediately follows that $\mathcal{B} = \hat{z} \times \mathcal{E}$ and $\mathcal{E} = -\hat{z} \times \mathcal{B}$. Therefore, for a given propagation direction, the behavior of the probe wave is determined by a single orthogonal vector formed from $\mathbf{E}_0$ and $\mathbf{B}_0$.

Equations (26) and (27) are one of the main results of this paper, giving the eigenstates of the probe in the effective medium generated by an arbitrary configuration of strong background fields. They form the basis of the analysis up to the end of Sec. IV (before nonlocal corrections induced by an axion field are considered in Sec. V).

C. Constant background fields: DC Kerr and Cotton-Mouton effects

As an illustrative example, we consider the simplest case where the background fields are constant, or at least slowly varying with respect to the wavelength and period of the probe. This case has already been analyzed in some detail in the literature [28,43–45], though with particular emphasis on the predictions of the EH model. We give a quick run-through of the various results here, showing that they are indeed reproduced by the effective-medium framework we have used, and paving the way for the analysis of a propagating wave as background in Sec. IV.

1. Refractive indices as eigenvalues

Since the background fields are slowly varying, the overbars in Eq. (26) are redundant, the 2×2 eigenvalue equation reducing to

$$\tfrac{1}{2}[\delta_1\mathcal{E}\mathcal{E}^{\mathrm{T}} + \delta_2\mathcal{B}\mathcal{B}^{\mathrm{T}} + \delta_3(\mathcal{E}\mathcal{B}^{\mathrm{T}} + \mathcal{B}\mathcal{E}^{\mathrm{T}})]\mathbf{e}_{\perp}^{(0)} = \delta n\mathbf{e}_{\perp}^{(0)}, \quad (28)$$

where the vectors $\mathcal{E}$ and $\mathcal{B}$ are now to be considered as constant. Since $\mathcal{E}$ and $\mathcal{B}$ have the same magnitude and are perpendicular to each other, we may use the orthonormal vectors $\hat{\mathcal{E}} = \mathcal{E}/|\mathcal{E}|$ and $\hat{\mathcal{B}} = \mathcal{B}/|\mathcal{B}|$ as a basis in the xy plane. The ordered vectors $\{\hat{\mathcal{E}}, \hat{\mathcal{B}}, \hat{z}\}$ form a right-handed orthonormal basis, $\hat{\mathcal{E}}$ and $\hat{\mathcal{B}}$ being analogous to the standard Cartesian basis vectors $\hat{x}$ and $\hat{y}$, respectively. In the $\{\hat{\mathcal{E}}, \hat{\mathcal{B}}\}$ basis, Eq. (28) can

---

[9]To see this, write Eq. (25) in the form $(\hat{z}\hat{z}^{\mathrm{T}} + \delta M)\mathbf{e}^{(0)} = 0$, where δM is of lowest order in δn and the $\delta\mathcal{L}$ matrices. At zeroth order, the longitudinal component $e_z^{(0)}$ vanishes, while the transverse component $\mathbf{e}_{\perp}^{(0)}$ is determined by the xy projection of δM.

be written in matrix notation as

$$\begin{bmatrix} \delta_1 & \delta_3 \\ \delta_3 & \delta_2 \end{bmatrix} \begin{bmatrix} e_{\mathcal{E}}^{(0)} \\ e_{\mathcal{B}}^{(0)} \end{bmatrix} = \frac{\delta n}{\frac{1}{2}|\mathcal{E}|^2} \begin{bmatrix} e_{\mathcal{E}}^{(0)} \\ e_{\mathcal{B}}^{(0)} \end{bmatrix}. \tag{29}$$

This is readily solved. There are two refractive indices (corresponding to two polarizations of the probe wave),

$$\delta n_{\pm} = \delta_{\pm} \tfrac{1}{2}|\mathcal{E}|^2, \tag{30}$$

where the coefficients $\delta_{\pm}$ are the eigenvalues of the matrix on the left-hand side of Eq. (29):

$$\delta_{\pm} = \tfrac{1}{2}(\delta_1 + \delta_2 \pm \sqrt{(\delta_1 - \delta_2)^2 + (2\delta_3)^2}). \tag{31}$$

By definition, $\delta_+ \geqslant \delta_-$ and $\delta n_+ \geqslant \delta n_-$. Equality only holds when $\delta_1 - \delta_2 = \delta_3 = 0$ (as in the BI model); otherwise the presence of the background fields makes the vacuum birefringent [44], a phenomenon referred to as the *DC Kerr effect* when the static external field is a pure electric field and the *Cotton-Mouton effect* when it is a pure magnetic field. The vacuum Cotton-Mouton effect is the basis for the BMV [19,20] and PVLAS [21,22] experiments. For a magnetic field oriented perpendicular to the direction of the probe wave, we have $|\mathcal{E}|^2 = B_0^2 = B_{0,\text{SI}}^2/\mu_0$, and the difference in the two refractive indices is

$$\Delta n = \sqrt{(\delta_1 - \delta_2)^2 + (2\delta_3)^2} \frac{B_{0,\text{SI}}^2}{2\mu_0}. \tag{32}$$

For a magnetic field of 1 T, this gives, for the EH and BI models,

$$\Delta n^{(\text{EH})} = 3.98 \times 10^{-24}, \quad \Delta n^{(\text{BI})} = 0. \tag{33}$$

The EH value is in agreement with the predictions of Refs. [19–22], while the vanishing of $\Delta n^{(\text{BI})}$ indicates the absence of birefringence in the BI model [26–28].

As noted in footnote 9, respecting causality requires the avoidance of a negative value of $\delta n_{\pm}$, or equivalently of $\delta_{\pm}$. It is straightforward to show that this implies the inequalities $\delta_1 \geqslant 0$, $\delta_2 \geqslant 0$, and $\delta_1 \delta_2 - \delta_3^2 \geqslant 0$ [in agreement with Eqs. (25) of [28]]. Interestingly, using the identification of the effective susceptibilities made in Eqs. (18), and using the simplifying assumption $\mathcal{F}_0 = \mathcal{G}_0 = 0$, we find

$$\chi_{ii}^{\text{e}} \chi_{jj}^{\text{m}} - (\alpha_{ij})^2 = (\delta_1 \delta_2 - \delta_3^2)(E_{0,i} E_{0,j} + B_{0,i} B_{0,j})^2. \tag{34}$$

So in this case the inequality $\delta_1 \delta_2 - \delta_3^2 \geqslant 0$, derived here from the requirement of causality, is equivalent to $\chi_{ii}^{\text{e}} \chi_{jj}^{\text{m}} - (\alpha_{ij})^2 \geqslant 0$, previously derived (for any optical medium) in [46] from the requirement of thermodynamic stability.

2. Anisotropy of the effective medium

While the factors $\delta_{\pm}$ are fixed by the post-Maxwellian parameters entering Eq. (6), the strength of the refractive index change is also proportional to $\frac{1}{2}|\mathcal{E}|^2$. This is simply quadratic in the background fields, but because of the projection and combination required to form $\mathcal{E}$ and $\mathcal{B}$, the dependence on relative orientation (between $\mathbf{E}_0$ and $\mathbf{B}_0$, as well as between these fields and the probe wave vector $\mathbf{k} \propto -\hat{z}$) can be rather complicated. After a bit of algebra, it can be shown that

$$\tfrac{1}{2}|\mathcal{E}|^2 = \tfrac{1}{2}(|\hat{z} \times \mathbf{E}_0|^2 + |\hat{z} \times \mathbf{B}_0|^2) + \hat{z} \cdot (\mathbf{E}_0 \times \mathbf{B}_0). \tag{35}$$
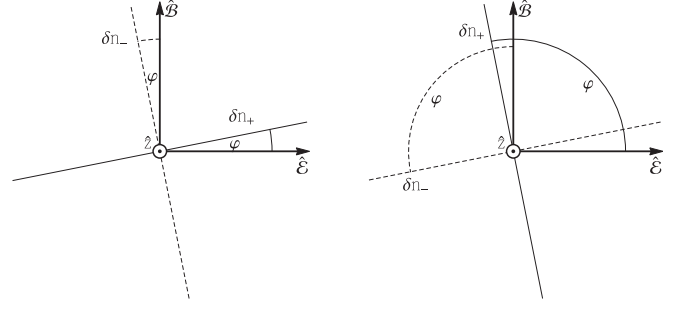

FIG. 1. Eigenpolarizations for the electric-field amplitude $\mathbf{e}^{(0)}$ of the probe, which propagates in the $-z$ direction (i.e., into the page). They lie in the xy plane, rotated with respect to $\hat{\mathcal{E}}$ and $\hat{\mathcal{B}}$ by the angle φ whose value is determined by Eqs. (37). The polarizations corresponding to refractive index changes δn_+ and δn_- of Eq. (30) are given by the left and right columns of $R(\varphi)$ of Eq. (36) and shown here in solid and dashed line, respectively. The left and right panels differ by a rotation of $\pi/2$, which (in effect) leaves the eigenpolarization directions invariant but switches the associated refractive indices. In the EH model, we find $\varphi = \pi/2$, so the δn_+ polarization is aligned with $\hat{\mathcal{B}}$ while the δn_- polarization is aligned with $\hat{\mathcal{E}}$.

The first term here is rather simple, in that $\mathbf{E}_0$ and $\mathbf{B}_0$ contribute separately, and with the squared magnitude of their projections onto the xy plane. The second term is more subtle, as it depends on the relative orientation of $\mathbf{E}_0$ and $\mathbf{B}_0$. Moreover, it is directionally dependent: whereas the first term depends only on the line along which the wave vector $\mathbf{k}$ lies (defined to be the z axis) and does not vary under the transformation $\mathbf{k} \to -\mathbf{k}$ (i.e., $\hat{z} \to -\hat{z}$), the second term changes sign under this transformation. In this sense the effective medium behaves as if it were moving with a velocity proportional to the "Poynting vector" $\mathbf{E}_0 \times \mathbf{B}_0$. This anisotropy (which is independent of the probe polarization as it stems only from the magnitude of the vector $\mathcal{E}$) was described in Ref. [45]. We shall see that it is also present when the background is a plane wave, with copropagating probe waves seeing no refractive index change while counterpropagating waves experience the strongest effect.

3. Eigenpolarizations

The eigenvectors of Eq. (29) give the two eigenpolarizations of $\mathbf{e}^{(0)}$. Since the matrix on the left-hand side is real and symmetric, the eigenvectors are necessarily real and (when normalized) form the columns of a two-dimensional rotation matrix,

$$R(\varphi) = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}. \tag{36}$$

The parameter φ is simply the angle through which the eigenvectors are rotated with respect to the basis $\{\hat{\mathcal{E}}, \hat{\mathcal{B}}\}$ (see Fig. 1), and is defined up to a multiple of π since a half-rotation simply flips the signs of the eigenvectors without changing their orientation. φ can thus be chosen to lie in the half-open interval $(-\pi/2, \pi/2]$, and the matrix on the left-hand side of Eq. (29) can be written as $R(\varphi)DR^{-1}(\varphi)$, where D is a diagonal matrix whose entries are the eigenvalues of Eq. (31).

We take δ_+ to be the first diagonal component of D, so that the left column of $R(\varphi)$ gives the polarization with the larger refractive index change, δn_+. A direct calculation shows that φ must satisfy

$$\delta_1 - \delta_2 = (\delta_+ - \delta_-)\cos(2\varphi), \quad 2\delta_3 = (\delta_+ - \delta_-)\sin(2\varphi). \quad (37)$$

There are three special cases. First, when $\delta_3 = 0$ and $\delta_1 - \delta_2 \neq 0$ (as in the EH model), we have $\varphi = 0$ or $\pi/2$ (depending on the sign of $\delta_1 - \delta_2$) and the polarizations are aligned with the vectors $\hat{\mathcal{E}}$ and $\hat{\mathcal{B}}$. Second, when $\delta_1 - \delta_2 = 0$ and $\delta_3 \neq 0$, we have $\varphi = \pm\pi/4$, so the polarizations are at 45° to $\hat{\mathcal{E}}$ and $\hat{\mathcal{B}}$. Third, when both $\delta_3 = 0$ and $\delta_1 - \delta_2 = 0$ (as in the BI model), φ is undefined, but this is not a problem as there is an absence of birefringence (and hence also of well-defined eigenpolarizations) in this case.

IV. OPTICAL KERR EFFECT

In this section we turn to the main focus of this paper: the refractive index change of vacuum engendered by an intense propagating wave or "pump," which provides the strong background fields described in Sec. III. We derive the dependence on the tilt angle (between the propagation directions of pump and probe), as well as the effect of elliptical polarization of the pump. Finally, we express the results in terms of the wave intensity (rather than the field strength) in order to extract the equivalent of the nonlinear index for vacuum by analogy with Eq. (1).

A. Fields of a monochromatic pump wave

Let us consider then the fields of a propagating beam, which we assume can be approximated as monochromatic over spacetime regions much larger than the wavelength and longer than the period of the probe. We may again use Eqs. (26) and (27), though now the overbars extracting the slowly varying components of the outer products in Eq. (26) will come into play. We write the pump fields in the form

$$\mathbf{E}_0 = \tfrac{1}{2}\mathbf{E}_0^{(0)}e^{i\mathbf{k}_0\cdot\mathbf{r}-i\omega_0 t} + \text{c.c.},$$
$$\mathbf{B}_0 = \tfrac{1}{2}\mathbf{B}_0^{(0)}e^{i\mathbf{k}_0\cdot\mathbf{r}-i\omega_0 t} + \text{c.c.}, \quad (38)$$

where, in order to satisfy the Maxwell equations (3) with $\mathbf{D} = \mathbf{E}$ and $\mathbf{H} = \mathbf{B}$ (as a single plane wave must, having $\mathcal{F}_0 = \mathcal{G}_0 = 0$), we have $\omega_0 = ck_0$ (where $k_0 = |\mathbf{k}_0|$), and

$$\mathbf{B}_0^{(0)} = \frac{\mathbf{k}_0}{k_0} \times \mathbf{E}_0^{(0)}. \quad (39)$$

Note that we do not specify the direction of $\mathbf{k}_0$, whose orientation with respect to the probe wave vector $\mathbf{k} = -k\hat{z}$ is taken to be arbitrary.

The next step is to work out the vectors $\mathcal{E}$ and $\mathcal{B}$ entering the matrix in Eq. (26), before application of the overbars. These will be oscillatory just as $\mathbf{E}$ and $\mathbf{B}$ are, and we may write

$$\mathcal{E} = \tfrac{1}{2}\mathcal{E}^{(0)}e^{i\mathbf{k}_0\cdot\mathbf{r}-i\omega_0 t} + \text{c.c.},$$
$$\mathcal{B} = \tfrac{1}{2}\mathcal{B}^{(0)}e^{i\mathbf{k}_0\cdot\mathbf{r}-i\omega_0 t} + \text{c.c.}, \quad (40)$$

where the amplitude vectors are given by

$$\mathcal{E}^{(0)} = \mathbf{E}_{0,\perp}^{(0)} - \hat{z} \times \mathbf{B}_{0,\perp}^{(0)} = -\hat{z} \times \left[\left(\hat{z} + \frac{\mathbf{k}_0}{k_0}\right) \times \mathbf{E}_0^{(0)}\right],$$

$$\mathcal{B}^{(0)} = \mathbf{B}_{0,\perp}^{(0)} + \hat{z} \times \mathbf{E}_{0,\perp}^{(0)} = -\hat{z} \times \left[\left(\hat{z} + \frac{\mathbf{k}_0}{k_0}\right) \times \mathbf{B}_0^{(0)}\right]. \quad (41)$$

Here, we have used Eq. (39), as well as standard identities concerning two successive applications of the cross product. These vectors evidently lie in the *xy* plane, and by construction we again have $\mathcal{B}^{(0)} = \hat{z} \times \mathcal{E}^{(0)}$ and $\mathcal{E}^{(0)} = -\hat{z} \times \mathcal{B}^{(0)}$. After some further algebra, the squared magnitude of $\mathcal{E}^{(0)}$ (and hence also of $\mathcal{B}^{(0)}$) can be shown to be

$$\mathcal{E}^{(0)\star} \cdot \mathcal{E}^{(0)} = \left(1 + \frac{\hat{z} \cdot \mathbf{k}_0}{k_0}\right)^2 \left|\mathbf{E}_0^{(0)}\right|^2$$
$$= (1 + \cos\theta)^2\left|\mathbf{E}_0^{(0)}\right|^2$$
$$= 4\cos^4\frac{\theta}{2}\left|\mathbf{E}_0^{(0)}\right|^2. \quad (42)$$

Here, we have introduced the *tilt angle* θ between the wave vectors of the pump and probe (illustrated in Fig. 2). This is defined to be zero when the pump and probe are exactly counterpropagating and $\pm\pi$ when they are exactly copropagating. Note that the magnitude of $\mathcal{E}^{(0)}$ vanishes in the latter case, this being consistent with the observation made after Eqs. (8) that there is no SPM in vacuum; on the other hand, it is maximal when the pump and probe are exactly counterpropagating.

B. Accounting for elliptical polarization of the pump

Inserting Eqs. (40) into Eq. (26), and implementing the overbars by dropping all rapidly oscillating terms, we are led to the following eigenproblem:

$$\tfrac{1}{4}\text{Re}\{\delta_1\mathcal{E}^{(0)\star}\mathcal{E}^{(0)\text{T}} + \delta_2\mathcal{B}^{(0)\star}\mathcal{B}^{(0)\text{T}} + \delta_3(\mathcal{E}^{(0)\star}\mathcal{B}^{(0)\text{T}}$$
$$+ \mathcal{B}^{(0)\star}\mathcal{E}^{(0)\text{T}})\}\mathbf{e}^{(0)} = \delta n\mathbf{e}^{(0)}. \quad (43)$$

Compared to the case of constant background fields studied in Sec. III C, we have here a complication in that the field amplitudes $\mathbf{E}_0^{(0)}$ and $\mathbf{B}_0^{(0)}$, and by extension the vectors $\mathcal{E}^{(0)}$ and $\mathcal{B}^{(0)}$, are generally complex. It is thus no longer convenient to use an orthonormal basis aligned with $\mathcal{E}^{(0)}$ and $\mathcal{B}^{(0)}$, since the matrix on the left-hand side of Eq. (43) also depends on $\mathcal{E}^{(0)\star}$ and $\mathcal{B}^{(0)\star}$, and in general $\mathcal{E}^{(0)\star} \neq \mathcal{E}^{(0)}$ and $\mathcal{B}^{(0)\star} \neq \mathcal{B}^{(0)}$. The issue is not with overall phases—it is clear that the matrix in question is invariant under equal overall phase rotations of $\mathcal{E}^{(0)}$ and $\mathcal{B}^{(0)}$—but with relative phases between the components of these vectors. Such relative phases are directly related to the degree of elliptical polarization of the pump wave. If we were to restrict our attention to a linearly polarized pump, there would be no such relative phase, $\mathcal{E}^{(0)}$ and $\mathcal{B}^{(0)}$ could be defined to be real, and the eigenproblem would be equivalent to that of Eq. (28) (except for an overall factor of $1/2$ stemming from the average over rapidly oscillating terms). We can thus expect to recover almost the same results as those of Sec. III C when the pump is linearly polarized. However, with a proper treatment of the complex vectors entering the left-hand side of Eq. (43), the effects of a general elliptical polarization can be fully included in the analysis.
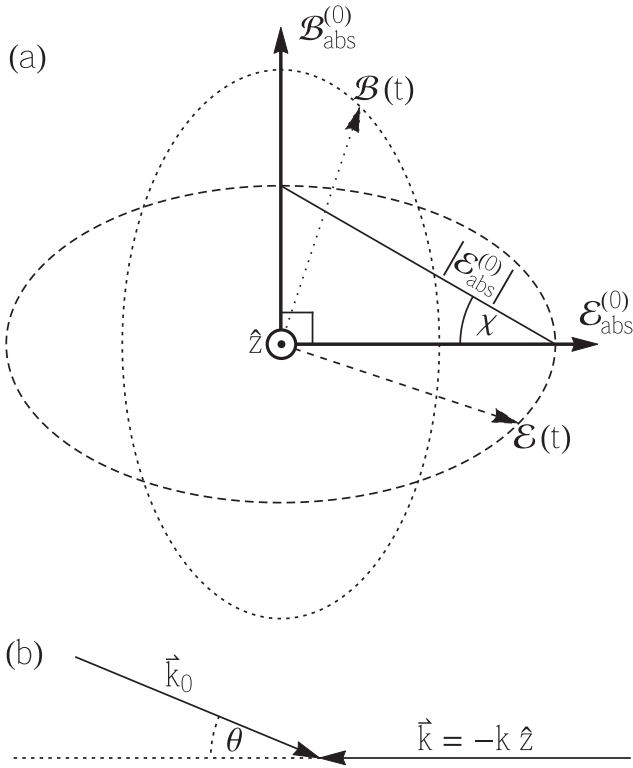
FIG. 2. Relative orientations for a monochromatic pump wave. In (a) are shown the vectors $\mathcal{E}(t)$ and $\mathcal{B}(t)$ (in dashed and dotted line, respectively) in the xy plane, with $\hat{z}$ pointing out of the page and the wave vector of the probe pointing into the page. These oscillate in time so as to trace out ellipses, rotated with respect to each other by $90°$. The constant real vectors $\mathcal{E}^{(0)}_{\text{abs}}$ and $\mathcal{B}^{(0)}_{\text{abs}}$ are defined to lie along the semimajor axes of their respective ellipses, with magnitude equal to the length of the hypotenuse of the right-angled triangle shown. The ellipticity angle χ is formed by the same right-angled triangle, though its sign is determined by the sense of rotation of $\mathcal{E}(t)$ and $\mathcal{B}(t)$. In (b) are shown the wave vectors of the pump and probe waves forming the tilt angle θ, equal to 0 for exactly counterpropagating waves and $\pm\pi$ for exactly copropagating waves.

Up to an overall phase, the complex field amplitudes $\mathbf{E}^{(0)}_0$ and $\mathbf{B}^{(0)}_0$ can always be written as a sum of two orthogonal real vectors, one (which for definiteness we take to be the larger in magnitude of the two) with a purely real coefficient, the other with a purely imaginary coefficient. These vectors necessarily lie in the plane perpendicular to the wave vector $\mathbf{k}_0$. They describe the principle directions of the polarization, i.e., the major and minor axes of the ellipse formed by the oscillating electric or magnetic field (see Fig. 2). The field lies along the minor axis exactly a quarter of a period after it lies along the major axis, and so the complex components of the field along these two directions appear with a relative phase of $\pi/2$. We may thus write the field amplitudes using the real vectors $\mathbf{E}^{(0)}_{\text{abs}}$ and $\mathbf{B}^{(0)}_{\text{abs}}$, and again utilizing the fact that $\mathbf{B}^{(0)}_0 = (\mathbf{k}_0/k_0) \times \mathbf{E}^{(0)}_0$:

$$\mathbf{E}^{(0)}_0 = e^{i\phi_0}\big(\cos\chi\, \mathbf{E}^{(0)}_{\text{abs}} + i\sin\chi\, \mathbf{B}^{(0)}_{\text{abs}}\big),$$
$$\mathbf{B}^{(0)}_0 = e^{i\phi_0}\big(\cos\chi\, \mathbf{B}^{(0)}_{\text{abs}} - i\sin\chi\, \mathbf{E}^{(0)}_{\text{abs}}\big), \tag{44}$$

where $\mathbf{B}^{(0)}_{\text{abs}} = (\mathbf{k}_0/k_0) \times \mathbf{E}^{(0)}_{\text{abs}}$ and $\mathbf{E}^{(0)}_{\text{abs}} = -(\mathbf{k}_0/k_0) \times \mathbf{B}^{(0)}_{\text{abs}}$. With this writing, $\mathbf{E}^{(0)}_0$ and $\mathbf{E}^{(0)}_{\text{abs}}$ have exactly the same magnitude. The angle $\chi \in [-\pi/4, \pi/4]$ is the so-called *ellipticity angle* of the polarization ellipse (also illustrated in Fig. 2). When $\chi = 0$, the wave is linearly polarized; when $\chi = \pm\pi/4$, it is circularly polarized, in which case the directions of $\mathbf{E}^{(0)}_{\text{abs}}$ and $\mathbf{B}^{(0)}_{\text{abs}}$ in the two-dimensional plane can be chosen arbitrarily. χ is defined to be positive when the sense of rotation of the fields is from $\mathbf{E}^{(0)}_{\text{abs}}$ towards $\mathbf{B}^{(0)}_{\text{abs}}$, corresponding to right-handed polarization of the pump.

By a straightforward application of Eqs. (41), we may now write

$$\mathcal{E}^{(0)} = e^{i\phi_0}\big(\cos\chi\, \mathcal{E}^{(0)}_{\text{abs}} + i\sin\chi\, \mathcal{B}^{(0)}_{\text{abs}}\big),$$
$$\mathcal{B}^{(0)} = e^{i\phi_0}\big(\cos\chi\, \mathcal{B}^{(0)}_{\text{abs}} - i\sin\chi\, \mathcal{E}^{(0)}_{\text{abs}}\big), \tag{45}$$

where we have defined

$$\mathcal{E}^{(0)}_{\text{abs}} = -\hat{z} \times \left[\left(\hat{z} + \frac{\mathbf{k}_0}{k_0}\right) \times \mathbf{E}^{(0)}_{\text{abs}}\right],$$
$$\mathcal{B}^{(0)}_{\text{abs}} = -\hat{z} \times \left[\left(\hat{z} + \frac{\mathbf{k}_0}{k_0}\right) \times \mathbf{B}^{(0)}_{\text{abs}}\right]. \tag{46}$$

It is again fairly straightforward to show that $\mathcal{B}^{(0)}_{\text{abs}} = \hat{z} \times \mathcal{E}^{(0)}_{\text{abs}}$ and $\mathcal{E}^{(0)}_{\text{abs}} = -\hat{z} \times \mathcal{B}^{(0)}_{\text{abs}}$, so that these are orthogonal and have equal magnitude. Therefore, analogously to Eqs. (44), Eqs. (45) define an ellipse of ellipticity angle χ in the xy plane. It is particularly interesting that the ellipticity angle remains χ when passing to $\mathcal{E}^{(0)}$ and $\mathcal{B}^{(0)}$, since these lie in the xy plane rather than the plane containing $\mathbf{E}^{(0)}_{\text{abs}}$ and $\mathbf{B}^{(0)}_{\text{abs}}$. One might naively have expected the shape of the ellipse to depend on the orientation of $\mathbf{k}_0$ with respect to $\hat{z}$ (as the orthogonal projections of the ellipses traced out by the electric and magnetic fields certainly do depend on the angle they are viewed from). Remarkably, however, it turns out that $\mathcal{E}^{(0)}$ and $\mathcal{B}^{(0)}$ combine the electric and magnetic fields of the pump wave in just the right way so that χ is invariant with respect to the relative orientation of pump and probe. Only the magnitude of $\mathcal{E}^{(0)}$ (and hence of $\mathcal{E}^{(0)}_{\text{abs}}$) varies with the orientation, as shown by Eq. (42).

C. Nonlinear index of vacuum

Finally, Eqs. (45) are inserted into Eq. (43), upon which it becomes the following eigenproblem:

$$\frac{1}{4}\big[\delta'_1 \mathcal{E}^{(0)}_{\text{abs}}\mathcal{E}^{(0)\text{T}}_{\text{abs}} + \delta'_2 \mathcal{B}^{(0)}_{\text{abs}}\mathcal{B}^{(0)\text{T}}_{\text{abs}} + \delta'_3\big(\mathcal{E}^{(0)}_{\text{abs}}\mathcal{B}^{(0)\text{T}}_{\text{abs}} + \mathcal{B}^{(0)}_{\text{abs}}\mathcal{E}^{(0)\text{T}}_{\text{abs}}\big)\big]\mathbf{e}^{(0)}$$
$$= \delta n\mathbf{e}^{(0)}, \tag{47}$$

where

$$\delta'_1 + \delta'_2 = \delta_1 + \delta_2,$$
$$\delta'_1 - \delta'_2 = (\delta_1 - \delta_2)\cos(2\chi),$$
$$\delta'_3 = \delta_3\cos(2\chi). \tag{48}$$

This is now in a form completely analogous to Eq. (28), with the vectors $\mathcal{E}^{(0)}_{\text{abs}}$ and $\mathcal{B}^{(0)}_{\text{abs}}$ being equal in magnitude and orthogonal to each other. They can thus be used to define basis vectors in the xy plane, $\hat{\mathcal{E}} = \mathcal{E}^{(0)}_{\text{abs}}/|\mathcal{E}^{(0)}_{\text{abs}}|$ and $\hat{\mathcal{B}} = \mathcal{B}^{(0)}_{\text{abs}}/|\mathcal{B}^{(0)}_{\text{abs}}|$, so that the vectors $\{\hat{\mathcal{E}}, \hat{\mathcal{B}}, \hat{z}\}$ form a right-handed orthonormal

basis. Restricting our attention to the xy plane in the basis $\{\hat{\mathcal{E}}, \hat{\mathcal{B}}\}$, Eq. (47) can be written in matrix form as follows:

$$
\begin{bmatrix} \delta_1' & \delta_3' \\ \delta_3' & \delta_2' \end{bmatrix} \begin{bmatrix} e_{\mathcal{E}}^{(0)} \\ e_{\mathcal{B}}^{(0)} \end{bmatrix} = \frac{\delta n}{\frac{1}{4}\left|\mathcal{E}_{\text{abs}}^{(0)}\right|^2} \begin{bmatrix} e_{\mathcal{E}}^{(0)} \\ e_{\mathcal{B}}^{(0)} \end{bmatrix}. \tag{49}
$$

The eigenvalues of the matrix on the left-hand side of this equation are readily found, and give the possible values of δn. Using Eq. (42) and the fact that $\mathcal{E}^{(0)}$ and $\mathcal{E}_{\text{abs}}^{(0)}$ are defined to have the same magnitude, we have

$$
\delta n_{\pm} = \delta_{\pm}' \cos^4 \frac{\theta}{2} \left|\mathbf{E}_0^{(0)}\right|^2, \tag{50}
$$

where

$$
\begin{aligned}
\delta_{\pm}' &= \tfrac{1}{2}\left(\delta_1' + \delta_2' \pm \sqrt{(\delta_1' - \delta_2')^2 + (2\delta_3')^2}\right) \\
&= \tfrac{1}{2}\left(\delta_1 + \delta_2 \pm \cos(2\chi)\sqrt{(\delta_1 - \delta_2)^2 + (2\delta_3)^2}\right). \tag{51}
\end{aligned}
$$

We thus see that the ellipticity angle χ directly affects the strength of the birefringence, which vanishes completely in the case of circular polarization. Moreover, Eqs. (37) still hold [being simply multiplied by an overall factor of $\cos(2\chi)$], so that the rotation angle φ of the eigenpolarizations with respect to the $\{\hat{\mathcal{E}}, \hat{\mathcal{B}}\}$ basis is independent of χ. Note that the probe eigenstates are linearly polarized no matter the polarization state of the pump, a direct result of the fact that the matrix on the left-hand side of Eq. (43) is real and symmetric. (In Sec. V we shall examine a model where this is no longer the case, allowing complex eigenvectors which encode states of elliptical polarization.)

We are now in a position to give explicit values for the corresponding nonlinear index n_2, defined [by analogy with Eq. (1)] such that the refractive index change $\delta n = n_2 I$, where I is the intensity of the pump wave. For definiteness and simplicity, we consider pump and probe to be exactly counterpropagating ($\theta = 0$), this being the optimal geometry[10] according to Eq. (50). The energy density of the pump (after averaging over rapidly oscillating terms) is $\left|\mathbf{E}_0^{(0)}\right|^2/2$, and its intensity is found upon multiplication by c. The corresponding values of n_2 are then simply $\delta_{\pm}' \times 2/c$. In the EH model we have, for a linearly polarized pump beam,

$$
\begin{aligned}
n_{2,\parallel}^{(\text{EH})} &\approx 0.888 \times 10^{-33}~\text{cm}^2/\text{W}, \\
n_{2,\perp}^{(\text{EH})} &\approx 1.555 \times 10^{-33}~\text{cm}^2/\text{W}, \tag{52}
\end{aligned}
$$

where the subscripts "$\parallel$" and "$\perp$" refer to the probe and pump fields being equally and orthogonally polarized, respectively. On the other hand, for a circularly polarized pump, n_2 no longer depends on the polarization of the probe and is simply the arithmetic mean of the two values given above:

$$
n_{2,\text{circ}}^{(\text{EH})} \approx 1.222 \times 10^{-33}~\text{cm}^2/\text{W}. \tag{53}
$$

It can thus be seen that the EH prediction for the nonlinear index of vacuum, though it depends on pump polarization and tilt angle, is on the order of $10^{-33}~\text{cm}^2/\text{W}$.

In the BI model instead, $\delta_{\pm}'$ of Eqs. (51) are both equal to $\delta^{(\text{BI})}$ and independent of the ellipticity angle χ. We thus have simply

$$
n_2^{(\text{BI})} = \frac{2}{cb^2}, \tag{54}
$$

where b^2 is the square of the critical field parametrizing the BI model (expressed in units of energy density). Using the value of $\delta^{(\text{BI})}$ in Eq. (10), with b being the field at the classical radius of the electron, this gives

$$
n_2^{(\text{BI})} \approx 0.229 \times 10^{-33}~\text{cm}^2/\text{W}. \tag{55}
$$

V. A DISPERSIVE MODEL: COUPLING TO AXIONS

Experiments in NLED have been considered as potentially enabling the detection of the axion [30–34], a hypothetical particle introduced as a possible explanation for strong CP invariance in quantum chromodynamics [11–14], and which has been proposed as a candidate for dark matter [47]. As far as electromagnetism is concerned, the axion field couples directly to $\mathcal{G} = \mathbf{E} \cdot \mathbf{B}$, and will thus contribute to the effective photon-photon interaction in NLED. However, there are compelling astrophysical [48] and cosmological [49] reasons to consider an axion mass which is significantly smaller than 1 eV, the energy scale of an optical photon. In this case, the Compton wavelength of the axion is long compared to the typical photon wavelength, and the assumption of purely local effective photon-photon interactions made in Sec. II is explicitly broken. Coupling to the axion field is thus not only of potential experimental relevance (though the question of experimental feasibility is beyond the scope of this paper), but is also of theoretical interest as it will lead to a dispersive model of NLED (as was recently illustrated in [50]).

In this section, starting from the electromagnetic Lagrangian of Eq. (6), we couple the electromagnetic field to an axion field. The analysis of previous sections is carried through in a similar manner; the post-Maxwellian parameters and the polarization of the pump are left unspecified, thereby generalizing the results of [50] (which considered the EH model and a linearly polarized pump). We shall assume that the various plane waves are infinite in extent and duration. This allows us to neglect retardation effects due to the non-instantaneous nature of the axion response, including photon-axion oscillations [34]. Instead, we here focus solely on the refractive index change induced by the axion coupling.

A. Lagrangian and constitutive relations

The total effective Lagrangian, including the coupling to axions, may be written

$$
\mathcal{L} = \mathcal{L}_{\text{EM}} + \mathcal{L}_{\text{ax}} + \mathcal{L}_{\text{int}}. \tag{56}
$$

Here, $\mathcal{L}_{\text{EM}}$ can be considered as the "local part" of the effective Lagrangian containing the electromagnetic fields alone, and is just that used in previous sections [and given in Eq. (6)].

---

[10]Reference [36] studied this particular geometry in the context of the EH model. Reference [38] studied a suboptimal geometry in which the waves have parallel polarizations and a tilt angle of $90°$, yielding an overall reduction factor of 7 in the value of n_2 with respect to the optimal value, $n_{2,\perp}^{(\text{EH})}$ in Eqs. (52).

The next term is the Lagrangian of the free axion field:

$$\mathcal{L}_{\text{ax}} = \tfrac{1}{2}(\partial_{ct}\phi)^2 - \tfrac{1}{2}(\nabla\phi)^2 - \tfrac{1}{2}k_{\text{ax}}^2\phi^2, \tag{57}$$

where $k_{\text{ax}} = m_{\text{ax}}c/\hbar$ is the wave vector associated to the Compton wavelength of the axion. Finally, the interaction between the axion and electromagnetic fields is described by

$$\mathcal{L}_{\text{int}} = -\eta\phi\mathcal{G}, \tag{58}$$

where the inverse square of the coupling constant, η^{-2}, has dimensions of energy per unit length.

Employing the separation into background and probe fields described in Sec. III A, and using a similar decomposition for the axion field $\phi = \phi_0 + \delta\phi$, we may write the total Lagrangian as

$$\begin{aligned}\mathcal{L} &= \mathcal{L}_{\text{EM}}(\mathbf{E}_0 + \mathbf{e}, \mathbf{B}_0 + \mathbf{b}) + \mathcal{L}_{\text{ax}}(\phi_0 + \delta\phi)\\ &\quad + \mathcal{L}_{\text{int}}(\mathbf{E}_0 + \mathbf{e}, \mathbf{B}_0 + \mathbf{b}, \phi + \delta\phi)\\ &\approx \mathcal{L}_0 + \mathcal{L}_{\text{probe}},\end{aligned} \tag{59}$$

where

$$\mathcal{L}_0 = \mathcal{L}_{\text{EM}}(\mathbf{E}_0, \mathbf{B}_0) + \mathcal{L}_{\text{ax}}(\phi_0) + \mathcal{L}_{\text{int}}(\mathbf{E}_0, \mathbf{B}_0, \phi_0) \tag{60}$$

is the Lagrangian associated with the background alone and, keeping only terms which are quadratic in the weak fields of the probe,

$$\begin{aligned}\mathcal{L}_{\text{probe}} &= \mathcal{L}_{\text{EM,probe}}(\mathbf{E}_0, \mathbf{B}_0; \mathbf{e}, \mathbf{b}) + \mathcal{L}_{\text{ax}}(\delta\phi)\\ &\quad - \eta(\phi_0\mathbf{e}\cdot\mathbf{b} + \delta\phi\mathbf{E}_0\cdot\mathbf{b} + \delta\phi\mathbf{B}_0\cdot\mathbf{e}).\end{aligned} \tag{61}$$

$\mathcal{L}_{\text{EM,probe}}$ is exactly the probe Lagrangian derived in Sec. III A, while $\mathcal{L}_{\text{ax}}$ (which is already purely quadratic) is again just the Lagrangian for a free axion field. The terms proportional to η describe the interplay between the electromagnetic and axion fields associated with the passage of the probe wave. Since we deal with a restricted class of background configurations,

we can simplify this term further. Noting that we deal either with plane-wave background fields, which satisfy $\mathcal{G}_0 = 0$ and hence $\phi_0 = 0$, or with constant background fields, in which case the term in $\mathbf{e}\cdot\mathbf{b}$ is a total derivative when expressed in terms of the vector potential [see the discussion following Eqs. (5), including footnote 4] to the probe fields, we have

$$\begin{aligned}\mathbf{d} &= \mathbf{d}_{\text{EM}} - \eta\delta\phi\mathbf{B}_0,\\ \mathbf{h} &= \mathbf{h}_{\text{EM}} + \eta\delta\phi\mathbf{E}_0,\end{aligned} \tag{62}$$

where $\mathbf{d}_{\text{EM}}$ and $\mathbf{h}_{\text{EM}}$ are due to $\mathcal{L}_{\text{EM,probe}}$ alone and are already defined in Eqs. (16). Our aim is to subject these relations to the same treatment as in Sec. III; in particular, to find the new forms of Eqs. (24) relating the amplitudes $\mathbf{d}^{(0)}$ and $\mathbf{h}^{(0)}$ directly to $\mathbf{e}^{(0)}$, so that a homogeneous linear equation analogous to Eq. (25) is obtained. Our first task, then, is to determine the axion field $\delta\phi$ generated by a probe wave of amplitude $\mathbf{e}^{(0)}$.

B. Response of axion field to passage of probe wave

The response of $\delta\phi$ to the presence of electromagnetic fields is determined by the following equation of motion:

$$\left[\partial_{ct}^2 - \nabla^2 + k_{\text{ax}}^2\right]\delta\phi = -\eta\delta\mathcal{G}, \tag{63}$$

where we have defined $\delta\mathcal{G} = \mathbf{E}_0\cdot\mathbf{b} + \mathbf{B}_0\cdot\mathbf{e}$. The probe-induced $\delta\mathcal{G}$ thus acts as a source for $\delta\phi$, with a simple relationship between their Fourier components:

$$\delta\phi_{\omega',k'} = \eta\frac{\delta\mathcal{G}_{\omega',k'}}{(\omega'/c)^2 - (k')^2 - k_{\text{ax}}^2}, \tag{64}$$

assuming of course that we are not at resonance, i.e., $(\omega'/c)^2 - (k')^2 - k_{\text{ax}}^2 \neq 0$. Using Eqs. (38) for the background or pump fields (noting that they reduce to constant fields when $\mathbf{k}_0$ and ω_0 vanish), we have, to lowest order,

$$\delta\mathcal{G} \approx \tfrac{1}{4}\left\{\left[\mathbf{B}_0^{(0)} + \hat{z}\times\mathbf{E}_0^{(0)}\right]e^{i\mathbf{k}_0\cdot\mathbf{r}-i\omega_0 t} + \left[\mathbf{B}_0^{(0)} + \hat{z}\times\mathbf{E}_0^{(0)}\right]^{\star}e^{-i\mathbf{k}_0\cdot\mathbf{r}+i\omega_0 t}\right\}\cdot\mathbf{e}^{(0)}e^{-ikz-i\omega t} + \text{c.c.}, \tag{65}$$

where we have used Eq. (22a), as well as the cyclic invariance of the vector triple product, to write $\mathbf{E}_0^{(0)}\cdot\mathbf{b}_0^{(0)} \approx (\hat{z}\times\mathbf{E}_0^{(0)})\cdot\mathbf{e}^{(0)}$. Using relation (64), we can immediately write down the generated axion field:

$$\begin{aligned}\delta\phi &\approx \frac{\eta}{4}\left\{\frac{\left[\mathbf{B}_0^{(0)} + \hat{z}\times\mathbf{E}_0^{(0)}\right]e^{i\mathbf{k}_0\cdot\mathbf{r}-i\omega_0 t}}{(\omega_0 + \omega)^2/c^2 - (\mathbf{k}_0 - k\hat{z})^2 - k_{\text{ax}}^2} + \frac{\left[\mathbf{B}_0^{(0)} + \hat{z}\times\mathbf{E}_0^{(0)}\right]^{\star}e^{-i\mathbf{k}_0\cdot\mathbf{r}+i\omega_0 t}}{(\omega_0 - \omega)^2/c^2 - (\mathbf{k}_0 + k\hat{z})^2 - k_{\text{ax}}^2}\right\}\cdot\mathbf{e}^{(0)}e^{-ikz-i\omega t} + \text{c.c.}\\ &\approx \frac{\eta}{4}\left\{\frac{\left[\mathbf{B}_0^{(0)} + \hat{z}\times\mathbf{E}_0^{(0)}\right]e^{i\mathbf{k}_0\cdot\mathbf{r}-i\omega_0 t}}{4\cos^2\frac{\theta}{2}\omega_0\omega/c^2 - k_{\text{ax}}^2} - \frac{\left[\mathbf{B}_0^{(0)} + \hat{z}\times\mathbf{E}_0^{(0)}\right]^{\star}e^{-i\mathbf{k}_0\cdot\mathbf{r}+i\omega_0 t}}{4\cos^2\frac{\theta}{2}\omega_0\omega/c^2 + k_{\text{ax}}^2}\right\}\cdot\mathbf{e}^{(0)}e^{-ikz-i\omega t} + \text{c.c.}\end{aligned} \tag{66}$$

In the second line, we have expanded the squares in the denominators, neglecting the variation of the refractive index here so that $k \approx \omega/c$, and used the fact that $\mathbf{k}_0\cdot\hat{z} = k_0\cos\theta$ where θ is the tilt angle between the pump and probe waves introduced in the previous section (and illustrated in Fig. 2).

C. Backreaction of axion field on probe fields

Substituting expression (66) for $\delta\phi$ back into the constitutive relations (62) removes the explicit dependence on the axion field, giving $\mathbf{d}$ and $\mathbf{h}$ in terms of the probe amplitude $\mathbf{e}^{(0)}$ alone. Since $\mathbf{d}_{\text{EM}}$ and $\mathbf{h}_{\text{EM}}$ are already given in Eqs. (24), we need focus here only on the additional terms $\delta\mathbf{d} = -\eta\delta\phi\mathbf{B}_0$ and $\delta\mathbf{h} = \eta\delta\phi\mathbf{E}_0$. Since $\delta\phi$, $\mathbf{E}_0$ and $\mathbf{B}_0$ are all generally oscillatory, the substitution generates rapidly oscillating terms that are far off-shell, much like when calculating $\mathbf{d}$ and $\mathbf{h}$ in Sec. III. As there, we retain only those terms whose oscillations are synchronized with those of the probe, in which case we may write $\delta\mathbf{d} = \text{Re}\{\delta\mathbf{d}^{(0)}\exp(-ikz - i\omega t)\}$ and $\delta\mathbf{h} = \text{Re}\{\delta\mathbf{h}^{(0)}\exp(-ikz - i\omega t)\}$, with amplitudes $\delta\mathbf{d}^{(0)}$ and $\delta\mathbf{h}^{(0)}$ that are linearly

related to $\mathbf{e}^{(0)}$:

$$\delta\mathbf{d}^{(0)} = \delta D_{\text{ax}}\mathbf{e}^{(0)}, \quad \delta\mathbf{h}^{(0)} = \delta H_{\text{ax}}\mathbf{e}^{(0)}. \tag{67}$$

δD_{ax} and δH_{ax} are 3×3 matrices, and direct substitution shows that

$$\delta D_{\text{ax}} = \frac{1}{8}\frac{\eta^2}{k_{\text{ax}}^2}\left\{\frac{\mathbf{B}_0^{(0)\star}[\mathbf{B}_0^{(0)} + \hat{z}\times\mathbf{E}_0^{(0)}]^{\text{T}}}{1 - 4\cos^2\frac{\theta}{2}\omega_0\omega/\omega_{\text{ax}}^2} + \frac{\mathbf{B}_0^{(0)}[\mathbf{B}_0^{(0)} + \hat{z}\times\mathbf{E}_0^{(0)}]^{\star\text{T}}}{1 + 4\cos^2\frac{\theta}{2}\omega_0\omega/\omega_{\text{ax}}^2}\right\},$$

$$\delta H_{\text{ax}} = -\frac{1}{8}\frac{\eta^2}{k_{\text{ax}}^2}\left\{\frac{\mathbf{E}_0^{(0)\star}[\mathbf{B}_0^{(0)} + \hat{z}\times\mathbf{E}_0^{(0)}]^{\text{T}}}{1 - 4\cos^2\frac{\theta}{2}\omega_0\omega/\omega_{\text{ax}}^2} + \frac{\mathbf{E}_0^{(0)}[\mathbf{B}_0^{(0)} + \hat{z}\times\mathbf{E}_0^{(0)}]^{\star\text{T}}}{1 + 4\cos^2\frac{\theta}{2}\omega_0\omega/\omega_{\text{ax}}^2}\right\}, \tag{68}$$

where $\omega_{\text{ax}} = ck_{\text{ax}} = m_{\text{ax}}c^2/\hbar$.

D. Axionic contribution to the refractive index

Writing Eqs. (67) as linear in $\mathbf{e}^{(0)}$ is particularly convenient as it immediately allows us to combine these results with those of Eqs. (24), and then to obtain the modification to Eq. (25). Indeed, it is straightforward to show that the matrix on the left-hand side of Eq. (25) is changed simply by addition of $\delta D_{\text{ax}} - \Omega_z\delta H_{\text{ax}}$, where [in accordance with the observations made just after Eq. (23)] Ω_z acting on δH_{ax} replaces the $\mathbf{E}_0^{(0)}$ and $\mathbf{E}_0^{(0)\star}$ of the second of Eqs. (68) by $\hat{z}\times\mathbf{E}_0^{(0)}$ and $\hat{z}\times\mathbf{E}_0^{(0)\star}$, respectively. Furthermore, as discussed after Eq. (25), we need only consider the xy projection of the matrix in Eq. (25) when working to first order in δn. The relevant combination of δD_{ax} and δH_{ax}, once projected onto the xy plane, depends only on the vector $\mathcal{B}^{(0)}$ of Eqs. (41), and not on $\mathbf{E}_0^{(0)}$ and $\mathbf{B}_0^{(0)}$ separately. In short, to the matrices on the left-hand sides of Eqs. (43) and (47) must be added the following:

$$(\delta D_{\text{ax}} - \Omega_z\delta H_{\text{ax}})|_{xy} = \frac{1}{8}\frac{\eta^2}{k_{\text{ax}}^2}\left\{\frac{\mathcal{B}^{(0)\star}\mathcal{B}^{(0)\text{T}}}{1 - 4\cos^2\frac{\theta}{2}\omega_0\omega/\omega_{\text{ax}}^2} + \frac{\mathcal{B}^{(0)}\mathcal{B}^{(0)\star\text{T}}}{1 + 4\cos^2\frac{\theta}{2}\omega_0\omega/\omega_{\text{ax}}^2}\right\}$$

$$\equiv \frac{1}{4}\left[\Delta_{2,\text{ax}}(\omega)\left\{\cos^2\chi\,\mathcal{B}_{\text{abs}}^{(0)}\mathcal{B}_{\text{abs}}^{(0)\text{T}} + \sin^2\chi\,\mathcal{E}_{\text{abs}}^{(0)}\mathcal{E}_{\text{abs}}^{(0)\text{T}}\right\} + i\Delta_{3,\text{ax}}(\omega)\sin(2\chi)\left\{\mathcal{E}_{\text{abs}}^{(0)}\mathcal{B}_{\text{abs}}^{(0)\text{T}} - \mathcal{B}_{\text{abs}}^{(0)}\mathcal{E}_{\text{abs}}^{(0)\text{T}}\right\}\right], \tag{69}$$

where in the last line we have used the decomposition of $\mathcal{B}^{(0)}$ given in Eqs. (45), and defined

$$\Delta_{2,\text{ax}}(\omega) = \frac{\eta^2}{k_{\text{ax}}^2}\frac{1}{1 - \cos^4\frac{\theta}{2}\left(4\omega_0\omega/\omega_{\text{ax}}^2\right)^2}, \quad \Delta_{3,\text{ax}}(\omega) = \frac{\eta^2}{k_{\text{ax}}^2}\frac{2\cos^2\frac{\theta}{2}\omega_0\omega/\omega_{\text{ax}}^2}{1 - \cos^4\frac{\theta}{2}\left(4\omega_0\omega/\omega_{\text{ax}}^2\right)^2}. \tag{70}$$

The eigenproblem thus becomes:

$$\frac{1}{4}\left[\delta_1''(\omega)\mathcal{E}_{\text{abs}}^{(0)}\mathcal{E}_{\text{abs}}^{(0)\text{T}} + \delta_2''(\omega)\mathcal{B}_{\text{abs}}^{(0)}\mathcal{B}_{\text{abs}}^{(0)\text{T}} + \delta_3''(\omega)\mathcal{E}_{\text{abs}}^{(0)}\mathcal{B}_{\text{abs}}^{(0)\text{T}} + \delta_3''^\star(\omega)\mathcal{B}_{\text{abs}}^{(0)}\mathcal{E}_{\text{abs}}^{(0)\text{T}}\right]\mathbf{e}^{(0)} = \delta n(\omega)\mathbf{e}^{(0)}, \tag{71}$$

where

$$\delta_1''(\omega) + \delta_2''(\omega) = \delta_1 + \delta_2 + \Delta_{2,\text{ax}}(\omega),$$

$$\delta_1''(\omega) - \delta_2''(\omega) = [\delta_1 - \delta_2 - \Delta_{2,\text{ax}}(\omega)]\cos(2\chi),$$

$$\delta_3''(\omega) = \delta_3\cos(2\chi) + i\Delta_{3,\text{ax}}(\omega)\sin(2\chi). \tag{72}$$

As a matrix equation, this is simply

$$\begin{bmatrix} \delta_1''(\omega) & \delta_3''(\omega) \\ \delta_3''^\star(\omega) & \delta_2''(\omega) \end{bmatrix}\begin{bmatrix} e_{\mathcal{E}}^{(0)} \\ e_{\mathcal{B}}^{(0)} \end{bmatrix} = \frac{\delta n(\omega)}{\frac{1}{4}\left|\mathcal{E}_{\text{abs}}^{(0)}\right|^2}\begin{bmatrix} e_{\mathcal{E}}^{(0)} \\ e_{\mathcal{B}}^{(0)} \end{bmatrix}. \tag{73}$$

The eigenvalues of the matrix on the left-hand side of this equation are calculated as before, and the refractive index changes $\delta n_\pm$ are again given by Eq. (50) with the now frequency-dependent factors $\delta_\pm''(\omega)$ taking the form

$$\delta_\pm''(\omega) = \frac{1}{2}\left(\delta_1''(\omega) + \delta_2''(\omega) \pm \sqrt{[\delta_1''(\omega) - \delta_2''(\omega)]^2 + |2\delta_3''(\omega)|^2}\right)$$

$$= \frac{1}{2}\left(\delta_1 + \delta_2 + \Delta_{2,\text{ax}}(\omega) \pm \sqrt{\{[\delta_1 - \delta_2 - \Delta_{2,\text{ax}}(\omega)]^2 + (2\delta_3)^2\}\cos^2(2\chi) + [2\Delta_{3,\text{ax}}(\omega)]^2\sin^2(2\chi)}\right). \tag{74}$$

In the EH model with a linearly polarized pump, the only change with respect to previous sections is that $\delta_2 \to \delta_2 + \Delta_{2,\text{ax}}(\omega)$. This is in agreement with Eq. (13) of [50] (an apparent difference by a factor of 4 arising only because of different definitions of the pump amplitude).

We briefly mention here that the group index $n_g = n + \omega dn/d\omega$, though generally complicated, can be fairly easily calculated in the limits of linear and circular polarization. Although $n < 1$ for some frequencies, we find (at least in these two limits) that $n_g > 1$ for *all* frequencies, so that relativistic causality is respected (as explained in footnote 9).

E. Eigenpolarizations

Generally speaking, the matrix on the left-hand side of Eq. (73) is Hermitian but *not* symmetric, thanks to the imaginary contribution to the off-diagonal component $\delta_3''(\omega)$. Unlike in previous sections, this means that its eigenvectors are generally complex, which in turn means that the eigenpolarizations are elliptically polarized. Two angles are required to describe these eigenpolarizations: an ellipticity angle $\psi \in [-\pi/4, \pi/4]$ giving the degree of elliptical polarization, and a rotation angle $\varphi \in (-\pi/2, \pi/2]$ giving the orientation of the major axes of the ellipses traced by $\mathbf{e}(t)$ and $\mathbf{b}(t)$ with respect to the basis $\{\hat{\mathcal{E}}, \hat{\mathcal{B}}\}$. The (normalized) eigenvectors thus form the columns of a unitary matrix $U(\psi, \varphi)$ which can be decomposed as follows:

$$U(\psi, \varphi) = E(\psi)R(\varphi) = \begin{bmatrix} \cos\psi & i\sin\psi \\ i\sin\psi & \cos\psi \end{bmatrix} \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}. \tag{75}$$

The matrix on the left-hand side of Eq. (73) can be written as $U(\psi, \varphi)DU^{-1}(\psi, \varphi)$, where D is a diagonal matrix whose entries are the eigenvalues δ_+'' and δ_-'' of Eq. (74). As before, we take δ_+'' to be the first diagonal component of D, so that the left column of U corresponds to the polarization with the larger refractive index change δn_+. Explicit calculation of this form of the matrix yields the following relations, which determine ψ and φ:

$$[\delta_1 - \delta_2 - \Delta_{2,\mathrm{ax}}(\omega)]\cos(2\chi) = [\delta_+''(\omega) - \delta_-''(\omega)]\cos[2\varphi(\omega)]\cos[2\psi(\omega)],$$

$$-2\Delta_{3,\mathrm{ax}}(\omega)\sin(2\chi) = [\delta_+''(\omega) - \delta_-''(\omega)]\cos[2\varphi(\omega)]\sin[2\psi(\omega)],$$

$$2\delta_3\cos(2\chi) = [\delta_+''(\omega) - \delta_-''(\omega)]\sin[2\varphi(\omega)]. \tag{76}$$

While in general the solutions $\psi(\omega)$ and $\varphi(\omega)$ are quite complicated, they become rather simple in the two polarization limits of the pump. When the pump is linearly polarized ($\chi = 0$), we find $\psi = 0$ so that the probe polarizations are also linearly polarized; the only nontriviality is in their rotation with respect to the pump fields, the rotation angle being given by Eqs. (37) with $\delta_2 \rightarrow \delta_2 + \Delta_{2,\mathrm{ax}}(\omega)$. On the other hand, for a circularly polarized pump ($\chi = \pm\pi/4$), we find $\psi = \pm\pi/4$ too, i.e., the eigenpolarizations of the probe are themselves circularly polarized. In the high-frequency regime where $\Delta_{3,\mathrm{ax}}(\omega) < 0$, it is straightforward to show[11] that the larger refractive index change δn_+ is felt by the probe rotating in the same sense as $\mathcal{E}(t)$ and $\mathcal{B}(t)$, or equivalently (since in Fig. 2 the probe propagates *into* the page) by the probe with opposite handedness to that of the pump.

F. Discussion of key axionic effects

The terms $\Delta_{2,\mathrm{ax}}(\omega)$ and $\Delta_{3,\mathrm{ax}}(\omega)$ describe two key effects induced by the coupling to axions. Moreover, each is dominant in a particular regime, for we have

$$\frac{\Delta_{3,\mathrm{ax}}(\omega)}{\Delta_{2,\mathrm{ax}}(\omega)} = 2\cos^2\frac{\theta}{2}\frac{\omega_0\omega}{\omega_{\mathrm{ax}}^2}. \tag{77}$$

Therefore, whether $\omega_0\omega/\omega_{\mathrm{ax}}^2$ is large or small compared to 1 determines which of the two terms is dominant. (Both can be considered large when $\omega_0\omega/\omega_{\mathrm{ax}}^2 \approx 1$, but then we are close to resonance, and the analysis performed here will cease to be valid.)

---

[11]There is a slight complication as there exists a degeneracy in φ and ψ [as defined by Eqs. (76)] when $\chi = \pm\pi/4$: either $\varphi = 0$ and $\psi = \chi$, or $\varphi = \pi/2$ and $\psi = -\chi$. However, plugging these solutions into Eq. (75), we find that switching between them simply amounts to multiplication of the columns of $U(\psi, \varphi)$ by $\pm i$, so that the eigenvectors remain unchanged (up to an unimportant phase).

1. Renormalization of δ_2

First, the parameter δ_2 entering the local part of the effective Lagrangian is renormalized in a frequency-dependent way to become $\delta_2 + \Delta_{2,\mathrm{ax}}(\omega)$. When $\omega_0\omega/\omega_{\mathrm{ax}}^2$ is very large, it is clear from Eqs. (70) that $\Delta_{2,\mathrm{ax}}(\omega) \rightarrow 0$. We may thus focus on the opposite limit, $\omega_0\omega/\omega_{\mathrm{ax}}^2 \rightarrow 0$, where the Compton wavelength of the axion is much smaller than the typical photon wavelengths (or indeed when the background fields are static). In this case, we have $\Delta_{2,\mathrm{ax}}(\omega) \rightarrow \eta^2/k_{\mathrm{ax}}^2$, with no residual dependence on frequency. This is the limit in which the effective photon-photon interaction becomes local; η^2/k_{ax}^2 can be absorbed into the definition of δ_2, and the analysis of Sec. IV carries through as before. We are thus provided with an explicit demonstration of the results of the local effective theory emerging in the correct limit.

For the first experimental probes of NLED, this renormalization is only expected to be significant if it is comparable to the "bare" EH value, $\delta_2^{(\mathrm{EH})}$ [see Eqs. (9)]. In units with $\hbar = c = 1$, this requires $\eta/m_{\mathrm{ax}} \gtrsim \sqrt{\delta_2^{(\mathrm{EH})}} \approx 3 \times 10^4$ GeV^{-2}. This has not yet been ruled out by current experimental tests of axions but, since values of η larger than about 10^{-10} GeV^{-1} have been essentially excluded (see Fig. 111.1 of [51]), this would require an axion mass smaller than a few times 10^{-6} eV. This makes it less likely that the condition $\omega_0\omega/\omega_{\mathrm{ax}}^2 \gg 1$ will be satisfied, unless the background field is essentially static. For the DeLLight experiment [25], which uses optical wavelengths with $\hbar\omega \sim 1$ eV for both the background and the probe fields, this renormalization of δ_2 is unlikely to be significant.

2. Elliptical birefringence

The second key axionic effect is the contribution $\propto \Delta_{3,\mathrm{ax}}(\omega)$ to the off-diagonal terms in Eq. (73). Interestingly, this contribution is imaginary, and leads to probe eigenstates which are elliptically polarized. It is clear that $\Delta_{3,\mathrm{ax}}(\omega) \rightarrow 0$ when $\omega_0\omega/\omega_{\mathrm{ax}}^2 \rightarrow 0$, as it should, since we then recover the results of the local effective theory where (as we have seen

in Sec. IV) the relevant matrix is purely real. In the opposite limit, we have already seen [in Eq. (77)] that $\Delta_{3,\mathrm{ax}}(\omega)$ becomes the dominant signature of the axion coupling, and yet, since it enters Eqs. (72) with a factor of $\sin(2\chi)$, it completely drops out when the pump is linearly polarized. We thus conclude that, when working in the limit $\omega_0\omega/\omega_{\mathrm{ax}}^2 \gg 1$, a circularly polarized pump is much more efficient than a linearly polarized one at inducing an axionic signature in the refractive index. This result is made all the more interesting by the fact that this axionic signature includes birefringence, since (as seen in Sec. IV C) models with local effective photon-photon interactions show no birefringence at all when the pump is circularly polarized.

Working in the appropriate limit where $\omega_0\omega/\omega_{\mathrm{ax}}^2$ is very large,

$$\Delta_{3,\mathrm{ax}}(\omega) \to -\frac{\eta^2 c^2}{8}\frac{1}{\cos^2\frac{\theta}{2}\omega_0\omega}, \tag{78}$$

and we see explicitly that this limiting case is independent of the actual value of m_{ax}. Moreover, in the optimized scenario with $\theta = 0$ and $\chi = \pm\pi/2$ (i.e., a counterpropagating probe and a circularly polarized pump), we have from Eq. (74) that the birefringence $\delta_+ - \delta_- = 2|\Delta_{3,\mathrm{ax}}(\omega)|$. It is appropriate to compare this with $\delta_2^{(\mathrm{EH})} - \delta_1^{(\mathrm{EH})}$, the birefringence predicted by the EH model (in the absence of axions) for a linearly polarized pump. In units with $\hbar = c = 1$, they are comparable when $\eta/\sqrt{\omega_0\omega} \sim 2\sqrt{\delta_2^{(\mathrm{EH})} - \delta_1^{(\mathrm{EH})}} \approx 4 \times 10^4~\mathrm{GeV}^{-2}$. At optical frequencies (as used in the DeLLight experiment [25]), and given that current exclusion plots indicate η is at most $\sim 10^{-10}~\mathrm{GeV}^{-1}$ (see Fig. 111.1 of [51]), this circular birefringence turns out to be at least six orders of magnitude smaller than the linear birefringence of the EH model.

VI. SUMMARY AND CONCLUSION

Starting from the most general Lagrangian for electromagnetic fields consistent with Lorentz invariance, locality of effective interactions, and weak nonlinearities (so that the lowest nonlinear contributions are sufficient to describe the physics), we have derived an effective-medium description for the propagation of weak probe waves in the presence of strong background fields, whose effects on the probe can be incorporated through well-defined electric, magnetic, and magnetoelectric susceptibility tensors. This description allows the assignment of a refractive index to the effective medium, though the index typically exhibits anisotropy and birefringence. In the case where the background fields are provided by an intense propagating wave or pump, there is a further dependence of the refractive index on the degree of elliptical polarization of the pump. The effects of wave vector direction and polarization turn out to be neatly separated: the misalignment of the wave vectors of pump and probe is equivalent to an overall reduction in the pump intensity by a factor of $\cos^4(\theta/2)$, where θ is the tilt angle between the two wave vectors; and the ellipticity angle χ enters into the strength of the birefringence with a factor of $\cos(2\chi)$, being maximum for a linearly polarized pump and vanishing for a circularly polarized one. Finally, factoring out the intensity of the pump allows us to extract the nonlinear index

of vacuum, which in the Euler-Heisenberg model derived from QED is typically on the order of $10^{-33}~\mathrm{cm}^2/\mathrm{W}$, around 18 orders of magnitude smaller than in nonlinear optical media.

Locality of effective photon-photon interactions was relaxed as a constraint through coupling the electromagnetic field to an axion field of unspecified mass. Generation of an axion field by the interaction between pump and probe, followed by backreaction of the axion field on the probe, yields a contribution to the effective photon-photon coupling which is dispersive when the axion's Compton wavelength is larger than or of the same order as the typical photon wavelength. In the case of a linearly polarized pump, this amounts to a straightforward renormalization of one of the post-Maxwellian parameters entering the Lagrangian, but when the pump is circularly polarized, there exists a residual birefringence that (as mentioned above) would vanish if the effective photon-photon interactions were purely local. Whether the axionic contribution to the birefringence is larger for a linearly or circularly polarized pump depends on the typical photon wavelength (defined via the geometric mean of the pump and probe frequencies): if it is much smaller than the Compton wavelength of the axion, the effect is larger for a circularly polarized pump; conversely, if the typical photon wavelength is much larger than the Compton wavelength of the axion, the effect is larger for a linearly polarized pump.

Concerning the DeLLight experiment [25], whose aim is to detect the deflection of a probe wave by the index variation induced by a tightly focused laser pulse, the results presented here confirm that an effect of this kind should indeed be seen, and any effects due to axions are expected to be subdominant. However, as discussed after Eqs. (15) (particularly in footnote 7), we have shown that only a single pump pulse is required, rather than two counterpropagating pump pulses as proposed in [25]. We conclude that the proposal presented in [25] can be simplified by keeping only the pump pulse which is counterpropagating with respect to the probe. We are currently performing numerical simulations of the simplified DeLLight experiment, whose results will be published in a future work. It is worth noting that the two-pump proposal of [25] would be an interesting case in which four-wave mixing processes could turn out to be important, as the two counterpropagating pump pulses would engender a stationary oscillation in space that could act as a diffraction grating for the probe [52]; this, however, is beyond the scope of the present work.

[1] M. Faraday, Philos. Mag. **28**, 294 (1846).

[2] N. Bloembergen, Rev. Mod. Phys. **54**, 685 (1982).

[3] N. Bloembergen, *Nonlinear Optics*, 4th ed. (World Scientific, Singapore, 1996).

[4] Y. R. Shen, *The Principles of Nonlinear Optics*, 1st ed. (Wiley, New York, 1984).

[5] R. Boyd, *Nonlinear Optics*, 3rd ed. (Academic, New York, 2008).

[6] G. P. Agrawal, *Nonlinear Fiber Optics*, 4th ed. (Academic, New York, 2007).

[7] J. Kerr, Philos. Mag. **50**, 337 (1875).

[8] H. Euler and B. Kockel, Naturwissenschaften **23**, 246 (1935).

[9] H. Euler, Ann. Phys. (Leipzig) **418**, 398 (1936).

[10] W. Heisenberg and H. Euler, Z. Phys. **98**, 714 (1936).

[11] R. D. Peccei and H. R. Quinn, Phys. Rev. Lett. **38**, 1440 (1977).

[12] R. D. Peccei and H. R. Quinn, Phys. Rev. D **16**, 1791 (1977).

[13] F. Wilczek, Phys. Rev. Lett. **40**, 279 (1978).

[14] S. Weinberg, Phys. Rev. Lett. **40**, 223 (1978).

[15] M. Born and L. Infeld, Proc. R. Soc. London, Ser. A **144**, 425 (1934).

[16] R. Battesti and C. Rizzo, Rep. Prog. Phys. **76**, 016401 (2013).

[17] M. Fouché, R. Battesti, and C. Rizzo, Phys. Rev. D **93**, 093020 (2016); **95**, 099902 (2017).

[18] M. Aaboud *et al.* (ATLAS Collaboration), Nat. Phys. **13**, 852 (2017).

[19] R. Battesti, B. Pinto Da Souza, S. Batut, C. Robilliard, G. Bailly, C. Michel, M. Nardone, L. Pinard, O. Portugall, G. Trénec, J.-M. Mackowski, G. L. J. A. Rikken, J. Vigué, and C. Rizzo, Eur. Phys. J. D **46**, 323 (2008).

[20] A. Cadène, P. Berceau, M. Fouché, R. Battesti, and C. Rizzo, Eur. Phys. J. D **68**, 16 (2014).

[21] G. Zavattini, U. Gastaldi, R. Pengo, G. Ruoso, F. Della Valle, and E. Milotti, Int. J. Mod. Phys. A **27**, 1260017 (2012).

[22] 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).

[23] R. V. Jones, Nature (London) **186**, 706 (1960).

[24] R. V. Jones, Proc. R. Soc. London, Ser. A **260**, 47 (1961).

[25] X. Sarazin, F. Couchot, A. Djannati-Ataï, O. Guilbaud, S. Kazamias, M. Pittman, and M. Urban, Eur. Phys. J. D **70**, 13 (2016).

[26] J. Plebański, *Lectures on Nonlinear Electrodynamics* (NORDITA, Copenhagen, 1970).

[27] G. Boillat, J. Math. Phys. **11**, 941 (1970).

[28] Z. Bialynicka-Birula and I. Bialynicki-Birula, Phys. Rev. D **2**, 2341 (1970).

[29] V. I. Denisov and I. P. Denisova, Dokl. Phys. **46**, 377 (2001).

[30] P. Sikivie, Phys. Rev. Lett. **51**, 1415 (1983).

[31] L. Maiani, R. Petronzio, and E. Zavattini, Phys. Lett. B **175**, 359 (1986).

[32] M. Gasperini, Phys. Rev. Lett. **59**, 396 (1987).

[33] K. Van Bibber, N. R. Dagdeviren, S. E. Koonin, A. K. Kerman, and H. N. Nelson, Phys. Rev. Lett. **59**, 759 (1987).

[34] G. Raffelt and L. Stodolsky, Phys. Rev. D **37**, 1237 (1988).

[35] E. Brezin and C. Itzykson, Phys. Rev. D **3**, 618 (1971).

[36] E. B. Aleksandrov, A. A. Ansel'm, and A. N. Moskalev, Zh. Eksp. Teor. Fiz. **89**, 1181 (1985) [J. Exp. Theor. Phys. **62**, 680 (1985)].

[37] N. N. Rozanov, Zh. Eksp. Teor. Fiz. **103**, 1996 (1993) [J. Exp. Theor. Phys. **76**, 991 (1993)].

[38] A. Ferrando, H. Michinel, M. Seco, and D. Tommasini, Phys. Rev. Lett. **99**, 150404 (2007).

[39] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, 4th ed. (Butterworth-Heinemann, Oxford, 1980).

[40] T.-D. Lee and C.-N. Yang, Phys. Rev. **104**, 254 (1956).

[41] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, Phys. Rev. **105**, 1413 (1957).

[42] F. Moulin and D. Bernard, Opt. Commun. **164**, 137 (1999).

[43] S. L. Adler, Ann. Phys. (NY) **67**, 599 (1971).

[44] G. L. J. A. Rikken and C. Rizzo, Phys. Rev. A **63**, 012107 (2000).

[45] G. L. J. A. Rikken and C. Rizzo, Phys. Rev. A **67**, 015801 (2003).

[46] W. F. Brown, Jr., R. M. Hornreich, and S. Shtrikman, Phys. Rev. **168**, 574 (1968).

[47] P. Arias, D. Cadamuro, M. Goodsell, J. Jaeckel, J. Redondo, and A. Ringwald, J. Cosmol. Astropart. Phys. **2012**, 013 (2012).

[48] G. G. Raffelt, in *Axions*, edited by M. Kuster, G. Raffelt, and B. Beltrán (Springer-Verlag, Berlin, Heidelberg, 2008), pp. 51–71.

[49] D. J. E. Marsh, Phys. Rep. **643**, 1 (2016).

[50] C. Baldenegro, S. Hassani, C. Royon, and L. Schoeffel, Phys. Lett. B **795**, 339 (2019).

[51] M. Tanabashi *et al.* (Particle Data Group), Phys. Rev. D **98**, 030001 (2018).

[52] A. Di Piazza, K. Z. Hatsagortsyan, and C. H. Keitel, Phys. Rev. Lett. **97**, 083603 (2006).