

Covariant electrodynamics in linear media: Optical metric

Robert T. Thompson*

Institute of Applied Physics, Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany

(Received 19 December 2017; published 2 March 2018)

While the postulate of covariance of Maxwell's equations for all inertial observers led Einstein to special relativity, it was the further demand of general covariance—form invariance under general coordinate transformations, including between accelerating frames—that led to general relativity. Several lines of inquiry over the past two decades, notably the development of metamaterial-based transformation optics, has spurred a greater interest in the role of geometry and space-time covariance for electrodynamics in ponderable media. I develop a generally covariant, coordinate-free framework for electrodynamics in general dielectric media residing in curved background space-times. In particular, I derive a relation for the spatial medium parameters measured by an arbitrary timelike observer. In terms of those medium parameters I derive an explicit expression for the pseudo-Finslerian optical metric of birefringent media and show how it reduces to a pseudo-Riemannian optical metric for nonbirefringent media. This formulation provides a basis for a unified approach to ray and congruence tracing through media in curved space-times that may smoothly vary among positively refracting, negatively refracting, and vacuum.

DOI: [10.1103/PhysRevD.97.065001](https://doi.org/10.1103/PhysRevD.97.065001)**I. INTRODUCTION**

The standard vectorial representation of Maxwell's equations,

$$\nabla \cdot \vec{B} = 0, \quad \nabla \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0, \quad (1a)$$

$$\nabla \cdot \vec{D} = \rho, \quad \nabla \times \vec{H} - \frac{\partial \vec{D}}{\partial t} = \vec{j}, \quad (1b)$$

has been the electrodynamics workhorse since it was codified by Heaviside well over a century ago. But this 3-vector representation harbors a hidden assumption: that space-time is Minkowskian. We may forgive Heaviside this trespass, coming 30 years before Einstein's seminal paper on general relativity, but no description of electrodynamics would be truly complete without incorporating the properties of the underlying space-time in which the fields exist.

While the relativistic nature of electrodynamics is fully acknowledged, the general space-time formulation is rarely employed in applications. Standard, widely regarded texts on electrodynamics or relativity typically only discuss tensorial electrodynamics of the vacuum [1–3]. But ponderable media do not displace space-time, as a whale displaces water. Rather, they are completely surrounded and permeated by the geometry of space-time—even the whale in water feels gravity. Thus no theory of macroscopic electrodynamics within media is truly complete without incorporating the properties of the underlying space-time in which the medium resides.

Here on Earth, most practical applications of electrodynamics do not require the heavy artillery of general relativity, but there are still good reasons to be interested in a space-time covariant formulation of electrodynamics within media. In particular, it is closely related to a convergence of ideas in transformation optics, analog gravity, premetric electrodynamics, and Lorentz violating space-times that have emerged over the past 15 years.

Transformation optics uses ideas about the similarity of the refractive properties of dielectric media with the light-bending properties of curved space-times to design optical media with unusual properties or functionality, such as negatively refracting media and invisibility cloaks [4–8]. The advent of structured composite metamaterials that possess some of the unnatural properties required to actually realize these unusual devices [9–11] provides an exciting demonstration that the geometrical, space-time manifold aspects of electrodynamics and light propagation are not just an academic abstraction, but can have real applications in engineering. Progress understanding and developing this new theoretical technology relies on the type of covariant formulation of electrodynamics in media studied here [12–20].

Analog models of curved space-times seek to replicate certain aspects of extreme gravitational systems, such as light propagation near black holes, in a nongravitational laboratory-accessible system, such as sound or surface waves in flowing fluids [21–24], or, more relevantly, light propagating through a suitable dielectric medium [25–34]. Once again it is the space-time covariant formulation of electrodynamics in media and the geometrical aspects of

*robert@cosmos.phy.tufts.edu

light propagation on manifolds that enables detailed study of dielectric analog space-times and their limitations [35].

Premetric electrodynamics is based on the observation that the space-time metric does not enter Maxwell's equations directly, but only through its associated Hodge dual [36,37]. By promoting the Hodge dual in Maxwell's equations to an independent structural field on a metric-free manifold, one may gain new insights into the structure of electrodynamics.

Last, Lorentz-violating space-times play a significant role in theories of quantum gravity and physics beyond the standard model, and it has been shown that Lorentz violation in effective field theories is connected to pseudo-Finsler geometries [38–40]. The Lorentzian nature of space-time was inspired by Maxwell's equations in vacuum, but it is known that the birefringence exhibited by general linear media is connected to pseudo-Finsler geometries [41–43], providing a natural setting for further musings on the nature of space-time.

In short, Finsler geometry is just Riemannian geometry without the quadratic restriction [44]. Riemannian metric geometry is based on a line element of the form

$$ds = F(x^1, \dots, x^n; dx^1, \dots, dx^n), \quad (2)$$

where F is a function on the tangent bundle such that

$$F^2 = g_{\alpha\beta}(x) dx^\alpha dx^\beta, \quad (3)$$

which provides the metric structure

$$g_{\alpha\beta}(x) = \frac{\partial F^2}{\partial x^\alpha \partial x^\beta}. \quad (4)$$

Finsler geometry relaxes this quadratic restriction on F , so that one may still define the metric in a similar way, but it is no longer independent of the cotangent basis. In other words, the metric depends not only on the point on the manifold but also on which direction you look [45].

A. Goal of this paper

The goal of this paper is to construct a covariant form of electrodynamics in linear media residing in a curved background space-time. Some elements of this formalism have been introduced and used in several papers studying transformation optics and dielectric analog space-times [16,31,46]. Here, those initial elements are extended, given a rigorous derivation, and combined into a comprehensive framework. In particular, the relationship between the four-dimensional $\binom{2}{2}$ material tensor χ and the usual spatial material parameters ϵ , μ , ${}^h\gamma$, and ${}^e\gamma$ is derived *ab initio*, and an expression for the pseudo-Finslerian optical metric is derived for quite general media and given in terms of the aforementioned usual spatial material parameters.

In keeping with the tenets of general covariance, the sought after formalism should be independent of the chosen coordinates in the sense that there exists a well-defined

method of shifting between coordinates, and therefore should be expressible entirely in terms of four-dimensional tensorial objects and the metric. The exterior calculus of differential forms is particularly well suited to the study of electrodynamics and will be adopted here.

Certainly this is not the first paper to study the covariant form of electrodynamics in media, which has a long history starting with Minkowski, Gordon, and Tamm in the early part of the last century [47–49], followed by Balazs, Quan, Plebanski, Post, and Ehlers [50–54] in the middle of the last century, and in more modern times by Perlick [41], Novello and Salim [55], Novello, Lorenci, Salim, and Klippert [56], Novello and Bittencourt [57], Visser, Barceló, and Liberati [58], Balakin and Zimdahl [59], and notably as part of the premetric community by Hehl, Obukhov, Rubilar, Lämmerzahl, and Itin [36,42,60]. Indeed, after more than 100 years it is still a fruitful area of research, with a very recent and thorough analysis by Schuster and Visser that is strongly related to the work pursued here [61], where they also consider a background metric and introduce the space-time splitting with respect to a timelike observer.

There are a number of mathematical similarities and common themes between these previous works and what is presented here. For example, differential forms are also used extensively by the premetric community, so some of the equations here have close counterparts in premetric electrodynamics.

It is quite common in the literature to adopt a 6-vector approach when dealing with electrodynamics in media; in other words, instead of the usual space-time formulation whereby (\vec{D}, \vec{H}) and (\vec{E}, \vec{B}) are select components of two-forms, they are put into two six-component vectors. While there is nothing invalid about such an approach, it does suffer from two drawbacks: (1) it subtly assumes an explicit $3 + 1$ decomposition of the space-time, which means that the field components have been rather rigidly defined, and makes the transformation properties somewhat less obvious, and (2) it makes the theory less compatible with other generally relativistic calculations one might want to perform, such as examining the Raychaudhuri equation for a congruence of light passing through ponderable media [46]. Instead, everything done here is fully covariant, four-dimensional, coordinate-free, and largely index-free. Of course, reference to the usual spatial parameters naturally requires a $3 + 1$ split, but here this is done with respect to a timelike vector field that need not align in any special way with a choice of coordinates. I believe this point of view provides additional insight into the structure of the theory by highlighting the role of the observer who must ultimately be making measurements on the fields and the medium, and is a mechanism also employed in Refs. [59,61].

The paper is organized as follows. Section II briefly reviews the crucial aspects of 3-vector electrodynamics in dielectric media that must be generalized to the four-dimensional setting. Section III introduces the mathematical

notation that will be used throughout the rest of the paper. Section IV introduces the tensorial version of electrodynamics in media and the differential and integral forms of Maxwell's equations, while Sec. V derives the relationship between the four-dimensional tensors and the "usual" spatial, or transverse, material tensors. Section VI presents the geometric optics limit of the theory, while Sec. VII derives the pseudo-Finslerian optical metric in this limit, and Sec. VIII shows how the pseudo-Finslerian optical metric becomes pseudo-Riemannian in certain circumstances. Finally, Sec. IX describes how all of these derivations relate to something tangible like ray tracing. I conclude with Sec. X.

II. 3-VECTOR CLASSICAL ELECTRODYNAMICS IN MEDIA

Since the goal of this paper is to study the covariant formulation of electrodynamics in dielectric media, it is useful to first review the familiar aspects of 3-vector electrodynamics in media that we wish to treat. In the three-dimensional Cartesian coordinates of flat Minkowski space-time, Maxwell's equations take the form of Eqs. (1) where \vec{E} and \vec{D} are, respectively, the electric field and electric flux density, and \vec{B} and \vec{H} are, respectively, the magnetic flux density and the magnetic field, and where the speed of light has been scaled to $c = 1$. The electric and magnetic fields \vec{E} and \vec{B} may be derived from scalar and vector potentials ϕ and \vec{A} by

$$\vec{B} = \nabla \times \vec{A}, \quad \vec{E} = -\nabla\phi - \frac{\partial \vec{A}}{\partial t}. \quad (5)$$

However, as far as finding solutions is concerned, the set of Eqs. (1) and (5) is incomplete. A supplemental condition is required on the set of fields $\{\vec{E}, \vec{B}, \vec{D}, \vec{H}\}$, which is typically provided through the basic constitutive relations

$$\vec{D} = \varepsilon \vec{E}, \quad \vec{B} = \mu \vec{H}. \quad (6)$$

In vacuum, $\varepsilon = \varepsilon_0$ and $\mu = \mu_0$ are constants (and for $c = 1$ we have $\varepsilon_0 = \mu_0 = 1$), but inside dielectric media the permittivity ε and permeability μ may not only be inhomogeneous but may not even be isotropic, described instead by matrix-valued functions. The reduction of dielectric media to a set of supplemental parameters in Maxwell's equations belies an underlying quantum field theoretic, or at least a microscopic, description of the interaction of electromagnetic fields and matter [62,63], but at the macroscopic level it is sufficient to use an effective theory that accounts for the average atomic response to the applied fields. The basic idea of this model is that an electron in an atom is slightly displaced by an applied electric field. The electron feels a restoring force determined by the nucleus. The exact form of the restoring force may be modulated by the presence of other electrons and

neighboring dipole moments, but in the "Lorentz" approximation it is modeled as being of Hooke's law variety—linear in the displacement. Thus the applied electric field induces a dipole moment \vec{p} , the magnitude of which depends, to lowest order, linearly on both the applied field and the effective spring constant. Sufficiently far from the dipole moment the dipole field is proportional to \vec{p} , and the net electric field is

$$\vec{E}_{\text{net}} = \vec{E}_{\text{applied}} + \vec{p} = \vec{E}_{\text{applied}} + \chi_E \vec{E}_{\text{applied}} = (1 + \chi_E) \vec{E}_{\text{applied}}, \quad (7)$$

where the electric susceptibility χ_E is linear in the average effective spring constant for the material. This may readily be extended to anisotropic and inhomogeneous material responses by changing the scalar χ_E into a position dependent matrix $\bar{\chi}_E(\vec{x})$. The matrix nature of $\bar{\chi}_E$ means the restoring force felt by the electron depends on the direction of displacement, and the value of $\bar{\chi}_E(\vec{x})$ represents an averaging over a small region around the point \vec{x} .

In the aggregate theory the electric flux density is related to the electric field via a constitutive relation that is typically written as

$$\vec{D} = \vec{E} + \vec{P}, \quad (8)$$

where the polarization density \vec{P} is found through a suitable averaging procedure of the dipole moment density. So far we have only considered \vec{P} as linear in \vec{E} for the simple Hooke's law model. More generally, one could also consider anharmonic dipole behavior and higher order multipole moment contributions to \vec{D} rather than just linearized dipoles, and each multipole contribution to the total polarization can be expanded in terms of \vec{E} so that [64–66]

$$D^i = \chi^i_j E^j + \gamma^{ij}_k \nabla_j E^k + \zeta^i_{jk} E^j E^k + \dots \quad (9)$$

Linear media for which γ is identically zero are said to be *linearly anisotropic*. In linearly anisotropic media the eigenstates of propagation are states of linear polarization. Linear media for which $\chi = 0$ and $\gamma \neq 0$ are said to be *circularly anisotropic*. In circularly anisotropic media the eigenstates of propagation are states of circular polarization. In a general medium with $\chi \neq 0$ and $\gamma \neq 0$ the eigenstates of propagation are states of elliptical polarization.

Similar to the electric response, ponderable media may exhibit a magnetic response to an applied field, which in the aggregate is usually expressed through the corresponding relation

$$\vec{B} = \vec{H} + \vec{M}, \quad (10)$$

where \vec{M} is the magnetization vector.

III. SOME REMARKS ON MATHEMATICAL NOTATION

The next step is to generalize the Heavisidian 3-vector mathematics of classical electrodynamics to the tensorial quantities that have meaning in a four-dimensional space-time. As much as possible, I will endeavor to use an index-free notation to eliminate unnecessary clutter and to ensure that the methodology is truly coordinate independent. The purpose of this section is to clarify the notation that is used throughout the following sections, and in particular to help make the notation accessible to those who have experience with electrodynamics from the 3-vector perspective but limited experience with general relativity. There are many excellent resources providing more detailed explanation of the mathematical concepts being used, such as Refs. [67,68]. A concise introduction to the required topics of differential geometry used for electrodynamics, with a focus on transformation optics, is available in Ref. [69] or Ref. [14].

In 3-vector notation, \vec{v} is the coordinate-free expression for an object that, relative to a coordinate basis such as the Cartesian basis $\{\hat{i}, \hat{j}, \hat{k}\}$, has the alternative expression $v^x \hat{i} + v^y \hat{j} + v^z \hat{k}$. If the basis is understood, then we can get away with just referring to the coefficients $\{v^x, v^y, v^z\}$. Relabeling $x, y, z \rightarrow 1, 2, 3$, then we can refer to either the coordinate-independent \vec{v} or the coordinate coefficients v^a , $a = 1, 2, 3$. In what follows, the coordinate-free notation for four-dimensional tensorial objects will be bold, e.g. \boldsymbol{v} or \boldsymbol{F} , and indices that range over all four space-time coordinates will be given Greek letters, e.g. v^α or $F_{\mu\nu}$. The position and ordering of indices carry meaning and are similar to the difference between row and column vectors, but indices can be raised and lowered with the space-time metric. The metric has two subscript indices, $\boldsymbol{g} = g_{\alpha\beta}$, while its inverse has two superscript indices, $\boldsymbol{g}^{-1} = g^{\alpha\beta}$. The types of indices carried by an object will be given when it is introduced. In the index-free notation I will denote index raising and lowering in two different ways.

- (1) When a single index is raised or lowered, I will indicate it with the metric or its inverse, for example

$$\boldsymbol{g}^{-1} \cdot \boldsymbol{F} = g^{\alpha\beta} F_{\beta\mu}, \quad \boldsymbol{F} \cdot \boldsymbol{g}^{-1} = F_{\beta\mu} g^{\mu\alpha}, \quad \boldsymbol{\mu} \cdot \boldsymbol{g} = \mu_\mu^\alpha g_{\alpha\beta}. \quad (11)$$

- (2) When all indices are raised or lowered (e.g. for single-index objects), I will use the musical isomorphisms, for example

$$\boldsymbol{u}^b = \boldsymbol{g} \cdot \boldsymbol{u} = g_{\alpha\beta} u^\beta = u_\alpha, \quad \boldsymbol{k}^\sharp = \boldsymbol{g}^{-1} \cdot \boldsymbol{k} = g^{\alpha\beta} k_\beta = k^\alpha. \quad (12)$$

Note that I have used a dot notation to indicate contraction on adjacent indices, similar to common matrix multiplication, e.g.

$$\boldsymbol{F} \cdot \boldsymbol{u} = F_{\alpha\beta} u^\beta, \quad \boldsymbol{u} \cdot \boldsymbol{F} = u^\alpha F_{\alpha\beta}, \quad \boldsymbol{\mu} \cdot \boldsymbol{B} = \mu_\mu^\nu B_\nu. \quad (13)$$

For more complicated contractions, the index expression will be given for clarity. Mixed index objects, such as $\boldsymbol{\mu}$ in the previous example, will be interpreted as linear endomorphisms on a vector space; e.g. $\boldsymbol{\mu}$ takes an element \boldsymbol{B} of the cotangent space and maps it to a new element of the cotangent space \boldsymbol{H} .

Tensorial objects can be multiplied together in two ways such that the number of indices increases.

- (1) The tensor product \otimes creates an $(m+n)$ -index tensor from an m -index tensor and an n -index tensor, e.g.

$$\boldsymbol{u} \otimes \boldsymbol{u} = u^\alpha u^\beta, \quad \boldsymbol{u}^b \otimes \boldsymbol{k}^\sharp = u_\alpha k^\beta. \quad (14)$$

- (2) The wedge product \wedge is an antisymmetric version of the tensor product, with the caveat that the tensors share the same index type. I will extend the usage of the wedge product to mixed-index objects, where the wedge shall only apply in the common exterior algebra, i.e. to indices of the same type, e.g.

$$\boldsymbol{u}^b \wedge \boldsymbol{E} = \boldsymbol{u}^b \otimes \boldsymbol{E} - \boldsymbol{E} \otimes \boldsymbol{u}^b = u_\alpha E_\beta - u_\beta E_\alpha = \delta_{\alpha\beta}^{\mu\nu} u_\mu E_\nu, \quad (15)$$

$$\boldsymbol{u}^b \wedge \boldsymbol{\delta} = \delta_{\alpha\beta}^{\mu\nu} u_\mu, \quad \boldsymbol{\varepsilon} \wedge \boldsymbol{u} = \delta_{\mu\nu}^{\sigma\rho} \varepsilon_\sigma^\mu u^\nu. \quad (16)$$

Last, two tensors next to each other with no symbol between them will indicate contraction on all possible indices, as in

$$\star \boldsymbol{F} = \star_{\alpha\beta}{}^{\mu\nu} F_{\mu\nu}, \quad (\boldsymbol{\delta} \wedge \boldsymbol{u}) \boldsymbol{F} = \delta_{\alpha\beta}^{\mu\nu} u^\beta F_{\mu\nu}. \quad (17)$$

IV. TENSORIAL ELECTRODYNAMICS

Returning to Maxwell's Eqs. (1), the apparent symmetry of the equations, and the fact that \vec{E} and \vec{B} may be obtained from the coupled set of potentials in Eqs. (5), hints at a deeper connection between the various fields. This connection may be better illuminated from the vantage of a unified space-time and the well defined operations of differential geometry on manifolds. The scalar and vector potentials may be combined into the four-dimensional potential 1-form \boldsymbol{A} , which in flat space-time is simply related to the usual scalar and vector potentials

$$\boldsymbol{A} = A_\mu = (-\phi, \vec{A}). \quad (18)$$

It turns out that electrodynamics is particularly well suited to description by exterior calculus. Turning to flat space-time and adopting Cartesian coordinates, the exterior derivative of \boldsymbol{A} can be expressed in matrix form as

$$(\mathbf{dA})_{\mu\nu} = A_{\nu,\mu} - A_{\mu,\nu} = \begin{pmatrix} 0 & \frac{\partial\phi}{\partial x} + \frac{\partial A_x}{\partial t} & \frac{\partial\phi}{\partial y} + \frac{\partial A_y}{\partial t} & \frac{\partial\phi}{\partial z} + \frac{\partial A_z}{\partial t} \\ -\frac{\partial\phi}{\partial x} - \frac{\partial A_x}{\partial t} & 0 & \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} & \frac{\partial A_z}{\partial x} - \frac{\partial A_x}{\partial z} \\ -\frac{\partial\phi}{\partial y} - \frac{\partial A_y}{\partial t} & \frac{\partial A_x}{\partial y} - \frac{\partial A_y}{\partial x} & 0 & \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \\ -\frac{\partial\phi}{\partial z} - \frac{\partial A_z}{\partial t} & \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} & \frac{\partial A_y}{\partial z} - \frac{\partial A_z}{\partial y} & 0 \end{pmatrix}, \quad (19)$$

the elements of which clearly show the relations Eqs. (5) in component form. Making the identifications suggested by Eqs. (5), we find that in the unified space-time approach the components of \vec{E} and \vec{B} are combined into a 2-form

$$\mathbf{F} = \mathbf{dA} \quad (20)$$

called the *field strength* tensor. In the four-dimensional description, \vec{E} and \vec{B} are no longer distinct objects, but are instead merely selected components of the tensor \mathbf{F} . The exterior derivative increases the number of tensor indices by one, so the exterior derivative of the 1-form A_μ is the 2-form $F_{\mu\nu}$, the coefficients of which were written as a 4×4 matrix in Eq. (19). The space of 1-forms has coordinate basis $\{dx^\alpha\}$, the space of 2-forms has basis $\{dx^\alpha \wedge dx^\beta\}$, the space of 3-forms has basis $\{dx^\alpha \wedge dx^\beta \wedge dx^\gamma\}$, etc., so the matrix expression of a 2-form represents the coefficients of a linear combination of basis elements

$$\mathbf{F} = \frac{1}{2} F_{\mu\nu} (dx^\mu \wedge dx^\nu) = E_a (dx^a \wedge dt) + \frac{1}{2} B_{ab} (dx^a \wedge dx^b). \quad (21)$$

Just as the component functions used to describe a vector differ among observers according to their independent coordinate systems, so too do the component functions of \mathbf{F} differ among observers. Thus while one observer may only see nonzero values for \mathbf{F} in the components labeled by E_a , another observer using different (space-time) coordinates might see nonzero components that she labels by B_{ab} .

The field strength tensor \mathbf{F} encodes information about the electric field strength E_a and the magnetic flux B_{ab} . It is no coincidence that we identify the “field” aspect with the time-space components of \mathbf{F} and the “flux” aspect with the space-space components; the rules for integration on manifolds lead naturally to the integral forms of Faraday’s law and Gauss’ law for magnetism, where \vec{B} has a clear interpretation as a flux through a spatial surface, as discussed in Sec. IV B. In a similar manner we need to encode information about the fields \vec{H} and \vec{D} , so it should come as no surprise that we let the *excitation tensor*

$$\mathbf{G} = -H_a (dx^a \wedge dt) + \frac{1}{2} D_{ab} (dx^a \wedge dx^b) \quad (22)$$

encode information about the magnetic field strength H_a and electric flux D_{ab} .

Comparing Eq. (22) with Eq. (21) it seems clear that if we want to conserve index type, then the simple vacuum relationships $\vec{D} = \vec{E}$ and $\vec{H} = \vec{B}$ need to be replaced by something else when talking about E_a and D_{ab} or H_a and B_{ab} . The 3-vector relations of Eqs. (6) must therefore be relating the coefficients of 2-forms to those of 1-forms without the explicit use of an operation that is designed for such a task. In fact, in vacuum the dual nature of these fields is related by just such an operator, called the Hodge dual,

$$\mathbf{G} = \star\mathbf{F}, \quad (23)$$

as discussed in more detail in the next subsection.

A. Differential form of Maxwell’s equations

An important property of the exterior derivative is that it is nilpotent, so $d^2\omega = 0$ for any differential form ω . Since $\mathbf{F} = \mathbf{dA}$, it immediately follows that

$$\mathbf{dF} = 0. \quad (24)$$

This single equation is equivalent to the pair of homogeneous Maxwell’s equations, which can be seen by explicitly calculating the exterior derivative of \mathbf{dA} in Minkowski space-time [1].

The homogeneous equation does not contain any information about the space-time and can be thought of as a constraint equation for the fields. Information about the space-time appears in the inhomogeneous equations, where it enters through the Hodge dual, \star . The Hodge dual is defined for any pseudo-Riemannian manifold that possesses a metric and a volume form, and it contains all the space-time information of the metric tensor. In particular, the fundamental electromagnetic fields are differential forms (antisymmetric tensors). Let $\Lambda^k T_p M$ and $\Lambda^k T_p^* M$ denote, respectively, the spaces of alternating tensor k -products of tangent and cotangent spaces of manifold M at point p . An alternating k -vector field (k -blade) is an element of $\Lambda^k TM = \bigcup_{p \in M} \Lambda^k T_p M$, while an alternating k -covector field (k -form) is an element of $\Lambda^k T^* M = \bigcup_{p \in M} \Lambda^k T_p^* M$. The metric (and its inverse) provides a map between tangent and cotangent spaces by raising and lowering indices, and is thus extensible to a map $\mathbf{g}: \Lambda^k TM \rightarrow \Lambda^k T^* M$, while the covariant volume form

(and its inverse) provides a map $\omega: \Lambda^k TM \rightarrow \Lambda^{(m-k)} T^*M$. The composition of these two maps defines the Hodge dual as a map $\star: \Lambda^k TM \rightarrow \Lambda^{(m-k)} T^*M$ or $\star: \Lambda^k T^*M \rightarrow \Lambda^{(m-k)} T^*M$, such that

$$\begin{array}{ccc} \Lambda^{(m-k)} T_p^*(M) & \xleftarrow{g} & \Lambda^{(m-k)} T_p(M) \\ \star \uparrow & \omega & \downarrow \star \\ \Lambda^k T_p^*(M) & \xleftarrow{g} & \Lambda^k T_p(M) \end{array} \quad (25)$$

is (modulo some numerical factors) commutative. For example, the Hodge dual of the 2-form F has the component expression

$$\star F = \frac{1}{2} (\omega \circ g^{-1}) F = \frac{1}{2} \sqrt{-g} \epsilon_{\alpha\beta\sigma\rho} g^{\sigma\mu} g^{\rho\nu} F_{\mu\nu}, \quad (26)$$

where $g = \det(g)$ and $\epsilon_{\alpha\beta\sigma\rho}$ is the completely antisymmetric Levi-Civita symbol. An important property of \star is that

$$\star\star\omega = (-1)^\ell \omega, \quad (27)$$

where the power depends on the dimension of the manifold and the degree of ω . For a 2-form in a four-dimensional space-time $\ell = 1$, and this is the only case we will be concerned with.

Consider now the Yang-Mills action

$$S = \int \frac{1}{2} F \wedge \star F - A \wedge J, \quad (28)$$

where $J = \star j$ is the charge-current 3-form source. Varying the action with respect to A provides the inhomogeneous Maxwell equation

$$d\star F = J. \quad (29)$$

In the presence of ponderable media, an incident field F can induce multipole moments of J that can contribute to the source term even if the total monopole (i.e. free) charge contribution to J may vanish. In such a case the solutions to Maxwell's equations include the particular solution

$$d\star P = J_{\text{multipole}}. \quad (30)$$

Then the general solution satisfies

$$dF = 0, \quad (31)$$

$$d\star(F + P) = dG = J_{\text{free}}, \quad (32)$$

and we now have a generalization of Eqs. (8) and (10) in the form of

$$G = \star(F + P). \quad (33)$$

Generalizing Eq. (9), the polarization tensor P should, in principle, be expandable in terms of F and its derivatives. Of interest here is when the polarization is linear in F , whence the previous constitutive relation simplifies to

$$G = \star\chi F, \quad (34)$$

or in indices

$$G_{\mu\nu} = \star_{\mu\nu}{}^{\alpha\beta} \chi_{\alpha\beta}{}^{\sigma\rho} F_{\sigma\rho}. \quad (35)$$

Recall that the constitutive relation in vacuum is simply $G = \star F$. The vacuum may therefore be considered a trivial linear dielectric medium for which $\chi_{\text{vac}} = id$.

B. Integral form of Maxwell's equations

We have seen the differential form version of Maxwell's equations, but in the 3-vector approach to electrodynamics one frequently wants to work with the integral form of the equations. How do the integral equations follow from this four-dimensional approach? Begin with the homogeneous equation $dF = 0$. Since dF is a 3-form, it can only be integrated over a three-dimensional submanifold $\Omega \subset \mathcal{M}$,

$$\int_{\Omega} dF = \int_{\partial\Omega} F = 0, \quad (36)$$

where the intermediate step follows from the generalized Stokes theorem. The integral of the 2-form F is over $\partial\Omega$, the two-dimensional boundary of Ω .

1. Faraday's law

Consider a simple example in flat space-time and using Cartesian coordinates. In a curved space-time one would have to pull back the integrand to a submanifold of a chart. To proceed with the integration one must first choose Ω , which we suppose to be the three-dimensional surface consisting of the xy plane extending into t . In other words, the restriction of dF to Ω is $(dF)_{txy} dt \wedge dx \wedge dy$. At first glance one may suppose to write the boundary terms as $F_{tx} dt \wedge dx + F_{ty} dt \wedge dy + F_{xy} dx \wedge dy$, but special attention must be paid to the orientation of $\partial\Omega$, which must be compatible with the orientation of Ω , which in turn must be compatible with the orientation of \mathcal{M} (the orientation of which is given by the volume form $\star 1$). In particular, the prescription for integration whereby

$$\int_{\partial\Omega} F dx^1 \wedge dx^2 \rightarrow \int_{\partial\Omega} F dx^1 dx^2 \quad (37)$$

introduces a certain blindness to the orientation since

$$\int_{\partial\Omega} -F dx^2 \wedge dx^1 \rightarrow - \int_{\partial\Omega} F dx^1 dx^2. \quad (38)$$

By comparing with Eq. (21) we see that \mathbf{F} restricted to the boundary selects the terms

$$\mathbf{F}|_{\partial\Omega} = F_{xt}dx \wedge dt + F_{yt}dy \wedge dt + F_{xy}dx \wedge dy. \quad (39)$$

The integral becomes

$$\begin{aligned} \int_{\partial\Omega} \mathbf{F} &= \int F_{xt}dxdt + \int F_{yt}dydt + \int F_{xy}dxdy \\ \Rightarrow \int E_x dx + E_y dy &= -\frac{\partial}{\partial t} \int B_z dxdy, \end{aligned} \quad (40)$$

which straightforwardly generalizes to

$$\int \vec{E} \cdot d\vec{r} = -\frac{\partial}{\partial t} \int \vec{B} \cdot \hat{n} dA. \quad (41)$$

2. Gauss' law for magnetic fields

This time choose a spatial volume for Ω . In this case we clearly have

$$\mathbf{F}|_{\partial\Omega} = F_{xy}dx \wedge dy + F_{yz}dy \wedge dz + F_{zx}dz \wedge dx, \quad (42)$$

and it follows that

$$\int_{\partial\Omega} \vec{B} \cdot \hat{n} dA = 0. \quad (43)$$

A similar treatment of the inhomogeneous equation $d\mathbf{G} = \mathbf{J}$ provides Ampere's law and Gauss' law for electric fields.

3. Vector potential source equations

Integrating the homogeneous and inhomogeneous Maxwell equations over appropriate submanifolds gave us the usual integral forms of Maxwell's equations. But these are not the only well-known results from electrodynamics. Where, for example, does the Biot-Savart law come from? The Biot-Savart law calculates \mathbf{F} by differentiating the vector potential \mathbf{A} . For sourceless free fields, one usually considers \mathbf{A} as given and the calculation $\mathbf{F} = d\mathbf{A}$ is straightforward. But how does one determine \mathbf{F} when \mathbf{A} is dynamically sourced by some charge-current distribution \mathbf{J} ? As a starting point, one could try to determine \mathbf{A} by integrating both sides of Maxwell's inhomogeneous equation

$$\int_{\Omega} d\star d\mathbf{A} = \int_{\Omega} \mathbf{J}. \quad (44)$$

The next expected step might be

$$\int_{\partial\Omega} \star d\mathbf{A} = \int_{\Omega} \mathbf{J}, \quad (45)$$

but then we can go no further and it seems we have reached an impasse. Instead of a direct integration, \mathbf{A} must be obtained from Maxwell's inhomogeneous equation with the assistance of Green's technique. In the Minkowski vacuum with only free charges, this Green function approach leads to Jefimenko's equations [1].

C. Properties of the constitutive relation

The constitutive quantity χ may be thought of as a map from 2-forms to 2-forms, which means that it can have at most 36 independent components. Because of this, it is possible to reformulate everything in terms of 2 six-dimensional vectors (\vec{E}, \vec{B}) and (\vec{D}, \vec{H}) , and a 6×6 constitutive matrix, and this approach is frequently taken in the literature. I do not follow this approach for three reasons: (1) the extraction of \vec{E} and \vec{B} is usually done by hand in a somewhat *ad hoc* way that typically relies on a Minkowski space-time decomposition; (2) the vectors \vec{B} and \vec{D} obscure the 2-form nature of these fields, which becomes more apparent in the integral formulation of Maxwell's equations, where \vec{B} and \vec{D} appear in integrals over a 2-surface; and (3) since modern computer algebra systems can easily handle large matrix calculations, there is no computational advantage to using a six-dimensional formulation, while at the same time a preferred use of the four-dimensional differential forms formulation is consistent with the rest of general relativity, and hence allows the possibility to incorporate the electrodynamics developed here in the standard relativistic framework.

We can demand that χ is independently antisymmetric on its covariant and contravariant indices. This symmetry condition reduces the number of free parameters to exactly the required 36. There are additional symmetries with respect to which it may sometimes be useful to decompose χ . In particular, consider the decomposition

$$\chi = {}^P\chi + {}^A\chi + {}^S\chi, \quad (46)$$

where

$${}^S\chi_{\alpha\beta\mu\nu} = \frac{1}{2}(\chi_{\alpha\beta\mu\nu} - \chi_{\mu\nu\alpha\beta}), \quad (47a)$$

$${}^A\chi_{\alpha\beta}{}^{\mu\nu} = f(\mathbf{x}) \star_{\alpha\beta}{}^{\mu\nu}, \quad (47b)$$

$${}^P\chi_{\alpha\beta\mu\nu} = \frac{1}{2}(\chi_{\alpha\beta\mu\nu} + \chi_{\mu\nu\alpha\beta}) - {}^A\chi_{\alpha\beta\mu\nu}, \quad (47c)$$

which ensure the symmetry properties

$${}^P\chi_{\alpha\beta\mu\nu} = {}^P\chi_{\mu\nu\alpha\beta}, \quad {}^A\chi_{\alpha\beta\mu\nu} = {}^A\chi_{\mu\nu\alpha\beta}, \quad {}^S\chi_{\alpha\beta\mu\nu} = -{}^S\chi_{\mu\nu\alpha\beta}. \quad (48)$$

Note that the symmetries here are defined with respect to the metric by lowering the second set of indices on χ , and is

therefore not a fundamental decomposition of χ on its own. In the premetric literature these are referred to as the principal, axion, and skewon parts [37,70,71]. Premetric electrodynamics replaces the Hodge dual appearing in the action of Eq. (28) with something like χ , so one should note that their χ has a somewhat different, although related, meaning to the χ used here. In particular, one could possibly identify $(\mathbf{g} \otimes \mathbf{g})\chi_{\text{premetric}} = \sqrt{-g}\chi$, but it should also be noted that the goals of premetric electrodynamics are somewhat different, and for electrodynamics in real media one should, strictly speaking, vary the action of Eq. (28) and subsequently identify particular solutions that are linear in the homogeneous field, as done in the derivation above, rather than as an *ansatz* in the action. The symmetric principal part ${}^P\chi$ has 20 components and is responsible for the usual macroscopic dielectric response parameters; the axion part ${}^A\chi$ has only one independent component; and the antisymmetric skewon part ${}^S\chi$ has 15 components. Additional symmetry conditions can therefore be imposed on χ that eliminate the skewon or axion parts, based, for example, on thermodynamic or energy conservation arguments, or by the lack of an observed directive effect in naturally occurring stationary media [53,72]. Since here I am interested in allowing the most general medium possible, I do not pursue such a decomposition.

V. RELATION TO 3-VECTOR QUANTITIES

Since most people have a more intuitive feeling for the three-dimensional vectors and medium parameters, it would be nice to explicitly show how χ is related to the usual permeability, permittivity, and magnetoelectric couplings. In the process, we may gain some better understanding of electrodynamics in media. One of the remarkable things about electrodynamics is that the fields are spacelike, but since there is no global definition of time, splitting space-time into space and time depends on the choice of observers, from which it follows that the identification of electric and magnetic fields is observer dependent. Indeed, the electric and magnetic fields are only defined relative to some observer. When dealing with tensorial electrodynamics, the electric and magnetic field components are often extracted from \mathbf{F} and \mathbf{G} by hand in some (locally, at least) Minkowski space-time, but we may gain some insight by explicitly retaining the role of the observer. Since the covariance between observers holds in curved space-times, where they may not necessarily be related by simple Lorentz transformations, this approach retains full covariance of the theory.

A. Space-time splitting

Consider an observer moving with 4-velocity \mathbf{u} with squared magnitude $\mathbf{g}(\mathbf{u}, \mathbf{u}) = (\mathbf{u}^b \cdot \mathbf{u})$ (or, strictly speaking, let \mathbf{u} be a hypersurface orthogonal timelike vector field

representing a family of observers). Relative to this observer, define the electric and magnetic 1-form fields by

$$\mathbf{E} = (\mathbf{u}^b \cdot \mathbf{u})^{-1} \mathbf{u} \cdot \mathbf{F} \quad \text{and} \quad \mathbf{B} = -(\mathbf{u}^b \cdot \mathbf{u})^{-1} \mathbf{u} \cdot \star \mathbf{F}, \quad (49)$$

which have the index expressions

$$E_\beta = (\mathbf{u}^b \cdot \mathbf{u})^{-1} F_{\alpha\beta} u^\alpha \quad \text{and} \quad B_\beta = -(\mathbf{u}^b \cdot \mathbf{u})^{-1} \star_{\alpha\beta}{}^{\mu\nu} F_{\mu\nu} u^\alpha. \quad (50)$$

Notice that by the antisymmetry of \mathbf{F} , we can express \mathbf{E} and \mathbf{B} with the somewhat more symmetric expressions

$$\mathbf{E} = -\frac{1}{2} (\mathbf{u}^b \cdot \mathbf{u})^{-1} (\delta \wedge \mathbf{u}) \mathbf{F} \quad \text{and} \quad \mathbf{B} = \frac{1}{2} (\mathbf{u}^b \cdot \mathbf{u})^{-1} (\delta \wedge \mathbf{u}) \star \mathbf{F}. \quad (51)$$

Although these last expressions are somewhat tautological, they will be useful shortly. It is immediately clear by the antisymmetry properties of \mathbf{F} and \star that $\mathbf{u} \cdot \mathbf{E} = \mathbf{u} \cdot \mathbf{B} = 0$. Thus the 1-form fields \mathbf{E} and \mathbf{B} are orthogonal to \mathbf{u} . We can reconstruct \mathbf{F} by

$$\mathbf{F} = \mathbf{u}^b \wedge \mathbf{E} + \star(\mathbf{u}^b \wedge \mathbf{B}), \quad (52)$$

which has the index expression

$$F_{\mu\nu} = 2(u_{[\mu} E_{\nu]} + \star_{\mu\nu}{}^{\alpha\beta} u_{[\alpha} B_{\beta]}). \quad (53)$$

The dual fields have a similar structure. According to observer \mathbf{u} , they are

$$\mathbf{D} = -(\mathbf{u}^b \cdot \mathbf{u})^{-1} \mathbf{u} \cdot \star \mathbf{G} \quad \text{and} \quad \mathbf{H} = -(\mathbf{u}^b \cdot \mathbf{u})^{-1} \mathbf{u} \cdot \mathbf{G}, \quad (54)$$

while

$$\mathbf{G} = -\mathbf{u}^b \wedge \mathbf{H} + \star(\mathbf{u}^b \wedge \mathbf{D}). \quad (55)$$

Clearly, \mathbf{D} and \mathbf{H} are also orthogonal to \mathbf{u} , $\mathbf{u} \cdot \mathbf{D} = \mathbf{u} \cdot \mathbf{H} = 0$.

B. Obtaining usual parameters from χ

We have seen that within linear dielectric media, the constitutive relation is $\mathbf{G} = \star \chi \mathbf{F}$. We can use this to connect to the usual concepts of permittivity, permeability, and magnetoelectric couplings in the following way. From Eq. (54) one finds

$$\begin{aligned} \mathbf{D} &= -(\mathbf{u}^b \cdot \mathbf{u})^{-1} \mathbf{u} \cdot \star \mathbf{G} \\ &= -(\mathbf{u}^b \cdot \mathbf{u})^{-1} \mathbf{u} \cdot (\star \star \chi \mathbf{F}) = (\mathbf{u}^b \cdot \mathbf{u})^{-1} \mathbf{u} \cdot \chi \mathbf{F} \\ &= (\mathbf{u}^b \cdot \mathbf{u})^{-1} \mathbf{u} \cdot \chi (\mathbf{u}^b \wedge \mathbf{E} + \star(\mathbf{u}^b \wedge \mathbf{B})) \\ &= -(\mathbf{u}^b \cdot \mathbf{u})^{-1} [2(\mathbf{u} \cdot \chi \cdot \mathbf{u}^b) \cdot \mathbf{E} + 2(\mathbf{u} \cdot \chi \star \cdot \mathbf{u}^b) \cdot \mathbf{B}] \\ &= \boldsymbol{\epsilon}^c \cdot \mathbf{E} + {}^b \boldsymbol{\gamma}^c \cdot \mathbf{B}, \end{aligned} \quad (56)$$

where the fourth line follows from the antisymmetry of χ in its second set of indices. Similarly, for \mathbf{H} one finds

$$\begin{aligned} \mathbf{H} &= -(\mathbf{u}^b \cdot \mathbf{u})^{-1} \mathbf{u} \cdot \mathbf{G} \\ &= -(\mathbf{u}^b \cdot \mathbf{u})^{-1} \mathbf{u} \cdot \star \chi \mathbf{F} \\ &= -(\mathbf{u}^b \cdot \mathbf{u})^{-1} \mathbf{u} \cdot \star \chi (\mathbf{u}^b \wedge \mathbf{E} + \star (\mathbf{u}^b \wedge \mathbf{B})) \\ &= (\mathbf{u}^b \cdot \mathbf{u})^{-1} [2(\mathbf{u} \cdot \star \chi \cdot \mathbf{u}^b) \cdot \mathbf{E} + 2(\mathbf{u} \cdot \star \chi \star \cdot \mathbf{u}^b) \cdot \mathbf{B}] \\ &= {}^e \gamma^c \cdot \mathbf{E} + \boldsymbol{\xi} \cdot \mathbf{B}. \end{aligned} \quad (57)$$

To summarize, we now have the parameters

$$\begin{aligned} \boldsymbol{\varepsilon}^c &= -2(\mathbf{u}^b \cdot \mathbf{u})^{-1} (\mathbf{u} \cdot \chi \cdot \mathbf{u}^b), \\ ({}^e \boldsymbol{\varepsilon}^c)_{\beta}^{\mu} &= -2(u^\lambda u_\lambda)^{-1} u^\alpha \chi_{\alpha\beta}{}^{\mu\nu} u_\nu, \end{aligned} \quad (58a)$$

$$\begin{aligned} \boldsymbol{\xi} &= 2(\mathbf{u}^b \cdot \mathbf{u})^{-1} (\mathbf{u} \cdot \star \chi \star \cdot \mathbf{u}^b), \\ \xi_{\beta}^{\mu} &= 2(u^\lambda u_\lambda)^{-1} u^\alpha \star_{\alpha\beta}{}^{\lambda\kappa} \chi_{\lambda\kappa}{}^{\sigma\rho} \star_{\sigma\rho}{}^{\mu\nu} u_\nu, \end{aligned} \quad (58b)$$

$$\begin{aligned} {}^b \boldsymbol{\gamma}^c &= -2(\mathbf{u}^b \cdot \mathbf{u})^{-1} (\mathbf{u} \cdot \chi \star \cdot \mathbf{u}^b), \\ ({}^b \boldsymbol{\gamma}^c)_{\beta}^{\mu} &= -2(u^\lambda u_\lambda)^{-1} u^\alpha \chi_{\alpha\beta}{}^{\lambda\kappa} \star_{\lambda\kappa}{}^{\mu\nu} u_\nu, \end{aligned} \quad (58c)$$

$$\begin{aligned} {}^e \boldsymbol{\gamma}^c &= 2(\mathbf{u}^b \cdot \mathbf{u})^{-1} (\mathbf{u} \cdot \star \chi \cdot \mathbf{u}^b), \\ ({}^e \boldsymbol{\gamma}^c)_{\beta}^{\mu} &= 2(u^\lambda u_\lambda)^{-1} u^\alpha \star_{\alpha\beta}{}^{\lambda\kappa} \chi_{\lambda\kappa}{}^{\mu\nu} u_\nu. \end{aligned} \quad (58d)$$

Note that Eqs. (58) are space-time expressions, so each of these is a 4×4 tensor rather than a 3×3 matrix. However, since $\mathbf{u} \cdot \mathbf{D} = 0$, one can readily see that we must have

$$\boldsymbol{\varepsilon}^c \cdot \mathbf{u}^b = 0, \quad u^\alpha (\boldsymbol{\varepsilon}^c)_{\alpha}^{\beta} = 0 \quad (59)$$

and similar for the other material tensors. Since \mathbf{u} is orthogonal to \mathbf{E} and \mathbf{B} in the domain of each material tensor, it follows that each of $\boldsymbol{\varepsilon}^c$, $\boldsymbol{\xi}$, ${}^b \boldsymbol{\gamma}^c$, and ${}^e \boldsymbol{\gamma}^c$ are linear automorphisms of the three-dimensional subspace of the cotangent bundle $T^*(M)$ orthogonal to \mathbf{u} ; thus it is also true that

$$\boldsymbol{\varepsilon}^c \cdot \mathbf{u}^b = 0, \quad ({}^e \boldsymbol{\varepsilon}^c)_{\alpha}^{\beta} u_{\beta} = 0 \quad (60)$$

and similar for $\boldsymbol{\xi}$, ${}^b \boldsymbol{\gamma}^c$, and ${}^e \boldsymbol{\gamma}^c$. For example, in a local Cartesian frame comoving with the observer, such that $u^\alpha = (1, 0, 0, 0)$, one would find

$$({}^e \boldsymbol{\varepsilon}^c)_{\alpha}^{\mu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \varepsilon_x^{cx} & \varepsilon_x^{cy} & \varepsilon_x^{cz} \\ 0 & \varepsilon_y^{cx} & \varepsilon_y^{cy} & \varepsilon_y^{cz} \\ 0 & \varepsilon_z^{cx} & \varepsilon_z^{cy} & \varepsilon_z^{cz} \end{pmatrix}. \quad (61)$$

In this comoving picture (equivalently the observer is ‘‘at rest’’), the 3-vector *matrix* relations are recovered simply by

taking the purely spatial part of the *tensor* relations in a local Cartesian frame.

Since the timelike vector field \mathbf{u} defines a spacelike foliation of the manifold, and electrodynamics takes place in the purely spatial three-dimensional subspace orthogonal to \mathbf{u} , then it is useful to define an operator

$$\mathbf{h} = \boldsymbol{\delta} - \frac{\mathbf{u}^b \otimes \mathbf{u}}{(\mathbf{u}^b \cdot \mathbf{u})}, \quad h_{\alpha}^{\beta} = \delta_{\alpha}^{\beta} - \frac{u_{\alpha} u^{\beta}}{u_{\mu} u^{\mu}} \quad (62)$$

that projects out the purely spatial part of a vector or 1-form, relative to \mathbf{u} . In other words, \mathbf{h} annihilates any component proportional to \mathbf{u} . Note that $\mathbf{h} \cdot \mathbf{h} = \mathbf{h}$ and that the Kronecker delta projects to $\mathbf{h} \cdot \boldsymbol{\delta} \cdot \mathbf{h} = \mathbf{h}$, so \mathbf{h} serves as the Kronecker delta on the subspace orthogonal to \mathbf{u} . Consequently, in this four-dimensional notation, the vacuum permeability and permittivity are $\boldsymbol{\mu} = \boldsymbol{\varepsilon} = \mathbf{h}$. By Eqs. (59) and (60) we can see that since $\boldsymbol{\varepsilon}^c$, $\boldsymbol{\xi}$, ${}^b \boldsymbol{\gamma}^c$, and ${}^e \boldsymbol{\gamma}^c$ are already orthogonal to \mathbf{u} , then $\mathbf{h} \cdot \boldsymbol{\varepsilon}^c \cdot \mathbf{h} = \boldsymbol{\varepsilon}^c$, etc.

C. Reconstruction of χ from transverse parameters

Now it is possible to reconstruct χ out of the transverse parameters. Acting on both sides of the constitutive relation $\mathbf{G} = \star \chi \mathbf{F}$ with \star , one finds

$$\begin{aligned} \chi \mathbf{F} &= -\star \mathbf{G} \\ &= \star (\mathbf{u}^b \wedge \mathbf{H}) + \mathbf{u}^b \wedge \mathbf{D} \\ &= \star (\mathbf{u}^b \wedge ({}^e \boldsymbol{\gamma}^c \cdot \mathbf{E} + \boldsymbol{\xi} \cdot \mathbf{B})) + \mathbf{u}^b \wedge (\boldsymbol{\varepsilon}^c \cdot \mathbf{E} + {}^b \boldsymbol{\gamma}^c \cdot \mathbf{B}) \\ &= \star (\mathbf{u}^b \wedge {}^e \boldsymbol{\gamma}^c) \cdot \mathbf{E} + \star (\mathbf{u}^b \wedge \boldsymbol{\xi}) \cdot \mathbf{B} \\ &\quad + (\mathbf{u}^b \wedge \boldsymbol{\varepsilon}^c) \cdot \mathbf{E} + (\mathbf{u}^b \wedge {}^b \boldsymbol{\gamma}^c) \cdot \mathbf{B}. \end{aligned} \quad (63)$$

Using Eqs. (51) for \mathbf{E} and \mathbf{B} , this becomes

$$\begin{aligned} \chi \mathbf{F} &= \frac{1}{2} (\mathbf{u}^b \cdot \mathbf{u})^{-1} [-\star (\mathbf{u}^b \wedge {}^e \boldsymbol{\gamma}^c \wedge \mathbf{u}) \\ &\quad + \star (\mathbf{u}^b \wedge \boldsymbol{\xi} \wedge \mathbf{u}) \star - (\mathbf{u}^b \wedge \boldsymbol{\varepsilon}^c \wedge \mathbf{u}) \\ &\quad + (\mathbf{u}^b \wedge {}^b \boldsymbol{\gamma}^c \wedge \mathbf{u}) \star] \mathbf{F}. \end{aligned} \quad (64)$$

Rearranging terms a little, the sought after identity is clearly

$$\begin{aligned} \chi &= \frac{1}{2} (\mathbf{u}^b \cdot \mathbf{u})^{-1} [-(\mathbf{u}^b \wedge \boldsymbol{\varepsilon}^c \wedge \mathbf{u}) + \star (\mathbf{u}^b \wedge \boldsymbol{\xi} \wedge \mathbf{u}) \star \\ &\quad - \star (\mathbf{u}^b \wedge {}^e \boldsymbol{\gamma}^c \wedge \mathbf{u}) + (\mathbf{u}^b \wedge {}^b \boldsymbol{\gamma}^c \wedge \mathbf{u}) \star]. \end{aligned} \quad (65)$$

The corresponding index expression is

$$\begin{aligned} \chi_{\gamma\delta}{}^{\sigma\rho} &= 2(u^\lambda u_\lambda)^{-1} [-u_{[\gamma} (\boldsymbol{\varepsilon}^c)_{\delta]}{}^{[\sigma} u^{\rho]} \\ &\quad + \star_{\gamma\delta}{}^{\alpha\beta} u_{[\alpha} (\boldsymbol{\mu}^{-1})_{\beta]}{}^{[\mu} u^{\nu]} \star_{\mu\nu}{}^{\sigma\rho} - \star_{\gamma\delta}{}^{\alpha\beta} u_{[\alpha} ({}^e \boldsymbol{\gamma}^c)_{\beta]}{}^{[\sigma} u^{\rho]} \\ &\quad + u_{[\gamma} ({}^b \boldsymbol{\gamma}^c)_{\delta]}{}^{[\mu} u^{\nu]} \star_{\mu\nu}{}^{\sigma\rho}]. \end{aligned} \quad (66)$$

So far, I have related χ to the set of transverse components $\{\epsilon^c, \xi, e\gamma^c, b\gamma^c\}$, corresponding to the pair of constitutive relations

$$\begin{aligned} \mathbf{D} &= \boldsymbol{\epsilon}^c \cdot \mathbf{E} + b\gamma^c \cdot \mathbf{B}, \\ \mathbf{H} &= \boldsymbol{\xi} \cdot \mathbf{B} + e\gamma^c \cdot \mathbf{E} \end{aligned} \quad (67)$$

that relate the set of fields $\{\mathbf{D}, \mathbf{H}\}$ to the set $\{\mathbf{E}, \mathbf{B}\}$. Although these relations follow most naturally from the four-dimensional constitutive relation $\mathbf{G} = \star\chi\mathbf{F}$, they are not the only way of expressing the transverse constitutive relations. In traditional electrodynamics, the constitutive relations usually relate the set of fields $\{\mathbf{D}, \mathbf{B}\}$ to the set $\{\mathbf{E}, \mathbf{H}\}$ via

$$\begin{aligned} \mathbf{D} &= \boldsymbol{\epsilon} \cdot \mathbf{E} + h\boldsymbol{\gamma} \cdot \mathbf{H}, \\ \mathbf{B} &= \boldsymbol{\mu} \cdot \mathbf{H} + e\boldsymbol{\gamma} \cdot \mathbf{E}. \end{aligned} \quad (68)$$

It is easy to see that these two representations are related by

$$\boldsymbol{\xi} = \bar{\boldsymbol{\mu}}, \quad \boldsymbol{\epsilon}^c = \boldsymbol{\epsilon} - h\boldsymbol{\gamma} \cdot \boldsymbol{\xi} \cdot e\boldsymbol{\gamma}, \quad e\boldsymbol{\gamma}^c = -\boldsymbol{\xi} \cdot e\boldsymbol{\gamma}, \quad b\boldsymbol{\gamma}^c = h\boldsymbol{\gamma} \cdot \boldsymbol{\xi}. \quad (69)$$

Of course, the matrix displayed in Eq. (61) clearly does not have an inverse, and neither does $\boldsymbol{\xi}$. However, the bar notation on $\bar{\boldsymbol{\mu}}$ denotes an inverse restricted to the three-dimensional transverse subspace. Since any of the transverse medium parameters, such as $\boldsymbol{\mu}$, are automorphisms of the transverse subspace, then an inverse is well defined by the requirement that

$$\mathbf{b} \cdot \bar{\mathbf{b}} = \bar{\mathbf{b}} \cdot \mathbf{b} = \mathbf{h} \quad (70)$$

for any transverse map \mathbf{b} . The restricted inverse is found from

$$\bar{b}_\alpha{}^\beta = \frac{3h_{\alpha\alpha_1\alpha_2}^{\beta\beta_1\beta_2}(b_{\beta_1}{}^{\alpha_1})(b_{\beta_2}{}^{\alpha_2})}{h_{\alpha_1\alpha_2\alpha_3}^{\beta_1\beta_2\beta_3}(b_{\beta_1}{}^{\alpha_1})(b_{\beta_2}{}^{\alpha_2})(b_{\beta_3}{}^{\alpha_3})}, \quad (71)$$

where

$$h_{\alpha\alpha_1\alpha_2}^{\beta\beta_1\beta_2} = (\mathbf{u}^b \cdot \mathbf{u})^{-1} u^\mu u_\nu \delta_{\mu\alpha\alpha_1\alpha_2}^{\nu\beta\beta_1\beta_2} \quad (72)$$

is the generalized Kronecker symbol on the transverse subspace.

The traditional representation of Eqs. (68) allows Eq. (65) for χ to be recast as

$$\begin{aligned} \chi &= \frac{1}{2}(\mathbf{u}^b \cdot \mathbf{u})^{-1} [-(\mathbf{u}^b \wedge \boldsymbol{\epsilon} \wedge \mathbf{u}) + \star(\mathbf{u}^b \wedge \bar{\boldsymbol{\mu}} \wedge \mathbf{u})\star \\ &\quad + \star(\mathbf{u}^b \wedge \bar{\boldsymbol{\mu}} \cdot e\boldsymbol{\gamma} \wedge \mathbf{u}) + (\mathbf{u}^b \wedge h\boldsymbol{\gamma} \cdot \bar{\boldsymbol{\mu}} \wedge \mathbf{u})\star \\ &\quad + (\mathbf{u}^b \wedge h\boldsymbol{\gamma} \cdot \bar{\boldsymbol{\mu}} \cdot e\boldsymbol{\gamma} \wedge \mathbf{u})]. \end{aligned} \quad (73)$$

This expression can actually be factored to

$$\begin{aligned} \chi &= \frac{1}{2}(\mathbf{u}^b \cdot \mathbf{u})^{-1} [-(\mathbf{u}^b \wedge \boldsymbol{\epsilon} \wedge \mathbf{u}) + \star(\mathbf{u}^b \wedge \mathbf{h}) + \mathbf{u}^b \wedge h\boldsymbol{\gamma}] \\ &\quad \cdot \bar{\boldsymbol{\mu}} \cdot [(\mathbf{h} \wedge \mathbf{u})\star + e\boldsymbol{\gamma} \wedge \mathbf{u}], \end{aligned} \quad (74)$$

which will be quite useful for obtaining the optical metric in the next sections.

VI. GEOMETRIC OPTICS LIMIT

The geometric optics limit of the wave equation for light is standard lore [73], and the presentation here essentially follows the standard method, albeit in an index-free tensorial notation. The wave propagation of light is described by a second order equation, but Maxwell's equations provide two first order equations. The simplest derivation of geometric optics is to first use $\mathbf{F} = d\mathbf{A}$ to write the inhomogeneous Maxwell equation as

$$d\star d\mathbf{A} = \mathbf{J} \quad (75)$$

and show that this is a wave equation for the 1-form potential \mathbf{A} . Operating on both sides of this equation with the Hodge dual, it may be rewritten in terms of the codifferential of a k -form on an m -dimensional space-time manifold

$$\delta = (-1)^{k(m+1)-1} \star d \star \quad (76)$$

as

$$\delta d\mathbf{A} = -\star \mathbf{J}. \quad (77)$$

While the exterior derivative increases the degree of a differential form by one, e.g. d of a 1-form results in a 2-form, the codifferential decreases the degree by one; so while $d\mathbf{A}$ is a 2-form, $\delta d\mathbf{A}$ is a 1-form.

Propagating fields must satisfy a hyperbolic partial differential equation, e.g. a wave equation. On curved manifolds, the hyperbolic operator on differential forms is the Laplace–de Rham operator [74]

$$\Delta = \delta d + d\delta. \quad (78)$$

Adding $d\delta\mathbf{A}$ to both sides of Eq. (77) we have the gauge-independent wave equation

$$\Delta\mathbf{A} = (\delta d + d\delta)\mathbf{A} = -\star \mathbf{J} + d\delta\mathbf{A}. \quad (79)$$

On the other hand, fixing the gauge to the Lorenz gauge

$$\delta\mathbf{A} = 0 \quad (80)$$

shows that Maxwell's inhomogeneous equation serves as a wave equation for \mathbf{A} in the Lorenz gauge, even in media with bound source contributions. Separating out the dipole contribution of \mathbf{J} , the form of Maxwell's equation

$$\star d\star\chi d\mathbf{A} = \delta\chi d\mathbf{A} = 0 \quad (81)$$

does indeed serve as the wave equation for \mathbf{A} in neutral, linear macroscopic media.

Since we are now dealing with a wave equation for \mathbf{A} , we can use a Jeffreys-Wentzel-Kramers-Brillouin (JWKB) type of approximation by assuming a plane wave solution of the form

$$\mathbf{A} = \hat{\mathbf{A}}(x^\mu) e^{-(i\lambda)^{-1}S(x^\mu)} \quad (82)$$

and retaining the leading order terms in the limit $\lambda \rightarrow 0$ [73]. In this scheme the amplitude $\hat{\mathbf{A}}$ is slowly varying compared to the phase function S [53]. Plugging the JWKB solution into the Lorenz gauge condition, one finds

$$\delta\mathbf{A} = \delta(\hat{\mathbf{A}} e^{-(i\lambda)^{-1}S}) = (i\lambda)^{-1} e^{-(i\lambda)^{-1}S} ((i\lambda)\delta\hat{\mathbf{A}} - \mathbf{g}(dS, \hat{\mathbf{A}})) = 0. \quad (83)$$

Keeping the leading order term in the limit $\lambda \rightarrow 0$ leaves

$$\mathbf{g}(\hat{\mathbf{A}}, dS) = 0. \quad (84)$$

The 1-form $\mathbf{k} = dS$, or $k_\nu = \partial_\nu S$, is the wave covector, and Eq. (84) is the usual result that the Lorenz condition requires $\hat{\mathbf{A}}$ to be orthogonal to the wave vector. Next, calculating

$$\begin{aligned} \delta\chi d\mathbf{A} &= (i\lambda)^{-1} \delta[e^{-(i\lambda)^{-1}S} \chi [(i\lambda)d\hat{\mathbf{A}} + \hat{\mathbf{A}} \wedge dS]] \\ &= (i\lambda)^{-2} e^{-(i\lambda)^{-1}S} [-\mathbf{g}(dS, \chi [(i\lambda)d\hat{\mathbf{A}} + \hat{\mathbf{A}} \wedge dS]) \\ &\quad + (i\lambda)\delta\chi [(i\lambda)d\hat{\mathbf{A}} + \hat{\mathbf{A}} \wedge dS]] = 0 \end{aligned} \quad (85)$$

and taking the limit $\lambda \rightarrow 0$, one finds

$$-\mathbf{g}(dS, \chi(\hat{\mathbf{A}} \wedge dS)) = 0. \quad (86)$$

In terms of \mathbf{k} , and exploiting the antisymmetry properties of χ , this may be written as

$$-(\mathbf{k}^\sharp \cdot \chi \cdot \mathbf{k}) \cdot \hat{\mathbf{A}} = 0, \quad (87)$$

or with index notation as

$$-g^{\alpha\sigma} \chi_{\alpha\beta}{}^{\mu\nu} k_\sigma k_\nu \hat{A}_\mu = 0. \quad (88)$$

Thinking of

$$\mathbf{X} = -\mathbf{k}^\sharp \cdot \chi \cdot \mathbf{k} \quad (89)$$

as a 4×4 matrix, it may be observed that the existence of a nontrivial solution to Eq. (87) requires

$$\det(\mathbf{X}) = 0. \quad (90)$$

In fact, this condition is satisfied identically. By the antisymmetry of the second set of indices on χ , $\hat{\mathbf{A}} \propto \mathbf{k}$ is already a trivial solution, so any nontrivial solution resides in the three-dimensional subspace orthogonal to \mathbf{k} , meaning the matrix is effectively only three dimensional. There are some different methods for dealing with this (see, for example, Ref. [36]); I follow a purely algebraic argument based on the classical adjugate matrix $\text{adj}(\mathbf{X})$ (used similarly by Refs. [46,60,61]). The adjugate is defined such that

$$\mathbf{X} \text{adj}(\mathbf{X}) = \det(\mathbf{X}) \mathbf{I} \quad (91)$$

and is closely related to the inverse; for if \mathbf{X} is invertible, then $\text{adj}(\mathbf{X}) \propto \mathbf{X}^{-1}$, but $\text{adj}(\mathbf{X})$ is defined even if \mathbf{X}^{-1} does not exist. Since $\det(\mathbf{X}) = 0$ identically, then it must be true that $\mathbf{X} \text{adj}(\mathbf{X}) = 0$. Since \mathbf{X} is nonzero and arbitrary, the subsidiary condition

$$\text{adj}(\mathbf{X}) = 0 \quad (92)$$

must be satisfied. Although this is a matrix condition, I show below that

$$\text{adj}(\mathbf{X}) = P(\mathbf{k} \otimes \mathbf{k}^\sharp), \quad (93)$$

where P is a scalar polynomial of fourth order in \mathbf{k} . Therefore, the condition for nontrivial solutions to Eq. (87) reduces to the scalar condition

$$P = 0. \quad (94)$$

To see this, begin by inserting some extra Kronecker \mathbf{h} into the expression for χ given in Eq. (74) to write

$$\begin{aligned} \chi &= \frac{1}{2} (\mathbf{u}^\flat \cdot \mathbf{u})^{-1} [-(\mathbf{u}^\flat \wedge \mathbf{h}) \cdot \boldsymbol{\varepsilon} \cdot (\mathbf{h} \wedge \mathbf{u}) \\ &\quad + [\star(\mathbf{u}^\flat \wedge \mathbf{h}) + (\mathbf{u}^\flat \wedge \mathbf{h}) \cdot {}^h\boldsymbol{\gamma}] \cdot \bar{\boldsymbol{\mu}} \cdot [(\mathbf{h} \wedge \mathbf{u}) \star \\ &\quad + {}^e\boldsymbol{\gamma} \cdot (\mathbf{h} \wedge \mathbf{u})]]. \end{aligned} \quad (95)$$

Next, with this expression for χ , compute

$$\begin{aligned} \mathbf{X} &= -\mathbf{k}^\sharp \cdot \chi \cdot \mathbf{k} \\ &= -\frac{1}{2} (\mathbf{u}^\flat \cdot \mathbf{u})^{-1} [-\mathbf{K} \cdot \boldsymbol{\varepsilon} \cdot \mathbf{L} + [\mathbf{M} + \mathbf{K} \cdot {}^h\boldsymbol{\gamma}] \cdot \bar{\boldsymbol{\mu}} \cdot [\mathbf{N} + {}^e\boldsymbol{\gamma} \cdot \mathbf{L}], \end{aligned} \quad (96)$$

where

$$\mathbf{K} = \mathbf{k}^\sharp \cdot (\mathbf{u}^b \wedge \mathbf{h}), \quad K_\delta^\beta = k^\gamma u_\alpha h_\kappa^\beta \delta_{\gamma\delta}^{\alpha\kappa}, \quad (97)$$

$$\mathbf{L} = (\mathbf{h} \wedge \mathbf{u}) \cdot \mathbf{k}, \quad L_{\mu\sigma} = \delta_{\kappa\nu}^{\sigma\rho} h_\mu^\kappa u^\nu k_\rho, \quad (98)$$

$$\mathbf{M} = \mathbf{k}^\sharp \cdot \star(\mathbf{u}^b \wedge \mathbf{h}), \quad M_\delta^\beta = -2k^\gamma \star_{\gamma\delta}^{\beta\alpha} u_\alpha, \quad (99)$$

$$\mathbf{N} = (\mathbf{h} \wedge \mathbf{u}) \star \cdot \mathbf{k}, \quad N_\mu^\sigma = -2u^\nu \star_{\nu\mu}^{\sigma\rho} k_\rho. \quad (100)$$

Note that these tensors only contain information about the observer \mathbf{u} , the wave covector \mathbf{k} , and the space-time metric (both in the \star and in the musical isomorphisms that raise and lower indices on \mathbf{u} and \mathbf{k}). Also note that the tensors \mathbf{M} and \mathbf{N} are eigentensors of \mathbf{K} and \mathbf{L} . In particular, one may verify that

$$\mathbf{K} \cdot \mathbf{M} = (\mathbf{k} \cdot \mathbf{u})\mathbf{M} \quad \text{and} \quad \mathbf{N} \cdot \mathbf{L} = (\mathbf{k} \cdot \mathbf{u})\mathbf{N}, \quad (101)$$

which allows the revision of Eq. (96) to

$$\begin{aligned} \mathbf{X} = & -\frac{1}{2}(\mathbf{u}^b \cdot \mathbf{u})^{-1} \mathbf{K} \cdot [-\boldsymbol{\varepsilon} + [(\mathbf{k} \cdot \mathbf{u})^{-1} \mathbf{M} + {}^h\boldsymbol{\gamma}]] \\ & \cdot \bar{\boldsymbol{\mu}} \cdot [(\mathbf{k} \cdot \mathbf{u})^{-1} \mathbf{N} + {}^e\boldsymbol{\gamma}]] \cdot \mathbf{L}. \end{aligned} \quad (102)$$

By the properties of the adjugate, we have

$$\begin{aligned} \text{adj}(\mathbf{X}) = & -\frac{1}{8}(\mathbf{u}^b \cdot \mathbf{u})^{-3} \text{adj}(\mathbf{L}) \cdot \text{adj}(-\boldsymbol{\varepsilon} + [(\mathbf{k} \cdot \mathbf{u})^{-1} \mathbf{M} + {}^h\boldsymbol{\gamma}]] \\ & \cdot \bar{\boldsymbol{\mu}} \cdot [(\mathbf{k} \cdot \mathbf{u})^{-1} \mathbf{N} + {}^e\boldsymbol{\gamma}]] \cdot \text{adj}(\mathbf{K}). \end{aligned} \quad (103)$$

One may readily verify that

$$\begin{aligned} \text{adj}(\mathbf{L}) &= (\mathbf{k} \cdot \mathbf{u})^2 \mathbf{k} \otimes \mathbf{u} \quad \text{and} \\ \text{adj}(\mathbf{K}) &= (\mathbf{k} \cdot \mathbf{u})^2 \mathbf{u}^b \otimes \mathbf{k}^\sharp, \end{aligned} \quad (104)$$

so that

$$\begin{aligned} \text{adj}(\mathbf{X}) = & -\frac{1}{8}(\mathbf{u}^b \cdot \mathbf{u})^{-3} (\mathbf{k} \cdot \mathbf{u})^4 (\mathbf{u} \cdot \text{adj}(-\boldsymbol{\varepsilon} + [(\mathbf{k} \cdot \mathbf{u})^{-1} \mathbf{M} + {}^h\boldsymbol{\gamma}]] \\ & \cdot \bar{\boldsymbol{\mu}} \cdot [(\mathbf{k} \cdot \mathbf{u})^{-1} \mathbf{N} + {}^e\boldsymbol{\gamma}]] \cdot \mathbf{u}^b) (\mathbf{k} \otimes \mathbf{k}^\sharp). \end{aligned} \quad (105)$$

Since each tensor in $\text{adj}(\dots)$ above is transverse, then the quantity $\mathbf{u} \cdot \text{adj}(\dots) \cdot \mathbf{u}^b$ is equivalent to the determinant restricted to the subspace orthogonal to \mathbf{u} . This enables the application of Sylvester's determinant theorem to bring out $\boldsymbol{\varepsilon}$,

$$\text{adj}(\mathbf{X}) = -\frac{1}{8}(\mathbf{u}^b \cdot \mathbf{u})^{-3} (\mathbf{k} \cdot \mathbf{u})^4 (\mathbf{u} \cdot \text{adj}(\boldsymbol{\varepsilon}) \cdot \text{adj}(\mathbf{Q}) \cdot \mathbf{u}^b) (\mathbf{k} \otimes \mathbf{k}^\sharp), \quad (106)$$

with

$$\mathbf{Q} = -\mathbf{h} + \bar{\boldsymbol{\mu}} \cdot [(\mathbf{k} \cdot \mathbf{u})^{-1} \mathbf{N} + {}^e\boldsymbol{\gamma}] \cdot \bar{\boldsymbol{\varepsilon}} \cdot [(\mathbf{k} \cdot \mathbf{u})^{-1} \mathbf{M} + {}^h\boldsymbol{\gamma}] \quad (107)$$

and where $\bar{\boldsymbol{\varepsilon}}$ is the restricted inverse of $\boldsymbol{\varepsilon}$, such that $\bar{\boldsymbol{\varepsilon}} \cdot \boldsymbol{\varepsilon} = \boldsymbol{\varepsilon} \cdot \bar{\boldsymbol{\varepsilon}} = \mathbf{h}$. Note that since $\boldsymbol{\varepsilon}$ is transverse, then $\boldsymbol{\varepsilon} = \mathbf{h} \cdot \boldsymbol{\varepsilon} \cdot \mathbf{h}$, which means that

$$\text{adj}(\boldsymbol{\varepsilon}) = \text{adj}(\mathbf{h} \cdot \boldsymbol{\varepsilon} \cdot \mathbf{h}) = \text{adj}(\mathbf{h}) \cdot \text{adj}(\boldsymbol{\varepsilon}) \cdot \text{adj}(\mathbf{h}). \quad (108)$$

But since

$$\text{adj}(\mathbf{h}) = (\mathbf{u}^b \cdot \mathbf{u})^{-1} \mathbf{u}^b \otimes \mathbf{u}, \quad (109)$$

then

$$\text{adj}(\boldsymbol{\varepsilon}) = (\mathbf{u}^b \cdot \mathbf{u})^{-2} (\mathbf{u} \cdot \text{adj}(\boldsymbol{\varepsilon}) \cdot \mathbf{u}^b) \mathbf{u}^b \otimes \mathbf{u}. \quad (110)$$

This brings us to the final form

$$\text{adj}(\mathbf{X}) = P(\mathbf{k} \otimes \mathbf{k}^\sharp) \quad (111)$$

as postulated, where

$$P = -\frac{1}{8}(\mathbf{u}^b \cdot \mathbf{u})^{-4} (\mathbf{k} \cdot \mathbf{u})^4 (\mathbf{u} \cdot \text{adj}(\boldsymbol{\varepsilon}) \cdot \mathbf{u}^b) (\mathbf{u} \cdot \text{adj}(\mathbf{Q}) \cdot \mathbf{u}^b). \quad (112)$$

Thus, I have shown that the requirement $\det(\mathbf{X}) = 0$, which is satisfied identically for wave solutions of Maxwell's equations, provides the scalar condition $P = 0$ with P given above. This subsection began with the form of $\boldsymbol{\chi}$ given in terms of the traditional parameters by Eq. (74). This was not necessary, and one could easily reexpress \mathbf{Q} in terms of the original set of transverse constitutive parameters derived from $\mathbf{G} = \star\boldsymbol{\chi}\mathbf{F}$, but the approach followed provides a connection to the set of parameters that are typically more familiar, and remarkably, they facilitate the factorization of P to determine the optical metric and hence the light cones of the medium, the subject of the next section.

VII. THE OPTICAL METRIC

The concept of the optical metric has a long history. From the early days of general relativity it was known that the light-deflecting properties of a massive spherical object could be obtained without general relativity if the space around the object were filled with an appropriately refracting medium [75]. Gordon reversed this observation by asking whether a refracting medium could be mathematically represented as a curved vacuum space-time, and found the optical metric for isotropic dielectric media moving in a background space-time [48]. Tamm studied anisotropic media residing in a Minkowski background space-time, but did not find the optical metric for such media [49]. Plebanski identified a mapping from a general curved space-time into a dielectric media residing in a flat Minkowski space-time [52], which essentially provides an avenue to map a medium into an optical metric. However, since Plebanski's derivation starts from the vacuum, it is restricted to nonbirefringent media. The

approach by Balakin and Zimdahl postulates two optical metrics and then seeks to match them with the material parameters [59], but does not appear to consider the background space-time in which the medium resides. The idea of the optical metric has received a great deal of attention over the years from different perspectives [36,37,42,60,61], many of which have a focus on the conditions under which the optical metric becomes pseudo-Riemannian. By contrast, the derivation here seeks a tractable expression for the optical metric in terms of the familiar transverse parameters: beginning with Maxwell's equations inside a general medium residing in a background space-time, and showing how the optical metric emerges in the geometrical optics limit.

As previously mentioned, P is a fourth order polynomial in \mathbf{k} , and wave solutions of Maxwell's equations must satisfy $P = 0$. For normalized observers in vacuum, $P = [\frac{1}{2}\mathbf{g}^{-1}(\mathbf{k} \otimes \mathbf{k})]^2 = [\frac{1}{2}g^{\mu\nu}k_\mu k_\nu]^2 = 0$, which shows a degeneracy in the solutions reflecting the fact that the vacuum is not birefringent and all polarization states see the same light cone. The strategy for media in a background space-time will be to show that P can be written in the form of two pseudo-Finslerian optical metrics

$$P \propto \left[\frac{1}{2}\mathbf{g}_+^{-1}(\mathbf{k} \otimes \mathbf{k}) \right] \left[\frac{1}{2}\mathbf{g}_-^{-1}(\mathbf{k} \otimes \mathbf{k}) \right] \quad (113)$$

that are degenerate for certain types of media and reduce to \mathbf{g}^{-1} in vacuum.

The remaining unknown quantity in P is $\mathbf{u} \cdot \text{adj}(\mathbf{Q}) \cdot \mathbf{u}^b$, a tensorial, index expression for which is

$$\mathbf{u} \cdot \text{adj}(\mathbf{Q}) \cdot \mathbf{u}^b = \frac{1}{6} u^\alpha u_\beta \delta_{\alpha\beta_1\beta_2\beta_3}^{\beta\alpha_1\alpha_2\alpha_3} Q_{\alpha_1}{}^{\beta_1} Q_{\alpha_2}{}^{\beta_2} Q_{\alpha_3}{}^{\beta_3}, \quad (114)$$

where $\delta_{\alpha\beta_1\beta_2\beta_3}^{\beta\alpha_1\alpha_2\alpha_3}$ is the generalized Kronecker delta. Thus we need to consider the cube of the complicated matrix expression for \mathbf{Q} in Eq. (107), and then apply some complicated combinatorics to each term. To do so with the most general possible χ can lead one in dizzying circles for months. Hence, to isolate the critical features of the optical metric I will make a simplifying assumption on the parameters but will try to keep them as general as possible. It turns out that most of the complexity in the optical metric comes from the magnetoelectric terms ${}^h\gamma$ and ${}^e\gamma$. To see this, consider the decomposition of ${}^h\gamma$ and ${}^e\gamma$ into antisymmetric, trace, and traceless-symmetric parts as

$${}^h\gamma = \gamma^\sharp \cdot \star(\mathbf{u}^b \wedge \mathbf{h}) + \text{Tr}({}^h\gamma)\mathbf{h} + {}^h\gamma^S \quad (115)$$

and

$${}^e\gamma = (\mathbf{h} \wedge \mathbf{u}) \cdot \star \gamma^e + \text{Tr}({}^e\gamma)\mathbf{h} + {}^e\gamma^S. \quad (116)$$

Here, γ^h and γ^e are 1-forms, while ${}^e\gamma^S$ and ${}^h\gamma^S$ are traceless and symmetric with respect to the background metric, i.e. when both indices are either down or up. By comparing

with the definitions of \mathbf{M} and \mathbf{N} in Eqs. (99) and (100), one can see that the antisymmetric parts of ${}^h\gamma$ and ${}^e\gamma$ will naturally combine with \mathbf{M} and \mathbf{N} in a special way, while the trace and traceless-symmetric parts will contribute in a much more complicated manner. Assume, therefore, that ${}^e\gamma$ and ${}^h\gamma$ are each purely antisymmetric and that

$$\text{Tr}({}^h\gamma) = \text{Tr}({}^e\gamma) = {}^e\gamma^S = {}^h\gamma^S = 0. \quad (117)$$

This assumption imposes 12 constraints on χ , effectively reducing the number of free parameters from 36 to 24, which is still reasonably general, but significantly reduces the complexity involved in finding the optical metric. One might think that this condition simply reduces the number of free parameters in χ by setting some of them to zero, for example that all the magnetoelectric contributions from the skewon part of χ , ${}^s\chi$ vanish. This is not quite the case because we are dealing with ${}^{(h/e)}\gamma$ rather than ${}^{(b/e)}\gamma^c$ which are related by μ , but thus far μ has had no symmetry conditions imposed on it.

With the preceding assumption on the magnetoelectric couplings, the complexity is reduced because it allows \mathbf{Q} to be written as

$$\mathbf{Q} = -\mathbf{h} + (\mathbf{k} \cdot \mathbf{u})^{-2} \bar{\mu} \cdot [(\mathbf{h} \wedge \mathbf{u}) \cdot \star (\mathbf{k} + (\mathbf{k} \cdot \mathbf{u}) \gamma^e)] \cdot \bar{\epsilon} \cdot [(\mathbf{k} + (\mathbf{k} \cdot \mathbf{u}) \gamma^h)^\sharp \cdot \star(\mathbf{u}^b \wedge \mathbf{h})]. \quad (118)$$

This can be rewritten with some more tractable notation. Let

$$\begin{aligned} \mathbf{Z}^e &= (\delta \wedge \mathbf{u}) \cdot \star (\mathbf{k} + (\mathbf{k} \cdot \mathbf{u}) \gamma^e), \\ \mathbf{Z}^h &= (\delta \wedge \mathbf{u}) \cdot \star (\mathbf{k} + (\mathbf{k} \cdot \mathbf{u}) \gamma^h), \end{aligned} \quad (119)$$

and

$$\mathbf{W} = \bar{\mu} \cdot \mathbf{Z}^e \cdot \bar{\epsilon} \cdot \mathbf{Z}^h. \quad (120)$$

Then the tensor \mathbf{Q} has the rather more appealing form

$$\mathbf{Q} = -\mathbf{h} + (\mathbf{k} \cdot \mathbf{u})^{-2} \mathbf{W} \quad (121)$$

from which it becomes much more straightforward to calculate the required adjugate. Expanding Eq. (114) leads to

$$\begin{aligned} \mathbf{u} \cdot \text{adj}(\mathbf{Q}) \cdot \mathbf{u}^b &= -\frac{1}{6} u^\alpha u_\beta \delta_{\alpha\beta_1\beta_2\beta_3}^{\beta\alpha_1\alpha_2\alpha_3} h_{\alpha_1}^{\beta_1} h_{\alpha_2}^{\beta_2} h_{\alpha_3}^{\beta_3} \\ &\quad + \frac{3}{6} (\mathbf{k} \cdot \mathbf{u})^{-2} u^\alpha u_\beta \delta_{\alpha\beta_1\beta_2\beta_3}^{\beta\alpha_1\alpha_2\alpha_3} h_{\alpha_1}^{\beta_1} h_{\alpha_2}^{\beta_2} (\mathbf{W})_{\alpha_3}{}^{\beta_3} \\ &\quad - \frac{3}{6} (\mathbf{k} \cdot \mathbf{u})^{-4} u^\alpha u_\beta \delta_{\alpha\beta_1\beta_2\beta_3}^{\beta\alpha_1\alpha_2\alpha_3} h_{\alpha_1}^{\beta_1} (\mathbf{W})_{\alpha_2}{}^{\beta_2} (\mathbf{W})_{\alpha_3}{}^{\beta_3}. \end{aligned} \quad (122)$$

The reason there is no $(\mathbf{k} \cdot \mathbf{u})^{-6}$ term is that

$$\begin{aligned}
& u^\alpha u_\beta \delta_{\alpha\beta_1\beta_2\beta_3}^{\beta\alpha_1\alpha_2\alpha_3} (\overset{h}{Z}_{\alpha_1}^{\beta_1}) (\overset{h}{Z}_{\alpha_2}^{\beta_2}) (\overset{h}{Z}_{\alpha_3}^{\beta_3}) \\
& = (\mathbf{u}^b \cdot \mathbf{u}) h_{\beta_1\beta_2\beta_3}^{\alpha_1\alpha_2\alpha_3} (\overset{h}{Z}_{\alpha_1}^{\beta_1}) (\overset{h}{Z}_{\alpha_2}^{\beta_2}) (\overset{h}{Z}_{\alpha_3}^{\beta_3}) = 0. \quad (123)
\end{aligned}$$

This is because the expression in the middle is proportional to the three-dimensional determinant of $\overset{h}{Z}$ in the transverse subspace, but it can readily be seen in Eq. (119) that $\overset{h}{Z}$ is annihilated by both \mathbf{u} and $\mathbf{k} + (\mathbf{k} \cdot \mathbf{u})\boldsymbol{\gamma}$ and hence is actually only two dimensional, so its three-dimensional determinant vanishes identically.

Next, consider the action of the generalized Kronecker delta on products of \mathbf{h} . One finds

$$u^\alpha u_\beta \delta_{\alpha\beta_1\beta_2\beta_3}^{\beta\alpha_1\alpha_2\alpha_3} h_{\alpha_1}^{\beta_1} h_{\alpha_2}^{\beta_2} h_{\alpha_3}^{\beta_3} = 6(\mathbf{u}^b \cdot \mathbf{u}), \quad (124a)$$

$$u^\alpha u_\beta \delta_{\alpha\beta_1\beta_2\beta_3}^{\beta\alpha_1\alpha_2\alpha_3} h_{\alpha_1}^{\beta_1} h_{\alpha_2}^{\beta_2} = 2(\mathbf{u}^b \cdot \mathbf{u}) h_{\beta_3}^{\alpha_3}, \quad (124b)$$

$$u^\alpha u_\beta \delta_{\alpha\beta_1\beta_2\beta_3}^{\beta\alpha_1\alpha_2\alpha_3} h_{\alpha_1}^{\beta_1} = (\mathbf{u}^b \cdot \mathbf{u}) h_{\beta_2\beta_3}^{\alpha_2\alpha_3}. \quad (124c)$$

With this, we now have

$$\begin{aligned}
\mathbf{u} \cdot \text{adj}(\mathbf{Q}) \cdot \mathbf{u}^b &= -(\mathbf{u}^b \cdot \mathbf{u}) \left[1 - (\mathbf{k} \cdot \mathbf{u})^{-2} \text{Tr}(\mathbf{W}) \right. \\
&\quad \left. - \frac{1}{2} (\mathbf{k} \cdot \mathbf{u})^{-4} (\text{Tr}(\mathbf{W} \cdot \mathbf{W}) - \text{Tr}(\mathbf{W})^2) \right]. \quad (125)
\end{aligned}$$

Completing the square, this can be factored to

$$\begin{aligned}
\mathbf{u} \cdot \text{adj}(\mathbf{Q}) \cdot \mathbf{u}^b &= -(\mathbf{u}^b \cdot \mathbf{u}) \left[1 - \frac{1}{2} (\mathbf{k} \cdot \mathbf{u})^{-2} \text{Tr}(\mathbf{W}) + (\mathbf{k} \cdot \mathbf{u})^{-2} \sqrt{\frac{1}{2} \text{Tr}(\mathbf{W} \cdot \mathbf{W}) - \frac{1}{4} \text{Tr}(\mathbf{W})^2} \right] \\
&\quad \times \left[1 - \frac{1}{2} (\mathbf{k} \cdot \mathbf{u})^{-2} \text{Tr}(\mathbf{W}) - (\mathbf{k} \cdot \mathbf{u})^{-2} \sqrt{\frac{1}{2} \text{Tr}(\mathbf{W} \cdot \mathbf{W}) - \frac{1}{4} \text{Tr}(\mathbf{W})^2} \right]. \quad (126)
\end{aligned}$$

Returning now to Eq. (112) for P , we have

$$P = \frac{1}{2} (\mathbf{u}^b \cdot \mathbf{u}) (\mathbf{u} \cdot \text{adj}(\boldsymbol{\varepsilon}) \cdot \mathbf{u}^b)^{-1} H_+ H_- \quad (127)$$

with

$$H_+ = \frac{1}{2} (\mathbf{u}^b \cdot \mathbf{u})^{-2} (\mathbf{u} \cdot \text{adj}(\boldsymbol{\varepsilon}) \cdot \mathbf{u}^b) \left[(\mathbf{k} \cdot \mathbf{u})^2 - \frac{1}{2} \text{Tr}(\mathbf{W}) + \sqrt{\frac{1}{2} \text{Tr}(\mathbf{W} \cdot \mathbf{W}) - \frac{1}{4} \text{Tr}(\mathbf{W})^2} \right] \quad (128)$$

and

$$H_- = \frac{1}{2} (\mathbf{u}^b \cdot \mathbf{u})^{-2} (\mathbf{u} \cdot \text{adj}(\boldsymbol{\varepsilon}) \cdot \mathbf{u}^b) \left[(\mathbf{k} \cdot \mathbf{u})^2 - \frac{1}{2} \text{Tr}(\mathbf{W}) - \sqrt{\frac{1}{2} \text{Tr}(\mathbf{W} \cdot \mathbf{W}) - \frac{1}{4} \text{Tr}(\mathbf{W})^2} \right]. \quad (129)$$

Recall that wave solutions of Maxwell's equations must satisfy $P = 0$. This condition can in turn be met through the satisfaction of either of the conditions $H_+ = 0$ or $H_- = 0$. In other words, the medium exhibits two wave-propagation eigenstates that follow different ray trajectories, i.e. birefringence. The H_\pm serve as pseudo-Finslerian structures on the manifold, and we may define the associated optical metrics

$$\mathbf{g}_\pm^{\mu\nu}(\mathbf{x}, \mathbf{k}) = \frac{\partial^2 H_\pm}{\partial k_\mu \partial k_\nu}, \quad (130)$$

where the arguments of \mathbf{g}_\pm^{-1} have been included explicitly to show that the optical metric depends not just on the location in the medium but also on the wave covector at any point. This dependence on the wave covector means that waves passing through the same point but propagating in different directions will in general see different optical metrics. This complicated dependence of the optical metrics on the wave covector potentially makes them a somewhat less useful concept than one might otherwise expect. The optical metric defines the light cone, but the residual dependence on \mathbf{k} may prove to make an algebraic determination of the light cone in practice rather difficult. Instead, all of the ray

content for each polarization is contained in H_{\pm} , as discussed in Sec. IX.

The expressions for H_{\pm} given above are not particularly illuminating. Ideally, we should be able to isolate the \mathbf{k} dependence in an expression of the form

$$H_{\pm} = a^{\alpha\beta} k_{\alpha} k_{\beta} \pm \sqrt{b^{\mu\nu\sigma\rho} k_{\mu} k_{\nu} k_{\sigma} k_{\rho}}, \quad (131)$$

where $a^{\alpha\beta}$ and $b^{\mu\nu\sigma\rho}$ depend in some way on the usual material parameters as functions of points on the manifold, but are independent of \mathbf{k} . Examining \mathbf{W} , one finds

$$\begin{aligned} \mathbf{W} &= \bar{\boldsymbol{\mu}} \cdot [(\boldsymbol{\delta} \wedge \mathbf{u}) \star (\mathbf{k} + (\mathbf{k} \cdot \mathbf{u}) \boldsymbol{\gamma}^e)] \\ &\quad \cdot \bar{\boldsymbol{\epsilon}} \cdot [(\mathbf{k} + (\mathbf{k} \cdot \mathbf{u}) \boldsymbol{\gamma}^h)^{\sharp} \star (\mathbf{u}^b \wedge \boldsymbol{\delta})] \\ W_{\alpha}^{\kappa} &= -(\mathbf{u}^b \cdot \mathbf{u}) h_{\psi\beta\varphi}^{\lambda\kappa\rho} g_{\lambda\tau} \bar{\boldsymbol{\epsilon}}^{\tau} g^{\sigma\psi} \bar{\boldsymbol{\mu}}_{\alpha}^{\beta} g^{\eta\varphi} \\ &\quad \times (\delta_{\rho}^{\mu} + \boldsymbol{\gamma}_{\rho}^e u^{\mu}) (\delta_{\eta}^{\nu} + \boldsymbol{\gamma}_{\eta}^e u^{\nu}) k_{\mu} k_{\nu} \\ &= W_{\alpha}^{\kappa\mu\nu} k_{\mu} k_{\nu} \end{aligned} \quad (132)$$

from which it follows that H_{\pm} attains the desired format

$$\begin{aligned} H_{\pm} &= \frac{1}{2} (\mathbf{u}^b \cdot \mathbf{u})^{-2} (\mathbf{u} \cdot \text{adj}(\boldsymbol{\epsilon}) \cdot \mathbf{u}^b) \left[\left(u^{\mu} u^{\nu} - \frac{1}{2} W_{\alpha}^{\alpha\mu\nu} \right) k_{\mu} k_{\nu} \right. \\ &\quad \left. \pm \sqrt{\left(\frac{1}{2} W_{\alpha}^{\beta\mu\nu} W_{\beta}^{\alpha\sigma\rho} - \frac{1}{4} W_{\alpha}^{\alpha\mu\nu} W_{\beta}^{\beta\sigma\rho} \right) k_{\mu} k_{\nu} k_{\sigma} k_{\rho}} \right]. \end{aligned} \quad (133)$$

With this explicit form of the dependence on \mathbf{k} , it is straightforward to show that the optical metrics may be computed as in Eq. (130), and one may subsequently show that

$$H_{\pm} = \frac{1}{2} \mathbf{g}_{\pm}^{\alpha\beta} k_{\alpha} k_{\beta} = 0 \quad (134)$$

provides the condition satisfied by light rays in dielectric media. Although the index expression Eq. (132) for $W_{\alpha}^{\pi\mu\nu}$ looks complicated, it is entirely expressed in terms of the ‘‘usual’’ transverse parameters $\boldsymbol{\epsilon}$, $\boldsymbol{\mu}$, ${}^h\boldsymbol{\gamma}$, and ${}^e\boldsymbol{\gamma}$, and the background metric, and it may easily be handled by any computer algebra system.

VIII. REDUCTION TO PSEUDO-RIEMANNIAN OPTICAL METRIC

A natural question of interest is, under what conditions does birefringence vanish? For nonbirefringent media, the independent light cones for each polarization become degenerate and $\mathbf{g}_{+} = \mathbf{g}_{-}$, or equivalently, $H_{+} = H_{-}$. This condition requires the square root term of Eqs. (128) and (129), or equivalently Eq. (133), to vanish, whence the two associated pseudo-Finslerian optical

metrics defined by Eq. (130) degenerate to a single, pseudo-Riemannian optical metric. For any given medium there may exist particular choices of \mathbf{k} for which the square root term vanishes, corresponding to optical axes of the medium. More generally, one would like to know which constraints on the medium must be imposed such that the square root vanishes for *all* \mathbf{k} , which is a more difficult question to answer. There are several possibilities that should be investigated, but from the complexity of the expression it seems clear that it will only vanish under quite restrictive conditions. In particular, consider the restrictions ${}^e\boldsymbol{\gamma} = {}^h\boldsymbol{\gamma} = \boldsymbol{\gamma}$, and $\boldsymbol{\epsilon}_{\alpha\beta}$ and $\boldsymbol{\xi}_{\alpha\beta}$ symmetric with respect to the background metric.

First, return to Eq. (132) for \mathbf{W} , and let

$$\mathbf{q} = \mathbf{k} + (\mathbf{k} \cdot \mathbf{u}) \boldsymbol{\gamma}; \quad (135)$$

then \mathbf{W} has the simplified expression

$$W_{\alpha}^{\kappa} = -(\mathbf{u}^b \cdot \mathbf{u}) (h_{\psi\beta\varphi}^{\lambda\kappa\rho} q_{\rho} q^{\varphi}) \bar{\boldsymbol{\epsilon}}_{\lambda}^{\psi} \bar{\boldsymbol{\mu}}_{\alpha}^{\beta} \quad (136)$$

(the metric has been absorbed by the symmetry of $\boldsymbol{\epsilon}$). Notice that now W_{α}^{κ} is actually orthogonal to *both* \mathbf{u} and \mathbf{q} , and is therefore effectively only two-dimensional. Define

$$\mathbf{j} = \mathbf{h} - \frac{(\mathbf{h} \cdot \mathbf{q}) \otimes (\mathbf{h} \cdot \mathbf{q})^{\sharp}}{(\mathbf{h} \cdot \mathbf{q}) \cdot (\mathbf{h} \cdot \mathbf{q})^{\sharp}} \quad (137)$$

as the projection operator from the subspace orthogonal to \mathbf{u} to the subspace orthogonal to both \mathbf{u} and \mathbf{q} . One may readily show that $\mathbf{j} \cdot \mathbf{j} = \mathbf{j}$ and that the Kronecker \mathbf{h} on the subspace orthogonal to \mathbf{u} is projected to $\mathbf{j} \cdot \mathbf{h} \cdot \mathbf{j} = \mathbf{j}$, and thus \mathbf{j} serves as the Kronecker tensor on the subspace orthogonal to both \mathbf{u} and \mathbf{q} . Similar to Eq. (72), one finds

$$h_{\tau\theta\varphi}^{\sigma\kappa\rho} q_{\rho} q^{\varphi} = ((\mathbf{h} \cdot \mathbf{q}) \cdot (\mathbf{h} \cdot \mathbf{q})^{\sharp}) j_{\tau\theta}^{\sigma\kappa}. \quad (138)$$

Equation (136) for W_{α}^{κ} now becomes

$$W_{\alpha}^{\kappa} = -(\mathbf{u}^b \cdot \mathbf{u}) ((\mathbf{h} \cdot \mathbf{q}) \cdot (\mathbf{h} \cdot \mathbf{q})^{\sharp}) j_{\psi\beta}^{\lambda\kappa} \bar{\boldsymbol{\epsilon}}_{\lambda}^{\psi} \bar{\boldsymbol{\mu}}_{\alpha}^{\beta}. \quad (139)$$

Calculating the argument of the square root in H_{\pm} in Eq. (133), one finds

$$\begin{aligned} &\frac{1}{2} \text{Tr}(\mathbf{W} \cdot \mathbf{W}) - \frac{1}{4} \text{Tr}(\mathbf{W})^2 \\ &= (\mathbf{u}^b \cdot \mathbf{u})^2 ((\mathbf{h} \cdot \mathbf{q}) \cdot (\mathbf{h} \cdot \mathbf{q})^{\sharp})^2 \\ &\quad \times \left[\frac{1}{2} j_{\tau\theta}^{\sigma\kappa} j_{\nu\beta}^{\mu\rho} \bar{\boldsymbol{\epsilon}}_{\sigma}^{\tau} \bar{\boldsymbol{\epsilon}}_{\mu}^{\nu} \bar{\boldsymbol{\mu}}_{\kappa}^{\rho} \bar{\boldsymbol{\mu}}_{\beta}^{\theta} - \frac{1}{4} (j_{\tau\theta}^{\sigma\kappa} \bar{\boldsymbol{\epsilon}}_{\sigma}^{\tau} \bar{\boldsymbol{\mu}}_{\kappa}^{\theta})^2 \right]. \end{aligned} \quad (140)$$

Expanding the generalized Kronecker tensors into products of \mathbf{j} , this becomes

$$\begin{aligned}
& \frac{1}{2}\text{Tr}(\mathbf{W} \cdot \mathbf{W}) - \frac{1}{4}\text{Tr}(\mathbf{W})^2 \\
&= (\mathbf{u}^b \cdot \mathbf{u})^2 ((\mathbf{h} \cdot \mathbf{q}) \cdot (\mathbf{h} \cdot \mathbf{q})^\#)^2 \\
&\quad \times \left[\frac{1}{2}\text{Tr}(\bar{\boldsymbol{\epsilon}}_j)^2 \text{Tr}(\bar{\boldsymbol{\mu}}_j^2) - \text{Tr}(\bar{\boldsymbol{\epsilon}}_j) \text{Tr}(\bar{\boldsymbol{\epsilon}}_j \cdot \bar{\boldsymbol{\mu}}_j^2) + \frac{1}{2}\text{Tr}(\bar{\boldsymbol{\epsilon}}_j^2 \cdot \bar{\boldsymbol{\mu}}_j^2) \right. \\
&\quad - \frac{1}{4}\text{Tr}(\bar{\boldsymbol{\epsilon}}_j)^2 \text{Tr}(\bar{\boldsymbol{\mu}}_j)^2 + \frac{1}{2}\text{Tr}(\bar{\boldsymbol{\epsilon}}_j) \text{Tr}(\bar{\boldsymbol{\mu}}_j) \text{Tr}(\bar{\boldsymbol{\epsilon}}_j \cdot \bar{\boldsymbol{\mu}}_j) \\
&\quad \left. - \frac{1}{4}\text{Tr}(\bar{\boldsymbol{\epsilon}}_j \cdot \bar{\boldsymbol{\mu}}_j)^2 \right], \tag{141}
\end{aligned}$$

where

$$\bar{\boldsymbol{\epsilon}}_j = \mathbf{j} \cdot \bar{\boldsymbol{\epsilon}} \cdot \mathbf{j}, \quad \bar{\boldsymbol{\mu}}_j = \mathbf{j} \cdot \bar{\boldsymbol{\mu}} \cdot \mathbf{j}. \tag{142}$$

This will vanish for two possible cases.

A. $\bar{\boldsymbol{\epsilon}}$ and $\bar{\boldsymbol{\mu}}$ proportional to \mathbf{h}

Consider first the case where $\bar{\boldsymbol{\epsilon}}$ and $\bar{\boldsymbol{\mu}}$ are both proportional to \mathbf{h} ; i.e. they are both isotropic so $\boldsymbol{\epsilon} = \epsilon \mathbf{h}$ and $\boldsymbol{\mu} = \mu \mathbf{h}$. Since $\mathbf{j} \cdot \mathbf{h} \cdot \mathbf{j} = j$, the sum over traces in Eq. (141) becomes

$$\begin{aligned}
& \epsilon^{-2} \mu^{-2} \left[\frac{1}{2}\text{Tr}(\mathbf{j})^2 \text{Tr}(\mathbf{j}^2) - \text{Tr}(\mathbf{j}) \text{Tr}(\mathbf{j} \cdot \mathbf{j}^2) + \frac{1}{2}\text{Tr}(\mathbf{j}^2 \cdot \mathbf{j}^2) \right. \\
& \quad \left. - \frac{1}{4}\text{Tr}(\mathbf{j})^2 \text{Tr}(\mathbf{j})^2 + \frac{1}{2}\text{Tr}(\mathbf{j}) \text{Tr}(\mathbf{j}) \text{Tr}(\mathbf{j} \cdot \mathbf{j}) - \frac{1}{4}\text{Tr}(\mathbf{j} \cdot \mathbf{j})^2 \right] \\
&= \epsilon^{-2} \mu^{-2} \left[\frac{1}{2}\text{Tr}(\mathbf{j})^2 \text{Tr}(\mathbf{j}) - \text{Tr}(\mathbf{j}) \text{Tr}(\mathbf{j}) + \frac{1}{2}\text{Tr}(\mathbf{j}) \right. \\
& \quad \left. - \frac{1}{4}\text{Tr}(\mathbf{j})^2 \text{Tr}(\mathbf{j})^2 + \frac{1}{2}\text{Tr}(\mathbf{j})^2 \text{Tr}(\mathbf{j}) - \frac{1}{4}\text{Tr}(\mathbf{j})^2 \right] \\
&= 0, \tag{143}
\end{aligned}$$

where the final equality follows from the fact that $\text{Tr}(\mathbf{j}) = 2$.

B. $\bar{\boldsymbol{\mu}} = \bar{\boldsymbol{\epsilon}}$ (equivalently $\boldsymbol{\mu} = \boldsymbol{\epsilon}$)

Next, consider the case in which $\boldsymbol{\mu} = \boldsymbol{\epsilon}$. In this case the sum of traces becomes

$$\begin{aligned}
& \frac{1}{4} [-\text{Tr}(\bar{\boldsymbol{\mu}}_j)^4 + 4\text{Tr}(\bar{\boldsymbol{\mu}}_j)^2 \text{Tr}(\bar{\boldsymbol{\mu}}_j^2) - 4\text{Tr}(\bar{\boldsymbol{\mu}}_j) \text{Tr}(\bar{\boldsymbol{\mu}}_j^3) \\
& \quad - \text{Tr}(\bar{\boldsymbol{\mu}}_j^2)^2 + 2\text{Tr}(\bar{\boldsymbol{\mu}}_j^4)]. \tag{144}
\end{aligned}$$

To see that this vanishes, rewrite it in the following clever way:

$$\begin{aligned}
& -\frac{2}{24} [\text{Tr}(\bar{\boldsymbol{\mu}}_j)^4 - 6\text{Tr}(\bar{\boldsymbol{\mu}}_j)^2 \text{Tr}(\bar{\boldsymbol{\mu}}_j^2) + 3\text{Tr}(\bar{\boldsymbol{\mu}}_j^2)^2 \\
& \quad + 8\text{Tr}(\bar{\boldsymbol{\mu}}_j) \text{Tr}(\bar{\boldsymbol{\mu}}_j^3) - 6\text{Tr}(\bar{\boldsymbol{\mu}}_j^4)] \\
& - \frac{1}{6} \text{Tr}(\bar{\boldsymbol{\mu}}_j) [\text{Tr}(\bar{\boldsymbol{\mu}}_j)^3 - 3\text{Tr}(\bar{\boldsymbol{\mu}}_j) \text{Tr}(\bar{\boldsymbol{\mu}}_j^2) + 2\text{Tr}(\bar{\boldsymbol{\mu}}_j^3)]. \tag{145}
\end{aligned}$$

Now, each of the bracketed terms is the trace expression for the determinant of a matrix, the first being the four-dimensional determinant and the second being the three-dimensional determinant; in other words, Eq. (145) is equivalent to

$$-2\text{Det}_4(\bar{\boldsymbol{\mu}}_j) - \text{Tr}(\bar{\boldsymbol{\mu}}_j) \text{Det}_3(\bar{\boldsymbol{\mu}}_j). \tag{146}$$

This vanishes because $\bar{\boldsymbol{\mu}}_j$, being the projection of $\bar{\boldsymbol{\mu}}$ into the two-dimensional subspace orthogonal to both \mathbf{u} and \mathbf{q} , is effectively only two dimensional, and hence has identically vanishing three- and four-dimensional determinants. We may conclude that a medium with $\bar{\boldsymbol{\mu}} = \bar{\boldsymbol{\epsilon}}$ and ${}^e\boldsymbol{\gamma} = {}^h\boldsymbol{\gamma}$ is also not birefringent, and the pseudo-Finslerian optical metric becomes pseudo-Riemannian in these two cases.

C. Expression for pseudo-Riemannian optical metric

Now that the conditions under which H_\pm admit a pseudo-Riemannian optical metric have been identified, we may find a more explicit expression for it in terms of the usual material parameters. Consider first the conditions adopted in Case 2 above: that $\boldsymbol{\epsilon}$ and $\boldsymbol{\mu}$ are equal and symmetric with respect to the background space-time metric, and that ${}^e\boldsymbol{\gamma}$ and ${}^h\boldsymbol{\gamma}$ are antisymmetric with respect to the background space-time metric and defined as described in Eqs. (115) and (116) with ${}^e\boldsymbol{\gamma} = {}^h\boldsymbol{\gamma} = \boldsymbol{\gamma}$. Returning to the expression for \mathbf{W} given by Eq. (136) and setting $\bar{\boldsymbol{\epsilon}} = \bar{\boldsymbol{\mu}}$,

$$\text{Tr}(\mathbf{W}) = W_\alpha^\alpha = -(\mathbf{u}^b \cdot \mathbf{u}) h_{\psi\beta\phi}^{\lambda\alpha\rho} \bar{\mu}_{\lambda\psi} \bar{\mu}_\alpha^\beta q_\rho q^\phi. \tag{147}$$

From Eq. (71) for the three-dimensional inverse on the transverse subspace, it may be seen that

$$(\mathbf{u}^b \cdot \mathbf{u}) h_{\psi\beta\phi}^{\lambda\alpha\rho} \bar{\mu}_{\lambda\psi} \bar{\mu}_\alpha^\beta = 2(\mathbf{u} \cdot \text{adj}(\bar{\boldsymbol{\mu}}) \cdot \mathbf{u}^b) \mu_\phi^\rho \tag{148}$$

and thus

$$\begin{aligned}
\text{Tr}(\mathbf{W}) &= -2(\mathbf{u} \cdot \text{adj}(\bar{\boldsymbol{\mu}}) \cdot \mathbf{u}^b) \mu_\psi^\theta g^{\psi\rho} q_\rho q_\theta \\
&= -2(\mathbf{u} \cdot \text{adj}(\bar{\boldsymbol{\mu}}) \cdot \mathbf{u}^b) (g^{\psi\rho} \mu_\psi^\theta k_\rho k_\theta + g^{\psi\rho} \mu_\psi^\nu \gamma_\nu k_\rho u^\theta k_\theta \\
&\quad + u^\rho k_\rho \gamma_\mu g^{\mu\mu} \mu_\psi^\theta k_\theta + \gamma_\mu g^{\mu\mu} \mu_\psi^\nu \gamma_\nu u^\rho u^\theta k_\rho k_\theta) \\
&= -2(\mathbf{u} \cdot \text{adj}(\bar{\boldsymbol{\mu}}) \cdot \mathbf{u}^b) (g^{\psi\rho} \mu_\psi^\theta + g^{\psi\rho} \mu_\psi^\nu \gamma_\nu u^\theta \\
&\quad + u^\rho \gamma_\mu g^{\mu\mu} \mu_\psi^\theta + \gamma_\mu g^{\mu\mu} \mu_\psi^\nu \gamma_\nu u^\theta) k_\rho k_\theta. \tag{149}
\end{aligned}$$

Writing it in an index-free form

$$\begin{aligned}
\text{Tr}(\mathbf{W}) &= -2(\mathbf{u} \cdot \text{adj}(\bar{\boldsymbol{\mu}}) \cdot \mathbf{u}^b) [g^{-1} \cdot \boldsymbol{\mu} + (g^{-1} \cdot \boldsymbol{\mu} \cdot \boldsymbol{\gamma}) \otimes \mathbf{u} \\
&\quad + \mathbf{u} \otimes (\boldsymbol{\gamma} \cdot g^{-1} \cdot \boldsymbol{\mu}) + (\boldsymbol{\gamma} \cdot g^{-1} \cdot \boldsymbol{\mu} \cdot \boldsymbol{\gamma}) \mathbf{u} \otimes \mathbf{u}] (\mathbf{k} \otimes \mathbf{k}). \tag{150}
\end{aligned}$$

Finally, returning to Eqs. (128) and (129) and using the fact that

$$(\mathbf{u}^b \cdot \mathbf{u})^{-2} (\mathbf{u} \cdot \text{adj}(\boldsymbol{\mu}) \cdot \mathbf{u}^b) (\mathbf{u} \cdot \text{adj}(\bar{\boldsymbol{\mu}}) \cdot \mathbf{u}^b) = 1, \quad (151)$$

one has

$$\begin{aligned} H_{\pm} = & \frac{1}{2} [((\mathbf{u}^b \cdot \mathbf{u})^{-2} (\mathbf{u} \cdot \text{adj}(\boldsymbol{\mu}) \cdot \mathbf{u}^b) \\ & + (\boldsymbol{\gamma} \cdot \mathbf{g}^{-1} \cdot \boldsymbol{\mu} \cdot \boldsymbol{\gamma})) \mathbf{u} \otimes \mathbf{u} \\ & + (\mathbf{g}^{-1} \cdot \boldsymbol{\mu} \cdot \boldsymbol{\gamma}) \otimes \mathbf{u} + \mathbf{u} \otimes (\boldsymbol{\gamma} \cdot \mathbf{g}^{-1} \cdot \boldsymbol{\mu}) \\ & + \mathbf{g}^{-1} \cdot \boldsymbol{\mu}] (\mathbf{k} \otimes \mathbf{k}). \end{aligned} \quad (152)$$

Comparing with Eq. (134), it is clear that the optical metrics of the medium are degenerate, $\mathbf{g}_+^{-1} = \mathbf{g}_-^{-1} = \mathbf{g}^{-1}$, and

$$\begin{aligned} \mathbf{g}^{-1} = & ((\mathbf{u}^b \cdot \mathbf{u})^{-2} (\mathbf{u} \cdot \text{adj}(\boldsymbol{\mu}) \cdot \mathbf{u}^b) + (\boldsymbol{\gamma} \cdot \mathbf{g}^{-1} \cdot \boldsymbol{\mu} \cdot \boldsymbol{\gamma})) \mathbf{u} \otimes \mathbf{u} \\ & + (\mathbf{g}^{-1} \cdot \boldsymbol{\mu} \cdot \boldsymbol{\gamma}) \otimes \mathbf{u} + \mathbf{u} \otimes (\boldsymbol{\gamma} \cdot \mathbf{g}^{-1} \cdot \boldsymbol{\mu}) + \mathbf{g}^{-1} \cdot \boldsymbol{\mu}. \end{aligned} \quad (153)$$

Although it required a rather lengthy derivation to obtain this expression, it may be verified by making a comparison between P computed with Eq. (127) using this optical metric and P computed directly from the adjugate of \mathbf{X} as in Eq. (93). Relative to the observer \mathbf{u} , the purely spatial part of the inverse metric is just $\boldsymbol{\mu}$ ($=\boldsymbol{\varepsilon}$) with its first index raised by the background space-time metric. The time-space components of \mathbf{g}^{-1} are somewhat more complicated, but simplify when one considers \mathbf{g} , where they become simply proportional to the magnetoelectric coupling 1-form $\boldsymbol{\gamma}$.

It is straightforward to check that the optical metric reduces to the background space-time metric in vacuum. Indeed, setting $\boldsymbol{\mu} = \mathbf{h}$ and $\boldsymbol{\gamma} = 0$, and making use of Eq. (109) for the adjugate of \mathbf{h} , one has

$$\begin{aligned} \mathbf{g}^{-1} = & (\mathbf{u}^b \cdot \mathbf{u})^{-1} \mathbf{u} \otimes \mathbf{u} + \mathbf{g}^{-1} \cdot \mathbf{h} \\ = & \mathbf{g}^{-1} \cdot \left(\frac{\mathbf{u}^b \otimes \mathbf{u}}{(\mathbf{u}^b \cdot \mathbf{u})} + \mathbf{h} \right) = \mathbf{g}^{-1} \cdot \boldsymbol{\delta} = \mathbf{g}^{-1}. \end{aligned} \quad (154)$$

It is also straightforward to check that Gordon's optical metric is recovered for the case of isotropic, nonmagneto-electrically coupled media comoving with the observer [48]. Indeed, setting $\boldsymbol{\varepsilon} = \epsilon \mathbf{h}$, $\boldsymbol{\mu} = \mu \mathbf{h}$, and $\boldsymbol{\gamma} = 0$, and calculating H_{\pm} , one finds

$$H_{\pm} = \frac{1}{2} [\epsilon (\mathbf{u}^b \cdot \mathbf{u})^{-1} \mathbf{u} \otimes \mathbf{u} + \mu^{-1} \mathbf{g}^{-1} \cdot \mathbf{h}] (\mathbf{k} \otimes \mathbf{k}). \quad (155)$$

Exploiting the conformal invariance of the metric we may rescale by μ and expand \mathbf{h} to obtain

$$\mathbf{g}^{-1} = \mathbf{g}^{-1} + (\mathbf{u}^b \cdot \mathbf{u})^{-1} (\epsilon \mu - 1) \mathbf{u} \otimes \mathbf{u}. \quad (156)$$

Setting the normalization to $(\mathbf{u}^b \cdot \mathbf{u}) = -1$, the index expression

$$\mathbf{g}^{\alpha\beta} = g^{\alpha\beta} - (\epsilon \mu - 1) u^{\alpha} u^{\beta} \quad (157)$$

is identical to the optical metric found by Gordon [48], [Eq. (16)].

IX. RAY TRACING AND KINEMATICS

To put the previous sections into a useful context and tie up the discussion of geometric optics, I turn now briefly to the question of ray tracing and the kinematics of congruences. This topic was recently covered in some detail in Ref. [46], where we studied the kinematics of light beams traversing dielectric media in curved background spacetimes, and derived the generalized Raychaudhuri equation associated with them. The Raychaudhuri equation describes how the cross section of a beam evolves along the length of the beam and can essentially be thought of as describing the focus of the beam. In Ref. [46] we started with the proposition that P is factorizable, deferring the proof for the present paper, and that the associated optical metrics exist in the form derived here.

In the geometric optics limit, solution data consist of space-time points p together with a wave-(co)vector \mathbf{k} at p . In other words, solutions consist of curves in the eight-dimensional cotangent bundle $\mathcal{C}: \mathbb{R} \rightarrow T^*M$, i.e. parametrized curves in phase space

$$\mathcal{C}(\tau) = (p(\tau), \mathbf{k}(\tau)) = (x^{\mu}(\tau), k_{\mu}(\tau)). \quad (158)$$

Thus, the geometric optics limit discards the tensorial nature of the fields and mathematically reduces the wave equation to the propagation of scalar particles. A physical ray trajectory $\tilde{\mathcal{C}}$ is the projection of a solution curve \mathcal{C} from the cotangent bundle into the manifold M with the projection operator

$$\pi: T^*M \rightarrow M \quad (159)$$

such that

$$\tilde{\mathcal{C}}(\tau) = \pi(\mathcal{C}(\tau)) = p(\tau). \quad (160)$$

Obtaining these physical trajectories on the manifold is the goal of ray tracing, but it should be borne in mind that one must actually solve for the solution curve \mathcal{C} in the cotangent bundle, and then project this to the manifold.

In the previous sections it was shown that solution curves must satisfy the condition $P = 0$, and, at least under certain conditions on the magnetoelectric coupling, that $P \propto H_+ H_-$. Thus any wave solution of Maxwell's equations in the geometrical optics limit must everywhere satisfy $H = 0$ for either $H = H_+$ or $H = H_-$. This means that $H = 0$ anywhere along a solution curve \mathcal{C} parametrized by τ , and thus

$$\frac{dH}{d\tau} = 0. \quad (161)$$

Since $H(p, \mathbf{k})$ is a function on the cotangent bundle, then

$$\frac{dH}{d\tau} = \left(\frac{dx^\mu}{d\tau}\right) \frac{\partial H}{\partial x^\mu} + \left(\frac{dk_\mu}{d\tau}\right) \frac{\partial H}{\partial k_\mu} = \dot{x}^\mu \frac{\partial H}{\partial x^\mu} + \dot{k}_\mu \frac{\partial H}{\partial k_\mu} = 0, \quad (162)$$

which implies Hamilton's canonical equations

$$\dot{x}^\mu = \frac{\partial H}{\partial k_\mu}, \quad \dot{k}_\mu = -\frac{\partial H}{\partial x^\mu}. \quad (163)$$

The solution to this set of coupled equations determines the trajectories in the phase space, and the $x^\mu(\tau)$ component gives the ray trajectories on the manifold. With H as found in the previous sections, Hamilton's equations allow for ray tracing through dielectric media in curved space-times, where the medium may smoothly vary among positively refracting, negatively refracting, and vacuum. Since the frequency component is traced over together with the spatial components of \mathbf{k} , they should automatically account for gravitational redshift and for frequency-shifting media [76].

Let

$$v^\mu = \dot{x}^\mu \quad (164)$$

denote the tangent to \tilde{C} . Since $H = \frac{1}{2} \mathbf{g}^{-1}(\mathbf{k} \otimes \mathbf{k})$ by Eq. (134), then by the first of Hamilton's equations it follows that

$$\mathbf{v} = \mathbf{g}^{-1} \cdot \mathbf{k}. \quad (165)$$

Thus the ray direction is related to the wave vector through the optical metric. In a vacuum space-time or in isotropic media one finds that \mathbf{v} and \mathbf{k} are “parallel.” The typical statement that \mathbf{v} and \mathbf{k} are *not* parallel in more general media belies the fact that despite having the light cone determined by the optical metric \mathbf{g} , measures of angle and distance are still understood as being made with respect to the background metric \mathbf{g} . With this realization, one may go further than simple ray tracing to consider the behavior of congruences of light, i.e. beams in media [46]. By analyzing such congruences, it has been shown that dielectric analog space-times—the idea of mimicking the behavior of light propagation in a curved space-time with a corresponding dielectric residing in flat space-time—are unfaithful in that although one may be able to mimic some idea of the ray trajectory of light, one cannot simultaneously replicate the behavior of a congruence [35].

It should be emphasized that this result requires access to the background metric of the space-time in which the medium is embedded, which is of course the natural setting

for the many ongoing experimental efforts in dielectric analog space-times. In a premetric setting, the optical metric is the only metric available and the analysis in that case is less clear, for while the optical metric of the vacuum would be identical to the optical metric of the background-free analog, the kinematics of congruences and the Raychaudhuri equation depend on covariant derivatives that in the vacuum are taken with respect to the background-metric compatible connection. Presumably, one could by fiat impose a covariant derivative in a background-free analog that is compatible with the optical metric there and which might give agreement with the vacuum kinematics, at least in certain cases, e.g. where the initial manifold is vacuum and one chooses a trivial projection. But if, for example, the initial manifold contains a nonvacuum contribution to the optical metric, such as from a refractive dust accretion disk in a Kerr background, then the distinction between background and optical metrics becomes immediately relevant in a determination of the kinematics—a distinction that will *not* carry through to the optical metric of a background-free analog. Such background-free constructions can therefore, in special cases, be consistent and distortion-free from a mathematical perspective, but are nearly always inconsistent from the perspective of real-world observers existing in background space-times making measurements with real instruments, and who may want to make other relevant (and in particular, timelike) measurements.

This approach to ray tracing in dielectric media in curved space-times may also be useful in astrophysical settings such as light propagating through an accretion disk that may possess some refractive properties.

X. SUMMARY AND CONCLUSIONS

The goal of this paper was to study the covariant electrodynamics of general linear dielectric media within a curved space-time. A $(3+1)$ space-time splitting was provided by a timelike vector field interpreted as a family of observers moving in the curved space-time. This allowed the discussion to be formulated in terms of the familiar spacelike material parameters that would be measured by these observers. Through the presence of the observers, the correspondence between χ and the transverse $\boldsymbol{\epsilon}$, $\boldsymbol{\mu}$, ${}^h\boldsymbol{\gamma}$, and ${}^e\boldsymbol{\gamma}$ was obtained through an explicit derivation that ensures true coordinate independence.

Of particular interest is the geometric optics limit and the derivation of the optical metric. I followed a completely *ab initio* derivation starting from the differential-forms expression of Maxwell's equations in linear media and obtained an expression for the pseudo-Finslerian optical metrics in terms of the familiar material parameters. I have shown that these pseudo-Finslerian optical metrics reduce to a pseudo-Riemannian optical metric for media obeying a curved space-time generalization of the Plebanski conditions, or for isotropic media. For such media in curved

space-times I have obtained an explicit formula for the optical metric, Eq. (153), and have shown that Gordon's optical metric is recovered exactly for isotropic media.

The formulation pursued here is particularly suited to transformation optics [16], and it is anticipated that the structure studied here will provide a framework for expanding the transformation optics theory, for example to nonlinear, dispersive, and lossy media. Furthermore, the ability to conduct ray tracing and study the kinematics of congruences through media in curved space-times could have application not only in transformation optics and dielectric

analog space-times but also in astrophysical settings such as light (or radio wave) propagation through accretion disks around massive compact objects, or in dark matter, e.g. MACHO, surveys.

ACKNOWLEDGMENTS

I thank Mohsen Fathi, Jonathan Gratus, Paul Kinsler, and Martin McCall for useful discussions. This work was supported in part by the Royal Society of New Zealand Marsden Fund Fast Start Grant No. UOO1219.

-
- [1] J.D. Jackson, *Classical Electrodynamics* 3rd ed. (Wiley, New York, 1998).
 - [2] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (W.H. Freeman and Co., San Francisco, 1973).
 - [3] J. Baez and J. P. Muniain, *Gauge Fields, Knots and Gravity* (World Scientific, Singapore, 1994).
 - [4] J. B. Pendry, Negative Refraction Makes a Perfect Lens, *Phys. Rev. Lett.* **85**, 3966 (2000).
 - [5] J. B. Pendry, D. Schurig, and D. R. Smith, Controlling electromagnetic fields, *Science* **312**, 1780 (2006).
 - [6] U. Leonhardt, Optical conformal mapping, *Science* **312**, 1777 (2006).
 - [7] U. Leonhardt and T. G. Philbin, General relativity in electrical engineering, *New J. Phys.* **8**, 247 (2006).
 - [8] M. W. McCall, A. Favaro, P. Kinsler, and A. Boardman, A spacetime cloak, or a history editor, *J. Opt.* **13**, 024003 (2011).
 - [9] D. R. Smith and N. Kroll, Negative Refractive Index in Left-Handed Materials, *Phys. Rev. Lett.* **85**, 2933 (2000).
 - [10] D. R. Smith, W. J. Padilla, D. C. Vier, S. C. Nemat-Nasser, and S. Schultz, Composite Medium with Simultaneously Negative Permeability and Permittivity, *Phys. Rev. Lett.* **84**, 4184 (2000).
 - [11] D. Schurig, J. J. Mock, B. J. Justice, S. A. Cummer, J. B. Pendry, A. F. Starr, and D. R. Smith, Metamaterial electromagnetic cloak at microwave frequencies, *Science* **314**, 977 (2006).
 - [12] F. L. Teixeira, Differential form approach to the analysis of electromagnetic cloaking and masking, *Microwave Opt. Technol. Lett.* **49**, 2051 (2007).
 - [13] L. Bergamin, Generalized transformation optics from triple spacetime metamaterials, *Phys. Rev. A* **78**, 043825 (2008).
 - [14] U. Leonhardt and T. G. Philbin, *Transformation Optics and the Geometry of Light* (Elsevier, New York, 2009), pp. 69–152.
 - [15] L. Bergamin, P. Alitalo, and S. A. Tretyakov, Nonlinear transformation optics and engineering of the Kerr effect, *Phys. Rev. B* **84**, 205103 (2011).
 - [16] R. T. Thompson, S. A. Cummer, and J. Frauendiener, A completely covariant approach to transformation optics, *J. Opt.* **13**, 024008 (2011).
 - [17] R. T. Thompson, S. A. Cummer, and J. Frauendiener, Generalized transformation optics of linear materials, *J. Opt.* **13**, 055105 (2011).
 - [18] O. Paul and M. Rahm, Covariant description of transformation optics in nonlinear media, *Opt. Express* **20**, 8982 (2012).
 - [19] R. T. Thompson and M. Fathi, Shrinking cloaks in expanding space-times: The role of coordinates and the meaning of transformations in transformation optics, *Phys. Rev. A* **92**, 013834 (2015).
 - [20] J. Gratus, P. Kinsler, M. W. McCall, and R. T. Thompson, On spacetime transformation optics: Temporal and spatial dispersion, *New J. Phys.* **18**, 123010 (2016).
 - [21] W. G. Unruh, Experimental Black-Hole Evaporation?, *Phys. Rev. Lett.* **46**, 1351 (1981).
 - [22] R. Schützhold and W. G. Unruh, Gravity wave analogues of black holes, *Phys. Rev. D* **66**, 044019 (2002).
 - [23] S. Weinfurter, E. W. Tedford, M. C. J. Penrice, W. G. Unruh, and G. A. Lawrence, Measurement of Stimulated Hawking Emission in an Analogue System, *Phys. Rev. Lett.* **106**, 021302 (2011).
 - [24] C. Barceló, S. Liberati, and M. Visser, Analogue gravity, *Living Rev. Relativity* **14**, 3 (2011).
 - [25] F. Felice, On the gravitational field acting as an optical medium, *Gen. Relativ. Gravit.* **2**, 347 (1971).
 - [26] B. Reznik, Origin of the thermal radiation in a solid-state analogue of a black hole, *Phys. Rev. D* **62**, 044044 (2000).
 - [27] R. Schützhold, G. Plunien, and G. Soff, Dielectric Black Hole Analogs, *Phys. Rev. Lett.* **88**, 061101 (2002).
 - [28] A. Greenleaf, Y. Kurylev, M. Lassas, and G. Uhlmann, Electromagnetic Wormholes and Virtual Magnetic Monopoles from Metamaterials, *Phys. Rev. Lett.* **99**, 183901 (2007).
 - [29] T. G. Philbin, C. Kuklewicz, S. Robertson, S. Hill, F. Knig, and U. Leonhardt, Fiber-optical analog of the event horizon, *Science* **319**, 1367 (2008).
 - [30] E. E. Narimanov and A. V. Kildishev, Optical black hole: Broadband omnidirectional light absorber, *Appl. Phys. Lett.* **95**, 041106 (2009).

- [31] R. T. Thompson and J. Frauendiener, Dielectric analog space-times, *Phys. Rev. D* **82**, 124021 (2010).
- [32] I. I. Smolyaninov, Metamaterial multiverse, *J. Opt.* **13**, 024004 (2011).
- [33] D. Bini, A. Geralico, and M. Haney, Refraction index analysis of light propagation in a colliding gravitational wave spacetime, *Gen. Relativ. Gravit.* **46**, 1644 (2014).
- [34] I. Fernández-Núñez and O. Bulashenko, Anisotropic metamaterial as an analogue of a black hole, *Phys. Lett. A* **380**, 1 (2016).
- [35] M. Fathi and R. T. Thompson, Cartographic distortions make dielectric spacetime analog models imperfect mimickers, *Phys. Rev. D* **93**, 124026 (2016).
- [36] F. W. Hehl, Y. N. Obukhov, and G. F. Rubilar, Light propagation in generally covariant electrodynamics and the Fresnel equation, *Int. J. Mod. Phys. A* **17**, 2695 (2002).
- [37] F. W. Hehl and Y. N. Obukhov, *Foundations of Classical Electrodynamics: Charge, Flux, and Metric* (Birkhauser, Boston, 2003).
- [38] V. A. Kostelecký and S. Samuel, Spontaneous breaking of Lorentz symmetry in string theory, *Phys. Rev. D* **39**, 683 (1989).
- [39] F. Girelli, S. Liberati, and L. Sindoni, Planck-scale modified dispersion relations and Finsler geometry, *Phys. Rev. D* **75**, 064015 (2007).
- [40] V. A. Kostelecký, Riemann-Finsler geometry and Lorentz-violating kinematics, *Phys. Lett. B* **701**, 137 (2011).
- [41] V. Perlick, *Ray Optics, Fermat's Principle, and Applications to General Relativity* (Springer, Berlin, 2000).
- [42] C. Lämmerzahl and F. W. Hehl, Riemannian light cone from vanishing birefringence in premetric vacuum electrodynamics, *Phys. Rev. D* **70**, 105022 (2004).
- [43] J. Skákala and M. Visser, Birefringence in pseudo-Finsler spacetimes, *J. Phys. Conf. Ser.* **189**, 012037 (2009).
- [44] S.-S. Chern, Finsler geometry is just Riemannian geometry without the quadratic restriction, *Not. Amer. Math. Soc.* **43**, 959 (1996).
- [45] M. F. Dahl, Licentiate thesis, Helsinki University of Technology, 2005.
- [46] R. T. Thompson and M. Fathi, Covariant kinematics of light in media and a generalized Raychaudhuri equation, *Phys. Rev. D* **96**, 105006 (2017).
- [47] H. Minkowski, Die grundgleichungen für die elektromagnetischen vorgänge in bewegten körpern, *Math. Ann.* **68**, 472 (1910).
- [48] W. Gordon, Zur lichtfortpflanzung nach der relativitätstheorie, *Ann. Phys. (Berlin)* **377**, 421 (1923).
- [49] I. E. Tamm, The electrodynamics of anisotropic media in the special theory of relativity, *Zh. Rus. Fiz.-Khim. Obshchestva, Otd. Fiz* **56**, 248 (1924).
- [50] N. L. Balazs, The propagation of light rays in moving media, *J. Opt. Soc. Am.* **45**, 63 (1955).
- [51] P. M. Quan, Le principe de Fermat en reletivité générale, in *Les théories relativistes de la gravitation, Royaumont 21-27 June 1959*, edited by M. A. Tonnelat and M. A. Lichnerowics (Colloques internationaux du centre national de la recherche scientifique, Paris, 1962).
- [52] J. Plebanski, Electromagnetic waves in gravitational fields, *Phys. Rev.* **118**, 1396 (1960).
- [53] E. J. Post, *Formal Structure of Electromagnetics* (North-Holland Publishing Co., Amsterdam, 1962).
- [54] J. Ehlers, Generalized electromagnetic null fields, and geometrical optics, in *Perspectives in Geometry and Relativity, essays in honor of Václav Hlavatý*, edited by B. Hoffmann (Indiana University Press, Bloomington, 1966), pp. 127.
- [55] M. Novello and J. M. Salim, Effective electromagnetic geometry, *Phys. Rev. D* **63**, 083511 (2001).
- [56] M. Novello, V. A. De Lorenci, J. M. Salim, and R. Klippert, Geometrical aspects of light propagation in nonlinear electrodynamics, *Phys. Rev. D* **61**, 045001 (2000).
- [57] M. Novello and E. Bittencourt, Gordon metric revisited, *Phys. Rev. D* **86**, 124024 (2012).
- [58] M. Visser, C. Barceló, and S. Liberati, Bi-refringence versus bi-metricity, in *Inquiring the Universe: Essays to Celebrate Professor Mario Novello Jubilee*, edited by J. M. Salim, S. E. P. Bergliffa, L. A. Oliveira, and V. A. De Lorenci (Frontier Group, Rio de Janeiro, 2003).
- [59] A. B. Balakin and W. Zimdahl, Optical metrics and birefringence of anisotropic media, *Gen. Relativ. Gravit.* **37**, 1731 (2005).
- [60] Y. Itin, On light propagation in premetric electrodynamics: The covariant dispersion relation, *J. Phys. A* **42**, 475402 (2009).
- [61] S. Schuster and M. Visser, Effective metrics and a fully covariant description of constitutive tensors in electrodynamics, *Phys. Rev. D* **96**, 124019 (2017).
- [62] R. Loudon, *The Quantum Theory of Light* (Oxford University Press, New York, 2000).
- [63] S. R. De Groot and L. G. Suttrop, Covariant derivation of the Maxwell equations: Multipole expansion of the polarization tensor to all orders, *Physica (Amsterdam)* **31**, 1713 (1965).
- [64] E. U. Condon and F. Seitz, Lorentz double refraction in the regular system, *J. Opt. Soc. Am.* **22**, 393 (1932).
- [65] R. A. Satten, Effects of atomic quadrupole moments upon the index of refraction, *J. Chem. Phys.* **26**, 766 (1957).
- [66] O. L. de Lange and R. E. Raab, Completion of multipole theory for the electromagnetic response fields D and H , *Proc. R. Soc. A* **459**, 1325 (2003).
- [67] L. W. Tu, *Differential Geometry* (Springer-Verlag, Berlin, 2017).
- [68] L. W. Tu, *An Introduction to Manifolds* (Springer-Verlag, Berlin, 2008).
- [69] R. T. Thompson and S. A. Cummer, Transformation optics, in *Advances in Imaging and Electron Physics*, edited by P. W. Hawkes (Elsevier, New York, 2012), Vol. 171, pp. 195–295.
- [70] Y. Itin, Wave propagation in axion electrodynamics, *Gen. Relativ. Gravit.* **40**, 1219 (2008).
- [71] Y. Itin, Skewon no-go theorem, *Phys. Rev. D* **88**, 107502 (2013).
- [72] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon Press, Oxford, 1960), translated from Russian by J. B. Sykes and J. S. Bell.

- [73] M. Born and E. Wolf, *Principles of Optics* (Cambridge University Press, Cambridge, 1999), p. 986.
- [74] T. Frankel, *The Geometry of Physics: An Introduction* 2nd ed. (Cambridge University Press, Cambridge, 2004).
- [75] A. S. Eddington, *Space, Time, and Gravitation* (Cambridge University Press, Cambridge, 1920).
- [76] S. A. Cummer and R. T. Thompson, Frequency conversion by exploiting time in transformation optics, *J. Opt.* **13**, 024007 (2011).