Classical and quantum aspects of electric-magnetic duality rotations in curved spacetimes

Ivan Agullo,[1,*] Adrian del Rio,[2,3,†] and Jose Navarro-Salas[2,‡]

[1]*Department of Physics and Astronomy, Louisiana State University, Baton Rouge, Louisiana 70803-4001, USA*

[2]*Departamento de Fisica Teorica, IFIC, Centro Mixto, Universitat de Valencia-CSIC, Valencia 46100, Spain*

[3]*Centro de Astrofísica e Gravitação-CENTRA, Departamento de Física, Instituto Superior Técnico-IST, Universidade de Lisboa, Lisboa 1049, Portugal*

(Received 30 October 2018; published 4 December 2018)

It is well known that the source-free Maxwell equations are invariant under electric-magnetic duality rotations, $\mathrm{F} \to \mathrm{F}\cos\theta + {}^\star\mathrm{F}\sin\theta$. These transformations are indeed a symmetry of the theory in the Noether sense. The associated constant of motion is the difference in the intensity between self-dual and anti-self-dual components of the electromagnetic field or, equivalently, the difference between the right and left circularly polarized components. This conservation law holds even if the electromagnetic field interacts with an arbitrary classical gravitational background. After reexamining these results, we discuss whether this symmetry is maintained when the electromagnetic field is quantized. The answer is in the affirmative in the absence of gravity but not necessarily otherwise. As a consequence, the net polarization of the quantum electromagnetic field fails to be conserved in curved spacetimes. This is a quantum effect, and it can be understood as the generalization of the fermion chiral anomaly to fields of spin one.

DOI: 10.1103/PhysRevD.98.125001

I. INTRODUCTION

Symmetries play an important role in many areas of science. They are widely considered as guiding principles for constructing physical theories, and their connection with conservation laws found by Noether one century ago [1] is a cornerstone of modern physics. An interesting example is given by Maxwell's theory of electrodynamics, whose invariance under Poincaré transformations leads to conservation of energy, linear, and angular momentum. (The invariance extends, in fact, to the full conformal group.) The theory is also invariant under gauge transformations when the electromagnetic potential is introduced, and when it is coupled to matter fields, the symmetry is related to the conservation of electric charge. Furthermore, in the absence of charges and currents, this theory enjoys a peculiar symmetry (in four spacetime dimensions). It is a simple exercise to check that Maxwell's equations, and also the stress-energy tensor, are invariant under the "exchange" of the electric and magnetic fields $\vec{E} \to \vec{B}$, $\vec{B} \to -\vec{E}$, as first noticed after the introduction of Maxwell's equations. This discrete $\mathbb{Z}_2$ operation is commonly known as a duality transformation. But the invariance of Maxwell's equations extends to $SO(2)$ rotations $\vec{E} \to \vec{E}\cos\theta + \vec{B}\sin\theta$, $\vec{B} \to \vec{B}\cos\theta - \vec{E}\sin\theta$, of which the duality transformation is just the particular case with $\theta = \pi/2$. Although apparently innocuous, this *continuous* transformation has revealed, in more recent times, interesting consequences.

In the mid-1960s, Calkin pointed out that these transformations leave Maxwell's action invariant, and he identified the associated conserved charge as the difference between the intensity of the right- and left-handed circularly polarized components of the electromagnetic field [2]. This conservation law was studied in more detail by Deser and Teitelboim in [3,4], and proved to remain true in curved spacetimes. This quantity is sometimes known as the optical helicity [5], and it also agrees with the V-Stokes parameter. Henceforth, besides conservation of energy and momentum, the polarization of electromagnetic radiation will also be a constant of motion as long as no electromagnetic sources are present, courtesy of the symmetry under electric-magnetic rotations.

A natural question now is to analyze whether this symmetry continues to hold in *quantum* electrodynamics. If j^μ_D is the Noether current associated with electric-magnetic rotations, this task reduces to checking if the vacuum expectation value $\langle \nabla_\mu j^\mu_D \rangle$ vanishes. In contrast to the classical theory, this is a nontrivial calculation that involves appropriate renormalization of ultraviolet divergences. It is well known that quantum fluctuations produce

*agullo@lsu.edu
†adriandelrio@tecnico.ulisboa.pt
‡jnavarro@ific.uv.es

off-shell contributions to physical quantities that might spoil classical symmetries. When this occurs, one says that there is a quantum anomaly in the theory.

Historically, the issue of quantum anomalies first appeared in the seminal works by Adler, Bell, and Jackiw, as a result of solving the pion decay puzzle [6,7]. They found that the chiral symmetry of the action of a massless Dirac field breaks down at the quantum level when the fermionic field interacts with an electromagnetic background. Namely, they obtained the celebrated chiral or axial anomaly $\langle \nabla_\mu j_A^\mu \rangle = -\frac{\hbar q^2}{8\pi^2} \mathrm{F}_{\mu\nu} {}^\star\mathrm{F}^{\mu\nu}$, where j_A^μ is the fermionic chiral current, $\mathrm{F}_{\mu\nu}$ the field strength of the background electromagnetic field, ${}^\star\mathrm{F}_{\mu\nu}$ its dual, and q the charge of the fermion. Later, a similar anomaly was found when the massless Dirac field is immersed in a classical gravitational background [8–10], $\langle \nabla_\mu j_A^\mu \rangle = \frac{\hbar}{192\pi^2} R_{\mu\nu\alpha\beta} {}^\star R^{\mu\nu\alpha\beta}$, where $R_{\mu\nu\alpha\beta}$ is the Riemann tensor. These discoveries led to an outbreak of interest in anomalies both in quantum field theory and mathematical physics, leading to further examples and a connection with the well-known index theorems in geometric analysis [11–13]. The existence of anomalies has important physical implications. Besides the prediction of the neutral pion decay rate to two photons, these anomalies have applications in studies of the matter-antimatter asymmetry of the Universe, the U(1) and strong CP problems in QCD, and provide a deeper understanding of the Standard Model via anomaly cancelation [14]. These cancelations have played a major role in string theories and supergravity too (for a detailed account see, for instance, Ref. [12] and references therein). A decade after the discovery of the chiral anomaly, the nature of quantum anomalies was further clarified by Fujikawa, using the language of path integrals [15,16]. He found that the existence of anomalies can also be understood as the failure of the measure of the path integral to respect the symmetries of the action. Fujikawa's arguments provided an alternative and elegant way of computing anomalies.

In this paper, we prove that electric-magnetic rotations are also anomalous, provided the electromagnetic field propagates in a sufficiently nontrivial spacetime. To meet our goal, we write Maxwell's theory in terms of self-dual and anti-self-dual variables, which will make the structure of the theory significantly more transparent, particularly in the absence of charges and currents. In fact, in these variables duality rotations look mathematically—and physically—similar to chiral transformations of massless spin-1/2 Dirac fields, and in this sense, our result can be understood as the spin-1 generalization of the fermionic chiral anomaly. We derive our result by using two complementary methods, namely, by directly computing $\langle \nabla_\mu j_D^\mu \rangle$ using the method of heat-kernel renormalization and by Fujikawa's path-integral approach.

This paper is organized as follows. In Sec. II we review the analysis of the classical duality symmetry in source-free electrodynamics and derive the associated Noether charge and current, both in the Lagrangian and Hamiltonian frameworks. In Sec. III we introduce self-dual and anti-self-dual variables and emphasize their advantages in the source-free theory. We show how Maxwell's equations can be conveniently written as first-order equations, either for fields or potentials, that are analogous to Weyl's equations for spin-1/2 fields. In Sec. IV, we derive a first-order action for Maxwell electrodynamics in self-dual and anti-self-dual variables, which makes the theory formally analogous to Dirac's theory of massless fermions. Section V discusses the quantum theory and the derivation of the quantum electromagnetic duality anomaly by using the two methods mentioned above. We finally give some concluding remarks in Sec. VI. To simplify the main text of the article, we have moved many of the mathematical details and calculations to Appendices A–G.

A shorter version of this work appeared in [17]. Here we provide further details and alternative avenues for arriving at the final result, and correct some minor errors that translate into a different numerical factor in the result for $\langle \nabla_\mu j_D^\mu \rangle$.

We follow the convention $\epsilon^{0123} = 1/\sqrt{-g}$ and metric signature $(+,-,-,-)$. More specifically, we follow the $(-,-,-)$ convention of [18]. We restrict ourselves to four-dimensional spacetimes and assume the Levi-Civita connection. We use Greek indices $\mu, \nu, \alpha, \ldots$ for tensors in curved spacetimes, while Latin indices $a, b, c, \ldots$ are used for tensors in Minkowski spacetime. Indices $I, J, K, \ldots$ or $\dot{I}, \dot{J}, \dot{K}, \ldots$ refer to tensors in an internal space associated with the spin-1 complex Lorentz representations. Unless otherwise stated, we assume all fields to be smooth and to have standard fall-off conditions at infinity. We use units for which $c = 1$.

II. CLASSICAL THEORY AND ELECTRIC-MAGNETIC ROTATIONS

A. Lagrangian formalism

In this paper we focus on free Maxwell's theory, i.e., electromagnetic fields in the absence of electric charges and currents, formulated on a globally hyperbolic spacetime $(M, g_{\mu\nu})$ with metric tensor $g_{\mu\nu}$. The classical theory is described by the action

$$S[A_\mu] = -\frac{1}{4} \int \mathrm{d}^4x \sqrt{-g} \mathrm{F}^{\mu\nu} \mathrm{F}_{\mu\nu}, \tag{2.1}$$

where F is a closed two-form ($\mathrm{dF} = 0$) defined in terms of its potential A as $\mathrm{F} = \mathrm{d}A$, or more explicitly, $\mathrm{F}_{\mu\nu} = \nabla_\mu A_\nu - \nabla_\nu A_\mu$. Maxwell's equations read $\Box A_\nu - \nabla^\mu \nabla_\nu A_\mu = 0$, where ∇ is the covariant derivative associated with $g_{\mu\nu}$ and $\Box \equiv g^{\mu\nu} \nabla_\mu \nabla_\nu$. When written in terms of the dual tensor ${}^\star\mathrm{F}$, these equations take the compact form $\mathrm{d}{}^\star\mathrm{F} = 0$ and, together with $\mathrm{dF} = 0$, make manifest that the field equations are invariant under electric-magnetic rotations

$$\mathrm{F} \to \mathrm{F}\cos\theta + {}^{\star}\mathrm{F}\sin\theta,$$
$$^{\star}\mathrm{F} \to {}^{\star}\mathrm{F}\cos\theta - \mathrm{F}\sin\theta. \tag{2.2}$$

For $\theta = \pi/2$ one has the more familiar duality transformation $\mathrm{F} \to {}^{\star}\mathrm{F}$ and ${}^{\star}\mathrm{F} \to -\mathrm{F}$. If this one-parameter family of transformations is a true symmetry of the action, then Noether's analysis must provide a conserved charge associated with it. We now analyze this problem. Our presentation simply rephrases, in a manifestly covariant way, the results of Ref. [3].

For the transformation (2.2) to be a symmetry, its infinitesimal version ($\delta\mathrm{F} = {}^{\star}\mathrm{F}\delta\theta$) must leave the action invariant or, equivalently, the Lagrangian density $\mathcal{L} = -1/4\sqrt{-g}\mathrm{F}_{\mu\nu}\mathrm{F}^{\mu\nu}$ must change by a total derivative, $\delta\mathcal{L} = \sqrt{-g}\nabla_\mu h^\mu$, for some current h^μ. This must be true even off shell, i.e., when F and *F do not satisfy the equations of motion. In analyzing if this is the case, one faces two issues. On the one hand, since F is a closed two-form (i.e., $\mathrm{dF} = 0$), for the transformation $\delta\mathrm{F} = {}^{\star}\mathrm{F}\delta\theta$ to be consistent, *F must also be closed, but this amounts to saying that equations of motion hold. In other words, the transformation (2.2) can only be consistently defined on shell.[1] Second, since the usual configuration variables of Maxwell's action are the vector potential A rather than the field F, to apply Noether's techniques we first need to rewrite (2.2) in terms of A. A convenient strategy to deal with these two issues is to define a more general transformation, which will agree with electric-magnetic rotations only on shell, as follows:

$$\delta A_\mu = Z_\mu \delta\theta, \tag{2.3}$$

where Z_μ is implicitly defined by $\mathrm{d}Z \equiv {}^{\star}\mathrm{F} + \mathrm{G}$, and G is a two-form that is subject to the following conditions, but arbitrary otherwise:

(1) G vanishes only when A_μ satisfies the equations of motion, $\mathrm{G}|_{\text{on shell}} = 0$. This ensures that $\mathrm{d}Z = {}^{\star}\mathrm{F}$ on shell, and then (2.3) reduces to the usual electric-magnetic transformation.

(2) G is not closed, $\mathrm{dG} \neq 0$—unless the equations of motion hold. This guarantees that *F is not closed (off shell).

(3) G has zero magnetic part relative to an arbitrary observer, i.e., $n^\nu {}^{\star}\mathrm{G}_{\mu\nu} = 0$, where n^ν is a timelike vector field, and *G is the dual of G. (This condition is equivalent to saying that the electric field relative to the observer satisfies Gauss's law.)

Note that Z_μ is a nonlocal functional of A_μ. However, as discussed in [3], this is not an impediment to applying Noether's formalism.

Under the transformation (2.3), we obtain (see Appendix A for more details)

$$\delta\mathcal{L} = -\delta\theta\frac{\sqrt{-g}}{2}\nabla_\mu[A_\nu {}^{\star}\mathrm{F}^{\mu\nu} - Z_\nu(\mathrm{d}Z)^{\mu\nu}] \equiv \sqrt{-g}\nabla_\mu h^\mu, \tag{2.4}$$

confirming that electric-magnetic rotations are a symmetry of source-free Maxwell's theory. The conserved Noether current j_D^μ associated with this symmetry is

$$j_D^\mu = \frac{1}{\sqrt{-g}}\frac{\partial\mathcal{L}}{\partial\nabla_\mu A_\nu}\delta A_\nu - h^\mu$$
$$= \frac{1}{2}[A_\nu {}^{\star}\mathrm{F}^{\mu\nu} - Z_\nu \mathrm{F}^{\mu\nu} - Z_\nu {}^{\star}\mathrm{G}^{\mu\nu}] \tag{2.5}$$

(we have dropped $\delta\theta$ from the definition of j_D^μ). This current is gauge dependent. But this is not a problem either, as long as the associated conserved charge is gauge invariant, which is in fact the case. When evaluated on shell (i.e., when $\mathrm{G} = 0$, and therefore $\mathrm{d}Z = {}^{\star}\mathrm{F}$)

$$j_D^\mu|_{\text{on shell}} = \frac{1}{2}[A_\nu {}^{\star}\mathrm{F}^{\mu\nu} - Z_\nu \mathrm{F}^{\mu\nu}]. \tag{2.6}$$

Now, if we foliate the spacetime using a one-parameter family of Cauchy hypersurfaces Σ_t, the quantity

$$Q_D = \int_{\Sigma_t} \mathrm{d}\Sigma_\mu j_D^\mu = -\frac{1}{2}\int_{\Sigma_t} \mathrm{d}\Sigma_3 (A_\mu B^\mu - Z_\mu E^\mu) \tag{2.7}$$

is a conserved charge, in the sense that it is independent of the choice of "leaf" Σ_t. In this expression, $\mathrm{d}\Sigma_3$ is the volume element in Σ_t, and $E^\mu = n_\nu \mathrm{F}^{\mu\nu}$ and $B^\mu = n_\nu {}^{\star}\mathrm{F}^{\mu\nu}$ are the electric and magnetic parts, respectively, of the electromagnetic tensor field F relative to the foliation Σ_t. The same expression for Q_D is obtained if $j_D^\mu|_{\text{on shell}}$ is used in place of j_D^μ in (2.7); hence, the conserved charge is insensitive to the extension of the transformation done above by the introduction of G.

One can check, by explicit computation, that $\nabla_\mu j_D^\mu = -Z_\nu \nabla_\mu F^{\mu\nu}$, and therefore $\nabla_\mu j_D^\mu = 0$ when the equations of motion $\nabla_\mu F^{\mu\nu} = 0$ hold. In the quantum theory, however, off-shell contributions of quantum origin may spoil the conservation of the current. The calculation of the expectation value of $\nabla_\mu j_D^\mu$ using the formalism derived in this section is complicated since it would involve the operator Z_μ, which is a (nonlocal) functional of the configuration

[1] This "difficulty" appears only in the second order formalism. If one uses a first-order Lagrangian, or a Hamiltonian formulation, the usual electric-magnetic rotations can be implemented off shell. This point has been emphasized in [4] and will be made explicit in the next subsection and in Sec. IV.

variable A_μ.[2] This difficulty can be alleviated by working in phase space, where one can treat Z_μ and A_μ as independent fields. This motivates the Hamiltonian analysis of the next subsection and the use of a first-order formalism in the rest of the paper. In particular, in Secs. III and IV we rederive j_D^μ in a first-order Lagrangian formalism using self-dual and anti-self-dual variables. This will make the derivation significantly more transparent. The physical interpretation of Q_D will also become more clear, and we postpone the discussion until then.

B. Hamiltonian formalism

The Hamiltonian formalism provides a complementary approach to the study of the electric-magnetic symmetry, and in this subsection we briefly summarize the derivation of Q_D following this framework. We restrict ourselves to Minkowski spacetime since the generalization to curved geometries using the standard vector potential and electric field as canonical coordinates becomes cumbersome.

Given an inertial frame in Minkowski spacetime, Maxwell's Lagrangian (2.1) takes the form

$$L = \int_{\mathbb{R}^3} \mathrm{d}^3x \mathcal{L} = \int_{\mathbb{R}^3} \mathrm{d}^3x \frac{1}{2}\left[(\dot{\vec{A}} - \vec{\nabla}A_0)^2 - (\vec{\nabla}\times\vec{A})^2\right], \quad (2.8)$$

where $\vec{\nabla}$ is the usual three-dimensional derivative operator. Our conventions are $\vec{A} \equiv (A_1, A_2, A_3)$, $\vec{E} \equiv (E_1, E_2, E_3)$, $E_i \equiv F_{i0}$, and $\vec{E}^2 \equiv E_1^2 + E_2^2 + E_3^2$. From this, we see that the canonically conjugate variable of $\vec{A}$ is the electric field $\frac{\delta L}{\delta \dot{A}_i} = E^i$, and the conjugate variable of A_0 vanishes since the Lagrangian does not involve $\dot{A}_0$. Then, A_0 is a Lagrange multiplier, and from its equation of motion, one obtains a constraint, the familiar Gauss's law $\vec{\nabla}\cdot\vec{E} = 0$. Then, the canonical phase space is made of pairs $(\vec{A}(\vec{x}), \vec{E}(\vec{x}))$, with a symplectic or Poisson structure given by $\{A_i(\vec{x}), E^j(\vec{x}')\} = \delta_i^j \delta^{(3)}(\vec{x} - \vec{x}')$. A Legendre transformation produces the Hamiltonian

$$H = \int_{\mathbb{R}^3} \mathrm{d}^3x \frac{1}{2}[\vec{E}^2 + (\vec{\nabla}\times\vec{A})^2 - A_0(\vec{\nabla}\cdot\vec{E})], \quad (2.9)$$

where we have disregarded a boundary term. In Dirac's terminology, $\vec{\nabla}\cdot\vec{E} = 0$ is a first-class constraint, and it tells us that there is gauge freedom in the theory, given precisely by the canonical transformations generated by $\vec{\nabla}\cdot\vec{E}$.

Hamilton's equations read

$$\dot{\vec{A}} = \{\vec{A}, H\} = -\vec{E} - \vec{\nabla}A_0$$
$$\dot{\vec{E}} = \{\vec{E}, H\} = \vec{\nabla}\times(\vec{\nabla}\times\vec{A}), \quad (2.10)$$

where $A_0(\vec{x})$ is now interpreted as an arbitrary function without dynamics, and the term proportional to it in the expression for $\dot{\vec{A}}$ corresponds precisely to the gauge flow. These six equations, together with the Gauss constraint, are equivalent to standard Maxwell's equations (once we define $\vec{B} \equiv \vec{\nabla}\times\vec{A}$).

Electric-magnetic rotations in phase space are given by

$$\delta\vec{E} = (\vec{\nabla}\times\vec{A}) \equiv \vec{B}, \qquad \delta\vec{A} = -(\vec{\nabla}\times)^{-1}\vec{E} \equiv \vec{Z}, \quad (2.11)$$

where $(\vec{\nabla}\times)^{-1}$ is the inverse of the curl; when acting on traverse fields—such as $\vec{E}$—it can be easily computed by using the relation $(\vec{\nabla}\times)^{-1} = -\nabla^{-2}\vec{\nabla}\times$. The presence of the operator $(\vec{\nabla}\times)^{-1}$ in (2.11) makes it evident that we are dealing with a transformation that is nonlocal in space.

Now, the generator of the transformation (2.11) can be easily obtained by computing the symplectic product of $(\vec{A}, \vec{E})$ and $(\delta\vec{A}, \delta\vec{E})$:

$$\begin{aligned} Q_D &= \Omega\left[(\vec{A}, \vec{E}), (\delta\vec{A}, \delta\vec{E})\right] \\ &= -\frac{1}{2}\int_{\mathbb{R}^3} d^3x \left[\vec{E}\cdot\delta\vec{A} - \vec{A}\cdot\delta\vec{E}\right] \\ &= \frac{1}{2}\int_{\mathbb{R}^3} d^3x \left[\vec{A}\cdot\vec{B} - \vec{Z}\cdot\vec{E}\right]. \end{aligned} \quad (2.12)$$

Here, Q_D is independent of A_0, and by integrating by parts, it is easy to show that only the transverse part of $\vec{A}$ and $\vec{Z}$ contributes to Q_D; hence, it is gauge invariant. It is also straightforward to check that Q_D is indeed the correct generator since $\delta\vec{A} = \{\vec{A}, Q_D\}$ and $\delta\vec{E} = \{\vec{E}, Q_D\}$ reproduce expressions (2.11). To finish, one can now check that $\dot{Q}_D = \{Q_D, H\} = 0$. Therefore, Q_D is a constant of motion. This implies that the canonical transformation generated by Q_D is a symmetry of the theory.

III. ELECTRODYNAMICS IN TERMS OF SELF-DUAL AND ANTI-SELF-DUAL VARIABLES

Many aspects of Maxwell's theory in the absence of charges and currents become more transparent when self-dual and anti-self-dual variables are used (see, e.g., Refs. [22–24]). Some of the advantages of these variables

[2] The first term in (2.5) and its quantum aspects have previously been discussed in [19] (see also [20]). However, this term by itself is not conserved classically (something that cannot be fixed by any gauge transformation), and in fact, its associated "charge" does not generate duality rotations in phase space (see Sec. II B). Therefore, the first term in (2.5) alone is not associated with the symmetry under electric-magnetic rotations. The fact that its vacuum expectation value does not vanish, although this is of physical interest in its own right, does not really prove the existence of an anomaly, as claimed in [19]. Other vacuum expectation values of physical interest have been computed in [21].

are well known and, in particular, they are commonly used in the spinorial formulation of electrodynamics [25]. For the sake of clarity, we introduce these variables first in Minkowski spacetime and extend the formalism later to curved geometries.

A. Minkowski spacetime

The self-dual and anti-self-dual components of the electromagnetic field are defined as $\vec{H}_\pm \equiv \frac{1}{\sqrt{2}}(\vec{E} \pm i\vec{B})$. We now enumerate the properties and interesting aspects of these complex variables.

(1) *Electric-magnetic rotations*.—The transformation rule of the electric and magnetic fields under electric-magnetic rotations,

$$\vec{E} \to \vec{E}\cos\theta + \vec{B}\sin\theta,$$
$$\vec{B} \to \vec{B}\cos\theta - \vec{E}\sin\theta, \tag{3.1}$$

translates to

$$\vec{H}_\pm = \frac{1}{\sqrt{2}}(\vec{E} \pm i\vec{B}) \to e^{\mp i\theta}\vec{H}_\pm. \tag{3.2}$$

An ordinary duality transformation $\vec{E} \to \vec{B}$, $\vec{B} \to -\vec{E}$ corresponds to $\theta = \pi/2$. Then, this operation produces[3] $i\vec{H}_\pm \to \pm\vec{H}_\pm$. It is for this reason that $\vec{H}_+$ and $\vec{H}_-$ are called the self-dual and anti-self-dual components of the electromagnetic field, respectively.

(2) *Lorentz transformations*.—The components of $\vec{E}$ and $\vec{B}$ mix with each other under a Lorentz transformation. For instance, under a boost of velocity v in the x direction,

$$\vec{E} = (E_x, E_y, E_z) \to [E_x, \gamma(E_y - vB_z), \gamma(E_z + vB_y)],$$
$$\vec{B} = (B_x, B_y, B_z) \to [B_x, \gamma(B_y + vE_z), \gamma(B_z - vE_y)], \tag{3.3}$$

where $\gamma = 1/\sqrt{1-v^2}$. This transformation does not correspond to any irreducible representation of the Lorentz group. However, when $\vec{E}$ and $\vec{B}$ are combined into $\vec{H}_\pm$, it is easy to see that the components of $\vec{H}_+$ and $\vec{H}_-$ no longer mix,

$$\vec{H}_\pm = (H_\pm^x, H_\pm^y, H_\pm^z)$$
$$\to [H_\pm^x, \gamma(H_\pm^y \pm ivH_\pm^z), \gamma(H_\pm^z \mp ivH_\pm^y)]. \tag{3.4}$$

These are the transformation rules associated with the two irreducible representations of the Lorentz group for fields of spin $s = 1$. They are the so-called (0,1) representation for $\vec{H}_+$ and the (1,0) one for $\vec{H}_-$. More generally, for any element of the restricted Lorentz group $SO^+(1,3)$ (rotations + boots), the infinitesimal transformation reads

$$H_\pm^J \to [D(\epsilon_{ab})]_{IJ} H_\pm^J = \left[\delta_{IJ} - \frac{1}{2}\epsilon_{ab}{}^{\pm}\Sigma^{ab}{}_{IJ}\right] H_\pm^J \tag{3.5}$$

(uppercase Latin indices $I, J, K, \ldots$ take values from 1 to 3), where δ_{IJ} is the Kronecker delta, ${}^{\pm}\Sigma^{ab}{}_{IJ}$ are the generators of the (0,1) and (1,0) representations,[4] and the antisymmetric matrix $\epsilon_{ab} = \epsilon_{[ab]}$ contains the parameters of the transformation. The use of self-dual and anti-self-dual fields $\vec{H}_\pm$ makes it more transparent that electrodynamics describes massless fields of spin $s = 1$.

(3) *Maxwell's equations*.—The equations of motions for $\vec{E}$ and $\vec{B}$,

$$\vec{\nabla} \cdot \vec{E} = 0, \qquad \vec{\nabla} \cdot \vec{B} = 0,$$
$$\vec{\nabla} \times \vec{E} = -\partial_t \vec{B}, \qquad \vec{\nabla} \times \vec{B} = \partial_t \vec{E}, \tag{3.6}$$

when written in terms of $\vec{H}_\pm$, take the form

$$\vec{\nabla} \cdot \vec{H}_\pm = 0, \qquad \vec{\nabla} \times \vec{H}_\pm = \pm i\partial_t \vec{H}_\pm. \tag{3.7}$$

Notice that, in contrast to $\vec{E}$ and $\vec{B}$, the self-dual and anti-self-dual fields are not coupled by the dynamics. The equations for $\vec{H}_-$ and $\vec{H}_+$ are related by complex conjugation.

Equations (3.7) are linear, and therefore the space of solutions has structure of vector space. It is spanned by positive- and negative-frequency solutions:

$$\vec{H}_\pm(t, \vec{x}) = \int_{\mathbb{R}^3} \frac{d^3k}{(2\pi)^3} [h_\pm(\vec{k}) e^{-i(kt - \vec{k}\cdot\vec{x})}$$
$$+ \bar{h}_\mp(\vec{k}) e^{i(kt - \vec{k}\cdot\vec{x})}] \hat{e}_\pm(\vec{k}), \tag{3.8}$$

where $k = |\vec{k}|$ and $h_\pm(\vec{k})$ are complex numbers that indicate the amplitude of the positive and negative frequency components of a particular solution (the "bar" denotes complex conjugation). The polarization vectors are given by $\hat{e}_\pm = \frac{1}{\sqrt{2}}(\hat{e}_1 \pm i\hat{e}_2)$, where $\hat{e}_1(\vec{k})$ and $\hat{e}_2(\vec{k})$ are two unit vectors that, together with $\hat{k}$, form an orthonormal triad of spacelike vectors, with orientation $\hat{e}_1 \times \hat{e}_2 = \hat{k}$.

[3] It is common to add the imaginary unit i because, in that way, this operation has real eigenvalues, and it can be represented by a self-adjoint operator in the quantum theory.

[4] They satisfy the algebra $[{}^{\pm}\Sigma^{ab}, {}^{\pm}\Sigma^{cd}] = (\eta^{ac\,\pm}\Sigma^{bd} - \eta^{ad\,\pm}\Sigma^{bc} + \eta^{bd\,\pm}\Sigma^{ac} - \eta^{bc\,\pm}\Sigma^{ad})$.

The explicit form (3.8) of a generic solution helps us to understand the relation between self-duality or anti-self-duality and helicity in Minkowski spacetime. By paying attention to the way the electric and magnetic parts (i.e., the real and imaginary parts of $\vec{H}_\pm$, respectively) rotate with respect to the direction of propagation $\hat{k}$ during the course of time, one finds the following relation:

(i) Positive-frequency Fourier modes $e^{-i(kt-\vec{k}\cdot\vec{x})}\hat{\epsilon}_\pm(\vec{k})$ have positive helicity (that corresponds to left-handed circular polarization) for self-dual fields and negative helicity for anti-self-dual fields.

(ii) For negative-frequency modes $e^{i(kt-\vec{k}\cdot\vec{x})}\hat{\epsilon}_\pm(\vec{k})$, the relation is inverse: They have negative helicity (right-handed circular polarization) for self-dual fields and positive helicity for anti-self-dual fields.

We see that duality and helicity are closely related concepts in Minkowski spacetime, although the relation is not trivial; one needs to distinguish between self-dual and anti-self-dual fields *and* positive and negative frequencies [26]. This is the analog of the familiar relation between chirality and helicity for massless spin-1/2 fermions. In this sense, duality is the chirality of photons.

Furthermore, in more general spacetimes where neither Fourier modes nor the notion of positive and negative frequency are useful, self-duality or anti-self-duality generalizes the concept of helicity, or handedness, of electromagnetic waves.

(4) *Self-dual and anti-self-dual potentials*.—The constraints $\vec{\nabla}\cdot\vec{H}_\pm = 0$ allow us to define the potentials $\vec{A}_\pm$ by

$$\vec{H}_\pm = \pm i\vec{\nabla}\times\vec{A}_\pm. \tag{3.9}$$

It is clear from this definition that the longitudinal part of $\vec{A}_\pm$ contains a gauge ambiguity consisting in adding the divergence of an arbitrary scalar function. Note also that no time derivatives have been involved in the definition of these potentials.

(5) *Maxwell's equations for potentials*.—Substituting (3.9) in the field equations (3.7) produces

$$\pm i\vec{\nabla}\times\vec{A}_\pm = -\partial_t\vec{A}_\pm + \vec{\nabla}A^0_\pm. \tag{3.10}$$

These equations by themselves are equivalent to Maxwell's equations. It may be surprising at first that Maxwell's theory can be written as first-order equations for potentials. This comes from the fact that in—and only in—the source-free theory, in addition to the standard potential $\vec{A}$ defined from $\vec{B} = \vec{\nabla}\times\vec{A}$, Gauss's law $\vec{\nabla}\cdot\vec{E} = 0$ allows us to define a second potential $\vec{Z}$, as $\vec{E}\equiv -\vec{\nabla}\times\vec{Z}$. Then, the first-order equations

$$\dot{\vec{A}} = \vec{\nabla}\times\vec{Z} + \vec{\nabla}A_0,$$
$$\dot{\vec{Z}} = -\vec{\nabla}\times\vec{A} + \vec{\nabla}Z_0 \tag{3.11}$$

are equivalent to Maxwell's equations (to see this, take the curl and use the relation between potentials and fields). Therefore, Maxwell's equations can be written as first-order equations for potentials at the expense of duplicating the number of potentials. The relation between the two sets of potentials is $A^\pm_a = \frac{1}{\sqrt{2}}(A_a \pm iZ_a)$.

(6) *Manifestly Lorentz-covariant equations*.—Equations (3.7) and (3.10) for fields and potentials can be rewritten in a more compact way as

$$\alpha_I^{ab}\partial_a H^I_+ = 0, \qquad \bar{\alpha}_I^{ab}\partial_a A^+_b = 0. \tag{3.12}$$

The equations for H_- and A_- are obtained by complex conjugation. In these expressions α_I^{ab} are three 4×4 matrices, for $I = 1, 2, 3$, and the bar over α_I^{ab} indicates complex conjugation. The components of these matrices in an inertial frame can be identified by comparing these equations with (3.7) and (3.10):

$$\alpha_1^{ab} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \end{pmatrix}, \quad \alpha_2^{ab} = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -i \\ 1 & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix},$$
$$\alpha_3^{ab} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \tag{3.13}$$

These matrices are antisymmetric ($\alpha_I^{ab} = \alpha_I^{[ab]}$), invariant under Lorentz transformations, and self-dual ($i^\star\alpha_I^{ab} = \alpha_I^{ab}$)—hence, $\bar{\alpha}_I^{ab}$ is anti-self-dual. As mentioned above, the equations for the potentials can be derived from the equations for the fields. The reverse is also true. Therefore, either set of equations completely describes the theory. Field equations similar to $\alpha_I^{ab}\partial_a H^I_+ = 0$ have been written before in [22,23]; our equations $\alpha_I^{ab}\partial_a H^I_+ = 0$ are also equivalent to Maxwell's equations in spinorial language [25].

(7) *Relation between $\vec{H}_\pm$ and the field strength* F_{ab}.—From the field strength F and its dual *F, we define the self-dual and anti-self-dual two-forms $\mathrm{F}_\pm = \frac{1}{\sqrt{2}}(\mathrm{F} \pm i^\star\mathrm{F})$, which satisfy $i^\star\mathrm{F}_\pm = \pm\mathrm{F}_\pm$. The relation between the field strength and $\vec{H}_\pm$ is then given by

$$\mathrm{F}^{ab}_+ = \alpha_I^{ab}H^I_+, \qquad \mathrm{F}^{ab}_- = \bar{\alpha}_{\bar{I}}^{ab}H^{\bar{I}}_-. \tag{3.14}$$

These relations imply that one can understand the three α_I^{ab} matrices as a basis for the three-dimensional complex vector space of self-dual tensors in Minkowski spacetime (see Appendix B). Then, H_+^I are simply the components of F_+^{ab} in this basis. Similarly, $\bar{\alpha}_I^{ab}$ provides a basis for anti-self-dual tensors.

On the other hand, by using the relations (3.14)—and the fact that the α_I-matrices are constant in spacetime, so they are transparent to derivatives—the field equations $\alpha_I^{ab}\partial_a H_+^I = 0$ and $\bar{\alpha}_I^{ab}\partial_a H_-^I = 0$ can be written as $\partial_a \mathrm{F}_+^{ab} = 0$ and $\partial_a \mathrm{F}_-^{ab} = 0$, which are equivalent Maxwell's equations in their more standard form.

(8) *Properties of α_I^{ab} matrices.*—Using the form of the α_I matrices (3.13), it is straightforward to check that they have the following properties:

(i) Anticommutation relations: $\{\alpha_I,\alpha_J\}\equiv\alpha^a{}_{bI}\alpha^{bc}{}_J+\alpha^a{}_{bJ}\alpha^{bc}{}_I=\delta_{IJ}\eta^{ac}$.

(ii) Commutation relations: $[\alpha_I,\alpha_J]\equiv\alpha^a{}_{bI}\alpha^{bc}{}_J-\alpha^a{}_{bJ}\alpha^{bc}{}_I={}^{+}\Sigma^{ac}{}_{IJ}$.

These properties can be thought of as the spin-1 analog of the familiar properties of the Pauli matrices $\sigma_i^{A\dot{A}}$.

To better understand these properties, and to generalize them to curved spacetimes (see next section), it is convenient to take a more geometric viewpoint and think about the field H_+^I as belonging to a complex, three-dimensional vector space V, which supports a (0,1) irreducible representation of the Lorentz group. This space is isomorphic to the space of self-dual tensors F_+ in Minkowski spacetime, and α_I^{ab} provides an isomorphism.

Furthermore, α_I^{ab} equips V with a product h_{IJ}, the image of the Minkowski metric[5] $h_{IJ}=\frac{1}{4}\eta_{ab}\eta_{cd}\alpha_I^{ac}\alpha_J^{bd}$, whose value turns out to be $h_{IJ}=-\delta_{IJ}$ in a Cartesian frame, and is obviously invariant under Lorentz transformations in V. This viewpoint makes clearer the analogy between the α_I^{ab} and the Pauli matrices $\sigma_i^{A\dot{A}}$ (recall that $\sigma_i^{A\dot{A}}$ provides an isometry between spatial vectors and spinors).

If $\vec{H}_+$ is an element of the complex vector space V, then $\vec{H}_-$ is an element of $\bar{V}$, the complex-conjugate space. Although naturally isomorphic, these two spaces are different, and from now on, we use dotted indices on elements of $\bar{V}\ni H_-^{\dot{I}}$. The properties of $\bar{\alpha}_{\dot{I}}^{ab}$ are obtained by complex conjugating the properties of α_I^{ab} written above. The anticommutation relations are identical. However, the conjugation changes the commutation relation to $[\bar{\alpha}_{\dot{I}},\bar{\alpha}_{\dot{J}}]={}^{-}\Sigma^{ab}{}_{\dot{I}\dot{J}}$, where now it is the generator of the (1,0) representation of the Lorentz group that enters in the equation. Appendix B contains further information about the properties of these tensors.

(9) *Second-order equations for the potentials A_a^+.*—We focus on A_a^+ since the derivation for A_a^- can be obtained from it by complex conjugation. The fastest way to obtain the familiar second-order differential equation for A_a^+ is to take the time derivative of (3.10), use commutativity between spatial and time derivatives, and then again (3.10) use to eliminate the first time derivative in favor of the curl. The result can then be written in covariant form as $\Box A_a^+ - \partial^b\partial_a A_b^+ = 0$.

Alternatively, we can use the following argument, which can be straightforwardly generalized to curved spacetimes. Notice that the equations of motion $\bar{\alpha}_{\dot{I}}^{ab}\partial_a A_b^+ = 0$ imply that the two-form $\partial_{[a}A_{b]}^+$ is self-dual. This is because, on the one hand, the antisymmetry of $\bar{\alpha}_{\dot{I}}^{ab}$ means that only the antisymmetric part of $\partial_a A_b^+$ contributes to the equations and, on the other, because contraction with $\bar{\alpha}_{\dot{I}}^{ab}$ extracts the anti-self-dual component of $\partial_{[a}A_{b]}^+$. Therefore, when the equations of motion hold, A_a^+ is the potential of a self-dual form, $\mathrm{F}_+ = \mathrm{d}A_+$. But if $\mathrm{d}A_+$ is self-dual, then the identity $\partial_{[a}\partial_b A_{c]}^+ = 0$ implies that $\partial^a\partial_{[a}A_{c]}^+ = 0$. These last equations are obviously equivalent to

$$\Box A_c^+ - \partial^a\partial_c A_a^+ = 0. \tag{3.15}$$

Therefore, the self-dual and anti-self-dual potentials $A_a^\pm$ satisfy the same second-order equations as the ordinary vector potential.

(10) *Conserved current and charge.*—In terms of self-dual and anti-self-dual variables, electric-magnetic rotations take the simple form

$$\begin{aligned} H_\pm^I(x) &\to e^{\mp i\theta}H_\pm^I(x),\\ A_a^\pm(x) &\to e^{\mp i\theta}A_a^\pm(x). \end{aligned} \tag{3.16}$$

And the on-shell current (2.6) takes the form

$$j_D^a|_{\text{on shell}} = -\frac{i}{2}[H_+^I\alpha_I^{ab}A_b^- - H_-^{\dot{I}}\bar{\alpha}_{\dot{I}}^{ab}A_b^+] \tag{3.17}$$

(note that this current is manifestly real). By using the form of the generic solution to the field equations (3.8), we find that the conserved charge

[5]In other words, given any two self-dual tensors ${}^{(1)}\mathrm{F}_+^{ab}$ and ${}^{(2)}\mathrm{F}_+^{ab}$, the isomorphism satisfies ${}^{(1)}\mathrm{F}_+^{ab\,(2)}\mathrm{F}_+^{cd}\eta_{ac}\eta_{bd} = {}^{(1)}H_+^I\,{}^{(2)}H_+^J 4h_{IJ}$, where ${}^{(i)}\mathrm{F}_+^{ab} = \alpha_I^{ab\,(i)}H_+^I$ for $i = 1, 2$.

$$Q_D = \int_{\mathbb{R}^3} d^3x j_D^0|_{\text{on shell}} = \int_{\mathbb{R}^3} \frac{d^3k}{(2\pi)^3 k}[|h_+(\vec{k})|^2 - |h_-(\vec{k})|^2] \quad (3.18)$$

is proportional to the difference in the intensity of the self-dual and anti-self-dual parts of the field or, equivalently, the difference between the right and left circularly polarized components—i.e., the net helicity. (Q_D has dimensions of angular momentum.) For this reason Q_D is often called the optical helicity or V-Stokes parameter.

B. Curved spacetimes

The generalization to curved spacetimes of the formalism just presented follows the strategy commonly used for Dirac spin-1/2 fields. Namely, one first introduces an orthonormal tetrad field, or vierbein, in spacetime $e_a^\mu(x)$.[6] With it, the curved spacetime α_I-matices are obtained from the flat space ones α_I^{ab} by

$$\alpha_I^{\mu\nu}(x) = e_a^\mu(x) e_b^\nu(x) \alpha_I^{ab}. \quad (3.19)$$

Furthermore, the Minkowski metric η_{ab} is replaced by $g_{\mu\nu}(x)$; η_{ab} is used to raise and lower flat-space indices $a, b, c, \ldots, g_{\mu\nu}(x)$ for indices in the tangent space of the spacetime manifold $\mu, \nu, \beta, \ldots$, and h_{IJ} and $h_{\dot{I}\dot{J}}$ for spin-1 indices. The matrices $\alpha_I^{\mu\nu}(x)$ satisfy algebraic properties analogous to the ones derived in Minkowski space

$$\{\alpha_I, \alpha_J\} \equiv \alpha^\mu{}_{\nu I} \alpha^{\nu\beta}{}_J + \alpha^\mu{}_{\nu J} \alpha^{\nu\beta}{}_I = -h_{IJ} g^{\mu\beta}, \quad (3.20)$$

$$[\alpha_I, \alpha_J] \equiv \alpha^\mu{}_{\nu I} \alpha^{\nu\beta}{}_J - \alpha^\mu{}_{\nu J} \alpha^{\nu\beta}{}_I = {}^+\Sigma^{\mu\beta}{}_{IJ}, \quad (3.21)$$

where ${}^+\Sigma^{\mu\beta}{}_{IJ} = e_a^\mu e_b^\beta {}^+\Sigma^{ab}{}_{IJ}$. The extension of the covariant derivative ∇_μ is also obtained by using standard arguments (see, e.g., Appendix A of [28]). Namely, the action of ∇_μ on indices I of fields $H_+^I \in V$ is uniquely determined by demanding compatibility with the isomorphism $\alpha_I^{\mu\nu}(x)$, $\nabla_\beta \alpha_I^{\mu\nu}(x) = 0$ (see Appendix B 2). The result, as one would expect, agrees with the usual expression for the covariant derivative acting on fields of spin s derived using group-theoretic methods, particularized to $s = 1$,

$$\nabla_\mu H_+^I = \partial_\mu H_+^I - \frac{1}{2}(w_\mu)_{ab} {}^+\Sigma^{abI}{}_J H_+^J,$$
$$\nabla_\mu H_-^{\dot{I}} = \partial_\mu H_-^{\dot{I}} - \frac{1}{2}(w_\mu)_{ab} {}^-\Sigma^{ab\dot{I}}{}_{\dot{J}} H_-^{\dot{J}}, \quad (3.22)$$

where ${}^\pm\Sigma$ are the generators of the (0,1) and (1,0) representations of the Lorentz algebra introduced in the previous section, and $(w_\mu)_{ab}$ is the standard one-form spin connection

$$(w_\mu)^a{}_b = e_\alpha^a \partial_\mu e_b^\alpha + e_\alpha^a e_b^\beta \Gamma_{\mu\beta}^\alpha, \quad (3.23)$$

where $\Gamma_{\mu\beta}^\alpha$ are the Christoffel symbols.

With this in hand, the generalization is straightforward.

(1) *Maxwell's equations for the fields.—*

$$\alpha_I^{\mu\nu} \nabla_\mu H_+^I = 0, \qquad \bar{\alpha}_{\dot{I}}^{\mu\nu} \nabla_\mu H_-^{\dot{I}} = 0. \quad (3.24)$$

Note the similarity with Dirac's equation. The relation between $H_\pm$ and the self-dual and anti-self-dual parts of the field strength F is given by $F_+^{\mu\nu} = \alpha_I^{\mu\nu} H_+^I$ and $F_-^{\mu\nu} = \bar{\alpha}_{\dot{I}}^{\mu\nu} H_+^{\dot{I}}$. With this, and keeping in mind that $\nabla_\mu \alpha_I^{\beta\nu}(x) = 0$, Eqs. (3.24) become $\nabla_\mu F_+^{\mu\nu} = 0 = \nabla_\mu F_-^{\mu\nu}$, which is manifestly equivalent to Maxwell's equations $\nabla_\mu F^{\mu\nu} = 0$ by recalling that $F_\pm = \frac{1}{\sqrt{2}}[F \pm i{}^\star F]$

(2) *Potentials $A_\mu^\pm$.—*The self-dual and anti-self-dual potentials satisfy the first-order equations:

$$\bar{\alpha}_{\dot{I}}^{\mu\nu} \nabla_\mu A_\nu^+ = 0, \qquad \alpha_I^{\mu\nu} \nabla_\mu A_\nu^- = 0. \quad (3.25)$$

These are equivalent to Maxwell's equations. This can be easily seen by using the same argument as we used in Minkowski spacetime, namely, by noticing that, because $\alpha_I^{\mu\nu}$ and $\bar{\alpha}_{\dot{I}}^{\mu\nu}$ project on self-dual and anti-self-dual forms, respectively, these two equations are simply the self-dual and anti-self-duality condition for the forms $F_{+\mu\nu} \equiv 2\nabla_{[\mu} A_{\nu]}^+$ and $F_{-\mu\nu} \equiv 2\nabla_{[\mu} A_{\nu]}^-$, respectively. This, in turns, implies that the identities $dF_+ = 0$, $dF_- = 0$ are equivalent to Maxwell's equations $\nabla_\mu F_+^{\mu\nu} = 0$, $\nabla_\mu F_-^{\mu\nu} = 0$ (see footnote in Appendix C). Additionally, $\nabla_\mu F_\pm^{\mu\nu} = 0$ is equivalent to the second-order equations $\nabla_\mu \nabla^{[\mu} A_\pm^{\nu]} = 0$.

The relation between $A_\mu^\pm$ and $H_\pm^I$ (before involving any equation of motion) requires foliation of spacetime in spatial Cauchy hypersurfaces Σ_t, in the same way as the relation between the electric and magnetic fields and the standard vector potential does. Given the foliation associated with the definition of $\alpha_I^{\mu\nu}$, $A_{+\mu}$ and H_+^I are related by means of the "curl":[7]

$$H_+^I = i\epsilon^{I\mu\nu} \nabla_\mu A_\nu^+ \quad (3.26)$$

[6]This noncoordinate orthonormal basis is defined by $g_{\mu\nu}(x) = \eta_{ab} e_\mu^a(x) e_\nu^b(x)$, with $\eta_{ab} = \text{diag}\{+1, -1, -1, -1\}$. We assume our spacetime admits such structure globally [27].

[7]Notice that this curl is independent of the connection ∇_μ, due to the antisymmetry of $\epsilon^{I\mu\nu}$ in μ and ν. It is useful to keep this in mind in manipulating expressions involving $H_\pm^I$ and $A_\mu^\pm$.

(and similarly for A^-_μ and $H^{\dot{I}}_-$), where $\epsilon^{I\mu\nu}$ is a "purely spatial" antisymmetric mixed tensor (see Appendix B 3 for its precise definition). As shown in Appendix C, one can easily see that if A^+_ν is a solution of (3.25), then H^I_+ defined by (3.26) satisfies the field equations (3.24). The reverse is also true.

IV. FIRST-ORDER LAGRANGIAN FORMALISM: DIRAC-TYPE FORMULATION

The goal of this section is to write a Lagrangian for electrodynamics in terms of self-dual and anti-self-dual variables. The similarity of Eqs. (3.24) and (3.25) to Dirac's equation motivates us to look for a first-order Lagrangian (i.e., linear in time derivatives) and write it in a form that will make Maxwell's theory manifestly analogous to Dirac's theory, where the mathematical structures associated with spin $s = 1/2$ will be replaced by their $s = 1$ analogs. This formulation will become very useful in the study of the electric-magnetic rotations in the quantum theory.

A. First-order Lagrangian

Consider the action

$$S[A_+,A_-] = -\frac{1}{2}\int \mathrm{d}^4x\sqrt{-g}\left[H^{\dot{I}}_-\bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu A^+_\nu + H^I_+\alpha^{\mu\nu}{}_I\nabla_\mu A^-_\nu\right]. \tag{4.1}$$

The Lagrangian density defined by the integrand differs in a total derivative from the standard Lagrangian $-\frac{1}{4}\sqrt{-g}\mathrm{F}_{\mu\nu}\mathrm{F}^{\mu\nu}$ (after passing from first to second-order formalism), thus leading to the same dynamics. The independent variables in this action are $A^\pm_\nu$, and therefore H^I_+ and $H^{\dot{I}}_-$ are understood as short-hand notation for $i\epsilon^{I\mu\nu}\nabla_\mu A^+_\nu$ and $-i\epsilon^{\dot{I}\mu\nu}\nabla_\mu A^-_\nu$, respectively. Note that this action is first order in time derivatives of $A^\pm_\mu$, and second order in spatial derivatives. Extremizing the action with respect to A^+_ν produces the desired equations of motion (see Appendix D for more details)

$$\frac{\delta S}{\delta A^+_\mu} = 0 \to \bar{\alpha}^{\mu\nu}_{\dot{I}}\nabla_\mu H^{\dot{I}}_- = 0, \tag{4.2}$$

and, as discussed above and proved in Appendix C, these last equations are equivalent to $\alpha^{\mu\nu}_I\nabla_\mu A^-_\nu = 0$. Similarly, from $\frac{\delta S}{\delta A^-_\mu} = 0$ one obtains $\bar{\alpha}^{\mu\nu}_{\dot{I}}\nabla_\mu A^+_\nu = 0$.

For the computations presented in the next section, it is convenient to fix the Lorenz gauge, $\nabla_\mu A^\mu_\pm = 0$. There is a remarkably simple way of incorporating this condition in the action (4.1). All we need to do is extend the domain of the indices I and $\dot{I}$ from $\{1,2,3\}$ to $\{0,1,2,3\}$, and define $\alpha^{\mu\nu}_0 = \bar{\alpha}^{\mu\nu}_0 \equiv -g^{\mu\nu}$. This is analogous to the familiar extension of the Pauli matrices $\vec{\sigma}$ by adding σ^0 (the identity), which commutes with all σ^i, $i = 1, 2, 3$. Algebraic properties of the $\alpha^{\mu\nu}_I$-matrices extended in this way appear in Appendix B 4.

To simplify the notation, we use the same name for the action and the tensors $\alpha^{\mu\nu}{}_I$, although from now on the index I is understood to run from 0 to 3. The equations of motion still take the same form,

$$\bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu A^+_\nu = 0, \qquad \alpha^{\mu\nu}{}_I\nabla_\mu A^-_\nu = 0, \tag{4.3}$$

but they now include the Lorenz condition as the equation for $I = 0$ ($\dot{I} = 0$),

$$g^{\mu\nu}\nabla_\mu A^+_\nu = 0, \qquad g^{\mu\nu}\nabla_\mu A^-_\nu = 0. \tag{4.4}$$

Note that the action now depends on two new variables $H^0_\pm$, but they have the sole role of acting as Lagrange multipliers to enforce Lorenz's condition.

Inspection of the action (4.1) reveals that, contrary to the standard Maxwell's Lagrangian, the Lagrangian density in (4.1) is manifestly invariant, $\delta\mathcal{L} = 0$, under electric-magnetic rotations $A^\mu_\pm \to e^{\mp i\theta}A^\mu_\pm$. It is now straightforward to derive the Noether's current (see Appendix E),

$$j^\mu_D|_{\text{on shell}} = (-g)^{-1/2}\left(\frac{\delta\mathcal{L}}{\delta\nabla_\mu A_{+\nu}}\delta A_{+\nu} + \frac{\delta\mathcal{L}}{\delta\nabla_\mu A_{-\nu}}\delta A_{-\nu}\right)\Big|_{\text{on shell}}$$
$$= \frac{i}{2}\left[H^{\dot{I}}_-\bar{\alpha}^{\mu\nu}{}_{\dot{I}}A^+_\nu - H^I_+\alpha^{\mu\nu}{}_I A^-_\nu\right]. \tag{4.5}$$

Using the relation between self-dual and anti-self-dual variables, and ordinary variables A_μ and $\mathrm{F}^{\mu\nu}$, it is straightforward to check that this expression agrees with $j^\mu_D|_{\text{on shell}}$ obtained in Sec. II, Eq. (2.6).

B. Dirac-type Lagrangian

The goal of this section is to rewrite the action (4.1) (including the Lorenz-gauge-fixing term) in a more convenient form that will make the theory formally similar to Dirac's theory of spin-1/2 fermions and will facilitate the computations in the next sections.

We first integrate by parts (4.1), so $A^\pm$ and $H_\pm$ appear in a more symmetric form,

$$S[A^+,A^-] = -\frac{1}{4}\int d^4x\sqrt{-g}[H^{\dot{I}}_-\bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu A^+_\nu - A^+_\nu\bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu H^{\dot{I}}_-$$
$$+ H^I_+\alpha^{\mu\nu}{}_I\nabla_\mu A^-_\nu - A^-_\nu\alpha^{\mu\nu}{}_I\nabla_\mu H^I_+]. \tag{4.6}$$

This action can now be written as

$$S[A^+,A^-] = -\frac{1}{4}\int d^4x\sqrt{-g}\bar{\Psi}i\beta^\mu\nabla_\mu\Psi \tag{4.7}$$

where we have defined[8]

$$\Psi = \begin{pmatrix} A^+ \\ H_+ \\ A^- \\ H_- \end{pmatrix}, \qquad \bar{\Psi} = (A^+, H_+, A^-, H_-),$$

$$\beta^\mu = i \begin{pmatrix} 0 & 0 & 0 & \bar{\alpha}^\mu \\ 0 & 0 & -\alpha^\mu & 0 \\ 0 & \alpha^\mu & 0 & 0 \\ -\bar{\alpha}^\mu & 0 & 0 & 0 \end{pmatrix}. \tag{4.8}$$

It is convenient to include, in the definition of Ψ, an arbitrary parameter ℓ^{-1} with dimensions of inverse length multiplying $A^\pm$, and compensate it by adding a global factor ℓ to the action. The action remains invariant, but the replacement $A^\pm \to \ell^{-1} A^\pm$ makes all the components of Ψ and $\bar{\Psi}$ have the same dimensions (namely, $\sqrt{\text{energy}/\text{length}^3}$). To simplify the notation, we will not write ℓ explicitly, but it should be taken into account in evaluating the dimensions of expressions containing Ψ and $\bar{\Psi}$.

The exact position of the indices in the components of Ψ and $\bar{\Psi}$ can be easily obtained by comparing (4.6)–(4.8). We have omitted them in the main body of this paper to simplify the notation, but the details can be found in Appendix F. Equation (4.7) is formally analogous to the action of a Majorana 4-spinor describing a field with zero electric charge, whose lower two components are complex conjugate from the upper ones.

From the algebraic properties of the extended α-matrices, (B31) and (B33), it is straightforward to check that β^μ satisfies the Clifford algebra Cliff (3,1),

$$\{\beta^\mu, \beta^\nu\} = 2g^{\mu\nu}\mathbb{I}. \tag{4.9}$$

We also have that $\nabla_\nu \beta^\mu(x) = 0$. These matrices can therefore be thought of as the spin-1 analog of the Dirac γ^μ matrices.

We now define the "chiral" matrix

$$\beta_5 \equiv \frac{i}{4!}\epsilon_{\alpha\beta\gamma\delta}\beta^\alpha\beta^\beta\beta^\gamma\beta^\delta = \begin{pmatrix} -\mathbb{I} & 0 & 0 & 0 \\ 0 & -\mathbb{I} & 0 & 0 \\ 0 & 0 & \mathbb{I} & 0 \\ 0 & 0 & 0 & \mathbb{I} \end{pmatrix}. \tag{4.10}$$

Some properties can be immediately checked:

$$\{\beta^\mu, \beta_5\} = 0, \qquad \beta_5^2 = \mathbb{I}. \tag{4.11}$$

Further details and properties can be found in Appendix F.

Although the basic variables in the action are the potentials $A_\mu^\pm$, at the practical level one can work by considering Ψ and $\bar{\Psi}$ as independent fields—note that this is the same as what one does when working with Majorana spinors. The equations of motion take the form

$$\frac{\delta S}{\delta \bar{\Psi}} = 0 \to i\beta^\mu \nabla_\mu \Psi = 0. \tag{4.12}$$

They contain four equations, one for each of the four components of Ψ. The upper two are the equations $\bar{\alpha}^{\mu\nu}{}_i \nabla_\mu A_\nu^+ = 0$ and $\alpha_I^{\mu\nu} \nabla_\mu H_+^I = 0$. The lower two are complex-conjugated equations.

Now, by acting on (4.12) with $(-i\beta^\alpha \nabla_\alpha)$, we obtain a second-order equation for Ψ:

$$\begin{aligned}(-i\beta^\alpha\nabla_\alpha) i\beta^\mu \nabla_\mu \Psi &= (\beta^{(\alpha}\beta^{\mu)} + \beta^{[\alpha}\beta^{\mu]})\nabla_\alpha\nabla_\mu\Psi \\ &= (\Box + \mathcal{Q})\Psi = 0,\end{aligned} \tag{4.13}$$

where we have used (4.9) and defined

$$\mathcal{Q}\Psi \equiv \frac{1}{2}\beta^{[\alpha}\beta^{\mu]} W_{\alpha\mu}\Psi \tag{4.14}$$

with

$$\begin{aligned} W_{\alpha\mu}\Psi &\equiv [\nabla_\alpha, \nabla_\mu]\Psi \\ &= \frac{1}{2} R_{\alpha\mu\sigma\rho} \begin{pmatrix} \Sigma^{\sigma\rho} & 0 & 0 & 0 \\ 0 & {}^+\Sigma^{\sigma\rho} & 0 & 0 \\ 0 & 0 & \Sigma^{\sigma\rho} & 0 \\ 0 & 0 & 0 & {}^-\Sigma^{\sigma\rho} \end{pmatrix}\Psi, \end{aligned} \tag{4.15}$$

where $\Sigma^{\sigma\rho}{}_{\alpha\beta} = \delta^\rho_\alpha \delta^\sigma_\beta - \delta^\rho_\beta \delta^\sigma_\alpha$ is the generator of the $(1/2, 1/2)$ (real) representation of the Lorentz group, while ${}^+\Sigma^{\sigma\rho}_{IJ}$ and ${}^-\Sigma^{\sigma\rho}_{\bar{i}\bar{j}}$ are the generators of the $(0,1) \oplus (0,0)$ and $(1,0) \oplus (0,0)$ representations, respectively.

Looking at the expression for $W_{\alpha\mu}\Psi$, we see that it contains real terms, $R_{\alpha\mu\sigma\rho}\Sigma^{\sigma\rho}$, as well as complex ones, $R_{\alpha\mu\sigma\rho}{}^\pm\Sigma^{\sigma\rho}$. The real terms come from the action of covariant derivatives on $A_\pm^\mu$. Since $A_\pm^\mu$ are vectors in spacetime, their covariant derivative includes a connection associated with the $(1/2, 1/2)$ representation of the Lorentz group.[9] The

[8]We could alternatively have defined a couple of fields with two "components," $\binom{A^+}{H_-}$ and $\binom{A^-}{H_+}$. Physical predictions would obviously be the same since we are just writing the same theory in different variables. However, the formal analogy with Dirac's theory is cleaner if we use the four-component object Ψ defined in (4.8).

[9]This does not mean, however, that $A_\pm^\mu$ transform according to the $(1/2, 1/2)$ representation of the Lorentz group; They do so only up to a gauge transformation [22]. See [29] for a more precise account of this issue.

complex terms in $W_{\mu\nu}\Psi$ originate from the (0,1) and (1,0) representations, with which $\vec{H}_{\pm}$ are associated.

The Poisson brackets for Ψ and $\bar{\Psi}$ can be easily derived from the canonical relations $\{A^{+}_{\mu}, H^{i}_{-}\} = \gamma^{i}_{\mu}\delta(\vec{x}, \vec{x}')$, in an analogous way as is usually done for Majorana spinors, with the being difference that, in the situation under consideration in this paper, the Poisson brackets must be promoted to commutation relations in the quantum theory. If anticommutators are used instead, one would find the quantum propagator to violate causality, as expected from the spin-statistics theorem. Therefore, in spite of the fermionlike appearance of the formulation used in this section, we are describing a theory of bosons.

1. Axial current

We now describe how the electric-magnetic symmetry and its associated conservation law look in the language introduced in this section. By using the chiral matrix β_5, the transformation reads

$$\Psi \to e^{i\theta\beta_5}\Psi, \qquad \bar{\Psi} \to \bar{\Psi}e^{i\theta\beta_5}. \tag{4.16}$$

Notice that this has the same form as a chiral transformation for fermions. Looking at the form of β_5 in Eq. (4.10), it is clear that the upper two components of Ψ, i.e., (A_+, H_+), represent the self-dual or positive chirality part of the field, while the lower two components (A_-, H_-) contain the anti-self-dual or the negative chiral part. The Lagrangian density (4.7) is manifestly invariant under these transformation, and in terms of Ψ the conserved current reads

$$j^{\mu}_{D} = \frac{1}{4}\bar{\Psi}\beta^{\mu}\beta_5\Psi. \tag{4.17}$$

The associated Noether charge is

$$Q_D = \int_{\Sigma_t} \mathrm{d}\Sigma_{\mu} j^{\mu}_{D} = \frac{1}{4}\int_{\Sigma_t} \mathrm{d}\Sigma_3 \bar{\Psi}\beta^0\beta_5\Psi, \tag{4.18}$$

where $\mathrm{d}\Sigma_3$ is the volume element of a spacelike Cauchy hypersurface Σ_t. This expression for Q_D is equivalent to the one obtained in previous sections [see Eq. (2.7)].

V. QUANTUM ANOMALY

In this section we analyze whether the classical symmetry under electric-magnetic rotations persists in the quantum theory. The most direct avenue to meet this goal is to compute the vacuum expectation value of the divergence $\nabla_{\mu} j^{\mu}_{D}$. A nonvanishing result would imply that the vacuum expectation value of the charge Q_D is not a constant of motion. For the sake of clarity, we perform the calculation using two different methods. First, we provide a direct computation of $\langle \nabla_{\mu} j^{\mu}_{D}\rangle$, in which the ultraviolet divergences are identified and subtracted in a covariant and self-consistent way, and then we reproduce the same result using Fujikawa's approach to anomalies based on path integrals. These two methods illuminate complementary aspects of the calculation.

A. Direct computation

Both j^{μ}_{D} and $\nabla_{\mu} j^{\mu}_{D}$ are operators quadratic in fields, and therefore the computation of their expectation values must include renormalization subtractions to eliminate potential divergences:

$$\langle \nabla_{\mu} j^{\mu}_{D}\rangle_{\mathrm{ren}} = \langle \nabla_{\mu} j^{\mu}_{D}\rangle - \langle \nabla_{\mu} j^{\mu}_{D}\rangle_{\mathrm{Ad}(4)}. \tag{5.1}$$

In this expression, $\langle \nabla_{\mu} j^{\mu}_{D}\rangle_{\mathrm{Ad}(4)}$ indicates renormalization terms of fourth adiabatic order that we will compute using the DeWitt-Schwinger asymptotic expansion. More precisely, this renormalization scheme works by writing $\langle \nabla_{\mu} j^{\mu}_{D}\rangle$ in terms of the Feymann two-point function $S(x, x') = -i\langle T\Psi(x)\bar{\Psi}(x')\rangle$, and then by replacing it by $[S(x, x') - S(x, x')_{\mathrm{Ad}(4)}]$, where $S(x, x')_{\mathrm{Ad}(4)}$ denotes the DeWitt-Schwinger subtractions up to fourth adiabatic order, and then taking the limit $x \to x'$.

A convenient way to regularize potential infrared divergences is by introducing a parameter $s > 0$ in the theory (that will be sent to zero at the end of the calculation), replacing the wave equation $D\Psi = 0$ by $(D + s)\Psi = 0$, where $D \equiv i\beta^{\mu}\nabla_{\mu}$ [30]. Therefore,

$$\begin{aligned} \nabla_{\mu} j^{\mu}_{D}(x) &= \nabla_{\mu}\left[\frac{1}{4}\bar{\Psi}(x)\beta^{\mu}\beta_5\Psi(x)\right] \\ &= \frac{-i}{4}[\bar{\Psi}(x)\overleftarrow{D}\beta_5\Psi(x) - \bar{\Psi}(x)\beta_5\overrightarrow{D}\Psi(x)] \\ &= \lim_{\substack{s\to 0\\ x\to x'}} \frac{-i}{2} s\bar{\Psi}(x)\beta_5\Psi(x') \\ &= \lim_{\substack{s\to 0\\ x\to x'}} \frac{-i}{2} s\mathrm{Tr}[\beta_5\Psi(x)\bar{\Psi}(x')], \end{aligned} \tag{5.2}$$

where we have used $\{\beta^{\mu}, \beta_5\} = 0$. If we now make a choice of vacuum state $|0\rangle$, we obtain[10]

$$\langle \nabla_{\mu} j^{\mu}_{D}\rangle = \lim_{\substack{s\to 0\\ x\to x'}} \frac{1}{2} s\mathrm{Tr}[\beta_5 S(x, x', s)]. \tag{5.3}$$

The renormalized expectation value is then given by

$$\langle \nabla_{\mu} j^{\mu}_{D}\rangle_{\mathrm{ren}} = \lim_{\substack{s\to 0\\ x\to x'}} \frac{1}{2} s\mathrm{Tr}[\beta_5(S(x, x', s) - S(x, x', s)_{\mathrm{Ad}(4)})]. \tag{5.4}$$

In this expression, $S(x, x', s)$ contains the information about the vacuum state, while the role of $S(x, x', s)_{\mathrm{Ad}(4)}$ is to

[10] We choose $x^0 > x'^0$ without loss of generality, so that $T\Psi(x)\bar{\Psi}(x') = \Psi(x)\bar{\Psi}(x')$.

remove the potential ultraviolet divergences, which are the same for all vacua. It is convenient to write $S(x,x',s)_{\mathrm{Ad}(4)} = [(D-s)G(x,x',s)]_{\mathrm{Ad}(4)}$ (D acts on the x-argument), where[11]

$$G(x,x',s) \sim \frac{\hbar\Delta^{1/2}(x,x')}{16\pi^2} \times \sum_{k=0}^{\infty} E_k(x,x') \int_0^{\infty} d\tau e^{-i(\tau s^2 + \frac{\sigma(x,x')}{2\tau})}(i\tau)^{(k-2)} \tag{5.5}$$

where $\sigma(x,x')$ is half of the geodesic distance squared between x and x', $\Delta^{1/2}(x,x')$ is the Van Vleck-Morette determinant, and the functions $E_k(x,x')$ are the DeWitt coefficients, which are geometric quantities, built from the metric and its first $2k$th derivatives. We only need the value of these coefficients when $x = x'$. For manifolds without boundaries, they are [31,32]

$$E_0(x) = \mathbb{I},$$

$$E_1(x) = \frac{1}{6}R\mathbb{I} - \mathcal{Q},$$

$$E_2(x) = \left[\frac{1}{72}R^2 - \frac{1}{180}R_{\mu\nu}R^{\mu\nu} + \frac{1}{180}R_{\alpha\beta\mu\nu}R^{\alpha\beta\mu\nu} - \frac{1}{30}\Box R\right]\mathbb{I} + \frac{1}{12}W_{\mu\nu}W^{\mu\nu} + \frac{1}{2}\mathcal{Q}^2 - \frac{1}{6}R\mathcal{Q} + \frac{1}{6}\Box\mathcal{Q},$$

where the expressions for $W_{\mu\nu} \equiv [\nabla_\mu, \nabla_\nu]$ and $\mathcal{Q}(x)$ are given in (4.15) and (4.14), respectively. Here, R, $R_{\mu\nu}$, and $R_{\alpha\beta\mu\nu}$ are the Ricci scalar, Ricci tensor, and Riemann curvature tensor.

Because of the symmetry of the classical action, the contribution of $S(x,x',s)$ to (5.4) vanishes for all choices of vacuum state. Therefore, $\langle \nabla_\mu j^\mu_D \rangle_{\mathrm{ren}}$ arises entirely from the subtraction terms, $S(x,x',s)_{\mathrm{Ad}(4)}$. This implies that $\langle \nabla_\mu j^\mu_D \rangle_{\mathrm{ren}}$ is independent of the choice of vacuum. Notice that the same occurs in the calculation of other anomalies, such as the fermionic chiral anomaly or the trace anomaly.

It turns our that only the terms with $k=2$ in (5.5) produce a nonvanishing contribution. Furthermore, we do not need to consider terms involving derivatives of $E_2(x,x')$ since they involve five derivatives of the metric and hence are of fifth adiabatic order. Taking into account that

$$\mathrm{Tr}[\beta_5 E_2(x,x))] = i\frac{1}{3}R_{\alpha\beta\mu\nu}{}^{\star}R^{\alpha\beta\mu\nu} \tag{5.6}$$

where ${}^{\star}R^{\alpha\beta\mu\nu} = \frac{1}{2}\epsilon^{\alpha\beta\sigma\rho}R_{\sigma\rho}{}^{\mu\nu}$ is the dual of the Riemann tensor, Eq. (5.4) produces

$$\langle \nabla_\mu j^\mu_D \rangle_{\mathrm{ren}} = -\frac{\hbar}{96\pi^2}R_{\alpha\beta\mu\nu}{}^{\star}R^{\alpha\beta\mu\nu}. \tag{5.7}$$

Appendix G contains details of the intermediate steps in this computation. A few comments are in order now:

(1) This result reveals that quantum fluctuations spoil the conservation of the axial current j^μ_D and break the classical symmetry under electric-magnetic (or chiral) transformations.

(2) The pseudoscalar $R_{\alpha\beta\mu\nu}{}^{\star}R^{\alpha\beta\mu\nu}$ is known as the Chern-Pontryagin density (its integral across the entire spacetime manifold is the Chern-Pontryagin invariant).

(3) It is important to notice the parallelism with the chiral anomaly for spin-1/2 fermions. The computations in that case would be very similar, except that one would have to use structures associated with spin-1/2 fields, rather than spin 1. This would change only the numerical coefficient in (5.7).

B. Path integral formalism

The functional integral for the theory under consideration is[12]

$$Z = \int D\bar{\Psi}D\Psi e^{i/\hbar S[\Psi,\bar{\Psi}]}. \tag{5.8}$$

The strategy of Fujikawa's approach to the computation of anomalies using path integrals is the following. The generating functional Z is invariant under the replacement $(\Psi,\bar{\Psi}) \to (\Psi' = e^{i\beta_5\theta}\Psi, \bar{\Psi}' = \bar{\Psi}e^{i\beta_5\theta})$ since this is just a change of variables and the path integral remains invariant under such a change. However, the two components of the integrand, the measure and the action, could change under the transformation. Noether's theorem—in the version in which one considers the parameter of the transformation $\theta(x)$ to be a spacetime function of compact support—tells us that $\delta S = -\int d^4x\sqrt{-g}\theta(x)\nabla_\mu j^\mu_D$. On the other hand, the integral measure $D\bar{\Psi}D\Psi$ could change by a nontrivial Jacobian, $D\bar{\Psi}D\Psi \to JD\bar{\Psi}'D\Psi'$. Then, the invariance of Z implies that these two changes must compensate each other; i.e., $J\cdot e^{-i/\hbar\int d^4x\sqrt{-g}\theta(x)\langle\nabla_\mu j^\mu_D\rangle}$ must be equal to 1. From this we see that quantum anomalies appear for those

[11]This expression for $G(x,x',s)_{\mathrm{Ad}(4)}$ is obtained by writing $G(x,x',s)_{\mathrm{Ad}}$ first in terms of its heat kernel $K(\tau,x,x')$, $G(x,x',s)_{\mathrm{Ad}} = i\hbar\Delta^{1/2}(x,x')\int_0^\infty d\tau e^{-i(\tau s^2+\frac{\sigma(x,x')}{2\tau})}K(\tau,x,x')$, and then by using the asymptotic expansion $K(\tau,x,x) \sim \frac{-i}{16\pi^2}\sum_{k=0}^\infty (i\tau)^{k-2}E_k(x)$ for $\tau\to 0$. See e.g., [31] for further details.

[12]As usual, the inclusion of the Lorentz gauge introduces two ghost scalar fields. These fields contribute to certain observables, such as the trace anomaly. However, one can check explicitly that they do not affect the computation of $\langle\nabla_\mu j^\mu_D\rangle$. It is for this reason that we have not written their contribution to the path integral.

classical symmetries that do not leave the measure of the path integral invariant, i.e., $J \neq 1$. The value of $\langle \nabla_\mu j^\mu_D \rangle$ can then be determined from J. The goal of this section is to compute these quantities.

The Jacobian J can be determined by using standard functional analysis techniques applied to the wave operator D^2, where $D = \beta^\mu \nabla_\mu$. Consider the space of square-integrable fields $\Psi(x)$ with respect to the product $\langle \Psi_1, \Psi_2 \rangle = \alpha \int d^4x \sqrt{-g} \Psi_1^\dagger \Psi_2$ [see Appendix F for further details, particularly the discussion around expressions (F8)], and $\alpha > 0$ is an arbitrary real parameter with dimensions of inverse action.[13] In terms of the original variables $A_\pm$ and $H_\pm$, the norm of $\Psi(x)$ reads $\langle \Psi, \Psi \rangle = \alpha \int d^4x \sqrt{-g} [2|A_+|^2 + 2|H_+|^2] \geq 0$.

It is easy to check that the operator D^2 is self-adjoint with respect to the product $\langle \Psi_1, \Psi_2 \rangle$. The self-adjointness of D^2 guarantees the existence of an orthonormal basis $\{\Psi_n\}$ made of eigenfunctions, $D^2 \Psi_n = \lambda_n^2 \Psi_n$. We will denote by a_n the components of a vector Ψ in this basis. An electric-magnetic rotation $\Psi \to \Psi' = e^{i\theta\beta_5}\Psi$ can now be expressed as a change of the components $a_n \to a'_n = \sum_m C_{nm} a_m$, with $C_{nm} = \langle \Psi_n, e^{i\theta\beta_5} \Psi_m \rangle$. With this, the Jacobian of the transformation reads

$$D\bar{\Psi} D\Psi \to J D\bar{\Psi}' D\Psi', \quad \text{with}$$
$$J = (\det C)^2 = e^{2\mathrm{Tr}[\ln C]} = e^{i2\sum_n \langle \Psi_n, \beta_5 \theta \Psi_n \rangle}. \qquad (5.9)$$

Then, the invariance of the path integral implies that, quantum mechanically,

$$\langle \nabla_\mu j^\mu_D \rangle_{\rm ren} = 2\hbar\alpha \sum_{n=0}^{\infty} \bar{\Psi}_n \beta_5 \Psi_n. \qquad (5.10)$$

To evaluate this expression we again use the heat-kernel approach. The kernel of the equation, $D^2\Psi = 0$, is [31][14]

$$K(\tau, x, x') = -4\alpha \sum_{n=0}^{\infty} e^{-i\tau\lambda_n^2} \Psi_n(x) \bar{\Psi}_n(x'). \qquad (5.11)$$

Then

$$\langle \nabla_\mu j^\mu_D \rangle_{\rm ren} = \frac{-1}{2} \hbar \lim_{\tau\to 0} \mathrm{Tr}[\beta_5 K(\tau, x, x)] = i \frac{\hbar}{32\pi^2} \mathrm{Tr}[\beta_5 E_2]$$
$$= -\frac{\hbar}{96\pi^2} R_{\alpha\beta\mu\nu} {}^\star R^{\alpha\beta\mu\nu}, \qquad (5.12)$$

where in the second equality we have used the expansion of $K(\tau, x, x')$ for $\tau \to 0$, written in footnote, and in the last equality we have used (5.6).

Recall that the path integral produces transition amplitudes for time-ordered products of operators between the "in" and "out" vacuum. However, the result for $\langle \nabla_\mu j^\mu_D \rangle_{\rm ren}$ comes entirely from the asymptotic terms in the heat kernel, which are the same for all vacua. Therefore, the result (5.12) agrees with the expectation value of $\nabla_\mu j^\mu_D$ in any vacuum state.

VI. CONCLUSIONS

The apparently trivial invariance of the source-free Maxwell's equations under duality transformations $F_{\mu\nu} \to {}^\star F_{\mu\nu}$ has interesting physical consequences. This mapping can be extended to a continuous "rotation" $F_{\mu\nu} \to \cos\theta F_{\mu\nu} + \sin\theta {}^\star F_{\mu\nu}$, which can be proven to be a symmetry of Maxwell's action both in flat and curved spacetimes. Noether's theorem then provides the existence of a conserved current and the associated constant of motion, which describes the polarization state of electromagnetic radiation. The main goal of this paper is to show that this conservation law does not survive the quantization in curved spacetimes, and an anomaly arises in the form of (5.12).

To meet our goal, we have rewritten Maxwell's theory by using self-dual and anti-self-dual variables. These fields transform under irreducible representations of the Lorentz group and describe the two chiral sectors of the theory. In this language, Maxwell's electric-magnetic rotations reduce to an ordinary chiral transformation, which in the absence of charges and currents becomes a symmetry of the classical theory. In this sense, our result can be understood as the spin-1 generalization of the spin-1/2 chiral anomaly.

Although anomalies arise mathematically as a consequence of taming ultraviolet divergences via regularization and renormalization, they have low-energy implications, as stressed, e.g., in [33]. To give some examples, in two-dimensional spacetimes the trace anomaly implies the Hawking effect [34], and the fermionic axial anomaly is closely related to the Schwinger pair creation effect [35]. Similarly, the electric-magnetic duality anomaly found in this paper is expected to have interesting physical applications in astrophysics, cosmology, and condensed matter systems. This paper has been devoted to discussing the details of theoretical formalism underlying the computation of this anomaly. A detailed analysis of its physical consequences will be the focus of future publications. In particular, we expect that gravitational dynamics will be able to produce net circular polarization on photons through asymmetric creation of right and left quanta. Some preliminary ideas were summarized in [36], where applications related to gravitational collapse and mergers in astrophysics were suggested.

[13] It is introduced in order to make the product dimensionless and, although $\alpha = \hbar^{-1}$ would be a natural choice, we leave it unspecified to make manifest that physical observables are independent of it; it cancels out in intermediate steps.

[14] The factor -4 appears as a consequence of the fact that the pair of spinor fields that are canonically conjugated are Ψ and $\bar{\Psi}$—and not Ψ and $\frac{\partial L}{\partial \partial_t \Psi} = -\frac{1}{4}\bar{\Psi}$.

ACKNOWLEDGMENTS

This work was supported by Grants No. FIS2017-91161-EXP and No. FIS2017-84440-C2-1-P, Project No. SEJI/2017/042 (Generalitat Valenciana), the COST action CA15117 (CANTATA) [supported by COST (European Cooperation in Science and Technology)], and NSF CAREER Grant No. PHY-1552603. A. d. R. was supported by the Spanish Ministry of Education Grant No. FPU13/04948, and acknowledges financial support provided under the ERC Consolidator Grant "Matter and strong-field gravity: New frontiers in Einstein's theory," No. MaGRaTh-646597. We thank A. Ashtekar, E. Bianchi, and J. Pulin for useful discussions.

APPENDIX A: NOETHER CURRENT

Here we provide a few more details about the variation of the Lagrangian density (2.4) under the infinitesimal transformation (2.3). We obtain

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial A_\nu}\delta A_\nu + \frac{\partial\mathcal{L}}{\partial\nabla_\mu A_\nu}\delta\nabla_\mu A_\nu = -\sqrt{-g}F^{\mu\nu}\nabla_\mu\delta A_\nu$$
$$= -\sqrt{-g}\mathrm{F}^{\mu\nu}\nabla_\mu Z_\nu. \qquad \text{(A1)}$$

The equality ${}^\star\mathrm{F}=\mathrm{d}Z+\mathrm{G}$ leads to $\mathrm{F}=-{}^\star\mathrm{d}Z-{}^\star\mathrm{G}$. Then ${}^\star\mathrm{G}_{\mu\nu}\mathrm{G}^{\mu\nu} = ({}^\star\mathrm{F}_{\mu\nu}-\mathrm{d}Z_{\mu\nu})(-({}^\star\mathrm{d}Z)^{\mu\nu}-\mathrm{F}^{\mu\nu}) = \mathrm{d}Z^{\mu\nu}({}^\star dZ)_{\mu\nu} - \mathrm{F}^{\mu\nu}{}^\star\mathrm{F}_{\mu\nu} + 2\mathrm{d}Z^{\mu\nu}\mathrm{F}_{\mu\nu}$, from which we get

$$\delta\mathcal{L} = -\sqrt{-g}\frac{1}{2}\nabla_\mu(A_\nu{}^\star\mathrm{F}^{\mu\nu} - Z_\nu{}^\star\mathrm{d}Z^{\mu\nu})$$
$$-\frac{1}{4}\sqrt{-g}{}^\star\mathrm{G}_{\mu\nu}\mathrm{G}^{\mu\nu}. \qquad \text{(A2)}$$

The last term is equal to the product of the electric and magnetic parts of G and, since the latter vanishes in one frame, ${}^\star\mathrm{G}_{\mu\nu}\mathrm{G}^{\mu\nu}=0$ in any frame. Then $\delta\mathcal{L}$ is the divergence of a current, $\delta\mathcal{L}=\sqrt{-g}\nabla_\mu h^\mu$, which implies that the action remains invariant.

The Noether current is then given by

$$j^\mu_D = \frac{1}{\sqrt{-g}}\frac{\partial\mathcal{L}}{\partial\nabla_\mu A_\nu}\delta A_\nu - h^\mu$$
$$= \frac{1}{2}[A_\nu{}^\star\mathrm{F}^{\mu\nu} - Z_\nu 2\mathrm{F}^{\mu\nu} - Z_\nu({}^\star\mathrm{d}Z)^{\mu\nu}], \qquad \text{(A3)}$$

which agrees with (2.5) after using $\mathrm{d}Z={}^\star\mathrm{F}+\mathrm{G}$. Acting now with the derivative operator on (A3), one finds

$$\nabla_\mu j^\mu_D = \frac{1}{2}[\nabla_\mu A_\nu{}^\star\mathrm{F}^{\mu\nu} - 2\nabla_\mu Z_\nu\mathrm{F}^{\mu\nu}$$
$$- 2Z_\nu\nabla_\mu\mathrm{F}^{\mu\nu} - \nabla_\mu Z_\nu({}^\star dZ)^{\mu\nu}]$$
$$= \frac{1}{2}[\nabla_\mu A_\nu{}^\star\mathrm{F}^{\mu\nu} - ({}^\star\mathrm{F}_{\mu\nu}+\mathrm{G}_{\mu\nu})\mathrm{F}^{\mu\nu} - 2Z_\nu\nabla_\mu\mathrm{F}^{\mu\nu}$$
$$-\frac{1}{2}({}^\star\mathrm{F}_{\mu\nu}+\mathrm{G}_{\mu\nu})(-\mathrm{F}^{\mu\nu}+{}^\star\mathrm{G}^{\mu\nu})]$$
$$= -Z_\nu\nabla_\mu\mathrm{F}^{\mu\nu}$$

(the Bianchi identity was used in the first equality) which vanishes on shell.

APPENDIX B: THE α_I^{ab} TENSOR

This Appendix contains additional properties of the α_I^{ab} tensors used in the main body of this paper. The properties for the tensors $\bar{\alpha}_I^{ab}$ are obtained by complex conjugation.

1. Definition and properties

Let $\{t^a, x^a, y^a, z^a\}$ be an inertial coordinate frame of contravariant vectors in 4D Minkowski spacetime. Consider the following set of complex, antisymmetric tensors:

$$\alpha_1^{ab} = -2(t^{[a}x^{b]} + iy^{[a}z^{b]}), \qquad \text{(B1)}$$

$$\alpha_2^{ab} = -2(t^{[a}y^{b]} + iz^{[a}x^{b]}), \qquad \text{(B2)}$$

$$\alpha_3^{ab} = -2(t^{[a}z^{b]} + ix^{[a}y^{b]}), \qquad \text{(B3)}$$

where the square brackets indicate antisymmetrization of indices. It is straightforward to check that they are self-dual, i.e., $i{}^\star\alpha_I^{ab} \equiv i\frac{1}{2}\epsilon^{ab}{}_{cd}\alpha_I^{cd} = \alpha_I^{ab}$. These three tensors form an orthogonal basis in the space of self-dual (complex) tensors in Minkowski spacetime. Given any such tensor F_+^{ab}, we can write it as

$$\mathrm{F}_+^{ab} = H_+^I\alpha_I^{ab}, \qquad \text{(B4)}$$

where H_+^I indicates the components of F_+^{ab} in this basis. This last equation can alternatively read as follows. Let V be a three-dimensional complex vector space, made of vectors H_+^I. Let $\{X_I, Y_I, Z_I\}$ be a basis of one-forms in the dual space V^*. Equation (B4) tells us that α_I^{ab} is an isomorphism between V and the space of self-dual tensors. An isomorphism can be obtained by the identifying basis

$$\alpha_I^{ab} \equiv \alpha_1^{ab}X_I + \alpha_2^{ab}Y_I + \alpha_3^{ab}Z_I. \qquad \text{(B5)}$$

This isomorphism can be used to endow V with a product $h_{IJ} = \frac{1}{4}\eta_{ab}\eta_{cd}\alpha_I^{ac}\alpha_J^{bd}$, which, in the basis we started with, has components equal to minus the Kronecker delta, $-\delta_{IJ}$. Spacetime indices $a, b, c, \cdots$ are raised and lowered with

Minkowski metric η_{ab}, while "internal" indices $I, J, K, \cdots$ are raised and lowered with h_{IJ}.

We collect here some useful properties of the tensors α_I^{ab}, which can be checked by direct computation:

$$\alpha_{abI}\alpha^{ab}{}_J = 4h_{IJ}, \tag{B6}$$

$$\alpha_{ab}{}^I\alpha_{cdI} = 4{}^+P_{abcd}, \tag{B7}$$

$$\alpha_{abI}\bar{\alpha}^{ab}{}_{\dot{J}} = 0, \tag{B8}$$

$$\alpha^a{}_{bI}\alpha^{cb}{}_J = h_{IJ}\eta^{ac} - [{}^+\Sigma_{IJ}]^{ac}. \tag{B9}$$

In property (B7), ${}^+P_{abcd} = \frac{1}{4}(\eta_{ac}\eta_{bd} - \eta_{ad}\eta_{bc} + i\epsilon_{abcd})$ is the projector on self-dual tensors in Minkowski spacetime, and $[{}^+\Sigma_{IJ}]^{ac}$ is the generator of the (0,1) representation of the Lorentz group, whose explicit form is $[{}^+\Sigma_{IJ}]^{ab} = -i\epsilon_{IJK}\alpha^{abK}$. Recall that, according to our sign conventions, we have $\eta_{ab} = t_a t_b - x_a x_b - y_a y_b - z_a z_b$ and $\epsilon^{abcd} = -4!t^{[a}x^b y^c z^{d]}$ in this basis. On the other hand, taking the symmetric and antisymmetric parts of (B9) yields the "commutation" and "anticommutation" properties of α_I^{ab}:

$$\alpha^{[a}{}_{bI}\alpha^{c]b}{}_J = -[{}^+\Sigma^{ac}]_{IJ}, \tag{B10}$$

$$\alpha^{(a}{}_{bI}\alpha^{c)b}{}_J = \eta^{ac}h_{IJ}. \tag{B11}$$

In a similar manner, the tensor

$$\bar{\alpha}_{\dot{I}}^{ab} \equiv \bar{\alpha}_1^{ab}X_{\dot{I}} + \bar{\alpha}_2^{ab}Y_{\dot{I}} + \bar{\alpha}_3^{ab}Z_{\dot{I}} \tag{B12}$$

provides an isomorphism between the vector space $\bar{V}$, the complex conjugate of V, and the space of anti-self-dual tensors in Minkowski spacetime,

$$\mathrm{F}_-^{ab} = H_-^{\dot{I}}\bar{\alpha}_{\dot{I}}^{ab}. \tag{B13}$$

The analogs of the properties (B6)–(B9) hold, replacing ${}^+P^{abcd}$ by the anti-self-dual projector ${}^-P^{abcd}$, which is simply the complex conjugate of ${}^+P^{abcd}$, and $[{}^+\Sigma_{IJ}]^{ac}$ by the generator of the (1,0) representations $[{}^-\Sigma_{\dot{I}\dot{J}}]^{ac}$.

The generalization to curved spacetimes is straightforward. Given a field of vierbeins $e_a^\mu(x)$, i.e., a field of orthonormal basis of tangent vectors in the spacetime manifold $(M, g_{\mu\nu})$, the $\alpha_I^{\mu\nu}$ tensor is constructed from the Minkowski space tensor α_I^{ab} by

$$\alpha_I^{\mu\nu}(x) = e_a^\mu(x)e_b^\nu(x)\alpha_I^{ab}. \tag{B14}$$

This makes it obvious that the properties (B6)–(B9) generalize to curved spacetimes by simply replacing the tensors η_{ab} and ϵ_{abcd} by their counterparts in curved geometries, $g_{\mu\nu}$ and $\epsilon_{\mu\nu\alpha\beta}$.

2. Covariant derivative operator

In this Appendix we provide some details regarding the extension of the action of the covariant derivative to indices $I, J, K, \ldots$.

Recall that the vierbein $e_a^\mu(x)$ at a given point of the spacetime manifold $(M, g_{\mu\nu})$ provides an isometry between the tangent space at x and Minkowski spacetime. The extension of the action of the covariant derivative ∇_μ on "internal" indices $a, b, c, \ldots$ is obtained by demanding $\nabla_\mu e_\nu^a(x) = 0$. This defines the connection one-form ω_μ,

$$\omega_\mu^{ab} = e_\nu^a\partial_\mu e^{b\nu} + \Gamma^\nu_{\mu\alpha}e_\nu^a e^{\alpha b}, \tag{B15}$$

where $\Gamma^\nu_{\mu\alpha}$ are the Christoffel symbols. Recall that ω_μ^{ab} is antisymmetric, $\omega_\mu^{ab} = \omega_\mu^{[ab]}$ (as a consequence of $\nabla_\mu g_{\alpha\beta} = 0$). To further extend the action of ∇_μ to the complex vector space V, we follow the standard strategy. Namely, by linearity the difference between any two possible extensions is characterized by

$$(\nabla_\mu - \bar{\nabla}_\mu)H_I = -C_{\mu I}{}^J H_J, \qquad H_I \in V^*. \tag{B16}$$

If we choose $\bar{\nabla}_\mu$ to be the ordinary derivative associated with a system of coordinates, $\bar{\nabla}_\mu = \partial_\mu$, we see that there are as many derivative operators as mixed tensors $C_{\mu I}{}^J$. The most natural condition to single out one of them is to demand that ∇_μ annihilates the isomorphism $\alpha_I^{\alpha\beta}(x)$,

$$0 \equiv \nabla_\mu\alpha_I^{\alpha\beta} = \partial_\mu\alpha_I^{\alpha\beta} + \Gamma^\alpha_{\mu\rho}\alpha_I^{\rho\beta} + \Gamma^\beta_{\mu\rho}\alpha_I^{\alpha\rho} - C_{\mu I}{}^J\alpha_J^{\alpha\beta}.$$

Now using $\alpha^{\alpha\beta}{}_I(x) = e_a^\alpha(x)e_b^\beta(x)\alpha^{ab}{}_I$, together with the properties of α_I^{ab}, we obtain from the previous equation the form of $C_{\mu I}{}^J$,

$$\begin{aligned} C_{\mu I}{}^J &= \frac{1}{2}e^a_\alpha(\partial_\mu e^\alpha_c)\alpha^J_{ab}\alpha^{cb}_I + \frac{1}{2}\alpha^J_{\alpha\beta}\Gamma^\alpha_{\mu\rho}\alpha^{\rho\beta}_I \\ &= \frac{1}{2}\alpha^J_{ab}[e^a_\alpha(\partial_\mu e^\alpha_c) + \Gamma^\nu_{\mu\alpha}e^a_\nu e^\alpha_c]\alpha^{cb}_I \\ &= \frac{1}{2}\alpha^J_{ab}\alpha_{Ic}{}^b\omega^{ac}_\mu = \frac{1}{2}\alpha^{[a}{}_b{}^J\alpha^{c]b}{}_I\omega_{\mu ac} \\ &= \frac{1}{2}\omega_\mu{}^{ab}[{}^+\Sigma_{ab}]_I{}^J \end{aligned} \tag{B17}$$

where $[{}^+\Sigma_{ab}]_I{}^J$ is the generator of the (0,1) representation of the Lorentz group. Therefore, the covariant derivative acting on the field H_I^+ is given by

$$\nabla_\mu H_I^+ = \partial_\mu H_I^+ - \frac{1}{2}\omega_\mu^{ab}[{}^+\Sigma_{ab}]_I{}^J H_J^+. \tag{B18}$$

Using the curved-space version of property (B7), one concludes that the condition $\nabla_\mu\alpha_I^{\alpha\beta} = 0$ in turn leads to the condition $\nabla_\mu\alpha^I_{\alpha\beta} = 0$ for the dual space, yielding

$$\nabla_\mu H^I_+ = \partial_\mu H^I_+ - \frac{1}{2}\omega^{ab}_\mu [{}^+\Sigma_{ab}]^I{}_J H^J_+. \tag{B19}$$

Further useful equalities can be found. Looking at property (B6) in curved space, the above conditions imply that $\nabla_\mu h_{IJ} = 0$. A similar derivation shows that the covariant derivative of the tensors h^{IJ} or $h^I_J = h^{IK}h_{KJ}$ also vanishes. Finally, the covariant derivative of the totally antisymmetric tensors ϵ_{IJK} is zero. This is readily seen by noting from (B9) that $\nabla_\mu {}^+\Sigma^{\alpha\beta}_{IJ} = 0$. Recalling that $[{}^+\Sigma_{IJ}]^{\alpha\beta} = -i\epsilon_{IJK}\alpha^{\alpha\beta K}$, then one concludes $\nabla_\mu \epsilon_{IJK} = 0$.

By complex conjugating (B18), we obtain

$$\nabla_\mu H^{\dot I}_- = \partial_\mu H^{\dot I}_- - \frac{1}{2}\omega^{ab}_\mu [{}^-\Sigma_{ab}]^{\dot I}{}_{\dot J} H^{\dot J}_-, \tag{B20}$$

where $[{}^-\Sigma_{ab}]_{\dot I}{}^{\dot J}$ is the generator of the (1,0) representation of the Lorentz group. The tensors $h_{\dot I\dot J}$, $h^{\dot I\dot J}$, $\delta^{\dot I}_{\dot J}$, and $\epsilon_{\dot I\dot J\dot K}$ are also annihilated by ∇_μ.

3. 3+1 spacetime decomposition

A globally hyperbolic spacetime can always be foliated by a one-parameter family of spatial hypersurfaces Σ_t, $M \simeq \mathbb{R}\times\Sigma_t$ [37]. If we denote by n^μ the unit timelike vector field everywhere orthogonal to Σ_t, then $h_{\mu\nu} = g_{\mu\nu} - n_\mu n_\nu$ is the induced spatial metric on Σ_t.

We can now use the isomorphism $\alpha^{\mu\nu}_I$ defined in (B5) to build the following mixed tensors:

(i) $\gamma^\mu_I \coloneqq n_\nu \alpha^{\mu\nu}{}_I$ provides an isomorphism between complex vectors in V and (spatial) vectors in the tangent space of Σ_t.

(ii) $\gamma^\mu_{\dot I} \coloneqq n_\nu \bar\alpha^{\mu\nu}{}_{\dot I}$ is similar to the previous map replacing V by its complex-conjugated space $\bar V$.

(iii) $\gamma^{\dot I}_I \coloneqq \gamma^\mu_I \gamma^{\dot I}_\mu$ provides an isomorphism between V and $\bar V$.

(iv) $\epsilon^{I\mu\nu} \coloneqq \gamma^I_\beta n_\alpha \epsilon^{\alpha\beta\mu\nu}$ defines a totally antisymmetric, "purely spatial" tensor with mixed indices.

From the last definition one can derive an identity that will be useful in later calculations,

$$i2\epsilon^{I\mu\nu} = \alpha^{\mu\nu I} - \bar\alpha^{\mu\nu\dot J}\gamma^I_{\dot J}. \tag{B21}$$

As we have already mentioned, $\alpha^{\mu\nu}_I$ provides a one-to-one correspondence between self-dual tensors $\mathrm{F}^{\mu\nu}_+$ and elements $H^I_+ \in V$. We can now also build an isomorphism between self-dual tensors $\mathrm{F}^{\mu\nu}_+$ and purely spatial vectors in spacetime constructed as $H^\mu_+ \equiv n_\nu \mathrm{F}^{\mu\nu}_+$. Indeed,

$$H^\mu_+ \equiv n_\nu \mathrm{F}^{\mu\nu}_+ = n_\nu \alpha^{\mu\nu}_I H^I_+ = \gamma^\mu_I H^I_+. \tag{B22}$$

From the above definitions, and using (B7) and (B9), one can easily verify the following properties,

$$\gamma^\nu_I \gamma^{\beta I} = \alpha^{\mu\nu}_I n_\mu \alpha^{\rho\beta I} n_\rho = -n^\nu n^\beta + g^{\nu\beta} = h^{\nu\beta}, \tag{B23}$$

$$\gamma^\nu_I \gamma_{\nu J} = \alpha^\mu{}_{\nu I} n_\mu \alpha^{\rho\nu J} n_\rho = \alpha^{(\mu}{}_{\nu I} n_\mu \alpha^{\rho)\nu J} n_\rho = h_{IJ}. \tag{B24}$$

This shows that γ^ν_I indeed provides an isometry between spatial complex vectors in Σ_t and elements of V. Notice that $\nabla_\mu \gamma^\nu_I \neq 0$, but the spatial derivative of H^I_+ satisfies $D_\mu H^\mu_+ = D_I H^I_+$:

$$\begin{aligned} D_\mu H^\mu_+ &= h^{\mu\nu}\nabla_\mu H^+_\nu = h^{\mu\nu}\nabla_\mu(\gamma^I_\nu H^+_I) \\ &= \gamma^{\mu I}\nabla_\mu H^+_I + h^{\mu\nu}(\nabla_\mu \gamma^I_\nu)H^+_I \\ &= \gamma^{\mu I}\nabla_\mu H^+_I + h^{\mu\nu}\alpha^\sigma{}_\nu{}^I(n_\mu a_\sigma + K_{\mu\sigma})H^+_I \\ &= \gamma^{\mu I}\nabla_\mu H^+_I \equiv D_I H^I_+, \end{aligned} \tag{B25}$$

where we have used $\nabla_\mu n_\sigma = n_\mu a_\sigma + K_{\mu\sigma}$, a_σ is the 4-acceleration of the vector field n_ν, and $K_{\mu\nu} = K_{(\mu\nu)}$ is the extrinsic curvature of the three-dimensional submanifold Σ_t. Furthermore, if $D_I H^I_+ = 0$, then H^I_+ can be written as the "curl" of a complex potential, $H^I_+ = i\epsilon^{I\mu\nu}\nabla_\mu A^+_\nu$. Indeed,

$$\begin{aligned} D_I H^I_+ &= D_\alpha H^\alpha_+ = iD_\alpha \epsilon^{\alpha\mu\nu}\nabla_\mu A^+_\nu = i\epsilon^{\alpha\mu\nu} D_\alpha \nabla_\mu A^+_\nu \\ &= i\epsilon^{\alpha\mu\nu}\nabla_\alpha\nabla_\mu A^+_\nu \\ &\propto \epsilon^{\alpha\mu\nu} R_{\alpha\mu\nu\beta} A^\beta_+ = 0, \end{aligned} \tag{B26}$$

where $R_{\alpha\mu\nu\beta}$ is the Riemann tensor, and we have used (B25), (B22), and (B23).

4. Extended α^I_{ab} and its algebraic properties

This section provides some details regarding the "extended $\alpha^{\mu\nu}_I$-tensors," defined by extending the range of the index I to run from 0 to 3, and setting $\alpha^{\mu\nu}_0 = -g^{\mu\nu}$.

We begin by defining α in Minkowski spacetime. Let $\hat V \equiv V \oplus \mathbb{C}$, equipped with a Lorentzian flat metric η_{IJ}, be our "internal" vector space, where the indices I, J run from 0 to 3. The complex three-dimensional vector space V, defined in Appendix B 1, is now a subspace of $\hat V$. Let n_I denote a unit timelike vector ($\eta_{IJ}n^I n^J = 1$) orthogonal to the V subspace (i.e., $n^I m^I \eta_{IJ} = 0$, for all $m^J \in V$). It spans a one-dimensional vector space. The metric tensor in $\hat V$ can be written as $\eta_{IJ} = n_I n_J + h_{IJ}$, where h_{IJ} is the metric tensor in V used in Appendix B 1. Let X_I, Y_I, Z_I, n_I be an orthonormal basis of $\hat V^*$, the dual space of $\hat V$, with $n_I = \eta_{IJ}n^J$. We now define the extended tensor $\alpha^{\mu\nu}_I$ by extending expression (B5) as follows:

$$\alpha^{ab}_I \equiv \alpha^{ab}_1 X_I + \alpha^{ab}_2 Y_I + \alpha^{ab}_3 Z_I - \eta^{ab} n_I. \tag{B27}$$

Therefore, we have

$$\alpha^{ab}_I = \alpha^{ab}_J h^J_I - n_I \eta^{ab}, \tag{B28}$$

where $\alpha_J^{ab} h_I^J$ (the projection of α_I^{ab} on V) is simply the α_I^{ab} tensor used in the previous subsection, before extending the range of indices $I, J, K, \ldots$.

The tensor α_I^{ab} defined in (B27) maps vectors in $\hat{V}$ to tensors in Minkowski spacetime of the form

$$\alpha_I\colon\ H_+^I \to \alpha_I^{ab} H_+^I = \mathrm{F}_+^{ab} - H_+^0 \eta^{ab}. \tag{B29}$$

Here, F_+^{ab} is an antisymmetric self-dual tensor that in Minkowski spacetime transforms under the (0,1) irreducible representation of the Lorentz group, while H_+^0 is a scalar function.

Thus, the extended tensors α_I^{ab} map vectors in $\hat{V}$ to tensors in Minkowski spacetime that transform under Lorentz under the $(0,1) \oplus (0,0)$ representation.

The properties (B6)–(B9) must be replaced by

$$\alpha_{abI} \alpha^{ab}{}_J = 4\eta_{IJ}, \tag{B30}$$

$$\alpha_{ab}{}^I \alpha_{cdI} = 4\,{}^+P^{abcd} + \eta^{ab}\eta^{cd} = 4\,{}^-P^{acbd} + \eta^{ac}\eta^{bd}, \tag{B31}$$

$$\alpha_{abI} \bar{\alpha}^{ab}{}_{\dot{J}} = 4 n_I n_{\dot{J}}, \tag{B32}$$

$$\alpha^a{}_{bI} \alpha^{cb}{}_J = \eta_{IJ}\eta^{ac} - {}^+M_{IJ}^{ac}, \tag{B33}$$

where ${}^+M_{IJ}^{ac} \equiv {}^+\Sigma_{IJ}^{ac} + 2\alpha^{ab}{}_K h^K_{(I} n_{J)} = -4\,{}^+P^{ac}{}_{bd} \gamma^b_{(I} \gamma^d_{J)}$,[15] and ${}^+\Sigma_{IJ}^{ac}$ is the generator of the $(0,1) \oplus (0,0)$ representation of the Lorentz group.[16] From (B33), we obtain the commutation and anticommutation relations

$$\alpha^{[a}{}_{bI} \alpha^{c]b}{}_J = -{}^+M_{IJ}^{ab}, \tag{B34}$$

$$\alpha^{(a}{}_{bI} \alpha^{c)b}{}_J = \eta^{ab}\eta_{IJ}. \tag{B35}$$

The generalization of these properties to curved spacetimes is done, again, by using a vierbein or orthonormal tetrad $e_a^\mu(x)$, to write the relation between the curved spacetime α_I-matrices and the flat spacetime ones,

$$\alpha_I^{\mu\nu}(x) = e_a^\mu(x) e_b^\nu(x) \alpha_I^{ab}. \tag{B36}$$

The covariant derivative acting on the extended indices $I, J, K, \ldots$ can be determined following the arguments in Appendix B 2, but now we demand that ∇_μ annihilates the extended tensor $\alpha_I^{\mu\nu}$, $\nabla_\alpha \alpha_I^{\mu\nu} = 0$. As expected, the result is

$$\nabla_\mu H_I^+ = \partial_\mu H_I^+ - \frac{1}{2}\omega_\mu^{ab+} \Sigma_{abI}{}^J H_J^+, \tag{B37}$$

where ${}^+\Sigma_{abI}{}^J$ is the generator of the $(0,1) \oplus (0,0)$ Lorentz representation. The properties of the conjugate tensors $\bar{\alpha}_{ab}{}^{\dot{I}}$ are obtained in a similar way.

We finish this Appendix by deriving a few useful relations. First, from (B30) we obtain $\nabla_\mu \eta_{IJ} = 0$. Second, by acting with $g_{\mu\nu}\nabla_\alpha$ on Eq. (B28), and by using $\nabla_\alpha \alpha_I^{\mu\nu} = 0$, we obtain $\nabla_\rho n_I = 0$ (since $\eta_{IJ} = n_I n_J + h_{IJ}$, we also conclude that $\nabla_\mu h_{IJ} = 0$). Recalling (B28) again, this last property implies that the covariant derivative defined in this section also annihilates the projection of $\alpha_I^{\mu\nu}$ into V, namely, $\nabla_\alpha(\alpha_I^{\mu\nu} h_J^I) = 0$.

APPENDIX C: MAXWELL EQUATIONS IN CURVED SPACETIME

This Appendix shows the equivalence between the equations of motion for the potentials (3.25) and the fields (3.24).

First, we show that the equation for the potential $\bar{\alpha}_{\dot{I}}^{\mu\nu} \nabla_\mu A_{+\nu} = 0$ implies the equation for the field $\alpha_I^{\mu\nu} \nabla_\mu H_+^I = 0$. (We focus on self-dual fields; the derivation for anti-self-dual fields can be obtained by complex conjugation.) To prove this, notice first that using the identity (B21), the equation for the potential implies that $2i\epsilon^{I\alpha\beta} \nabla_\alpha A_{+\beta} = \alpha^{\mu\nu I} \nabla_\mu A_{+\nu}$. Then, recalling that $H_+^I \equiv i\epsilon^{I\mu\nu} \nabla_\mu A_\nu^+$, we see that when $A_{+\nu}$ satisfies the equations of motion, the relation between the field and the potential can be rewritten as $H_+^I = \frac{1}{2}\alpha^{\mu\nu I} \nabla_\mu A_{+\nu}$. Acting now with $\alpha_I^{\delta\rho}\nabla_\rho$, we obtain

$$\alpha_I^{\delta\rho} \nabla_\rho H_+^I = \frac{1}{2} \nabla_\rho \alpha_I^{\delta\rho} \alpha^{\mu\nu I} \nabla_\mu A_{+\nu} = -2\nabla_\rho \nabla^{[\delta} A_+^{\rho]} = 0, \tag{C1}$$

where we have used the fact that $-4\alpha_I^{\delta\rho}\alpha^{\mu\nu I}$ is a projector on self-dual fields and, in the last equality, all solutions of $\bar{\alpha}_{\dot{I}}^{\mu\nu} \nabla_\mu A_{+\nu} = 0$ are also solutions of the second-order equations $\nabla_\rho \nabla^{[\delta} A_+^{\rho]} = 0$ [remember the discussion below Eq. (3.25)].

Next, we want to show the reverse; i.e., starting from $\alpha_I^{\mu\nu} \nabla_\mu H_+^I = 0$, we want to show that there exists a potential $A_{+\nu}$, related to H_+^I by $H_+^I = i\epsilon^{I\mu\nu} \nabla_\mu A_{+\nu}$, that satisfies $\bar{\alpha}_{\dot{I}}^{\mu\nu} \nabla_\mu A_{+\nu} = 0$.

We begin by noticing that the identity $\nabla_\mu \alpha_I^{\alpha\beta} = 0$ allows us to write the field equations as $\nabla_\mu(\alpha_I^{\mu\nu} H_+^I) = 0$. Because $\alpha_I^{\mu\nu} H_+^I$ is a self-dual tensor, this equation implies that the two-form defined by $\mathrm{F}_{+\mu\nu} \equiv \alpha_{\mu\nu I} H_+^I$ is closed,[17] $d\mathrm{F}_+ = 0$.

[15] We use the same notation for $\gamma_I^\mu \coloneqq n_\nu \alpha_I^{\mu\nu}$ with the extended alpha tensors, as well as the rest of the mixed tensors in Appendix B 3. Its distinction is clear from the context.

[16] Note that we denote the generator of the $(0,1) \oplus (0,0)$ representation with the same symbol as the (0,1) generator that we used in the previous subsections; the $(0,1) \oplus (0,0)$ generator has more components than the (0,1) one; namely, the components corresponding to I or J equal 0. However, these components are all equal to zero; hence, we find it appropriate to use the same name for the two generators.

[17] Notice that for self-dual or anti-self-dual two-forms, $\nabla_\mu w^{\mu\nu} = 0$ if and only if $\nabla_\mu {}^\star w^{\mu\nu} = 0$, the latter formula being equivalent to $dw = 0$.

This allows the introduction of a potential one-form $A_{+\mu}$, $\mathrm{F}_+ = \mathrm{d}A_+$. Then, $\mathrm{d}A_+$ is self-dual; that is to say, the contraction of $\bar{\alpha}_{\dot{I}}^{\alpha\beta}$ and $\mathrm{d}A_+$ vanishes. But this is precisely the equation of motion we are looking for, $\bar{\alpha}_{\dot{I}}^{\alpha\beta}\nabla_\alpha A_{+\beta} = 0$. What remains is to prove that H_+^I and $A_{+\mu}$ are related by means of a curl. To see this, we notice that since $\mathrm{F}_{+\mu\nu} \equiv \alpha_{\mu\nu I} H_+^I$, we have $\alpha_{\mu\nu I} H_+^I = 2\nabla_{[\mu} A_{+\nu]}$. Multiplying both sides by $\alpha^{\mu\nu J}$ produces $H_+^J = \frac{1}{2}\alpha^{\mu\nu J}\nabla_{[\mu}A_{+\nu]}$. Now using the relation (B21) and the equation for $A_{+\beta}$, this relation reduces to $H_+^J = i\epsilon^{J\mu\nu}\nabla_\mu A_{+\nu}$, which is what we wanted to prove.

APPENDIX D: DERIVING EQUATIONS OF MOTION FROM FIRST-ORDER ACTION

In this Appendix we derive the equation of motion from the first-order action, described in Sec. IV A. We begin with the derivation of Eq. (4.2) from the action (4.1). Recall that in this action we have not yet introduced the Lorentz gauge, and the indices $I, J, \ldots$ and $\dot{I}, \dot{J}, \ldots$ take values 1,2,3.

From the form of the action (4.1),

$$S_M[A_+, A_-] = -\frac{1}{2}\int d^4x\sqrt{-g}[H_-^{\dot{I}}\bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu A_\nu^+ + H_+^I\alpha^{\mu\nu}{}_I\nabla_\mu A_\nu^-], \tag{D1}$$

we have

$$0 = \frac{\delta S_M}{\delta A_\nu^+} = \frac{1}{2}\bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu H_-^{\dot{I}} + \nabla_\mu \frac{i}{2}\epsilon^{I\mu\nu}\alpha^{\alpha\beta}{}_I\nabla_\alpha A_\beta^-. \tag{D2}$$

We now use the identity (B21) (note that $\nabla_\mu\epsilon^{I\mu\nu} \neq 0$) and (B7) to write

$$\begin{aligned} 0 &= \frac{1}{2}\bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu H_-^{\dot{I}} + \frac{1}{4}\alpha^{\mu\nu I}\alpha^{\alpha\beta}{}_I\nabla_\mu\nabla_\alpha A_\beta^- \\ &\quad - \frac{i}{2}\nabla_\mu\bar{\alpha}^{\mu\nu\dot{I}}\epsilon^{\alpha\beta}{}_{\dot{I}}\nabla_\alpha A_\beta^- - \frac{1}{4}\bar{\alpha}^{\mu\nu\dot{I}}\bar{\alpha}^{\alpha\beta}{}_{\dot{I}}\nabla_\mu\nabla_\alpha A_\beta^- \\ &= \frac{1}{2}\bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu H_-^{\dot{I}} - \frac{i}{2}\nabla_\mu\bar{\alpha}^{\mu\nu\dot{I}}\epsilon^{\alpha\beta}{}_{\dot{I}}\nabla_\alpha A_\beta^- \\ &\quad + [{}^+P^{\mu\nu}]^{\alpha\beta}\nabla_\mu\nabla_\alpha A_\beta^- - [{}^-P^{\mu\nu}]^{\alpha\beta}\nabla_\mu\nabla_\alpha A_\beta^-. \end{aligned} \tag{D3}$$

Recalling that $H_-^{\dot{I}} = -i\epsilon^{\dot{I}\mu\nu}\nabla_\mu A_\nu^\pm$, and using the Bianchi identity $\epsilon^{abcd}R_{bcde} = 0$, we get

$$\begin{aligned} 0 &= \bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu H_-^{\dot{I}} + \frac{i}{2}\epsilon^{\mu\nu\alpha\beta}\nabla_\mu\nabla_\alpha A_\beta^- \\ &= \bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu H_-^{\dot{I}} + \frac{i}{4}\epsilon^{\mu\nu\alpha\beta}R_{\mu\alpha\beta\sigma}A_-^\sigma \\ &= \bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu H_-^{\dot{I}}. \end{aligned} \tag{D4}$$

Finally, as shown in Appendix C, these equations are equivalent to $\alpha_I^{\mu\nu}\nabla_\mu A_\nu^- = 0$. Similarly, by differentiating the action with respect to A_-, we obtain $\alpha^{\mu\nu}{}_I\nabla_\mu H_+^I = 0$, which implies $\bar{\alpha}_{\dot{I}}^{\mu\nu}\nabla_\mu A_\nu^+ = 0$.

Now we again derive the equations of motion, but starting from the action that incorporates the Lorenz-gauge condition. As explained in Sec. IV A, this gauge condition is incorporated by extending the range of the indices $I, J, \ldots$ and $\dot{I}, \dot{J}, \ldots$ to take values 0,1,2,3, by introducing Lagrange multipliers $H_\pm^0$, and defining $\alpha_0^{\mu\nu} = -g^{\mu\nu}$. In order to take advantage of the calculation done a few lines above, we now keep the indices $I, J, \ldots$ and $\dot{I}, \dot{J}, \ldots$ running from 1 to 3, and explicitly write the Lorenz-gauge fixing term in the action:

$$\begin{aligned} S[A_\pm, H_\pm^0] = &-\frac{1}{2}\int dx^4\sqrt{-g}[H_-^{\dot{I}}\bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu A_\nu^+ + H_+^I\alpha^{\mu\nu}{}_I\nabla_\mu A_\nu^- \\ &- H_-^0\nabla_\mu A_+^\mu - H_+^0\nabla_\mu A_-^\mu]. \end{aligned} \tag{D5}$$

Variation with respect to $H_\pm^0$ provides the Lorenz-gauge condition: $\nabla_\mu A_\pm^\mu = 0$. Variation with respect to A_ν^+ yields

$$0 = \frac{\delta S}{\delta A_\nu^+} = \bar{\alpha}_{\dot{I}}^{\mu\nu}\nabla_\mu H_-^{\dot{I}} - \frac{1}{2}\nabla^\nu H_-^0. \tag{D6}$$

Let us first focus on the 0-component of this equation with respect to the (arbitrary) spacetime decomposition used to relate $H_\pm$ and $A_\pm^\mu$. This is done by contracting (D6) with the timelike vector n_ν. The term involving $H_-^{\dot{I}}$ vanishes [$D_{\dot{I}}H_-^{\dot{I}} = 0$ by construction, see (B26)], and we obtain $n^\mu\nabla_\mu H_-^0 \equiv \partial_t H_-^0 = 0$. On the other hand, acting with ∇_ν on Eq. (D6), we get $\Box H_-^0 = 0$; the term involving $H_-^{\dot{I}}$ again vanishes because $\bar{\alpha}_{\dot{I}}^{\mu\nu}\nabla_\nu\nabla_\mu H_-^{\dot{I}} = 0$.[18] Now, if both $\Box H_-^0 = 0$ and $\partial_t H_-^0 = 0$ hold, then $D_I D^I H_-^0 = 0$ holds. Choosing that H_-^0 vanishes at spatial infinity, one gets $H_-^0 = 0$. With this, Eq. (D6) reduces to $\bar{\alpha}_{\dot{I}}^{\mu\nu}\nabla_\mu H_-^{\dot{I}} = 0$, which is the correct equation of motion. An identical reasoning can be applied for H_+^0. Following the same arguments as in Appendix C, we can write the (extended) first-order equations of motion for the potentials as

$$\alpha_I^{\mu\nu}\nabla_\mu A_\nu^- = 0, \tag{D8}$$

with $I = 0$, 1, 2, 3, the 0 component being the Lorentz-gauge fixing.

[18]This last formula can be checked as follows:

$$\begin{aligned} \bar{\alpha}_{\dot{I}}^{\mu\nu}\nabla_\nu\nabla_\mu H_-^{\dot{I}} &= \nabla_\mu\nabla_\nu\bar{\alpha}_{\dot{I}}^{\mu\nu}H_-^{\dot{I}} = \nabla_\mu\nabla_\nu F_-^{\mu\nu} \\ &= \frac{1}{2}R_{\mu\nu\alpha}{}^\mu F_-^{\alpha\nu} + \frac{1}{2}R_{\mu\nu\alpha}{}^\nu F_-^{\mu\alpha} \\ &= \frac{1}{2}R_{\nu\alpha}F_-^{\alpha\nu} - \frac{1}{2}R_{\mu\alpha}F_-^{\mu\alpha} = 0. \end{aligned} \tag{D7}$$

APPENDIX E: DERIVING THE NOETHER CURRENT FROM FIRST-ORDER ACTION

In this section we derive the Noether current in first-order formalism by working directly with the variables A_+ and A_- and the action functional (4.1).

The variations of the Lagrangian density $\mathcal{L} = \mathcal{L}[A_+, A_-]$ under an infinitesimal electric-magnetic rotation of the potentials, $\delta A_\pm = \mp i\delta\theta A_\pm$, produces

$$\begin{aligned}\delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial A^+_\mu}\delta A^+_\mu + \frac{\partial\mathcal{L}}{\partial\nabla_\mu A^+_\nu}\delta\nabla_\mu A^+_\nu + \text{c.c.}\\ &= -\frac{1}{2}H^{\dot{I}}_-\bar{\alpha}^{\mu\nu}{}_{\dot{I}}(-i\delta\theta)\nabla_\mu A^+_\nu\\ &\quad -\frac{1}{2}i\epsilon^{I\mu\nu}\alpha^{\rho\sigma}{}_I\nabla_\rho A^-_\sigma(-i\delta\theta)\nabla_\mu A^+_\nu + \text{c.c.}\\ &= \frac{i\delta\theta}{2}H^{\dot{I}}_-\bar{\alpha}^{\mu\nu}{}_{\dot{I}}\nabla_\mu A^+_\nu + \frac{i\delta\theta}{2}H^I_+\alpha^{\rho\sigma}{}_I\nabla_\rho A^-_\sigma + \text{c.c.} = 0. \end{aligned} \quad \text{(E1)}$$

We find that, unlike in second-order formalism, the duality rotation leaves the Lagrangian invariant. The Noether current is now constructed as

$$\begin{aligned} j^\mu_D &= \frac{\partial\mathcal{L}}{\partial\nabla_\mu A^+_\nu}\delta A^+_\nu + \text{c.c.}\\ &= -\frac{1}{2}H^{\dot{I}}_-\bar{\alpha}^{\mu\nu}{}_{\dot{I}}(-i\delta\theta)A^+_\nu\\ &\quad -\frac{1}{2}i\epsilon^{I\mu\nu}\alpha^{\rho\sigma}{}_I\nabla_\rho A^-_\sigma(-i\delta\theta)A^+_\nu + \text{c.c.}\\ &= \frac{i\delta\theta}{2}[H^{\dot{I}}_-\bar{\alpha}^{\mu\nu}{}_{\dot{I}}A^+_\nu - H^I_+\alpha^{\mu\nu}{}_IA^-_\nu]\\ &\quad + \left[-\frac{\delta\theta}{2}\epsilon^{I\mu\nu}A^+_\nu\alpha^{\rho\sigma}{}_I\nabla_\rho A^-_\sigma + \text{c.c.}\right]. \end{aligned} \quad \text{(E2)}$$

This expression agrees with the result obtained in Sec. II, Eq. (2.5). Note that the last term in (E2) does not contribute to the associated Noether charge, and it is proportional to the equations of motion, vanishing on shell. It is not difficult to find that it agrees exactly with the last term in (2.5).

APPENDIX F: DEFINITION OF Ψ AND β^μ AND THEIR PROPERTIES

In this Appendix we define the fields Ψ introduced in Sec. IV B, as well as the matrices β^μ and β_5, and discuss their properties.

Given the complex potentials $A^\pm_\mu$ and the self-dual and anti-self-dual fields, H^I_+ and $H^{\dot{I}}_-$, we define the object

$$\Psi = \begin{pmatrix} A_{+\nu}\\ H^I_+\\ A^\nu_-\\ H_{-\dot{I}} \end{pmatrix}. \quad \text{(F1)}$$

Note that all four components of this object are related: A^-_μ is the complex conjugate of A^+_μ, and $H^I_+ = i\epsilon^{I\mu\nu}\nabla_\mu A^+_\nu$, $H^{\dot{I}}_-$ is the conjugate of H^I_+. Therefore, Ψ is the spin-1 analog of a Majorana spinor, whose upper and lower components are related by complex conjugation (Majorana fields represent real spinors with zero electric charge).

If we denote by X the vector space of all Ψ, we define the linear map $\beta^\mu : X \to X$ by

$$\beta^\mu\Psi = i\begin{pmatrix} \bar{\alpha}^\mu{}_{\nu\dot{I}}H^{\dot{I}}_-\\ -\alpha^\mu{}_\nu{}^IA^\nu_-\\ \alpha^{\mu\nu}_IH^I_+\\ -\bar{\alpha}^{\mu\nu}_{\dot{I}}A^+_\nu \end{pmatrix}, \quad \text{(F2)}$$

which is well defined for all Ψ. In matrix notation

$$\beta^\mu = i\begin{pmatrix} 0 & 0 & 0 & \bar{\alpha}^\mu{}_{\nu\dot{I}}\\ 0 & 0 & -\alpha^\mu{}_\nu{}^I & 0\\ 0 & \alpha^{\mu\nu}_I & 0 & 0\\ -\bar{\alpha}^{\mu\nu}_{\dot{I}} & 0 & 0 & 0 \end{pmatrix}. \quad \text{(F3)}$$

Now define the product of two β^μ as the composite operation $\beta^\mu\beta^\nu : X \to X$, defined by $(\beta^\mu\beta^\nu)\Psi = \beta^\mu(\beta^\nu\Psi)$. This is linear, and it leads to

$$\beta^\mu\beta^\nu = \begin{pmatrix} \bar{\alpha}^\mu{}_\alpha{}^{\dot{J}}\bar{\alpha}^{\nu\beta}_{\dot{J}} & 0 & 0 & 0\\ 0 & \alpha^\mu{}_\alpha{}^I\alpha^{\nu\alpha}_J & 0 & 0\\ 0 & 0 & \alpha^{\mu\alpha}_I\alpha^\nu{}_\beta{}^I & 0\\ 0 & 0 & 0 & \bar{\alpha}^{\mu\alpha}_{\dot{I}}\bar{\alpha}^\nu{}_\alpha{}^{\dot{J}} \end{pmatrix}. \quad \text{(F4)}$$

Using the properties (B34) and (B35) of the α_I matrices, one can easily write the symmetric and antisymmetric parts of this expression in μ and ν; in particular, the symmetric part produces the anticommutation relations written in Eq. (4.9).

Define the "chiral" matrix β_5 as a linear map $\beta_5 : X \to X$ by $\beta_5 \equiv \frac{i}{4!}\epsilon_{\mu\nu\alpha\beta}\beta^\mu\beta^\nu\beta^\alpha\beta^\beta$. Manipulating this expression we obtain

$$\beta_5 = \frac{i}{4!}\epsilon_{\mu\nu\alpha\beta}\beta^{[\mu}\beta^{\nu]}\beta^{[\alpha}\beta^{\beta]}$$
$$= \frac{i}{6}\epsilon_{\mu\nu\alpha\beta}\begin{pmatrix} 4[{}^{+}P^{\mu\nu}]_{\sigma}{}^{\rho}[{}^{+}P^{\alpha\beta}]_{\rho}{}^{\delta} & 0 & 0 & 0 \\ 0 & \frac{1}{4}[{}^{+}M^{\mu\nu}]^{I}{}_{K}[{}^{+}M^{\alpha\beta}]^{K}{}_{J} & 0 & 0 \\ 0 & 0 & 4[{}^{-}P^{\mu\nu}]^{\sigma}{}_{\rho}[{}^{-}P^{\alpha\beta}]^{\rho}{}_{\delta} & 0 \\ 0 & 0 & 0 & \frac{1}{4}[{}^{-}M^{\mu\nu}]_{\dot{I}}{}^{\dot{K}}[{}^{-}M^{\alpha\beta}]_{\dot{K}}{}^{\dot{J}} \end{pmatrix}$$
$$= \begin{pmatrix} -g_{\sigma}{}^{\rho} & 0 & 0 & 0 \\ 0 & -h^{I}{}_{J} & 0 & 0 \\ 0 & 0 & g^{\sigma}{}_{\rho} & 0 \\ 0 & 0 & 0 & h_{\dot{I}}{}^{\dot{J}} \end{pmatrix}. \tag{F5}$$

The tensor ${}^{\pm}P$ is the projector on (anti-)self-dual tensors defined below (B9), and the tensors $[{}^{\pm}M^{\alpha\beta}]$ were defined below Eq. (B33). In the above calculation, we have used the self-duality property, $\pm i{}^{\star}P^{\pm} = P^{\pm}$. The map β_5 has the following properties:

$$\beta_5^2 = \begin{pmatrix} g_{\sigma}{}^{\rho} & 0 & 0 & 0 \\ 0 & h^{I}{}_{J} & 0 & 0 \\ 0 & 0 & g^{\sigma}{}_{\rho} & 0 \\ 0 & 0 & 0 & h_{\dot{I}}{}^{\dot{J}} \end{pmatrix}, \qquad \{\beta_5, \beta^{\mu}\} = 0. \tag{F6}$$

A duality transformation can be implemented by means of the linear operation $T_{\theta}: X \to X$, with $T_{\theta} = e^{i\theta\beta_5}$, $\theta \in \mathbb{R}$.

Let X^* be the dual space, the space of linear functionals over X. Given $\Psi \in X$ as in (F1), we define $\bar{\Psi} \in X^*$ by

$$\bar{\Psi} := (\, A_{+}^{\nu} \quad H_{I}^{+} \quad A_{\nu}^{-} \quad H_{-}^{\dot{I}}\,). \tag{F7}$$

The action functional (4.7) is thus a well-defined quantity. To construct a Hilbert space from X, we need to endow it with an inner product. Note that, while the product $\bar{\Psi}\Psi \in \mathbb{C}$ is well defined, it does not produce a positive real number. We can define a (positive-definite) inner product as follows:

$$\langle \Psi_1, \Psi_2 \rangle = \alpha \int d^4x \sqrt{-g}\bar{\Psi}_1 \delta \Psi_2, \tag{F8}$$

where α is an arbitrary positive real constant with dimensions of inverse action (see footnote 13), and $\delta: X \to X$ is a linear application defined by

$$\delta\Psi = \begin{pmatrix} A_{\nu}^{-} \\ H_{-}^{I} \\ A_{+}^{\nu} \\ H_{\dot{I}}^{+} \end{pmatrix} \tag{F9}$$

which in matrix notation reads

$$\delta = \begin{pmatrix} 0 & 0 & \delta_{\nu}^{\mu} & 0 \\ 0 & 0 & 0 & \gamma_{\dot{I}}^{I} \\ \delta_{\mu}^{\nu} & 0 & 0 & 0 \\ 0 & \gamma_{\dot{I}}^{I} & 0 & 0 \end{pmatrix} \tag{F10}$$

($\gamma_{I\dot{I}}$ was defined in Appendix B 3). This operation is useful since $\bar{\Psi}\delta\Psi \geq 0$. By expanding the fields as in (F7) and (F9), one checks that expression (F8) is real, and in particular $\langle \Psi_1, \Psi_2 \rangle = \langle \Psi_2, \Psi_1 \rangle$. Linearity with the second variable is trivial.

The analog of this product for the Dirac field is commonly written simply as $\langle \Psi_1, \Psi_2 \rangle = \alpha \int d^4x \sqrt{-g}\Psi_1^{\dagger}\Psi_2$, where the matrix $\delta (= \gamma^0)$ is implicit in $\Psi^{\dagger} \equiv \bar{\Psi}\delta$ to simplify the notation (see, e.g., Ref. [16]). Note, however, that the presence of δ is required in order to make the operation well defined regarding the position of indices. We use the product (F8) in Sec. V B.

APPENDIX G: DETAILS IN THE CALCULATION OF THE ELECTROMAGNETIC DUALITY ANOMALY

This Appendix provides details of the intermediate steps summarized in Sec. V regarding the computation of $\langle \nabla_{\mu} j_D^{\mu} \rangle_{\rm ren}$. In that section we needed to compute

$$\langle \nabla_{\mu} j_D^{\mu} \rangle = \lim_{\substack{s\to 0 \\ x\to x'}} \frac{1}{2} s \mathrm{Tr}[\beta_5 S(x, x', s)], \tag{G1}$$

where $S(x, x', s)_{\rm Ad(4)} = [(D_x - s)G(x, x', s)]_{\rm Ad(4)}$, and with the asymptotic expansion in (5.5). There is no need to explicitly find the asymptotic expansion of $\Delta^{1/2}(x, x')$, $\sigma(x, x')$, and $E_k(x, x')$ in the short-distance limit.

We show first that the derivative term, $D_x G(x, x', s)$, does not contribute to $\langle \nabla_{\mu} j_D^{\mu} \rangle$. From this contribution one only has to consider the $k=0$, 1 terms in the sum (5.5) since the term with $k=2$ is of adiabatic order five. The action of the derivative on $G(x,x',s)$ produces three contributions: one that goes with $\nabla^x_\mu \Delta^{1/2}(x,x')$, another with $\nabla^x_\mu \sigma(x,x')$, and another with $\nabla^x_\mu E_k(x,x')$. The first two are multiplied by $\mathrm{Tr}\{\beta^\mu \beta_5 E_k(x)\}$, and this quantity vanishes for both $k=0$, 1. Regarding the contribution of $\nabla^x_\mu E_k(x,x')$, it vanishes because of the limit $s\to 0$. To see this, notice that for $k=0$, 1, the factor $\nabla^x_\mu E_k(x,x')$ appears multiplied in the sum (5.5) by the following contributions, respectively,

$$\int_0^\infty d\tau e^{-i(\tau s^2+\frac{\sigma(x,x')}{2\tau})}(i\tau)^2 = \frac{2i}{\sigma(x,x')} + O(s^2) \quad \text{(G2)}$$

$$\int_0^\infty d\tau e^{-i(\tau s^2+\frac{\sigma(x,x')}{2\tau})}(i\tau) = 2i\log s + O(s^0) \quad \text{(G3)}$$

so the limit $s\to 0$ in (G1) vanishes.

We show now that the other term contributing to $S(x,x',s)_{\mathrm{Ad}(4)}$, $sG(x,x',s)_{\mathrm{Ad}(4)}$, only provides a nonzero result by means of the $k=2$ term in the asymptotic sum (5.5). First notice that the limit $x\to x'$ can be safely taken. On the other hand, higher values of k in (5.5) provide contributions of more than 4 derivatives of the metric to $S(x,x',s)$, so they are of higher adiabatic order. The $k=0$ case vanishes because it is proportional to $\mathrm{Tr}\{\beta_5 E_0(x)\} = \mathrm{Tr}\{\beta_5\} = 0$. The $k=1$ term does not contribute either because it is proportional to $\mathrm{Tr}\{\beta_5 E_1(x)\} = \mathrm{Tr}\{\beta_5 \mathcal{Q}\}$, and[19]

$$\begin{aligned}\mathrm{Tr}(\beta_5 \mathcal{Q}) &= -2iR_{\mu\nu\alpha\beta}\mathrm{TrIm}\left[{}^{+}\mathrm{P}^{\mu\nu}\Sigma^{\alpha\beta} - \frac{1}{4}{}^{+}\mathrm{M}^{\mu\nu-}\Sigma^{\alpha\beta}\right] \\ &= \frac{1}{2}iR_{\mu\nu\alpha\beta}\mathrm{TrIm}[{}^{+}\mathrm{M}^{\mu\nu-}\Sigma^{\alpha\beta}] \\ &= -2iR_{\mu\nu IJ}\epsilon^{\mu\nu a I} = 2iR_{\mu\nu\alpha\rho}\epsilon^{\mu\nu\alpha\sigma}n^\rho n_\sigma = 0.\end{aligned}$$

Then, it only remains to calculate the $k=2$ term in the asymptotic sum (5.5),

$$\begin{aligned}\langle \nabla_\mu j^\mu_D\rangle &= \frac{i\hbar}{32\pi^2}\mathrm{Tr}(\beta_5 \mathrm{E}_2) \\ &= \frac{i\hbar}{32\pi^2}\left[\frac{1}{12}\mathrm{Tr}(\beta_5 \mathrm{W}_{\mu\nu}\mathrm{W}^{\mu\nu}) + \frac{1}{2}\mathrm{Tr}(\beta_5 \mathcal{Q}^2)\right] \quad \text{(G4)}\end{aligned}$$

with $W_{\mu\nu} \equiv [\nabla_\mu, \nabla_\nu]$ given in (4.15) and

$$\begin{aligned}\mathcal{Q}\Psi &\equiv \frac{1}{2}\beta^{[\alpha}\beta^{\mu]}W_{\alpha\mu}\Psi \\ &= -\frac{1}{2}R_{\mu\nu\alpha\beta}\begin{pmatrix} -2\,{}^{+}P^{\mu\nu}\Sigma^{\alpha\beta} & 0 & 0 & 0 \\ 0 & \frac{1}{2}{}^{+}M^{\mu\nu+}\Sigma^{\alpha\beta} & 0 & 0 \\ 0 & 0 & -2\,{}^{-}P^{\mu\nu}\Sigma^{\alpha\beta} & 0 \\ 0 & 0 & 0 & \frac{1}{2}{}^{-}M^{\mu\nu-}\Sigma^{\alpha\beta}\end{pmatrix}\Psi\end{aligned}$$

where $\Sigma^{\alpha\beta}{}_{\mu\nu}$ is the generator of the $(1/2,1/2)$ representation of the Lorentz group, ${}^{+}\Sigma^{\alpha\beta}{}_{IJ}$ is the generator of the $(0,1)\oplus(0,0)$ representation, ${}^{-}\Sigma^{\alpha\beta}{}_{ij}$ is the generator of $(1,0)\oplus(0,0)$, and ${}^{\pm}P_{abcd} = \frac{1}{4}(\eta_{ac}\eta_{bd} - \eta_{ad}\eta_{bc} \pm i\epsilon_{abcd})$. A lengthy but straightforward computation produces

$$\mathrm{Tr}(\beta_5 W_{\mu\nu}W^{\mu\nu}) = -2iR_{\mu\nu\alpha\beta}{}^\star R^{\mu\nu\alpha\beta}, \quad \text{(G5)}$$

$$\mathrm{Tr}(\beta_5 \mathcal{Q}^2) = iR^{\mu\nu\alpha\beta\star}R_{\mu\nu\alpha\beta}. \quad \text{(G6)}$$

With this, we obtain

$$\mathrm{Tr}(\beta_5 E_2) = i\frac{1}{3}R^{\mu\nu\alpha\beta\star}R_{\mu\nu\alpha\beta}. \quad \text{(G7)}$$

[19] In this calculation we use the relation ${}^{+}P_{abcd} = -\frac{1}{4}[\Sigma_{abcd} + i({}^\star\Sigma_{ab})_{cd}]$ and the Bianchi identity $R_{\mu\nu\alpha\beta}\epsilon^{\mu\nu\alpha\rho} = 0$ (several times).

[1] E. Noether, Invariante variationsprobleme, *Nachr. v. d. Ges. d. Wiss. zu Gottingen, Math-Phys.* (Weidmann Bookstore, Berlin, 1918).

[2] M. G. Calkin, An invariance property of the free electromagnetic field, Am. J. Phys. **33**, 958 (1965).

[3] S. Deser and C. Teitelboim, Duality transformations of Abelian and non-Abelian gauge fields, Phys. Rev. D **13**, 1592 (1976).

[4] S. Deser, Off-shell electromagnetic duality invariance, J. Phys. A **15**, 1053 (1982).

[5] R. P. Cameron, S. M. Barnett, and A. M. Yao, Duplex symmetry and its relation to the conservation of optical helicity, Phys. Rev. A **86**, 013845 (2012).
[6] S. L. Adler, Axial-vector vertex in spinor electrodynamics, Phys. Rev. **177**, 2426 (1969).
[7] J. S. Bell and R. Jackiw, A PCAC puzzle: $\pi_0 \to \gamma\gamma$ in the σ-model, Il Nuovo Cimento A **60**, 47 (1969).
[8] T. Kimura, Divergence of axial-vector current in the gravitational field, Prog. Theor. Phys. **42**, 1191 (1969).
[9] R. Delbourgo and A. Salam, The gravitational correction to PCAC, Phys. Lett. B **40**, 381 (1972).
[10] T. Eguchi and P. G. O. Freund, Quantum Gravity and World Topology, Phys. Rev. Lett. **37**, 1251 (1976).
[11] T. Eguchi, P. B. Gilkey, and A. J. Hanson, Gravitation, gauge theories and differential geometry, Phys. Rep. **66**, 213 (1980).
[12] F. Bastianelli and P. van Nieuwenhuizen, *Path Integrals and Anomalies in Curved Space*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge, England, 2006).
[13] M. Nakahara, *Geometry, Topology and Physics*, Graduate Student Series in Physics (Hilger, Bristol, 2003), p. 505.
[14] M. D. Schwartz, *Quantum Field Theory and the Standard Model* (Cambridge University Press, Cambridge, England, 2014).
[15] K. Fujikawa, Path-Integral Measure for Gauge-Invariant Fermion Theories, Phys. Rev. Lett. **42**, 1195 (1979).
[16] K. Fujikawa, Path integral for gauge theories with fermions, Phys. Rev. D **21**, 2848 (1980).
[17] I. Agullo, A. del Rio, and J. Navarro-Salas, Electromagnetic Duality Anomaly in Curved Spacetimes, Phys. Rev. Lett. **118**, 111301 (2017).
[18] K. S. Thorne, C. W. Misner, and J. A. Wheeler, *Gravitation* (W. H. Freeman and Co., San Francisco, 1973).
[19] A. D. Dolgov, I. B. Khriplovich, A. I. Vainshtein, and V. I. Zakharov, Photonic chiral current and its anomaly in a gravitational field, Nucl. Phys. **B315**, 138 (1989).
[20] M. Reuter, Chiral anomaly of antisymmetric tensor fields, Phys. Rev. D **37**, 1456 (1988).
[21] A. Landete, I. Agullo, and J. Navarro-Salas, Electric-magnetic duality and renormalization in curved spacetimes, Phys. Rev. D **90**, 124067 (2014).
[22] S. Weinberg, Feynman rules for any spin. II. Massless particles, Phys. Rev. **134**, B882 (1964).
[23] J. S. Dowker and Y. P. Dowker, Particles of arbitrary spin in curved spaces, Proc. Phys. Soc. **87**, 65 (1966).
[24] R. Penrose, Zero rest-mass fields including gravitation: Asymptotic behaviour, Proc. R. Soc. A **284**, 159 (1965).
[25] R. Penrose and W. Rindler, *Spinors and Space-Time* Vol. 1, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge, England, 1984).
[26] A. Ashtekar, A note on helicity and self-duality, J. Math. Phys. (N.Y.) **27**, 824 (1986).
[27] R. Geroch, Spinor structure of spacetimes in general relativity. I, J. Math. Phys. (N.Y.) **9**, 1739 (1968).
[28] A. Ashtekar and R. S. Tate, *Lectures on Non-Perturbative Canonical Gravity*, Advanced Series in Astrophysics and Cosmology (World Scientific, Singapore, 1991).
[29] C. M. Bender, Free massless fields as infinite-dimensional representations of the Lorentz group, Phys. Rev. **168**, 1809 (1968).
[30] I. G. Moss and D. J. Toms, Invariants of the heat equation for non-minimal operators, J. Phys. A **47**, 215401 (2014).
[31] L. Parker and D. J. Toms, *Quantum Field Theory in Curved Spacetime: Quantized Fields and Gravity* (Cambridge University Press, Cambridge, England, 2009).
[32] D. V. Vassilevich, Heat kernel expansion: User's manual, Phys. Rep. **388**, 279 (2003).
[33] G. 't Hooft, Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking, Nato Sci. Ser. B **59**, 135 (1980).
[34] S. M. Christensen and S. A. Fulling, Trace anomalies and the Hawking effect, Phys. Rev. D **15**, 2088 (1977).
[35] N. H. Christ, A. S. Blaer, and J.-F. Tang, Anomalous Fermion Production by a Julia-Zee Dyon, Phys. Rev. Lett. **47**, 1364 (1981).
[36] I. Agullo, A. del Rio, and J. Navarro-Salas, Gravity and handedness of photons, Int. J. Mod. Phys. D **26**, 1742001 (2017).
[37] R. M. Wald, *General Relativity* (Chicago University Press, Chicago, IL, 1984).