

Quantization of general linear electrodynamics

Sergio Rivera and Frederic P. Schuller

Albert Einstein Institute,

Max Planck Institute for Gravitational Physics,

Am Mühlenberg 1, 14476 Potsdam, Germany

(Dated: January 4, 2011)

General linear electrodynamics allow for an arbitrary linear constitutive relation between the field strength two-form and induction two-form density if crucial hyperbolicity and energy conditions are satisfied, which render the theory predictive and physically interpretable. Taking into account the higher-order polynomial dispersion relation and associated causal structure of general linear electrodynamics, we carefully develop its Hamiltonian formulation from first principles. Canonical quantization of the resulting constrained system then results in a quantum vacuum which is sensitive to the constitutive tensor of the classical theory. As an application we calculate the Casimir effect in a bi-refrangent linear optical medium.

1. INTRODUCTION

Classical electromagnetism can be formulated on much more general optical backgrounds than the familiar ones described in terms of Lorentzian manifolds. From the point of view of electrodynamics, this is because one merely needs a constitutive law that links the electromagnetic field strength two-form F with the induction two-form density H , and thus closes the relations

$$(dF)_{\alpha\beta\gamma} = 0, \quad (dH)_{\alpha\beta\gamma} = \epsilon_{\alpha\beta\gamma\delta} j^\delta, \quad (1)$$

which any electromagnetic theory featuring charge conservation and no magnetic monopoles in four dimensions must satisfy in the presence of a current vector field density j . This point has been made most prominently and lucidly by [1]. Even if one restricts attention to linear constitutive laws [48], new the resulting electrodynamic theories will generically feature birefringence, meaning that distinguished polarizations of light will travel at different speeds. Now the most general action for an electromagnetic gauge potential that results in a linear constitutive law, and which we will carefully quantize in this paper, is

$$S[A, G] = -\frac{1}{8} \int dx^4 \omega_G \left[F_{\alpha\beta} F_{\gamma\delta} G^{\alpha\beta\gamma\delta} + j^\alpha A_\alpha \right], \quad (2)$$

where G is a smooth covariant rank four tensor field with the symmetries $G_{\alpha\beta\gamma\delta} = G_{\gamma\delta\alpha\beta}$ and $G_{\alpha\beta\gamma\delta} = -G_{\beta\alpha\gamma\delta}$, and which is invertible in the sense that there is a smooth contravariant tensor

field $G^{\alpha\beta\gamma\delta}$ so that $G^{\alpha\beta\rho\sigma}G_{\rho\sigma\gamma\delta} = 2(\delta_\gamma^\alpha\delta_\delta^\beta - \delta_\delta^\alpha\delta_\gamma^\beta)$ and there is a well-defined volume form ω_G for such area metric tensors [2]. The birefringence of such general linear electrodynamics is encoded in its dispersion relation, or equivalently the causal structure, of the associated field equations [3]. This dispersion relation is known [4] to be of higher polynomial order, and indeed the central challenge faced in this paper is to properly deal with this fact, both in the classical and quantum analysis. The importance of understanding Maxwell theory on such general linear backgrounds is that the latter comprehensively describe all linear optical backgrounds ranging from fundamental spacetime geometries beyond Lorentzian geometry [3, 5–8] over the effective spacetime structure seen by photons to first order quantum corrections in a curved Lorentzian spacetime [9] to all non-dissipative linear optical media available in the laboratory [8].

The present article develops the canonical quantization of these most general linear electrodynamics from first principles, and arrives at an explicit calculation of the quantum vacuum of the theory. We show that the related Casimir effect detects deviations from a non-birefringent background with an amplification which in principle is limited only by technological constraints. Arriving at these results requires special care when obtaining the Hamiltonian formulation of the classical theory that precedes the actual quantization. While quite generally the Dirac-Bergmann quantization procedure of course also applies to these gauge field dynamics, the key issue is the question of which hypersurfaces provide viable initial data surfaces on which the canonical phase space variables can be defined and evolved by the Hamiltonian. It is precisely this question that makes the problem of quantization of the dynamics (2) so subtle, and requires the conceptually robust understanding of its causal structure developed in [3] and concisely summarized in section 2. Only when using the insights gained there, can the formulation of the Hamiltonian picture in section 3 and the canonical quantization in section 7 proceed as usual, based on the derivation of the Dirac brackets in section 4 and the diagonalization of the Hamiltonian in section 6, which is particularly simple for the area metrics in a neighborhood of Lorentzian metric geometries, as shown in section 5. However, having gone through the laborious quantization procedure, one is rewarded in section 8 with the said method to measure deviations from a metric-induced background through the Casimir effect in particular, and a demonstration of how to quantize field theories with higher-order polynomial dispersion relations [10–27] in general.

For simplicity, we restrict attention to flat area metric manifolds throughout the paper. Analogous to any other geometric structure on a smooth manifold, an area metric is called flat if there exists a set of charts covering the underlying smooth manifold such that the components of the area metric tensor are constant within each such chart.

2. CAUSAL STRUCTURE OF LINEAR ELECTRODYNAMICS

The Hamiltonian formulation of the dynamics (2), on which the canonical quantization will be built, hinges on several key results of the associated causality theory. Here we summarize the central results of practical importance. For a detailed derivation of these results we refer the reader to [3]. A necessary condition for Maxwell theory on a four-dimensional area metric background to be predictive is that the following polynomial [4] on covectors k ,

$$P(k) = -\frac{1}{24}\epsilon_{\rho\sigma\tau\epsilon}\epsilon_{\mu\nu\omega\vartheta}G^{\rho\sigma\mu\alpha}G^{\beta\tau\nu\gamma}G^{\delta\epsilon\omega\vartheta}k_{\alpha}k_{\beta}k_{\gamma}k_{\delta}, \quad (3)$$

is hyperbolic. This means that there is at least one covector h with $P(h) \neq 0$ such that for every covector q the polynomials

$$P_{q,h}(\lambda) = P(q + \lambda h) \quad (4)$$

have only real roots λ , in which case h is said to be a hyperbolic covector with respect to P . That hyperbolicity is a necessary criterion for a well-posed initial value problem is a central result of the theory of partial differential equations [28, 29]. For the flat area metric manifolds discussed here, the hyperbolicity of P is even sufficient for the predictivity of the theory [28]. Initial data, given on a hypersurface whose normal covectors are all hyperbolic with respect to P , are then uniquely evolved away from the hypersurface. Thus a Hamiltonian formulation of the dynamics, which deals precisely with the evolution between initial data surfaces, must be based on a foliation $\{t, x^a\}$ of the manifold whose leaves $t = \text{const}$ are hypersurfaces with hyperbolic conormal.

However, this requirement needs to be sharpened if one requires that the actual initial data can be collected by observers. The definition of observers now hinges on the so-called dual polynomial $P^{\#}$ [30], which for those P that arise from area metrics by virtue of (3) and which admit hyperbolic covectors, can be calculated explicitly and takes the deceptively simple form

$$P^{\#}(X) = -\frac{1}{24}\epsilon^{\rho\sigma\tau\epsilon}\epsilon^{\mu\nu\omega\vartheta}G_{\rho\sigma\mu\alpha}G_{\beta\tau\nu\gamma}G_{\delta\epsilon\omega\vartheta}X^{\alpha}X^{\beta}X^{\gamma}X^{\delta}, \quad (5)$$

which sends any tangent vector X to a real number. That the dual polynomial $P^{\#}$ can be calculated analytically at all, and takes such a comparatively simple form, is only due to an interplay of the area metric structure underlying it and the necessary hyperbolicity of P . While the hyperbolic covectors of P distinguish admissible initial data surfaces, admissible observers are distinguished by their worldline tangent vectors being hyperbolic vectors of $P^{\#}$. In other words, the very existence of observers restricts the admissible area metric geometries further to those where also $P^{\#}$ is

hyperbolic. But exactly this hyperbolicity of $P^\#$ then allows to make a choice of time-orientation, which in turn implies a choice of positive energy. More precisely, a time-orientation is chosen by picking one connected set of all hyperbolic tangent vectors, a so-called hyperbolicity cone $C^\#$, out of the several such connected components defined by $P^\#$. But then the covectors q for which all future-directed observers measure positive energy, $q(X) > 0$ for all $X \in C^\#$, themselves constitute a cone $(C^\#)^+$ in cotangent space, which thus deserves to be called the positive energy cone with respect to the chosen time-orientation. The latter, in turn, selects the (open and convex) cone C of hyperbolic covectors of P that lie within the positive energy cone $(C^\#)^+$. For technical convenience we require, without loss of generality, that P be positive on all of C ; indeed, from (3) it is clear that this always can be arranged for by switching the overall sign of G .

Besides the hyperbolicity of P and $P^\#$, one finally needs to require that there exists a time orientation such that any non-zero P -null covector lies either in $(C^\#)^+$ or $-(C^\#)^+$. In other words, the energy of any massless momentum is to have a definite sign upon which all observers agree. If and only if this bi-hyperbolicity and energy distinguishing properties are met, it is justified to call the underlying area metric manifold an area metric *spacetime*, and we will consider only such. For an illustration in a typical case, see figure 1, and for a detailed exposition of these concepts, see [3].

The final piece of technology concerns the duality map between covectors and vectors in an area metric spacetime. The map

$$L : C \rightarrow L(C), \quad L(q) = \frac{P(q, q, q, \cdot)}{P(q, q, q, q)} \quad (6)$$

is shown in [3] to be a well-defined and invertible Legendre map precisely because P is assumed to be bi-hyperbolic and energy-distinguishing. Spacelike hypersurfaces are meaningfully defined as those having tangent directions that are purely spatial with respect to some observer. More precisely, the spacelike hypersurfaces are those whose conormals lie in $L^{-1}(C^\#)$. But since it can be shown that $L^{-1}(C^\#)$ always lies within C , the condition that a hypersurface be spacelike (and thus initial data on it accessible by local observers) further sharpens the condition for a feasible initial data surface for the dynamics (2) we identified before. Thus only a foliation (t, x^1, x^2, x^3) of an area metric spacetime into spacelike hypersurfaces $t = \text{const}$ (which then contain all vectors that are annihilated by the covector $L^{-1}(\partial/\partial t)$, see figure 1) provides an appropriate temporal-spatial split for the Hamiltonian formulation. For the flat area metric spacetimes considered here, one may further choose the coordinates such that $L^{-1}(\partial/\partial t)$ is conormal to the spacelike hypersurfaces. In other words, one can choose trivial shift and lapse in the flat case.

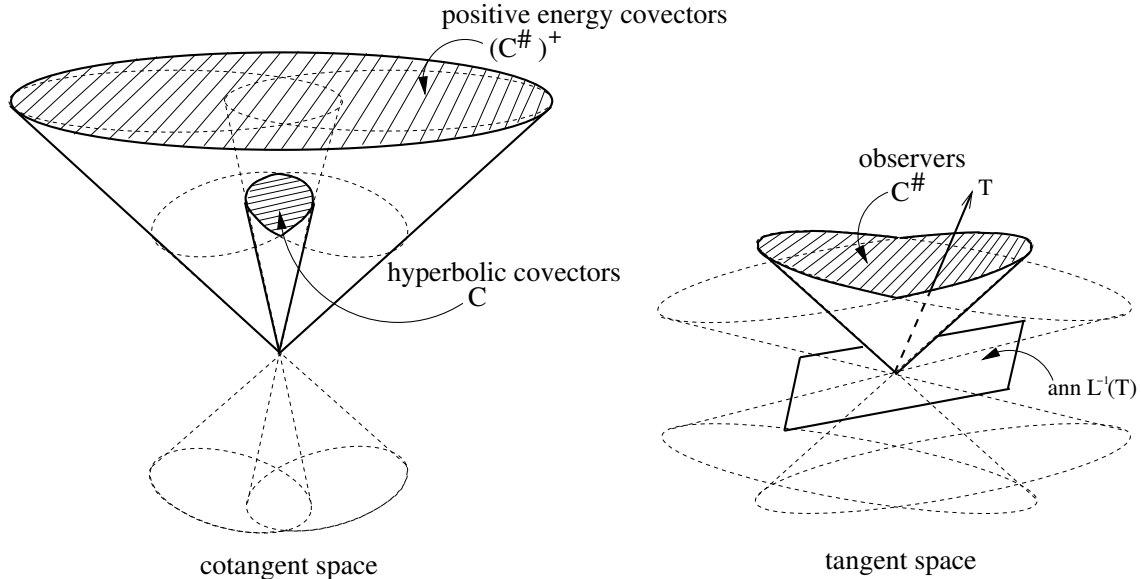


FIG. 1: Causal structure of a typical bi-hyperbolic and energy-distinguishing polynomial dispersion relation. Dotted surfaces indicate sets of covectors and vectors that are null with respect to the polynomial P and its dual $P^\#$, respectively. The cone C of hyperbolic covectors and the cone $C^\#$ of observers both arise as hyperbolicity cones. Purely spatial directions, as seen by an observer with worldline tangent T , are those vectors annihilated by the preimage of T under the Legendre map L .

The reader may find it helpful to get a feel for these seemingly abstract conditions for the special case where the area metric is induced by a metric g by virtue of $G_{\alpha\beta\gamma\delta} = g_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}g_{\beta\gamma}$. Precisely the same conceptual steps force one then to take the metric g to be of Lorentzian signature (otherwise Maxwell theory would not be well-posed). Since in this metric-induced case $P(k) = (g^{\alpha\beta}k_\alpha k_\beta)$ as usual, we have $L(C) = C^\#$, and thus one recovers the standard Lorentzian notions of observers and spacelike hypersurfaces. However, the general construction presented before does not justify itself from this reduction to the metric case. The general treatment rather demonstrates the appropriateness and consistency of the standard Lorentzian definitions from a conceptual point of view.

3. HAMILTONIAN FORMULATION

With the appropriate foliation (t, x^1, x^2, x^3) of the area metric spacetime with spacelike leaves for constant time t and conormals given by $L^{-1}(\partial/\partial t)$, as constructed in the previous section, we are now in a position to develop the Hamiltonian formulation of the dynamics encoded in the action (2). For a flat area metric spacetime, one can choose coordinates not only such that

the area metric has constant components throughout those charts, but also that additionally the components of the volume form ω_G featuring in the action are numerically identical to those of the totally antisymmetric Levi-Civita symbol $\epsilon_{\alpha\beta\gamma\delta}$ defined by $\epsilon_{0123} = +1$. In such a coordinate system, we obtain the canonical momenta associated with the field variables (A_0, A_i) from the Lagrangian density \mathcal{L} of the action (2) as

$$\begin{aligned}\pi^0 &= \frac{\delta\mathcal{L}}{\delta(\partial_0 A_0)} = 0, \\ \pi^i &= \frac{\delta\mathcal{L}}{\delta(\partial_0 A_i)} = -G^{0i0j}\partial_0 A_j - G^{0ij0}\partial_j A_0 - G^{0ijk}\partial_j A_k,\end{aligned}\tag{7}$$

where here, and for the remainder of the paper, latin indices range from 1 to 3, while greek indices continue to range from 0 to 3. In the language of the theory for constrained systems [31, 32], we thus identify $\phi_1 = \pi^0 \approx 0$ as a primary constraint of the dynamics. Defining the matrix M_{ij} with the property $M_{ij}G^{0i0j} = \delta_j^k$, whose existence is guaranteed if the differential equations coming from (2) are hyperbolic (see appendix A), and using (7) to express $\partial_0 A_i$ in terms of the canonical momenta π^i , we find the total Hamiltonian density

$$\begin{aligned}\mathcal{H} &= -\frac{1}{2}M_{js}\pi^j\pi^s - A_0\partial_j\pi^j - \pi^j M_{ij}G^{0irk}\partial_r A_k \\ &\quad + \frac{1}{2}G^{ijk r}\partial_i A_j\partial_k A_r - \frac{1}{2}M_{ir}G^{0ijk}G^{0rmn}\partial_m A_n\partial_j A_k + u_1(x)\pi^0(x).\end{aligned}\tag{8}$$

Following the Dirac-Bergmann algorithm [32] for obtaining the Hamiltonian formulation of systems with constraints, we now compute the commutator $\{\pi^0, \mathcal{H}\}$. If this commutator is not zero, we need to impose $\{\pi^0, \mathcal{H}\} \approx 0$ as a secondary constraint, in order to ensure that the primary constraint $\phi_1 \approx 0$ is preserved under time evolution. Indeed, one obtains $\{\pi^0, \mathcal{H}\} = -\partial_j\pi^j$. Thus we impose $\phi_2 = \partial_j\pi^j \approx 0$ as a secondary constraint, which must be added to (8) with a corresponding Lagrange multiplier. The total Hamiltonian now reads

$$\mathcal{H} = \mathcal{H}_0 + u_1(x)\pi^0(x) + (u_2(x) - A_0)\partial_j\pi^j,\tag{9}$$

with

$$\mathcal{H}_0 = -\frac{1}{2}M_{js}\pi^j\pi^s - \pi^j M_{ij}G^{0irk}\partial_r A_k + \frac{1}{2}(G^{m n j k} - M_{ir}G^{0ijk}G^{0rmn})\partial_m A_n\partial_j A_k.\tag{10}$$

Now we find $\{\phi_2, \mathcal{H}\} = 0$, so that the Dirac-Bergmann algorithm ends here and $\phi_1 \approx 0$ and $\phi_2 \approx 0$ exhaust the constraints. However, $\{\phi_1(t, \vec{x}), \phi_2(t, \vec{y})\} = 0$, so that ϕ_1 and ϕ_2 are first class constraints, implying that the multipliers $u_1(x)$ and $u_2(x)$ are completely undetermined. The

infinitesimal gauge transformations induced by (ϕ_1, ϕ_2) on the canonical variables (A_α, π^α) are

$$\delta A_\alpha(t, \vec{x}) = \int d^3y \epsilon^I(t, \vec{y}) \{A_\alpha(t, \vec{x}), \phi_I(t, \vec{y})\} = \epsilon^1(t, \vec{x}) \delta_\alpha^0 - \delta_\alpha^i \partial_i \epsilon^2(t, \vec{x}), \quad (11)$$

$$\delta \pi^\alpha(t, \vec{x}) = \int d^3y \epsilon^I(t, \vec{y}) \{\pi^\alpha(t, \vec{x}), \phi_I(t, \vec{y})\} = 0, \quad (12)$$

with $I = 1, 2$ and $\epsilon^I(t, \vec{x})$ being the infinitesimal parameters of the transformations. Knowledge of these generators of gauge transformations allows us to identify classical observables of the theory as those functionals that are invariant under gauge transformations. Equivalently, observables commute with the constraints $\{O, \phi_I\} \approx 0$. In the present case, it can be checked that the electromagnetic inductions

$$D^a = -G^{0a0b} F_{0b} - \frac{1}{2} G^{0abk} F_{bk} \quad (13)$$

$$= -G^{0a0b} \partial_0 A_b - G^{0ab0} \partial_j A_0 - G^{0abk} \partial_b A_k,$$

$$H_a = -\frac{1}{2} \epsilon_{0abc} \left[G^{bcm0} F_{m0} + \frac{1}{2} G^{bcmn} F_{mn} \right] \quad (14)$$

$$= -\frac{1}{2} \epsilon_{0abc} \left[G^{bcm0} (\partial_m A_0 - \partial_0 A_m) + G^{bcmn} \partial_m A_n \right],$$

defined with respect to the chosen foliation of spacetime into spacelike hypersurfaces, indeed commute with the constraints, so that they can be used as observables. Thus we are finally able to write the Hamiltonian (10) for our system in terms of gauge-invariant observables D^a and H_a as

$$\mathcal{H}_0 = \frac{1}{2} U_{al} D^a D^l + \frac{1}{2} V^{al} H_a H_l, \quad (15)$$

where the matrices U and V are given as

$$U_{al} = -M_{al} + \frac{1}{8} T_{pqjk} G^{0sjk} G^{0tpq} M_{s(l} M_{a)t} \quad (16)$$

$$V^{al} = -\frac{1}{8} \epsilon^{0jk(a} \epsilon^{l|pq} T_{pqjk}, \quad (17)$$

with T_{pqjk} defined such that

$$(G^{pqmn} - G^{0rpq} G^{0amn} M_{ar}) T_{pqtu} = -8 \delta_{[t}^m \delta_{u]}^n. \quad (18)$$

The existence of T is guaranteed due to the invertibility properties of area metrics; indeed it can be written explicitly in terms of the block matrices constituting the area metric tensor, see again appendix A.

4. GAUGE FIXING AND DIRAC BRACKETS

In order to determine the Dirac brackets associated with our system, one needs to remove the indeterminacy in the Lagrange multipliers by fixing a gauge. This is achieved here by manually

imposing two further constraints $\phi_3 \approx 0, \phi_4 \approx 0$ such that $\det\{\phi_I(\vec{x}), \phi_J(\vec{y})\} \neq 0$, with $I, J = 1, \dots, 4$, so that the new set of constraints ϕ_I is now of second class. In our case, the Euler-Lagrange equations for the gauge field A obtained from the action (2) are given by

$$G^{\alpha\beta\gamma\delta} \partial_\beta \partial_\delta A_\gamma = 0, \quad (19)$$

which is conveniently split into one temporal equation

$$G^{0a0b} \partial_a \partial_b A_0 + \left[G^{0abc} \partial_a \partial_c - G^{0a0b} \partial_0 \partial_a \right] A_b = 0 \quad (20)$$

and three spatial equations

$$\left[G^{0bla} \partial_a \partial_b - G^{0l0m} \partial_0 \partial_m \right] A_0 + \left[G^{0l0m} \partial_0^2 - 2G^{0(lm)a} \partial_0 \partial_a + G^{lamd} \partial_a \partial_d \right] A_m = 0. \quad (21)$$

As the third constraint we impose the Glauber gauge

$$\phi_3 = A_0(\vec{x}) - \int d^3x' G(\vec{x}, \vec{x}') G^{0abc} \partial'_a \partial'_c A_b(\vec{x}') \approx 0 \quad (22)$$

with $-G^{0a0b} \partial_a \partial_b G(\vec{x}, \vec{x}') = \delta(\vec{x} - \vec{x}')$, or more explicitly,

$$G(\vec{x}, \vec{x}') = -\frac{1}{4\pi \sqrt{-M_{ab}(x^a - x'^a)(x^b - x'^b)}}. \quad (23)$$

The expression under the square root is non-negative ultimately due to the energy distinguishing property (see appendix A). Consistency of the gauge (22) with the temporal equation (20) requires that the last constraint

$$\phi_4 = G^{0a0b} \partial_a A_b \approx 0. \quad (24)$$

In summary, our constraints ϕ_I are given by

$$\begin{aligned} \phi_1 = \pi^0 \approx 0, & & \phi_3 = A_0(\vec{x}) - \int d^3x' G(\vec{x}, \vec{x}') G^{0abc} \partial'_a \partial'_c A_b(\vec{x}') \approx 0, \\ \phi_2 = \partial_a \pi^a \approx 0, & & \phi_4 = G^{0a0b} \partial_a A_b \approx 0, \end{aligned} \quad (25)$$

and satisfy

$$\{\phi_I(t, \vec{x}), \phi_J(t, \vec{y})\} = \int \frac{d^3k}{(2\pi^3)} \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -G^{0a0b} k_a k_b \\ 1 & 0 & 0 & 0 \\ 0 & G^{0a0b} k_a k_b & 0 & 0 \end{bmatrix} e^{i\vec{k} \cdot (\vec{x} - \vec{y})}. \quad (26)$$

The matrix above $\{\phi_I(t, \vec{x}), \phi_J(t, \vec{y})\}$ is invertible, so that the constraints ϕ_I are now of second class and the gauge freedom is gone. Its inverse $(\{\phi(\vec{x}), \phi(\vec{y})\}^{-1})^{IJ}$, defined through

$$\int d^3y \{\phi_I(\vec{x}), \phi_J(\vec{y})\} (\{\phi(\vec{y}), \phi(\vec{z})\}^{-1})^{JM} = \delta_I^M \delta(\vec{x} - \vec{z}), \quad (27)$$

is simply given as

$$\{\phi_I(t, \vec{x}), \phi_J(t, \vec{y})\}^{-1} = \int \frac{d^3 k}{(2\pi^3)} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{1}{G^{0a0b} k_a k_b} \\ -1 & 0 & 0 & 0 \\ 0 & -\frac{1}{G^{0a0b} k_a k_b} & 0 & 0 \end{bmatrix} e^{i\vec{k} \cdot (\vec{x} - \vec{y})}. \quad (28)$$

Equipped with equation (28) we can now follow Dirac's procedure and replace the standard Poisson bracket $\{, \}$ by the Dirac bracket $\{, \}_D$, which is defined as

$$\{A(\vec{x}), B(\vec{y})\}_D = \{A(\vec{x}), B(\vec{y})\} - \int d^3 z d^3 w \{A(\vec{x}), \phi_I(\vec{z})\} (\{\phi(\vec{z}), \phi(\vec{w})\}^{-1})^{IJ} \{\phi_J(\vec{w}), B(\vec{y})\}. \quad (29)$$

Thus we arrive at the fundamental Dirac brackets of our system, with respect to which the theory must be quantized

$$\begin{aligned} \{A_\alpha(t, \vec{x}), \pi^\beta(t, \vec{y})\}_D &= \int \frac{d^3 k}{(2\pi)^3} \left[\delta_\alpha^\beta - \delta_\alpha^0 \delta_0^\beta - \frac{\delta_\alpha^m \delta_n^\beta k_a k_m G^{0a0n}}{G^{0p0q} k_p k_q} - \frac{\delta_\alpha^0 \delta_b^\beta G^{0abc} k_a k_c}{G^{0p0q} k_p k_q} \right] e^{i\vec{k} \cdot (\vec{x} - \vec{y})}, \\ \{A_\alpha(t, \vec{x}), A_\beta(t, \vec{y})\}_D &= 0, \\ \{\pi^\alpha(t, \vec{x}), \pi^\beta(t, \vec{y})\}_D &= 0, \end{aligned} \quad (30)$$

and the dynamics of the system is simply generated by the Hamilton equations

$$\begin{aligned} \partial_t A_\alpha(t, \vec{x}) &\approx \int d^3 y \{A_\alpha(t, \vec{x}), \mathcal{H}_0(\vec{y})\}_D, \\ \partial_t \pi^\alpha(t, \vec{x}) &\approx \int d^3 y \{\pi^\alpha(t, \vec{x}), \mathcal{H}_0(\vec{y})\}_D, \end{aligned} \quad (31)$$

where, due to the use of Dirac brackets, only \mathcal{H}_0 is involved.

5. BIHYPERBOLIC AREA METRICS CLOSE TO LORENTZIAN METRICS

The preceding Hamiltonian analysis and calculation of Dirac brackets made only implicit use of the requirement that the area metric background be bi-hyperbolic and energy-distinguishing, namely in the abstract constructions underlying the definition of spacetime foliations into spacelike leaves. But now we need to explicitly solve the field equations (21) with the gauge imposed by (22), and this requires to restrict attention to concrete bi-hyperbolic and energy-distinguishing area metric backgrounds. Moreover, for actual calculations it is most convenient to choose a coordinate frame in which the area metric takes a simple normal form. The normal form theory of area metrics in four dimensions has been developed in [8], and used in [3] to show that the area metric cannot

be bi-hyperbolic unless the endomorphism J on the space of two-forms defined through

$$J_{\gamma\delta}{}^{\alpha\beta} = G^{\gamma\delta\mu\nu}\omega_{\mu\nu\alpha\beta} \quad (32)$$

has a complex eigenvalue structure (Segré type) of the form $[1\bar{1}1\bar{1}1\bar{1}]$, $[2\bar{2}1\bar{1}]$, $[3\bar{3}]$, $[1\bar{1}1\bar{1}11]$, $[2\bar{2}11]$, $[1\bar{1}1111]$ or $[111111]$. However, four-dimensional area metrics that are induced by a Lorentzian metric automatically lie in the first class, $[1\bar{1}1\bar{1}1\bar{1}]$, and moreover the continuous dependence of the eigenvalues of an endomorphism on the components of a representing matrix implies that any area metric in the neighborhood of such a metric-induced area metric is equally of class $[1\bar{1}1\bar{1}1\bar{1}]$. Thus area metrics of immediate phenomenological relevance are clearly those of this first class, and it can be shown that by $GL(4)$ frame transformations these can always be brought to the form

$$G^{[ab][cd]} = \begin{bmatrix} -\alpha & 0 & 0 & \rho & 0 & 0 \\ 0 & -\beta & 0 & 0 & \sigma & 0 \\ 0 & 0 & -\gamma & 0 & 0 & \tau \\ \rho & 0 & 0 & \alpha & 0 & 0 \\ 0 & \sigma & 0 & 0 & \beta & 0 \\ 0 & 0 & \tau & 0 & 0 & \gamma \end{bmatrix} \quad \begin{array}{l} \text{for real } \rho, \sigma, \tau \text{ and} \\ \text{real positive } \alpha, \beta, \gamma, \end{array} \quad (33)$$

where for notational purposes, G is considered here as a bilinear form on the space of two-forms for convenience, and the representing matrix shown above is with respect to the obvious induced basis in the order $[01], [02], [03], [23], [31], [12]$. The positivity of α, β, γ follows from our convention that P is positive on the hyperbolicity cone C . If and only if the area metric is induced by a Lorentzian metric, do the real scalars assume the values $\alpha = \beta = \gamma = 1$ and $\rho = \sigma = \tau = 0$. So any finite (but not too large) deviation from a Lorentzian metric is encoded in these scalars and the frame that brings about this normal form.

It is straightforward to show that if one chooses $\rho = \sigma = \tau$, the polynomial

$$\begin{aligned} P(q) &= \alpha\beta\gamma(q_0^4 + q_1^4 + q_2^4 + q_3^4) \\ &\quad + \alpha(\beta^2 + \gamma^2)(q_2^2q_3^2 - q_0^2q_1^2) \\ &\quad + \beta(\alpha^2 + \gamma^2)(q_1^2q_3^2 - q_0^2q_2^2) \\ &\quad + \gamma(\alpha^2 + \beta^2)(q_1^2q_2^2 - q_0^2q_3^2) \end{aligned} \quad (34)$$

associated with an area metric of this class is hyperbolic with respect to $h = L^{-1}(\partial/\partial t)$. This is most efficiently verified in the normal frame by observing that for $h = (1, 0, 0, 0)$, the real symmetric Hankel matrix $H_1(P_{q,h})$ associated with the polynomial $P_{q,h}$ is positive definite for any covector q ,

which implies that P is hyperbolic [33, 34]. The dual polynomial $P^\#$ takes precisely the same shape in the normal form frame employed here, and thus is also seen to be hyperbolic. Finally also the energy-distinguishing property is easily checked. Finally, note that for area metrics with polynomial (34), we have

$$G^{0abc} = \rho \epsilon^{0abc} \quad (35)$$

in this normal form frame, which significantly simplifies the field equations (20) and (21) whose solutions we will now be able to obtain, orthogonalize appropriately, and thus obtain a diagonalization of the Hamiltonian.

It is worth noting that the hyperbolic polynomial (34) only factorizes if at least two of the scalars α, β, γ coincide, so that area metrics with a bi-metric dispersion relation merely present a subset of measure zero within the set of area metrics neighboring Lorentzian metrics. Indeed, for the generic case of mutually different scalars, the polynomial P is irreducible. Thus theories trying to account for birefringence in linear electrodynamics by some sort of bi-metric geometry fail to parametrize almost all relevant geometries near Lorentzian metric ones.

6. DIAGONALIZATION OF THE HAMILTONIAN

In order to diagonalize the Hamiltonian (10) for bi-hyperbolic and energy-distinguishing general linear electrodynamics with a higher-order polynomial dispersion relation given by (34), we first need to find the solutions of the classical field equations (20) and (21). After choosing the Glauber gauge (25), the first equation is trivially satisfied, and the second one reduced to

$$\left[G^{0l0m} \partial_0^2 + G^{lamd} \partial_a \partial_d \right] A_m(t, \vec{x}) = 0, \quad (36)$$

due to (35). Moreover, these field equations are completely equivalent to the field equations arising from (31). Specifically, we look for plane wave solutions

$$A_a(t, \vec{x}) = \int \frac{d^3 p}{(2\pi)^3} e^{-i(\omega t + \vec{p} \cdot \vec{x})} f_a(\vec{p}), \quad (37)$$

so that introducing (37) into (36) we observe that the equation

$$\left[G^{0l0m}(\omega)^2 + G^{lamd} p_a p_d \right] f_m(\vec{p}) = 0 \quad (38)$$

must be satisfied if (37) is indeed a solution. Equation (38) has non-trivial solutions only if

$$\det \left(G^{0l0m}(\omega)^2 + G^{lamd} p_a p_d \right) = 0. \quad (39)$$

The non-zero frequencies ω for which this is the case are precisely the solutions of $P(\omega, \vec{p}) = 0$, compare (A6). From the energy distinguishing condition of an area metric spacetime it follows that this frequencies are non-zero unless $\vec{p} = 0$, and real because of the hyperbolicity of P . It is then further immediate from (34) that if some (without loss of generality positive) $\omega(\vec{p})$ is a solution for some given \vec{p} in our normal frame, then so is $-\omega(\vec{p})$, and that $\omega(\vec{p}) = \omega(-\vec{p})$. Thus we have four non-zero energy solutions $\pm\omega^I(\vec{p})$ labeled by $I = 1, 2$, two positive and two negative ones, for each spatial momentum \vec{p} . Therefore any solution of the field equations for the real gauge potential A can be expanded as

$$A_a(t, \vec{x}) = \sum_{I=1,2} \int_{N^{\text{smooth}}} \frac{d^3p}{(2\pi)^3} \left(e^{-i(\omega^I(\vec{p})t + \vec{p}\cdot\vec{x})} f_a^I(\vec{p}) + e^{i(\omega^I(\vec{p})t + \vec{p}\cdot\vec{x})} f_a^{*I}(\vec{p}) \right), \quad (40)$$

where strictly speaking, the integral is to be taken only over spatial momenta \vec{p} for which the roots ω of $P(\omega, \vec{p})$ are non-degenerate, so that the elementary plane wave solutions are linearly independent. However, the set of covectors for which these zeros are degenerate is of measure zero [3], so that this restriction of the integral domain can be technically disregarded. It may be worth emphasizing that the standard appearance of this expansion is somewhat deceptive, since the ω^I appearing here are solutions of (39), rather than the standard Lorentzian dispersion relation.

Having obtained a basis of solutions of the classical field equations, we now identify an inner product that is preserved under time evolution and positive definite for positive energy solutions. To this end, consider solutions $A_a(\vec{p})(t, \vec{x})$ and $\tilde{A}_a(\vec{q})(t, \vec{x})$ of the field equation for specific spatial covectors \vec{p} and \vec{q} , respectively. Using the field equation (36), it can be shown that the continuity equation

$$\partial_0 \left(G^{0a0b} A_a^*(\vec{p}) \overleftrightarrow{\partial}_0 \tilde{A}_b(\vec{q}) \right) + \partial_m \left(-G^{a(mn)b} A_a^*(\vec{p}) \overleftrightarrow{\partial}_n \tilde{A}_b(\vec{q}) \right) = 0 \quad (41)$$

is satisfied. This implies that we have a conserved charge Q given by

$$Q = \int d^3x G^{0a0b} A_a^*(\vec{p}) \overleftrightarrow{\partial}_0 \tilde{A}_b(\vec{q}). \quad (42)$$

The above defined charge Q can be used to define a scalar product in the space of solutions, which then by definition is conserved under time evolution and is defined as $(A(\vec{p}), \tilde{A}(\vec{q})) = -iQ$. It satisfies the following properties

$$\begin{aligned} (A(\vec{p}), \lambda \tilde{A}(\vec{q})) &= \lambda (A(\vec{p}), \tilde{A}(\vec{q})) \\ (\lambda A(\vec{p}), \tilde{A}(\vec{q})) &= \lambda^* (A(\vec{p}), \tilde{A}(\vec{q})) \\ (A(\vec{p}), \tilde{A}(\vec{q})) &= (A(\vec{p}), \tilde{A}(\vec{q}))^* = -(A^*(\vec{p}), \tilde{A}^*(\vec{q})). \end{aligned} \quad (43)$$

Hence, if we define for our different frequency solutions

$$F_a^I(\vec{p})(t, \vec{x}) = e^{-i(\omega^I(\vec{p})t + \vec{p} \cdot \vec{x})} f_a^I(\vec{p}), \quad (44)$$

we find that $(F^I(\vec{p}), F^{*J}(\vec{q})) = 0$ and

$$(F^I(\vec{p}), F^J(\vec{q})) = -(F^{*I}(\vec{p}), F^{*J}(\vec{q})) = -2\omega^I(\vec{p})G^{0a0b} f_a^{I*}(\vec{p}) f_a^I(\vec{p}) \delta^{IJ} \delta(\vec{p} - \vec{q}). \quad (45)$$

In the derivation of the above results we used charge conservation to find that for $I \neq J$

$$G^{0a0b} f_a^{I*}(\vec{p}) f_b^{J*}(-\vec{p}) = G^{0a0b} f_a^{I*}(\vec{p}) f_b^J(\vec{p}) = 0. \quad (46)$$

Moreover, since G^{0a0b} is negative definite due to (33), equation (45) shows that the positive energy solutions can be positively normalized, implying in turn that the negative energy solutions are negatively normalized. This indefiniteness of the scalar product is responsible for creation and annihilation processes. Choosing, without loss of generality, $f_a^I(\vec{p}) = \frac{a_a^I(\vec{p})}{\sqrt{2\omega^I(\vec{p})}}$, we finally have

$$(F^I(\vec{p}), F^J(\vec{q})) = -(F^{*I}(\vec{p}), F^{*J}(\vec{q})) = -G^{0a0b} a_a^{I*}(\vec{p}) a_a^I(\vec{p}) \delta^{IJ} \delta(\vec{p} - \vec{q}), \quad (47)$$

and our general solution reads

$$A_a(t, \vec{x}) = \sum_{I=1,2} \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega^I(\vec{p})}} \left(e^{-i(\omega^I(\vec{p})t + \vec{p} \cdot \vec{x})} a_a^I(\vec{p}) + e^{i(\omega^I(\vec{p})t + \vec{p} \cdot \vec{x})} a_a^{*I}(\vec{p}) \right). \quad (48)$$

Now that we have the general solution (48), we can use it to write the Hamiltonian evaluated at a solution in diagonal form,

$$\begin{aligned} H_0 &= \int d^3x \mathcal{H}_0(\vec{x}) = -\frac{1}{2} \int d^3x G^{0a0b} (\partial_0 A_a \partial_0 A_b - A_a \partial_0^2 A_b) \\ &= \frac{1}{2} \sum_{I,J} \int \frac{d^3p}{(2\pi)^3} \frac{d^3q}{(2\pi)^3} \omega^J(\vec{q}) [(F^I(\vec{p}), F^J(\vec{q})) + (F^J(\vec{p}), F^I(\vec{q}))] \\ &= -\frac{1}{2} \sum_{I=1,2} \int \frac{d^3p}{(2\pi)^3} \omega^I(\vec{p}) G^{0a0b} [a_a^{*I}(\vec{p}) a_b^I(\vec{p}) + a_a^I(\vec{p}) a_b^{*I}(\vec{p})]. \end{aligned} \quad (49)$$

The last expression shows that the classical Hamiltonian is positive because G^{0a0b} is negative definite.

7. QUANTIZATION

Equipped with the results developed so far, we are now ready to quantize the electromagnetic field. First, notice that if we multiply equation (38) by p_l then the amplitude eigenvectors $a_b^I(\vec{p})$ satisfy

$$G^{0a0b} p_a a_b^I(\vec{p}) = 0, \quad (50)$$

such that the constraints $G^{0a0b}\partial_a A_b \approx 0$ and $\partial_a \pi^a \approx 0$ are satisfied. Now it can be shown [3] that for almost all spatial momenta \vec{p} , the two associated positive energies do not coincide, $\omega^{I=1}(\vec{p}) \neq \omega^{I=2}(\vec{p})$, so that the covectors $a_b^{I=1}(\vec{p})$ and $a_b^{I=2}(\vec{p})$ are determined up to scale, linearly independent and thus forming a basis for the space of all purely spatial covectors v for which $G^{0a0b}p_a v_b = 0$, and diagonalize the Hamiltonian (49). Thus the only one freedom we have is a choice of normalization, which we choose such that any solution $a_b^I(\vec{p})$ is expressed as $a_b^I(\vec{p}) = a^I(\vec{p})\epsilon_b^I(\vec{p})$ with the covectors $\epsilon_b^I(\vec{p})$ normalized with respect to our scalar product, i.e.,

$$-G^{0a0b}\epsilon_a^{I*}(\vec{p})\epsilon_b^I(\vec{p}) = 1, \quad (51)$$

where there is no summation over I . Furthermore, p_a and any $a_b^I(\vec{p})$ are clearly linearly independent, such that the set of covectors

$$\left\{ \epsilon_b^{I=1}(\vec{p}), \epsilon_b^{I=2}(\vec{p}), \frac{\vec{p}}{\sqrt{-G^{0a0b}p_a p_b}} \right\} \quad (52)$$

constitute a basis for V , which is orthonormalized with respect to the scalar product (43). Hence, they satisfy the completeness relation

$$-G^{0i0j} \sum_{I=1,2} \epsilon_j^{I*}(\vec{p})\epsilon_b^I(\vec{p}) = \delta_b^i - \frac{p_m p_b G^{0m0i}}{G^{0r0s} p_r p_s}. \quad (53)$$

Notice that the normalized covectors $\epsilon_b^I(\vec{p})$ satisfy the orthogonality identities (46). Now the general solution (48) takes the form

$$A_a(t, \vec{x}) = \sum_{I=1,2} \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\sqrt{2\omega^I(\vec{p})}} \left(e^{-i(\omega^I(\vec{p})t + \vec{p}\cdot\vec{x})} a^I(\vec{p})\epsilon_a^I(\vec{p}) + e^{i(\omega^I(\vec{p})t + \vec{p}\cdot\vec{x})} a^{I*}(\vec{p})\epsilon_a^{*I}(\vec{p}) \right), \quad (54)$$

where the coefficients $a^I(\vec{p})$ correspond to the amplitudes of the solutions and depend on the initial values that one considers for a specific problem in the classical approach. At the quantum level, these amplitudes are precisely the mathematical objects that should be promoted to operators, such that the corresponding quantum field reads

$$\hat{A}_a(t, \vec{x}) = \sum_{I=1,2} \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\sqrt{2\omega^I(\vec{p})}} \left(e^{-i(\omega^I(\vec{p})t + \vec{p}\cdot\vec{x})} \hat{a}^I(\vec{p})\epsilon_a^I(\vec{p}) + e^{i(\omega^I(\vec{p})t + \vec{p}\cdot\vec{x})} \hat{a}^{I\dagger}(\vec{p})\epsilon_a^{*I}(\vec{p}) \right). \quad (55)$$

Using this quantum solution and the expressions for the energy and spatial momentum (which can be obtained by calculating the energy-momentum tensor) we find that the quantum Hamiltonian and quantum spatial momentum operators are given by

$$\hat{H}_0 = \frac{1}{2} \sum_{I=1,2} \int \frac{d^3 p}{(2\pi)^3} \omega^I(\vec{p}) \left[\hat{a}^I(\vec{p})\hat{a}^{I\dagger}(\vec{p}) + \hat{a}^{I\dagger}(\vec{p})\hat{a}^I(\vec{p}) \right], \quad (56)$$

$$\hat{P}_i = \frac{1}{2} \sum_{I=1,2} \int \frac{d^3 p}{(2\pi)^3} p_i \left[\hat{a}^I(\vec{p})\hat{a}^{I\dagger}(\vec{p}) + \hat{a}^{I\dagger}(\vec{p})\hat{a}^I(\vec{p}) \right]. \quad (57)$$

Hence, if we identify the operators $\hat{a}^I(\vec{p})$, $\hat{a}^{I\dagger}(\vec{p})$ with annihilation and creation operators respectively, a condition for the Hamiltonian to be positive definite is that these operators obey the bosonic commutation relations

$$\begin{aligned} [\hat{a}^I(\vec{p}), \hat{a}^{J\dagger}(\vec{q})] &= (2\pi)^3 \delta^{IJ} \delta(\vec{p} - \vec{q}), \\ [\hat{a}^I(\vec{p}), \hat{a}^J(\vec{q})] &= [\hat{a}^{I\dagger}(\vec{p}), \hat{a}^{J\dagger}(\vec{q})] = 0. \end{aligned} \quad (58)$$

Hence, the quantum Hamiltonian operator can be written as

$$\hat{H}_0 = \sum_{I=1,2} \int \frac{d^3p}{(2\pi)^3} \omega^I(\vec{p}) \hat{a}^{I\dagger}(\vec{p}) \hat{a}^I(\vec{p}) + \sum_{I=1,2} \frac{1}{2} \int d^3p \omega^I(\vec{p}) \delta(0), \quad (59)$$

from which expression we identify the energy of the electromagnetic vacuum, which was calculated here for plane wave solutions without any boundary conditions, as

$$E_{\text{vac}}(\text{no boundaries}) = \sum_{I=1,2} \frac{1}{2} \int d^3p \omega^I(\vec{p}) \delta(0). \quad (60)$$

In the next section, we will calculate how this expression changes if one imposes boundary conditions. Finally, by using the completeness relation (53) one confirms that

$$\left[\hat{A}_a(t, \vec{x}), \hat{\pi}^b(t, \vec{y}) \right] = i \int \frac{d^3p}{(2\pi)^3} \left[\delta_a^b - \frac{p_m p_a G^{0m0b}}{G^{0r0s} p_r p_s} \right] e^{i\vec{p} \cdot (\vec{x} - \vec{y})}, \quad (61)$$

which shows the consistency of the quantization procedure with the Dirac brackets (30), since the latter reduce to the above form due to (35).

8. APPLICATION: CASIMIR EFFECT IN A BIREFRINGENT LINEAR OPTICAL MEDIUM

The Hamiltonian (59) shows that the quantization of general linear electrodynamics leads to a modified quantum vacuum compared to standard non-birefringent Maxwell theory. In fact, local physical phenomena which only depend on the quantum vacuum can be used to test and bound the non-metricity of spacetime. In this section we analyze one such phenomenon, namely the Casimir effect; similar studies can be conducted for the Unruh effect and spontaneous emission.

The Casimir effect [35] arises because of the energy cost incurred by imposing boundary conditions on the electromagnetic field strength. Physically, such boundary conditions arise for instance by introducing perfectly conducting metal plates into the spacetime. For two infinitely extended plates parallel to the 1-2-plane, and this is the configuration we will study here for general linear electrodynamics, the electromagnetic field strength must satisfy the boundary conditions

$$F_{01}|_{\text{plates}} = F_{02}|_{\text{plates}} = F_{12}|_{\text{plates}} = 0 \quad (62)$$

everywhere on either plate; this follows, by Stokes' theorem and thus independent of the geometric background, from the physical assumption that the plates are ideal conductors inside of which the field strength must vanish.

Now the key point is that having, or not having, boundary conditions for the vacuum amounts to an energy difference, the so-called Casimir energy

$$E_{\text{Casimir}} = E_{\text{vac}}(\text{plate boundaries}) - E_{\text{vac}}(\text{no boundaries}). \quad (63)$$

But both energies on the right hand side diverge and need to be regularized such that their difference is independent of the regulator. This is most easily achieved by first considering boundary conditions analogous to (62), but for all six faces of a finite rectangular box with faces parallel to the coordinate planes, and separated by coordinate distances L_1, L_2, L_3 . In a second step we will then push all faces a very large coordinate distance L apart in order to obtain an expression for $E_{\text{vac}}(\text{no boundaries})$ regularized by L , and similarly push all but two faces in order to obtain a corresponding regularized expression for $E_{\text{vac}}(\text{plate boundaries})$. The difference of these two regulated quantities will indeed turn out to be finite per unit area and be independent of the regulator L .

Now more precisely, a basis of solutions of general linear electrodynamics satisfying the box boundary conditions is labeled by a triple (n_1, n_2, n_3) of non-negative integers and a polarization $I = 1, 2$ and takes the form

$$\begin{aligned} A_x(\vec{x}) &= a_x^I(n_1, n_2, n_3) \cos(n_1\pi \frac{x}{L_1}) \sin(n_2\pi \frac{y}{L_2}) \sin(n_3\pi \frac{z}{L_3}), \\ A_y(\vec{x}) &= a_y^I(n_1, n_2, n_3) \sin(n_1\pi \frac{x}{L_1}) \cos(n_2\pi \frac{y}{L_2}) \sin(n_3\pi \frac{z}{L_3}), \\ A_z(\vec{x}) &= a_z^I(n_1, n_2, n_3) \sin(n_1\pi \frac{x}{L_1}) \sin(n_2\pi \frac{y}{L_2}) \cos(n_3\pi \frac{z}{L_3}), \end{aligned} \quad (64)$$

where the $a_m^I(n_1, n_2, n_3)$ are solutions to equation (38) for $\omega^I(n_1\pi/L_1, n_2\pi/L_2, n_3\pi/L_3)$, which always exist if the dispersion relation is bi-hyperbolic and energy distinguishing. The vacuum energy in the presence of the box boundary conditions is thus given by the discrete sum

$$E_{\text{vac}}(\text{box boundaries}) = \frac{1}{2} \sum_{\vec{n}=0}^{\infty} \sum_{I=1,2} \omega^I(\pi \frac{n_1}{L_1}, \pi \frac{n_2}{L_2}, \pi \frac{n_3}{L_3}). \quad (65)$$

Removing appropriate faces to a coordinate distance L one finds from this, in the very large L limit, the L -regularized expression for the vacuum energy without boundary conditions

$$E_{\text{vac}}^L(\text{no boundaries}) = \frac{L^3}{2\pi^3} \sum_{I=1,2} \int_0^{\infty} d^3p \omega^I(\vec{p}), \quad (66)$$

and the L -regularized expression for the vacuum energy in the presence of two plates parallel to the 1-2-plane and separated by a coordinate distance d

$$E_{\text{vac}}^L(\text{plate boundaries}) = \frac{L^2}{2\pi^2} \sum_{I=1,2} \sum_{n'} \int_0^\infty dp_x dp_y \omega^I \left(p_x^2, p_y^2, \left(\frac{n\pi}{d}\right)^2 \right), \quad (67)$$

where the prime in the summation symbol n means that a factor $1/2$ should be inserted if this integer is zero, for then we have just one independent polarization. Hence we find for the physical vacuum Casimir energy $U(d) = (E_{\text{vac}}(\text{plate boundaries}) - E_{\text{vac}}(\text{no boundaries}))/L^2$ per unit area

$$U(d) = \frac{1}{2\pi^2} \sum_{I=1,2} \left[\sum_{n'} \int_0^\infty dp_x dp_y \omega^I \left(p_x^2, p_y^2, \left(\frac{n\pi}{d}\right)^2 \right) - \frac{d}{\pi} \int_0^\infty dp_x dp_y dp_z \omega^I(p_x^2, p_y^2, p_z^2) \right]. \quad (68)$$

In principle, the execution of the above integrals can proceed as in the standard case. However, with the frequencies ω^I now being solutions to a quartic, rather than quadratic, dispersion relation, these integrals are much harder particularly due to the absence of rotational invariance. Fortunately, the fact that contributions from the two different polarizations $I = 1, 2$ are simply added in the above expression allows for an analytic study of the case where the polynomial P is reducible. In terms of the scalars $\alpha, \beta, \gamma, \rho$ defining the area metric in a normal form frame, this is the case if and only if two of the scalars α, β, γ coincide, and we may take $\alpha = \beta$, for instance. Even in this simplest of non-trivial cases, the Casimir energy crucially depends on the birefringence properties of the underlying general linear electrodynamics. More precisely, the polynomial in (34) factorizes into two Lorentzian metrics,

$$P(p) = \alpha(\alpha p_0^2 - \alpha p_3^2 - \gamma(p_1^2 + p_2^2))(\gamma p_0^2 - \gamma p_3^2 - \alpha(p_1^2 + p_2^2)), \quad (69)$$

so that we immediately obtain the positive energy solutions

$$\omega^{I=1} = \left[\frac{1}{\alpha} (\alpha p_3^2 + \gamma p_2^2 + \gamma p_1^2) \right]^{1/2} \quad \text{and} \quad \omega^{I=2} = \left[\frac{1}{\gamma} (\gamma p_3^2 + \alpha p_1^2 + \alpha p_2^2) \right]^{1/2}, \quad (70)$$

turning (68) into a sum of integrals as they appear in the standard Casimir problem on a Lorentzian background. Thus from here on the standard calculation of the Casimir effect [36] can be followed for each of these integrals separately, and one finally obtains the Casimir energy (68)

$$U(d) = -\frac{1}{2} \left(\frac{\alpha}{\gamma} + \frac{\gamma}{\alpha} \right) \frac{\pi^2}{720d^3}. \quad (71)$$

This energy difference of course results in a Casimir force

$$F(d) = -U'(d) = -\frac{1}{2} \left(\frac{\alpha}{\gamma} + \frac{\gamma}{\alpha} \right) \frac{\pi^2}{240d^4} \quad (72)$$

between the plates. The standard Casimir force is recovered if and only if $\alpha = \beta = \gamma$, and irrespective of the value of the scalar ρ . This in turn is equivalent to the absence of classical bi-refringence [37]. Note that the amplification of any bi-refringence is limited only by the technological constraint of how small the separation d between the plates can be made in any realistic set-up. In contrast to classical bi-refringence tests, which usually require accumulative effects over large distances (with all the uncertainties present in such non-local measurements), one sees here that the Casimir force allows for a detection of bi-refringence by way of a highly local measurement. Conversely, of course, experimental measurements of the Casimir force agreeing with the standard prediction within the given technological constraints can be used to put stringent bounds on the non-metricity of the spacetime region where the measurement is conducted.

9. CONCLUSIONS

The canonical quantization of general linear electrodynamics, as undertaken in this article, required the solution of several, and in themselves challenging, questions.

First, from the classical field theory point of view, it had to be clarified which general linear electrodynamics are predictive on the one hand and physically interpretable in terms of quantities measurable by observers on the other hand. The answer to both questions is encoded in the polynomial dispersion relation of the field theory, and amounts to the simple algebraic conditions that the latter be bi-hyperbolic and energy-distinguishing. Further down the road, these conditions turned out to be crucial in ensuring the existence of a Glauber gauge, which allowed to define a time-conserved scalar product in the space of classical solutions, on which all further developments were based.

Second, and closely related, is the construction of a Hamiltonian formulation of general linear electrodynamics. The causal structure encoded in the higher-order polynomial dispersion relation of this theory required a revision of the construction of suitable spacetime foliations that underlie a Hamiltonian formulation. The key point here was that the leaves of the foliation must be such that initial data provided on them must be causally evolved by the field equations and at the same time be accessible to observers. It turned out that bi-hyperbolic and energy-distinguishing area metric manifolds provide precisely the structure to ensure both, and ultimately render the classical Hamiltonian positive.

Third, the quantum Hamiltonian operator is positive definite. For a theory with a higher-order polynomial dispersion relation this is far from trivial, and again only due to bi-hyperbolicity and the

energy-distinguishing property. The positive definiteness of the quantum Hamiltonian operator is inherited from the positivity of the classical Hamiltonian because the positive energy solutions have positive norm with respect to the scalar product identified before. This is of course synonymous with the stability of the quantum vacuum, and thus the physical relevance of the Casimir effect we derived from it.

The wider lesson learnt from our study consists in this being a prototypical, and rather non-trivial example for the quantization of a field theory with a modified dispersion relation. Such theories are discussed extensively throughout the literature with a number of motivations, but usually disregarding the fundamental consistency conditions that were instrumental in this work. In particular, the classically inevitable condition that the dispersion relation be given by a bi-hyperbolic and energy distinguishing polynomial proved inevitable also at virtually every step of the quantization process.

Actual calculations were made tractable by employing the fact that the dispersion relation of general linear electrodynamics is ultimately determined by a fourth rank area metric tensor for which a complete algebraic classification and associated normal forms are available for the phenomenologically directly relevant case of four spacetime dimensions. This normal form theory was also used to ensure that the birefringent optical backgrounds for which we calculated the Casimir effect (and which owe their physical relevance to their parametrizing the neighborhood of non-birefringent optical media) indeed are bi-hyperbolic and energy-distinguishing. While it is possible to directly exclude 16 out of a total of 23 algebraic classes of four-dimensional area metrics as admissible spacetime structures, a complete and simple characterization of all area metric manifolds that *are* bi-hyperbolic and energy-distinguishing however remains an open problem. The high interest that would attach to a comprehensive solution of this problem is clearly underlined by the pivotal role we saw these conditions to play for the classical and quantum theory alike.

Another open, albeit well-defined, problem is the coupling of fermions to general linear electrodynamics. The issue is the very definition of spinors in the presence of a higher-degree polynomial dispersion relation, rather than one given by a Lorentzian metric. For rather than satisfying the standard binary Dirac algebra, generalized Dirac matrices that intertwine spacetime and spinor indices must now satisfy a quaternary algebra determined by the fourth-degree polynomial associated with a four-dimensional area metric spacetime structure. Even employing the normal form theory, representations of this quaternary algebra appear hard to find in any other but the case of a reducible dispersion relation satisfying the relevant conditions (which then leads to a sixteen-dimensional spinor representation with an associated refined Dirac equation for this special

bi-metric case). Once a representation in the general case is obtained, the canonical quantization can proceed exactly along the now clearly defined path for such theories, and complete a full theory of general linear quantum electrodynamics including charges.

Concluding, we see that the results of this article open up the arena for comprehensive, and above all conceptually watertight, studies of quantum effects brought about by birefringence. Indeed, beyond the Casimir force we calculated here explicitly, any other effect rooting in the quantum vacuum of electrodynamics can be directly calculated now on the basis of the technical findings of this paper. This includes for instance the Unruh effect or the spontaneous emission of photons from quantized point particles. Once spinor fields are included, the range of effects of course extends to the full spectrum of processes discussed in standard quantum electrodynamics with charged fermions. Far from being merely academic musings, however interesting, these findings are of immediate relevance to physicists with interests ranging from fundamental theory to material science. Indeed, while on the one hand directly testable in birefringent optical media in laboratory experiments [49], the constructions of this paper on the other hand also put phenomenological studies of modified dispersion relations [10–27] (or, equivalently, Lorentz-violating spacetime structures [38–47]), as they now abound in the literature, on a solid theoretical footing.

Appendix A: area metrics and hyperbolicity

An area metric in four dimensions takes the following block matrix Petrov form [8]

$$G^{[\alpha\beta][\gamma\delta]} = \begin{bmatrix} M & K \\ K^T & N \end{bmatrix}, \quad (\text{A1})$$

where the antisymmetric index pairs [01], [02], [03], [23], [31], [12] label, in this order, the basis in which the matrix is given. The matrices M, K, N are 3×3 matrices are related to the area metric through

$$\begin{aligned} M^{ab} &= G^{0a0b}, \\ K_b^a &= \frac{1}{2} \epsilon_{0bmn} G^{0amn}, \\ N_{ab} &= \frac{1}{4} \epsilon_{0amn} \epsilon_{0bpq} G^{mnpq}, \end{aligned} \quad (\text{A2})$$

where ϵ is the totally antisymmetric Levi-Civita symbol. If M in the expression above is invertible, which at the end of this section we will see to be the case if the corresponding area metric leads to well-posed field equations, then $\det(N - K^T M K) \neq 0$. This ensures the existence of the object

T_{pqtu} defined in (18), which can be shown to be explicitly given by

$$T_{abtu} = -2\epsilon_{tuf}\epsilon_{abm} \left((N - K^T M K)^{-1} \right)^{fm}. \quad (\text{A3})$$

The principal symbol of the linear field equations governing the dynamics for the electromagnetic field on area metric backgrounds was found in [8] as the determinant of the 6×6 matrix

$$A^\alpha p_\alpha = \begin{bmatrix} G^{0m0n} p_0 - 2G^{0(mn)a} p_a & -\frac{1}{2}\epsilon_{nef} G^{efma} p_a \\ \epsilon^{mna} p_a & \delta_m^n p_0 \end{bmatrix}. \quad (\text{A4})$$

Using that for any $n \times n$ matrices A, B, C, D

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(AD - BC) \quad \text{if } CD = DC, \quad (\text{A5})$$

we can write the determinant of (A4) as the determinant of a 3×3 matrix as

$$\det(A^\alpha p_\alpha) = \det \left(G^{0m0n} p_0^2 - 2G^{0(mn)a} p_a p_0 + G^{manu} p_a p_u \right) = -p_0^2 P(p), \quad (\text{A6})$$

where P is precisely the polynomial given in equation (3). Using (A6), the polynomial P can now be expressed in terms of the constitutive matrices M, K, N . After calculation one finds

$$P_G(p_0, \vec{p}) = a p_0^4 + b(\vec{p}) p_0^3 + c(\vec{p}) p_0^2 + d(\vec{p}) p_0 + e(\vec{p}), \quad (\text{A7})$$

with coefficients

$$\begin{aligned} a &= -\det(M^{ab}), \\ b(\vec{p}) &= \epsilon_{abc} G_M^{aclm} K_l^b p_m, \\ c(\vec{p}) &= -\left(N_{rf} G_M^{lrnf} + \epsilon_{abc} \epsilon^{tmr} K_t^b K_r^a M^{cl} + 2K_a^e K_e^{[m} M^{a]l} + 2K_a^l K_e^{[a} M^{m]e} \right) p_l p_m, \\ d(\vec{p}) &= -2\epsilon^{ren} \left(2N_{ea} K_r^{[m} M^{a]l} - K_r^m K_a^l K_e^a \right) p_l p_m p_n, \\ e(\vec{p}) &= -N_{bf} \epsilon^{bmr} \epsilon^{fns} \left(\frac{1}{2} M^{uv} N_{rs} - K_r^u K_s^v \right) p_u p_m p_v p_n, \end{aligned} \quad (\text{A8})$$

where $G_M^{ambn} = M^{ab} M^{mn} - M^{an} M^{mb}$. Using this expression, we find that a necessary condition for the well-posedness of the initial value problem is that the matrix M^{ab} is invertible. Moreover, for an observer's frame, in which the coefficient $b(\vec{p}) = 0$ as explained in [3], the matrices K and M satisfy

$$K_f^{[m} M^{n]f} = 0. \quad (\text{A9})$$

In this case, equations (A8) reduce to

$$\begin{aligned}
a &= -\det(M^{ab}), \\
b(\vec{p}) &= 0, \\
c(\vec{p}) &= -\left(N_{rf} G_M^{lrmf} + \epsilon_{abc} \epsilon^{tmr} K_t^b K_r^a M^{cl} + 2K_a^e K_e^{[m} M^{a]l}\right) p_l p_m, \\
d(\vec{p}) &= -2\epsilon^{ren} \left(2N_{ea} K_r^{[m} M^{a]l} - K_r^m K_a^l K_e^a\right) p_l p_m p_n, \\
e(\vec{p}) &= -N_{bf} \epsilon^{bmr} \epsilon^{fns} \left(\frac{1}{2} M^{uv} N_{rs} - K_r^u K_s^v\right) p_u p_m p_v p_n.
\end{aligned} \tag{A10}$$

For area metrics for which there exists a frame such that $G^{0abc} = \rho \epsilon^{0abc}$, such as those considered from section (5) onwards, or equivalently $K_b^a = \phi \delta_b^a$, the polynomial $P_G(p)$ is further reduced to

$$\begin{aligned}
P_G(p_0, \vec{p}) &= -p_0^4 \det(M) - p_0^2 (N_{mn} G_M^{imjn}) p_i p_j - \frac{1}{2} (M^{ij} p_i p_j) (G^{bksl} N_{bs} p_k p_l) \\
&= -p_0^4 \det(M) - p_0^2 (N_{mn} G_M^{imjn}) p_i p_j - \det(N) (M^{ij} p_i p_j) (N^{-1})^{kl} p_k p_l.
\end{aligned} \tag{A11}$$

From the energy distinguishing property of the area metric spacetimes considered here it follows that $P(p_0, \vec{p}) = 0$ does not have any solutions $p_0 = 0$ unless $\vec{p} = 0$. But then the matrix M^{ab} must be of definite signature. For suppose that this is not the case, then one could find $\vec{p} \neq 0$ such that $M^{ij} p_i p_j = 0$. That would imply extra zero solutions for p_0 , in contradiction to the energy distinguishing condition. The same holds for the matrix N . Thus, without loss of generality, we assume that M is negative definite; then using Descartes's rule of signs, hyperbolicity of P implies that N must be positive definite. The opposite definiteness of M and N can be shown to be also sufficient for the hyperbolicity of (A11). This is indeed the case for class I area metrics (33) with $\rho = \sigma = \tau$.

-
- [1] F. Hehl, Y. Obukhov, and I. Obukhov, *Foundations of classical electrodynamics: Charge, flux, and metric*. Birkhauser, 2003.
 - [2] R. Punzi, F. Schuller, and M. Wohlfarth, "Area metric gravity and accelerating cosmology," *Journal of High Energy Physics*, vol. 2007, p. 030, 2007.
 - [3] D. Raetzl, S. Rivera, and F. P. Schuller, "Geometry of physical dispersion relations," 2010. Temporary entry.
 - [4] F. W. Hehl, Y. N. Obukhov, and G. F. Rubilar, "Light propagation in generally covariant electrodynamics and the fresnel equation," *Int. J. Mod. Phys.*, vol. A17, pp. 2695–2700, 2002.
 - [5] F. Schuller and M. Wohlfarth, "Geometry of manifolds with area metric," *Nucl. Phys.*, vol. B747, pp. 398–422, 2006.

- [6] F. Schuller and M. Wohlfarth, “Canonical differential geometry of string backgrounds,” *Journal of High Energy Physics*, vol. 2006, no. 02, p. 059, 2006.
- [7] R. Punzi, F. Schuller, and M. Wohlfarth, “Propagation of light in area metric backgrounds,” *Classical and Quantum Gravity*, vol. 26, no. 3, p. 035024, 2009.
- [8] F. Schuller, C. Witte, and M. Wohlfarth, “Causal structure and algebraic classification of non-dissipative linear optical media,” *Annals of Physics*, 2010.
- [9] I. T. Drummond and S. J. Hathrell, “Qed vacuum polarization in a background gravitational field and its effect on the velocity of photons,” *Phys. Rev. D*, vol. 22, pp. 343–355, Jul 1980.
- [10] M. Rinaldi, “Modified dispersion relations and trans-planckian physics,” *ArXiv e-prints*, Nov. 2007.
- [11] Z. Chang and X. Li, “Lorentz invariance violation and symmetry in randers-finsler spaces,” *Phys. Lett.*, vol. B663, pp. 103–106, 2008.
- [12] C. Lämmerzahl, D. Lorek, and H. Dittus, “Confronting finsler spacetime with experiment,” *General Relativity and Gravitation*, vol. 41, pp. 1345–1353, 2009. 10.1007/s10714-008-0710-9.
- [13] C. Lämmerzahl, A. Macías, and H. Müller, “Lorentz invariance violation and charge (non)conservation: A general theoretical frame for extensions of the maxwell equations,” *Phys. Rev. D*, vol. 71, p. 025007, Jan 2005.
- [14] S. Liberati, T. A. Jacobson, and D. Mattingly, “High energy constraints on lorentz symmetry violations,” 2001.
- [15] V. Perlick, “On the hyperbolicity of Maxwell’s equations with a local constitutive law,” *Arxiv preprint arXiv:1011.2536*, 2010.
- [16] Y. Ling, B. Hu, and X. Li, “Modified dispersion relations and black hole physics,” *Physical Review D*, vol. 73, no. 8, p. 87702, 2006.
- [17] L. Sindoni, “Higgs mechanism in Finsler spacetimes,” *Physical Review D*, vol. 77, no. 12, p. 124009, 2008.
- [18] L. Sindoni, F. Girelli, and S. Liberati, “Emergent gravitational dynamics in Bose-Einstein condensates,” in *To appear in the proceedings of the XXV Max Born Symposium*, vol. 3, 2009.
- [19] L. Sindoni, “A note on particle kinematics in Horava-Lifshitz scenarios,” *Arxiv preprint arXiv:0910.1329*, 2009.
- [20] C. Barcelo, A. Cano, G. Jannes, and L. J. Garay, “Probing effects of modified dispersion relations with bose-einstein condensates,” *PoS*, vol. QG-PH, p. 007, 2007.
- [21] R. Garattini, “Modified dispersion relations and black hole entropy,” *Physics Letters B*, vol. 685, no. 4-5, pp. 329–337, 2010.
- [22] J. Bazo, M. Bustamante, A. Gago, and O. Miranda, “High energy astrophysical neutrino flux and modified dispersion relations,” *International Journal of Modern Physics A*, vol. 24, no. 31, pp. 5819–5829, 2009.
- [23] M. Gregg and S. Major, “On modified dispersion relations and the chandrasekhar mass limit,” *International Journal of Modern Physics D*, vol. 18, no. 6, pp. 971–982, 2009.

- [24] G. W. Gibbons, J. Gomis, and C. N. Pope, “General very special relativity is finsler geometry,” *Phys. Rev.*, vol. D76, p. 081701, 2007.
- [25] J. Lukierski, H. Ruegg, and W. J. Zakrzewski, “Classical and quantum mechanics of free [kappa]-relativistic systems,” *Annals of Physics*, vol. 243, no. 1, pp. 90 – 116, 1995.
- [26] J. Kowalski-Glikman, “Observer-independent quantum of mass,” *Physics Letters A*, vol. 286, no. 6, pp. 391 – 394, 2001.
- [27] F. Girelli, S. Liberati, R. Percacci, and C. Rahmede, “Modified dispersion relations from the renormalization group of gravity,” *Classical and Quantum Gravity*, vol. 24, no. 16, p. 3995, 2007.
- [28] L. Gårding, “Linear hyperbolic partial differential equations with constant coefficients,” *Acta Mathematica*, vol. 85, no. 1, pp. 1–62, 1964.
- [29] M. Atiyah, R. Bott, and L. Gårding, “Lacunae for hyperbolic differential operators with constant coefficients I,” *Acta Mathematica*, vol. 124, no. 1, pp. 109–189, 1970.
- [30] B. Hassett, *Introduction to algebraic geometry*. Cambridge Univ. Press, 2007.
- [31] A. Hanson, T. Regge, C. Teitelboim, and A. dei Lincei, “Constrained Hamiltonian systems,” 1976.
- [32] K. Sundermeyer, “Constrained Dynamics (Lecture Notes in Physics 169),” *Springer-Verlag, Berlin*, 1982.
- [33] F. Gantmacher, “The theory of matrices. Vols. 1, 2,” *Chelsea, New York*, 1959.
- [34] S. Basu, R. Pollack, and M. Roy, *Algorithms in real algebraic geometry*. Springer-Verlag New York Inc, 2006.
- [35] H. Casimir, “On the attraction between two perfectly conducting plates,” in *Proc. Akad. Wet. Amsterdam*, vol. 51, pp. 793–795, 1948.
- [36] P. Milonni and C. Eberlein, *The quantum vacuum: an introduction to quantum electrodynamics*. Academic press Boston, 1994.
- [37] A. Favaro and L. Bergamin, “The non-birefringent limit of all linear, skewonless media and its unique light-cone structure,” *Arxiv preprint arXiv:1008.2343*, 2010.
- [38] T. Jacobson, S. Liberati, and D. Mattingly, “Lorentz violation at high energy: concepts, phenomena, and astrophysical constraints,” *Annals of Physics*, vol. 321, no. 1, pp. 150–196, 2006.
- [39] V. Kostelecký and S. Samuel, “Spontaneous breaking of Lorentz symmetry in string theory,” *Physical Review D*, vol. 39, no. 2, pp. 683–685, 1989.
- [40] R. Gambini and J. Pullin, “Nonstandard optics from quantum spacetime,” *Phys. Rev.*, vol. D59, p. 124021, 1999.
- [41] J. Alfaro, H. A. Morales-Tecotl, and L. F. Urrutia, “Quantum gravity corrections to neutrino propagation,” *Phys. Rev. Lett.*, vol. 84, pp. 2318–2321, 2000.
- [42] D. Sudarsky, L. Urrutia, and H. Vucetich, “New observational bounds to quantum gravity signals,” *Phys.Rev.Lett.*, vol. 89, p. 231301, 2002.
- [43] R. C. Myers and M. Pospelov, “Ultraviolet modifications of dispersion relations in effective field theory,” *Phys. Rev. Lett.*, vol. 90, p. 211601, May 2003.

- [44] J. Magueijo and L. Smolin, “Lorentz invariance with an invariant energy scale,” *Phys. Rev. Lett.*, vol. 88, p. 190403, Apr 2002.
- [45] N. E. Mavromatos, “String quantum gravity, lorentz-invariance violation and gamma-ray astronomy,” 2010. * Temporary entry *.
- [46] M. Bojowald, H. A. Morales-Tecotl, and H. Sahlmann, “On loop quantum gravity phenomenology and the issue of Lorentz invariance,” *Phys.Rev.*, vol. D71, p. 084012, 2005.
- [47] S. Hossenfelder, “Experimental Search for Quantum Gravity,” 2010. * Temporary entry *.
- [48] Born-Infeld theory presents the most prominent example of electrodynamics with a non-linear constitutive relation between the induction and field strength.
- [49] For not truly continuous optical media, such as any material in the laboratory, the calculations made in this paper will only apply to the same approximation to which the medium can be modeled as an area metric spacetime. This will be the case for wavelengths that are well above the average distance between the atoms constituting the material but also well below wavelengths of the order of the separation between the plates. More sophisticated calculations taking into account these cut-offs, as well as finitely extended plates or different geometric configurations should now be feasible for the interested specialist, based on the results derived in this paper.